

Business Longitudinal Analysis Data Environment

Alan Hering
Data Integration Operations
Australian Bureau of Statistics
Informing Australia's important decisions



Outline

- ▶ What is BLADE?
- ▶ How is BLADE being used?
- ▶ How you can get access to BLADE?

- 1) Firstly a quick introduction as to what BLADE is (and how it might be helpful for research projects)
- 2) How is BLADE being used?
- 3) And how you can access BLADE data for your research projects

What is BLADE?

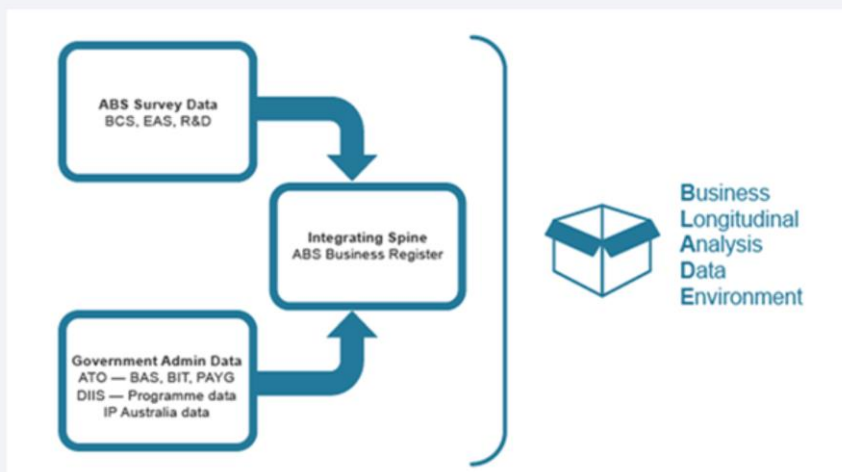
21/02/2020

What is BLADE

Not this guy!!



What is BLADE?

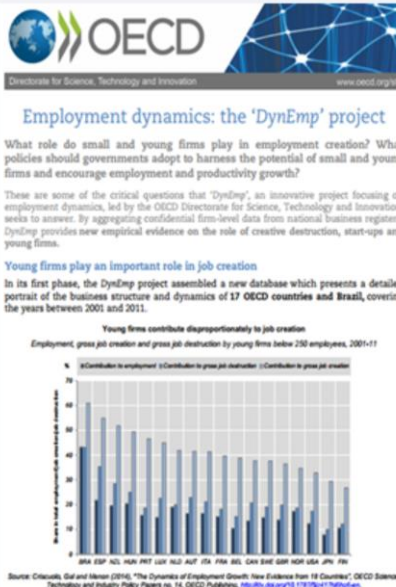


BLADE is a methodology that brings together micro-data from different sources and links them to a longitudinal integrating spine

The linked datasets enables analysis of a wide range of information on businesses over time such as business dynamics, performance and innovation.

The BIRTH of BLADE

- ▶ Joint project between the ABS and Dept. of Industry, Innovation and Science
- ▶ Enabled Australia's participation in the OECD 'DynEmp' project
 - Understanding employer dynamics



21/02/2020

Originally established as a joint project between the ABS and Dept. of Industry, Innovation and Science. The first version of BLADE linked a small set of ABS Survey Data to Tax data and some DIIS program data.

It enabled Australia's participation in the OECD 'DynEmp' project on understanding employer dynamics, particularly for small and young businesses.

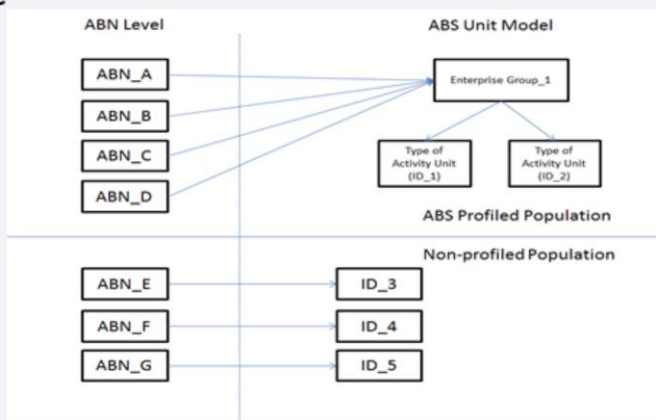
Since then it has expanded to include more ABS Survey Data, Tax data and other Economic Activity administrative data.

What is BLADE?

- ▶ Spine contains all businesses (ABNs) that have been active at any time between 2001/02 and 2018/19
 - Sourced from the ABS Business Register
 - Updated Annually
- ▶ BLADE brings together:
 - Taxation Information
 - ABS Survey Data
 - Economic Activity data such as Imports and Exports
 - Intellectual Property information
 - Government Program Administrative data

The BLADE integrating Spine is sourced from the ABS Business Register. It contains all businesses that were active at any time between 2001/01 and 2018/19. It is updated annually which adds in new businesses, identifies ceased businesses and changes to the structure of the complex businesses through the ABS Profiling processes.

- ▶ BLADE Spine has the ABS Economic Units model as its structure



The BLADE Spine is based on the ABS Units Model and is sourced from the ABS Business Register.

It has two main populations.

Non-Profiled Population where the structure of businesses are considered simple and there is a direct 1-1 link between the ABN and the BLADE Unit.

The second population is the Profiled population.

This represents the larger and complex businesses. The ABS undertakes a regular profiling of these businesses to understand their complex structures. The BLADE unit in this population is what is called the TAU or Type of Activity Unit.

It represents a significant activity undertaken by the overall group of ABNs. ABNs can be linked to one or many TAUs within these structures.

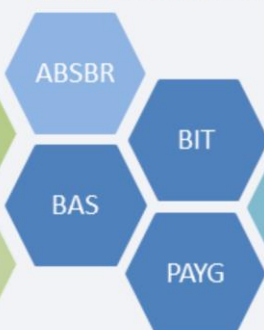
Having the ABS Units models as the basis for the BLADE Spine enable analysis to be undertaken on both simple and complex units.

BLADE Data Modules

ABS Survey Data Modules



BLADE Taxation Modules



Other Modules



- All datasets can be linked by a common key
- Full data item list available upon request

**** This slide has animations ****

The structure of BLADE can be best described as a series of modules that can be linked together by a common key. Client data can also be integrated in order to link to these modules via a key.

These modules fall into 3 distinct groups:

BLADE Core datasets:

Indicative data items such as State, Division, Type of Legal Organisation etc (ABS Business Register), 2001-02 to 2016-17

Australian Taxation Office data-

Business Activity Statements (BAS), 2001-02 to 2016-17

Business Income Tax (BIT), 2001-02 to 2016-17

Pay As You Go (PAYG), 2001-02 to 2016-17

ABS Survey Data

Business Characteristics Survey (BCS) 2005-06 to 2016-17

Management Capabilities Module (MCM) 2015-16

Business Expenditure on Research and Development (BERD) 2004-05 to 2015-16

Private Non-Profit Expenditure on Research and Development (PNPERD) 2006-07 to 2016-17

Government Expenditure on Research and Development (GovERD), 2006-07 to 2016-17

Economic Activity Survey (EAS) 2004-05 to 2016-17

These first two sets of data are at and ABS Units model level (or BLADE ID level)

Other modules

IP Australia: Intellectual Property Longitudinal Research Data (IPLORD), 2001-02 to 2016-17

Department of Home Affairs: Merchandise Exports Data, 2003-04 to 2017-18

Department of Home Affairs: Merchandise Imports Data, 2003-04 to 2017-18

These datasets and others that are integrated for one off projects are provided in DataLab at the level they are provided with a concordance file to provide the link to the BLADE unit(s).

More modules are in the process of being added to the BLADE asset such as the Energy, Environment and Water Survey (EWES)

Taxation Data



- ▶ **ABS Business Register (ABSBR) data**
 - ANZSIC, SISC, State, Postcode, Birthdate
- ▶ **Business Activity Statement (BAS) data**
 - Turnover, Wages & Salaries, GST Instalments, Capital Purchases
- ▶ **Business Income Taxation (BIT) data**
 - Business Income, Expenses, Assets, Liabilities, Profit and Loss
- ▶ **Pay As You Go (PAYG) data**
 - Full Time Equivalent (derived), Headcount

The 3 sources of tax data are sourced directly from the ATO.

All data is annualised to a financial year basis and integrated to the BLADE Spine

- ▶ **Economic Activity Survey (EAS)**
 - Income, Expenses, Industry Value Added, Labour costs, Inventories, earnings before interest and tax (EBIT), gross fixed capital formation
- ▶ **Business Characteristics Survey (BCS, MCM)**
 - Various variables related to innovation, expenditure on innovation, nature and extent of business collaboration, extent of use of IT
- ▶ **Expenditure on R&D (BERD, GOVERD, PNPERD)**
 - Breakdown of R&D, Effort in R&D, Sources of R&D funding

ABS Survey Data is only a sample of the population is contains detailed information on the businesses selected.

When combined with the population wide administrative data, such as taxation data, it enables analysts to developed detailed and complex models on business performance and dynamics that can be applied to the rest of the population not included in the survey sample.

ABS Surveys are selected from frames sourced from ABSBR making it easy to integrated to the BLADE Spine.

Economic Activity Survey (EAS) 2004-05 to 2016-17

Annual frequency, approx 18,000 businesses

Business Characteristics Survey (BCS) 2005-06 to 2016-17

Annual frequency; approx. 7,500 businesses

Management Capabilities Module (MCM) 2015-16

One off survey, 14,500 businesses.

Business Expenditure on Research and Development (BERD) 2004-05 to 2015-16

Annual frequency from 2004-05 through 2013-14, then moves to biennial* approx 4k businesses

Private Non-Profit Expenditure on Research and Development (PNPERD) 2006-07 to 2016-17

Approx biennial# approx 100 businesses

Government Expenditure on Research and Development (GovERD), 2006-07 to 2016-17

Approx biennial# approx 120 - 150 businesses

Economic Activity Survey (EAS) 2004-05 to 2016-17

Annual frequency, approx 18,000 businesses

- ▶ Agriculture Survey and Census data
 - Integrated for a specific project
 - Farm level data 2005/06 – 2015/16
 - Linked climate data to geography of farms

Approximately 100k farms
Detail commodity information
Use of data is subject to Data Custodian approval

- ▶ Commodity level information at the ABN level for 2001-2018
- ▶ Detailed information includes:
 - Commodity (8 digit AHECC)
 - Month of Import/Export
 - Value, Gross Weight, Quantity
 - Port of Discharge, Country of Destination, mode of transport

The Merchandise Imports and Exports data is sourced from the Dept. Of Home Affairs.

Provided at the detailed commodity transaction level for each business on a monthly basis.

Data integrated to the BLADE spine based on the ABN but the analysis data set is kept at the Export declaration/transaction level, providing a very rich level of information available for analysts.

Exports Data, 2003-04 to 2017-18

Monthly reference periods presented in FY datasets, between 30-50k businesses, 1.7m to 4.3m transactions

Merchandise Imports Data, 2003-04 to 2017-18

Monthly reference periods presented in FY datasets, 85-243k businesses, 2.8m to 3.5m transactions

Other Government Data



- ▶ Intellectual Property Australia
 - IPLORD - annual snapshot of the stocks and flows of IP rights for Australian applicants
- ▶ Regularly integrate government program data for specific projects
 - Dept. of Industry, Innovation, and Science - Grants and Services data
 - Used for program/support evaluation, business performance post-support
 - Identify new sectors requiring support

Value add of each type of data

- ▶ **Survey Data:**
 - The survey data is richer and has diverse variables
 - Originally collected for statistical purposes
 - Is a sample and can be periodic
- ▶ **Admin Data:**
 - Essentially a census of Australian firms
 - Rich in financial variables
 - Not originally collected for statistical purposes
 - Longer time dimension

How is BLADE being used?

How is BLADE being used?

- ▶ BLADE is being used for a wide variety of research
 - Evaluation of government programs
 - Understanding how business perform or are impacted by different real world economic and environmental events
 - Understanding which businesses are innovative or undertake R&D and how this impacts on their performance
 - The list goes on....
- ▶ You'll hear about a few in the next session

Areas of research using BLADE



Some examples of BLADE Research Projects:



Wage subsidies and business dynamics
(DEPARTMENT OF EMPLOYMENT, SKILLS, SMALL AND FAMILY BUSINESS)



Climate variability on the productivity and profitability of Australian farms
(DEPARTMENT AGRICULTURE)



Indigenous Environmental Programs: Social-Economic Analysis
(DEPARTMENT OF THE ENVIRONMENT AND ENERGY)



How do businesses respond to shocks and change
(RESERVE BANK OF AUSTRALIA)



Credit Constraints and Entrepreneurship Decline
(DEPARTMENT OF INDUSTRY, INNOVATION AND SCIENCE)



Impact Assessment of Austrade's services
(AUSTRADE)

18



Visit the [BLADE Research Projects](https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+BLADE+Research+Projects) page for the full list of projects

Link:

<https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+BLADE+Research+Projects>

Currently BLADE is used by over 150 researchers and analysts across Australia on over 60 projects.

This slide provides some examples of approved research projects that use BLADE data.

- **Wage subsidies and business dynamics**

This project uses integrated data from the Department of Jobs and Small Business' Employment Services System (ESS) and BLADE to identify the characteristics of businesses that have used wage subsidies provided through the Government's employment services program to investigate how utilisation of wage subsidies is linked to business dynamics.

- **Indigenous Environmental Programs: Social-Economic Analysis**

This project will assess, discover and synthesise new evidence in existing government datasets that highlights the return on investment from Australian Government support for Indigenous environmental programs.

- **Credit Constraints and Entrepreneurship Decline**

In view of the sudden decline in the number of firm entries in Australia between 2005 and 2011, this project will examine the link between entrepreneurship dynamics in Australia and measures of financing constraints. The project will test

whether a deterioration in access to credit was behind the decline in entrepreneurship and the increasing risk of exit faced by them. Established methods from the finance literature will be applied to infer financing constraints from firm-level data and use this as the basis of determining key explanatory variable(s). The project aims to demonstrate the extent to which financial constraints and access to finance was driving the observed decline in entrepreneurship. The results will also provide guidance for future policies to counter another decline in firm entry.

- **Climate variability on the productivity and profitability of Australian farms**

This study will develop a better understanding of the effects of climate variability on the productivity and profitability of Australian farms, establishing the relationships between climate conditions (i.e., rainfall and temperature) and farm production and financial outcomes.

- **How do businesses respond to shocks and change**

This project seeks to investigate how Australian businesses and labour markets respond to macroeconomic shocks and longer-term structural change.

- **Impact Assessment of Austrade's services**

This project assists Austrade to establish a set of consistent, objective methodologies designed to provide robust quantitative estimates of the impact of many of Austrade's services and programs aimed at promoting Australia's trade and investment. The outcomes can provide evidence for how Austrade assists in strengthening Australia's export base and investment, while shedding light on the areas requiring further attention.

The Access process for BLADE Research Projects is the same as for MADIP Research Projects. All projects that use BLADE data must go through a rigorous assessment and approval process, managed by the ABS. Only authorised researchers will be granted access to BLADE data for policy analysis, research, and statistical purposes. Access is to data with all direct identifiers removed.

All projects are assessed under the [Five Safes Framework](#). For a project to be approved, the ABS and the data custodians (the agencies that collect the data) must agree to the proposed use of the data. The project must be assessed as being in the public interest and be in accordance with the legislation of the relevant agencies. All users are legally obliged to use data responsibly for approved purposes, comply with the conditions of access, and maintain confidentiality of data.

BLADE Enhancements to enable research



► Business Location

- At present only the mailing address location of a business is available on BLADE
- ABS project to identify Business Activity Location
 - Actual location of activity and multiple activity
 - Available end of June 2020

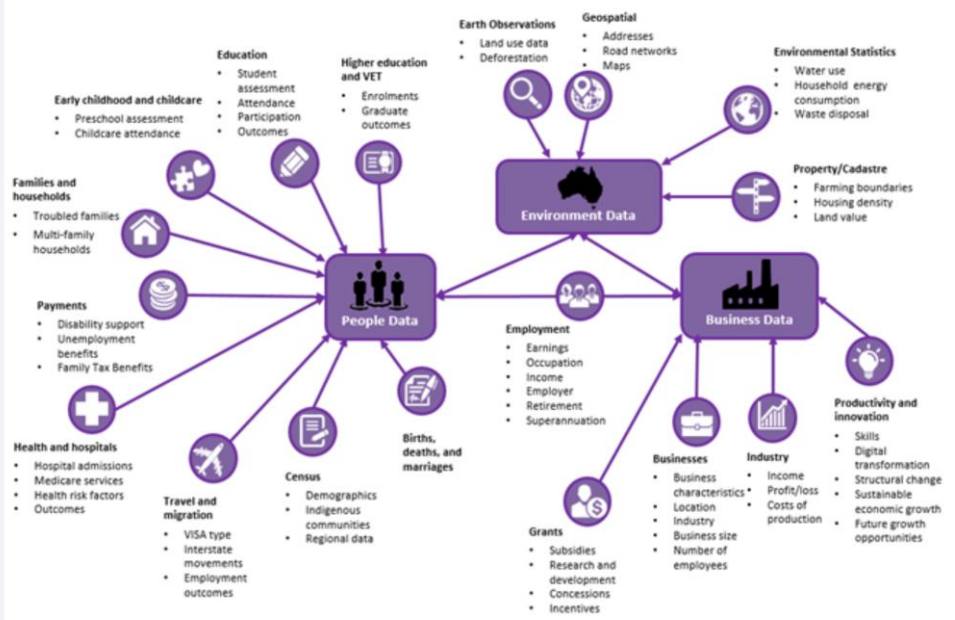
At present the location information that is available on BLADE is the post code of the mailing address of the business.

In most cases this is based on the information received from the ATO or the mailing address used for ABS surveys.

A project is currently underway that is using new and multiple sources to identify the locations of where activity is for each active business on the ABS Business register. This will include businesses undertaking activity at multiple locations.

This information is expected to be available on BLADE from the end of June 2020.

Future Directions – Australian Integrated Data Assets



20

21/02/2020

One of the main goals of the ABS Data Integration Program is to enable the research and analysis of issues that cut across multiple dimensions of Australian society, namely Economical, Social and Environmental.

As mention in 2019, 12 years of Agriculture Survey and Census data was integrated to BLADE. As we had farm location on the ag data we were also able to link climate data from the Bureau of Meteorology.

Also, the ABS is in the process on integrating a number of BLADE data modules to MADIP to enable researchers to investigate if past migration programs have achieved their intended outcomes, and provide insights into potential effects of future policy changes (including on macroeconomic outcomes and the Commonwealth fiscal balance).

ABS is investigation bringing these current assets together in a safe way under as Australian Integrated Data Assets.


How you can get access to BLADE?

How is BLADE Accessed?




- ▶ From 2019 ABS micro-data products are now accessible for approved projects to researchers from:
 - Government (Commonwealth, State and Territory)
 - Government contractors
 - Academic researcher and
 - Public Policy Institutes
- ▶ Access virtually to ABS DataLab
 - Access to unidentified micro data






Government, government contractors and individuals can apply to access BLADE data for statistical and research purposes (i.e. not for compliance or regulatory purposes).




Five Safes Framework



Five Safes Framework

 Safe people	Is the researcher authorised to access and use the data appropriately?
 Safe projects	Is the data to be used for an appropriate purpose?
 Safe settings	Does the access environment prevent unauthorised use?
 Safe data	Has appropriate and sufficient protection been applied to the data?
 Safe output	Are the statistical results non-disclosive?

23


Ultimately the ABS objective is maintaining the balance between the utility of the micro-data and the privacy and security of the data.

- The way in which we do this is the ABS uses the internationally recognised [Five Safes Framework](#) for managing the safe and secure release of BLADE microdata. The ABS DataLab is the environment through which it is accessed.
- The Five Safes are the basis of the Data Sharing Principles published by the Office of the National Data Commissioner earlier this year which to guide safe sharing of Commonwealth data.
- The Five Safes Framework enables use of data for policy analysis, research, and statistical purposes, and, at the same time, maintain the privacy and confidentiality of information.
- The five safes are controls or levers, each of which is considered independently on a sliding scale, as well as jointly, to evaluate whether or not a particular method of data access meets confidentiality and privacy requirements.
 - **Safe People** – Safe Researcher on-boarding (including training about responsible use of data and maintaining data security), researcher paperwork and Responsible officer undertaking.
 - **Safe Projects** – All projects are for statistical or research purposes (not for compliance), are of public value, and have been approved by the relevant governance committee, data custodians, and ABS Senior Manager.
 - **Safe Settings** – Various security settings on the DataLab including no capacity for the researchers to import or export data.
 - **Safe Data** – The datasets to be accessed by the researchers are de-identified and prepared in accordance with ABS Disclosure Review Committee (DRC) advice.
 - **Safe Outputs** – ABS apply a rigorous output checking and approval process prior to any data leaving the ABS environment.

These Five Safes considerations as a whole ensure safe access to data.



BLADE Access Process

- ▶ ABS Integrated Data project proposal
- ▶ 5 Safes Assessment
- ▶ Data Custodian Approval sought
- ▶ Integration of data (if required)
- ▶ DataLab Training
- ▶ ACCESS!!!

Contact: aida@abs.gov.au or alan.herning@abs.gov.au

Questions?