New Zealand Agricultural & Resource Economics Society (Inc.)

Analysis of Agriculture Production Survey and Annual Enterprise Survey Data: Findings and Lessons

Stephen J. Murray[1], Meena Satishkumar, & Alice Marfell-Jones

Ministry for Primary Industries, Wellington, New Zealand

[1] Stephen.Murray@mpi.govt.nz

Paper presented at the 2013 NZARES Conference

Lincoln University – Canterbury, New Zealand. August 28-30, 2013

Analysis of Agriculture Production Survey and Annual Enterprise Survey data:

Findings and lessons

Stephen J. Murray*, Meena Satishkumar, and Alice Marfell-Jones

Ministry for Primary Industries, Wellington, New Zealand

*Contact details

Email: Stephen.Murray@mpi.govt.nz

Phone: +64 4 8940023

ABSTRACT

The Agriculture Production Survey and Annual Enterprise Survey were analysed as part of a redesign of the Ministry for Primary Industries' Farm Monitoring Programme. These data sets are hosted by Statistics New Zealand, and access is possible for accredited researchers. This paper is intended to 1) provide examples of the types of analyses possible using these data sets, and 2) describe key lessons from our experiences.

Key words: Primary industries, research, financial, Statistics New Zealand

INTRODUCTION

In 2012 an Information Development Programme (IDP) was initiated at the Ministry for Primary Industries (MPI, 2013a). The IDP is focused on 1) improving the quality of MPI's data; 2) reducing duplication in data collection; and 3) increasing access to data for additional analyses. A major focus for the programme is revising MPI's Farm Monitoring programme (MPI, 2013a).

Farm Monitoring

Farm Monitoring has been run since 1978 (MPI, 2004). Current reports are designed to represent the financials and production, both current and forecast, of a range of farm types throughout New Zealand (MPI, 2013b). However, recent internal reviews of the FM programme and user information needs assessments indicated the models were only partially meeting information needs. For instance, the structure of the industry has changes since Farm Monitoring began. In addition, the programme was not providing value-for-money.

Statistics New Zealand data

We explored whether the Agriculture Production Survey (APS) and the Annual Enterprise Survey (AES) could be used as a core data source for Farm Monitoring. As these data sets are held by Statistics New Zealand the analyses were conducted in the Data Lab (Statistics New Zealand, 2013). A key part of the Data Lab is the Prototype Longitudinal Business

Database (Fabling, 2009). The Datasets section contains a summary of the AES, APS, and Prototype Longitudinal Business Database.

Paper Outline

This paper presents the initial analysis of the AES and APS, as well as follow-up analyses on business structure and on-farm activities. The purpose is to provide examples of the sorts of outputs that can be produced and to outline lessons from the analysis process.

The paper is divided into four sections.

1. Background information on the datasets used
2. The initial analyses
3. The follow-up analyses
4. General lessons

DATASETS

Annual Enterprise Survey

The AES is a financial survey of economically significant businesses operating in New Zealand (Statistics New Zealand, 2013d; see *Coverage* from Statistics New Zealand, 2013e, for a definition of economically significant). Most of the data is obtained from Inland Revenue Department IR10 forms (around 66% in 2010; Statistics New Zealand, 2013d). For the sectors covered by Farm Monitoring all financial information comes from the IR10 form. IR10 data is available in the business database as the:

1. raw IR10 variables
2. IR10 variables transformed into the standard AES variables
3. data in '2' weighted for use in Statistics New Zealand 'Hot off the Press' publications.

The IR10 is non-compulsory (Inland Revenue Department, 2013, p. 3), and some businesses submit their accounts directly to IRD. These businesses are not captured in the non-imputed AES data.

Agriculture Production Survey

The APS is an annual survey of "all businesses engaged in 'agricultural production activity' (including livestock, cropping, horticulture and forestry)" (Statistics New Zealand, 2012). Each year approximately half of all eligible businesses are surveyed, with a census including all businesses being run every five years. Pastoral, arable, and forestry information is collected yearly. Horticulture questions are included every second year and in census years.

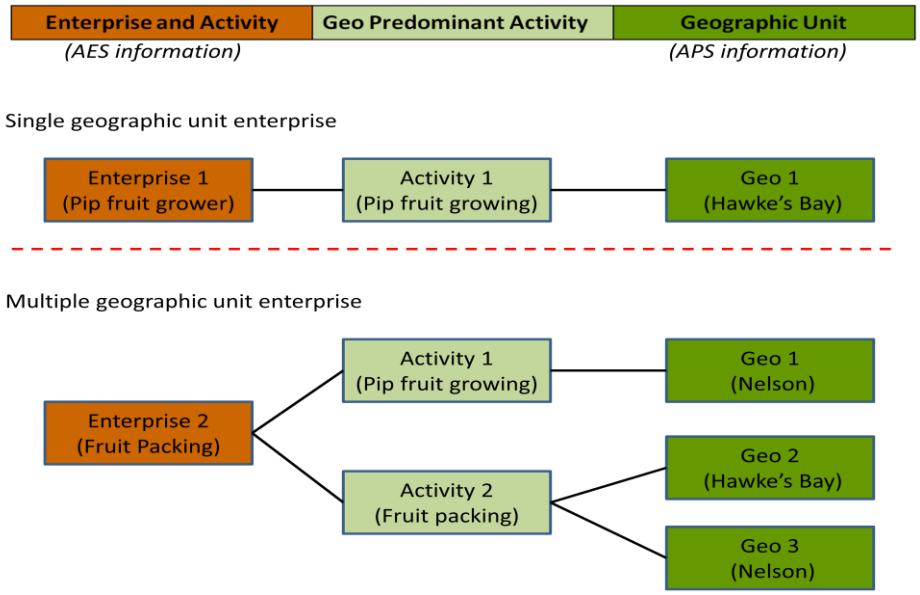
The Prototype Longitudinal Business Database

The Prototype Longitudinal Business Database is built around the Longitudinal Business Frame (Seyb, 2003), a record of most businesses in New Zealand. Each business unit has a unique enterprise number, and one or more associated physical locations (geographic or geo units; see Figure 1). Most businesses follow the *Single Geographic Unit Enterprise*

structure (Seyb, 2003). The geo unit is recorded at the smallest statistical unit level used by Statistics New Zealand, the meshblock (Statistics New Zealand, 2013b).

Activities at the geo unit level are classified using the Australia New Zealand Standard Industrial Classification (ANZSIC) 2006 (Australian Bureau of Statistics, 2006). The activity producing the highest value of final goods or services less the cost of goods or services used in production (Value Added; Statistics New Zealand, 2013c) is defined as the predominant activity. The predominant activity for an enterprise is based on its associated geo units.

Figure 1. Structure of the Longitudinal Business Frame (modified from Seyb, 2003).



INITIAL ANALYSES: THE COMBINED AES AND APS DATASETS

Sample selection method

Five sectors were selected for the initial analysis: Pipfruit, kiwifruit, grapes, dairy, and deer. For each sector the sample consisted of enterprises with:

1. the relevant ANZSIC code (e.g. A016000 for Dairy),
2. non-imputed AES data,
3. non-imputed APS data,
4. one geo location.

These criteria were chosen as we wanted to represent activities and financials at the farm level. Including imputed data or multi-geo enterprises were considered to add too much complexity to the analyses (see *Lessons from the initial analyses* below).

Initial findings

Table 1 provides the results of an early analysis for South Island deer farming using 2010 AES and APS data. The table should be viewed as an illustration of possible analyses rather than concrete figures about deer (see *Overall Lessons from our Experience*), especially

regarding financial variables. More information on business types can be found from Business.Govt.NZ (2013).

For this sample companies have substantially more grazing land than sole traders or partnerships. Differences in deer numbers by business type match the trend for land area. However, there is a lot of variability within business types (compare means to medians).

Table 1. Land, stock levels, and financials per hectare for South Island deer farms, split by business type, for 2010. Figures represent means, with medians in brackets. Please note that financial figures are unreliable as they are a combination of weighted and unweighted data.

	Business Type			
Variables	**Sole trader (n = 18)**	**Partnership (n = 78)**	**Company n = 36)**	**Overall (n = 135)**
Land information (ha)				
Total grazing land	466 (172)	273 (177)	1251 (413)	553 (200)
Total land	509 (191)	345 (221)	1597 (459)	693 (236)
Deer numbers (head)				
Total deer on farm	844 (657)	953 (816)	2097 (1024)	1222 (876)
Total R1 hinds and stags	286 (231)	382 (296)	897 (425)	498 (296)
Total R2 hinds	102 (65)	91 (70)	175 (80)	113 (73)
Total R2 stags	49 (18)	38 (11)	58 (16)	44 (14)
Total R3+ stags	45 (14)	89 (31)	134 (65)	93 (30)
Financial variables ($)				
Profit / ha	-43 (24)	255 (112)	-240 (4)	78 (61)
Income / ha	1259 (1123)	1735 (1370)	1680 (906)	1618 (1231)
Expenses / ha	1155 (947)	1465 (1240)	1780 (984)	1475 (1125)
Assets / ha	9167 (5638)	7788 (5771)	7749 (4479)	7789 (5525)
Liabilities / ha	3039 (2402)	2169 (1430)	5662 (1804)	3168 (1699)
Wages / ha	41 (7)	34 (0)	177 (57)	72 (9)
Farms with other livestock types (%)				
At least 1 dairy cattle	17	28	19	24
At least 1 beef cattle	83	60	69	64
At least 1 sheep	56	53	67	56
Other livestock numbers (head)				
Total dairy cattle on farm)	c. (0)	87 (0)	78 (0)	c. (0)
Total beef cattle on farm	59 (36)	55 (17)	192 (55)	91 (28)
Total sheep on farm	675 (10)	615 (20)	1514 (86)	849 (32)

c. = confidentialised data

At least 64% of the sample run at least some form of other livestock. Deer farms are more likely to run beef or sheep compared with dairy cattle. In contrast, current Farm Monitoring deer models represent farms that only run deer (Ministry for Primary Industries, 2012a, b). Note also stock unit means and medians are based on all deer farms not just those running the specific livestock type.

Lessons from the initial analyses

The data was considered useful for developing a higher level picture of the sectors, but not suitable for a core farm monitoring dataset. Higher-level analyses are presented in the *Follow-Up Analyses* section. Reasons the data were considered unsuitable for farm monitoring are presented below.

- The level of detail in the AES and APS data is not sufficient for farm-level monitoring

Some information in both the AES and APS datasets was not collected at sufficient detail for the monitoring programme. For instance, specific on-farm expenses such as pruning or shearing are grouped under "wages/salaries" for the IR10. The APS data also lacked key production data, such as milk solids and carcass weight. This information is essential for examining farm level productivity.

- The sample may not be representative of the population

Table 1 represents 135 of 954 South Island deer farms (total figure obtained from NZ.Stat, 2013). This result means approximately 86% of deer farms failed to meet at least one of selection criteria 2 to 4. These criteria may not be independent of farm characteristics. For instance, businesses which did not have IR10 information tended to employ more people than those which had IR10s (unpublished analyses). More large than small deer farms may therefore be excluded from the sample.

- Multiple-geo enterprises make determining farm-level financials difficult

It is easy to match financial information to farm information for single-geo enterprises. For a multi-geo enterprises the' financials need to be split across the different geos. There was not a clearly defined way to split finances at the time of running the analyses.

- Some business structures are not represented in the business frame

Other business structures cannot be determined from the data. For example, for a dairy farm the cattle may be owned by a sole trader with the land owned by a trust. Both the trust and the sole trader could be the same person. For this farm financial information would only be available for the cattle-side of the business. There is no way to link two different enterprises to one physical location using the information available in the Data Lab.

FOLLOW-UP ANALYSES: HIGH LEVEL BUSINESS STRUCTURE AND ON-FARM ACTIVITY

Follow-up sample selection method

The business structure sample was selected from the business frame based on ANZSIC 2006 code. All businesses coded to a specific sector (e.g. dairy) on June 2010 were sampled.

The sample for on-farm activities was drawn from the 2007 APS census data. All geo unit including any level of a specific activity (e.g. at least one dairy cow, 1 ha of pipfruit) were selected.

Business Structure findings

General findings regarding the business structure of pipfruit orchards are presented in Table 2. The majority are companies, followed by partnerships. There are generally fewer employees than geo units, with the exception of the two key pipfruit regions. This finding is

possible as workers who do not draw a salary, such as some sole traders and volunteers, are not counted in the employment statistics.

Table 2. Number of geographic units and employment figures for pipfruit enterprises, 2010, split by region and business type. All figures have been rounded so totals may not equal the sum of the business types or regions.

Region	Sole Trader	Partnership	Company	Trusts/ Estates	Other*	Total
Number of Geographic Units (**Total employment including working proprietors**)						
North Island*	-	-	-	-	3 (**6**)	3 (**6**)
Hawke's Bay	33 (**15**)	69 (**15**)	150 (**330**)	18 (**6**)	-	267 (**380**)
Other North Island	24 (**9**)	60 (**20**)	42 (**40**)	6 (**0**)	-	132 (**70**)
South Island*	-	-	-	-	6 (**18**)	6 (**18**)
Nelson/Tasman/ Marlborough	30 (**18**)	66 (**45**)	57 (**140**)	0 (**0**)	-	156 (**190**)
Other South Island	6 (**0**)	36 (**3**)	18 (**12**)	0 (**0**)	-	60 (**12**)
Total NZ	96 (**45**)	228 (**85**)	270 (**520**)	27	9	627

* Note: Due to small numbers of "Other" business types these values have been aggregated at the North and South Island level.

We also attempted to examine vertical integration using the business frame data. The predominant activity of the enterprise attached to each sampled geo was examined. Nine enterprises representing 18 geo units had predominant activities indicating fruit packing or retail. Discussion with the business frame team at Statistics New Zealand revealed constraints around examining vertical integration. If you are interested in more information about these constraints please contact the lead author.

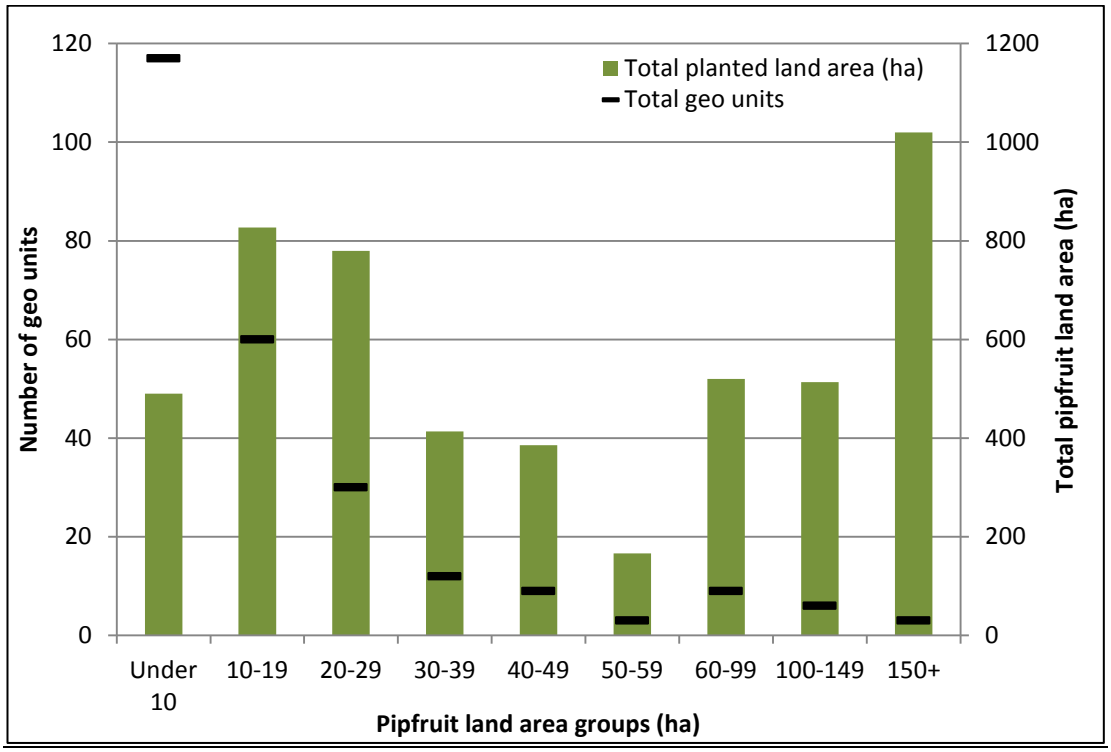
Pipfruit: Number of orchards and planted area

We conducted an analysis of pipfruit area using 2012 APS data. Figure 2 represents Hawke's Bay pipfruit orchards. The figure is useful for showing the difference between the actual sector structure and existing Farm Monitoring models. There are a large range of orchard sizes in the Hawke's Bay. For comparison, the Hawke's Bay pipfruit model (MPI, 2012c) was an orchard with 18 ha of pipfruit. There were therefore many orchards that were not included in the sample frame for Farm Monitoring.

Intended approach for the on-farm activity analysis

An outline of a possible output is shown in Table 3 to provide a guide to the intended analyses. The initial analyses used 2007 APS census data, which is out-of-date. At the time of publication these analyses had not been run on the 2012 data.

Figure 2. Number of Hawke's Bay geo units growing pipfruit and total land area split by land are groups.



These outputs can be produced by region, business type, and geo unit ANZSIC code. For instance, the difference in range of activities for core pipfruit orchards versus other business which also grow pipfruit can be examined. Please contact the lead author if you are interested in these outputs once they are produced.

Table 3. Potential variables and descriptive statistics for the on-farm activity analyses

Variable	N	Mean	Std Dev	Percentiles (5, 10, 25, 50, 75, 90, 95)
Specific land uses (ha) *E.g. Apples, pears, other fruit, outdoor vegetables, arable crops*				
General land areas (ha) *E.g. Total land, Horticulture, grass land, grain seed fodder, exotic trees, native bush*				
Livestock (head) *E.g. Dairy cattle, beef cattle, sheep, deer, pigs, other*				

OVERALL LESSONS FROM OUR EXPERIENCE

The AES and APS datasets are useful for examining on-farm activity and business structure, the latter illustrated by the pipfruit planted area analyses (Figure 2). Further analyses of the 2012 APS data are currently underway. For further information about ongoing analyses please contact the lead author.

- There is a lot of potential for additional analyses using the AES and APS data

The linked AES and APS datasets can be used to answer some interesting questions about farming in New Zealand. However, the lead authors do not have the background to suggest what these questions may be. An indication of the sorts of analyses that can be conducted can be obtained by examining APS survey forms and either the IR10 form or the AES metadata.

- Factor in the time to develop an understanding of the Data Lab data sets when planning research projects.

There are a large number of data sets that are available in the Data Lab. Understanding the linkages between these data sets will take time. In addition, within each data set there can be a combination of raw and derived and/or weighted and unweighted variables.

For instance, Table 1 represents unweighted farm-level data matched with weighted AES data. This error occurred due to the different ways financial data is stored in the data lab. Raw IR10 data is available, as is the IR10 data is transformed into the specific AES variables. The transformed data is available in both unweighted and weighted forms. For Table 1 we used weighted AES variables which were included with unweighted AES variables in the same dataset.

- Caution is advised when interpreting initial results

Following on from the previous lesson, there is likely to be a need to conduct additional work to test the meaningfulness of any results. In our original analyses we produced figures that did not make sense (e.g., unrealistic number of cattle per hectare). For instance, assuming the results in Table 2 are sufficient is a mistake. Additional steps to ensure the data is meaningful include comparing results with other available data sets, follow-up data analyses, talking to the appropriate teams at Statistics New Zealand, and engaging with industry experts and groups.

- Other researchers and specific business teams in Statistics New Zealand are useful resources

The complexity of the Data Lab can be partially addressed through other researchers and Statistics New Zealand business teams. Other researchers are likely to have a good understanding of the linkages between the data sets they work with. For specific queries about a data set the relevant business group can usually help. Our work benefitted from both sources of information.

Disclaimer

This report was undertaken while the author was on secondment to Statistics New Zealand.

The opinions, findings, recommendations and conclusions expressed in this report are those of the author(s). Statistics NZ takes no responsibility for any omissions or errors in the information contained here.

Access to the data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular business or organisation. The results in this paper have been confidentialised to protect individual businesses from identification.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information must be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes. Any person who has had access to the unit-record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to privacy and confidentiality. Any discussion of data limitations or weaknesses is not related to the data's ability to support Inland Revenue's core operational requirements.

REFERENCES

Australian Bureau of Statistics (2006). Chapter 4: Methods of classification. In Australian and New Zealand Standard Industrial Classification (ANZSIC) 2006, pp 21-24. Retrieved from http://www.stats.govt.nz/surveys_and_methods/methods/classifications-and-standards/classification-related-stats-standards/industrial-classification.aspx

Business.Govt.NZ (2013). Understanding business structures. Retrieved August 23, 2013, from http://www.business.govt.nz/starting-and-stopping/business-structures/understanding-business-structures

Fabling, R. (2009). A Rough Guide to New Zealand's Longitudinal Business Database. Global COE Hi-Stat Discussion Paper Series gd09-103, Institute of Economic Research, Hitotsubashi University.

Inland Revenue Department (2013). IR10 Guide (April 2013). Retrieved from http://www.ird.govt.nz/resources/0/c/0c69ec804ee4ec74aa05eab6c72de7d5/ir10g-2013.pdf

Ministry for Primary Industries (2004). Farm Monitoring Review (*internal document*). Ministry for Primary Industries: Wellington.

Ministry for Primary Industries (2012a). Farm Monitoring 2012: North Island Deer. Wellington. Ministry for Primary Industries.

Ministry for Primary Industries (2012b). Farm Monitoring 2012: South Island Deer. Wellington. Ministry for Primary Industries.

Ministry for Primary Industries (2012c). Farm Monitoring Report 2012 - Horticulture Monitoring: Pipfruit. Wellington. Ministry for Primary Industries.

Ministry for Primary Industries (2013a). Improving MPI's Sector Analysis. Retrieved August 22, 2013, from http://www.mpi.govt.nz/news-resources/statistics-forecasting/improving-mpis-sector-analysis.aspx

Ministry for Primary Industries (2013b). Statistical Publications. Retrieved August 22, 2013, from http://www.mpi.govt.nz/news-resources/statistics-forecasting/statistical-publications

NZ.Stat (2013). Detailed industry by region. Retrieved August 22, 2013, from http://nzdotstat.stats.govt.nz/wbos/Index.aspx?DataSetCode=TABLECODE2310#

Seyb, A. (2003). The Longitudinal Business Frame. Statistics New Zealand, Wellington.

Statistics New Zealand (2012). Information about the Agricultural Production Census – 30 June 2012. Retrieved August 5, 2013, from http://www2.stats.govt.nz/domino/external/omni/omni.nsf/bf3c32a862ae07cd4c25656e000031c6/60e3a3304fce89f2cc257acd0012021b?OpenDocument

Statistics New Zealand (2013a). Data Lab. Retrieved September 18, 2013, from http://www.stats.govt.nz/tools_and_services/microdata-access/data-lab.aspx.

Statistics New Zealand (2013b). Meshblock. Retrieved September 11, 2013, from http://www.stats.govt.nz/surveys_and_methods/methods/classifications-and-standards/classification-related-stats-standards/meshblock.aspx

Statistics New Zealand (2013c). Measuring value added in the education and health industries. Wellington: Statistics New Zealand

Statistics New Zealand (2013d). Annual Enterprise Survey. Retrieved August 22, 2013, from http://datainfoplus.stats.govt.nz/Item/nz.govt.stats/36809771-984d-4e6b-89a1-576f2118b05b

Statistics New Zealand (2013e). New Zealand Demography Statistics: At February 2012. Retrieved September 18, 2013, from http://www.stats.govt.nz/browse_for_stats/businesses/business_characteristics/BusinessDemographyStatistics_HOTPFeb12/Data%20Quality.aspx