Proceedings of Statistics Canada Symposium 2001
Achieving Data Quality in a Statistical Agency: A Methodological Perspective

ELECTRONIC DATA REPORTING – RECENT DEVELOPMENTS AT THE AUSTRALIAN BUREAU OF STATISTICS

Rob Burnside and Emma Farrell[1]

ABSTRACT

This paper is a summary of recent Australian Bureau of Statistics (ABS) methodological and other developments and experience with electronic data reporting (EDR). It deals particularly with that part of EDR loosely defined as 'e-forms' or screen-based direct collection instruments where all or most data are entered manually by the respondent. In this context it covers recent ABS experience and current work, but does not revisit the historical EDR work or cover other developments in Australia outside the ABS.

KEY WORDS: Electronic data reporting; e-forms; electronic data capture; on-line forms; electronic forms; mode effects; form design

1. INTRODUCTION

The ABS has now completed initial testing and evaluation of electronic 'forms' in the three ABS surveys identified in an earlier country paper presented by Edwin Aplin - *Statistical EDR: The Australian Experience* - at the Statistics Canada EDR Ottawa Workshop 25-27/9/2000 (Local Government, Internet Service Provides and Business Expectations). Brief details are given in the first part of this paper [a more detailed description of the findings of this testing are discussed in detail on the ABS Website (1)]. The second part of this paper discusses wider EDR issues and their influence on ABS thinking and directions.

2. ABS FORMAL EDR INSTRUMENT ASSESSMENTS TO DATE

2.1 Internet Activity Survey (IAS)

2.2 The instrument

The IAS is a quarterly census of Internet Service Providers (ISPs). It collects activity information about ISPs' subscribers, services offered (by type of service), and point of presence (POP) details. After in-house testing and development of an EDR instrument constructed using a dedicated electronic form tool (Jetform's FormFlow99), visits were made to eight ISP in three iterations of testing.

It was by far the most complex EDR instrument tested out of the three discussed here. It included a

[1] Rob Burnside, Australian Bureau of Statistics, PO Box 10, Belconnen, ACT, 2617 Australia
rob.burnside@abs.gov.au

Emma Farrell, Australian Bureau of Statistics, PO Box 10, Belconnen, ACT, 2617 Australia
emma.farrell@abs.gov.au

range of in-line (immediate) and on-completion edits, extensive help and definitions, and a combined navigation and completion status indicator (see attachment 1).

The questions required responses in a number of formats - multiple answer check boxes, mutually exclusive radio buttons, text and numeric data entry boxes and a table. This provided an opportunity to test several different on-screen input formats. By the time the visits were complete we were comfortable with the design and functionality of the instrument, though there was clearly an expectation from many testers that we would use a Web-style 'e-form' rather than an application.

Getting acceptable security working and tested for secure data transmission along with distribution and installation issues delayed a planned live trial. We developed an HTML form (see below) during this time, and our thinking shifted against using the FormFlow e-form application so the development and live pilot test was halted.

2.3 Outcome of testing

Heavily requested features during the visits were:

help - a keyword search facility, context sensitive help, use of hyperlinks as well as the tree navigator and help button to access help;

'tooltips' for all buttons;

a 'come back later' button and flag/icon for skipped questions;

a default 'ok' for accepting address and contact details; and

a fully printable form and/or printable summary of questions.

3. BUSINESS EXPECTATIONS SURVEY (BES)

3.1 The instrument

The (BES) collects mainly qualitative information on expected levels of key economic activity.

After an introductory screen the main instrument was implemented as a single scrolling page with a permanently visible toolbar holding buttons for key actions. It was a server-based Web-page style instrument and was not resumable - ie providers had to completed it in one session. A 'Help' function for data explanatory notes (as opposed to application help) was implemented as a new window drawn on top of the main screen. It was accessed by either a button in the tool bar or from the question text which also acted as hyperlink. No secure server was available for initial testing, and the instrument was distributed on CD ahead of the testing visits along with a request to have a look at it before we arrived.

Two versions of the instrument were tested as part of a longer term aim of exploring the effect and suitability of different design features and devices. The first version (A) used drop down answer category boxes ('list boxes'), the related short and medium term expectations questions were visible side by side on a single screen, and immediate edits were executed on exit from a field. (see attachments 2 and 3)

The second version (B) used radio buttons for the answer categories for questions, and, because of the extra space taken up by having a number range visible for each possible response, used more screen space. It was therefore necessary to present the medium term answer table underneath the short term one. There were no immediate edits.

Other than the differences mentioned above, the two forms were the same, with all edits being run on completion of the instrument.

3.2 Outcome of testing

There was an overwhelming preference for version B. Although some respondents liked the more compact side-by-side layout of version A, all but one respondent preferred the radio buttons of version B to the drop down answer categories of version A. The main reason given was that the radio button approach used less keystrokes and the full range of possible responses was visible.

Navigation - no problems were observed in navigating around the form during the visits, and all but one provider used mousing in preference to tabbing.

Edits - the lack of immediate edits in version B was not a concern, although those respondents who tested version A were not bothered by the presence of immediate edits;

Help - most providers gave very positive feedback in relation to the help function, which was identical in both versions. Some accessed it through the question links and some used the 'explanatory notes' button; three respondents commented that the button label should simply say 'help'.

4. PUBLIC FINANCE - LOCAL GOVERNMENT

4.1 The instruments

The two instruments discussed above were developed specifically for the tests. In this third case, electronic reporting using spreadsheets was already in use for local government collections in two Australian States (Victoria and New South Wales). These closely reflect the format used in agency and local authority financial statements and align closely with the data items and definitions used in public sector accounting standards and by the Federal Grants Commission.

The focus of these visits therefore was not to develop and test an instrument, but to investigate the provider behaviour and data sourcing aspects of EDR. This is particularly relevant because of a perception that reporting was or could be moved beyond manual completion of the electronic collection instruments to automatic extraction of the data sought.

Both the Victorian and NSW instruments are EXCEL spreadsheets. They use minimal design and formatting with separate worksheets for sections and navigation, and are visually very dense. Though basically similar, two interesting differences are that the NSW spreadsheet is for an annual collection and includes more reconciliation checks, whereas the Victorian instrument is quarterly, having a column for each quarter, and is sent out once at the start of the financial year for repeated use in reporting each quarter. Previous quarterly data are built up in the spreadsheet through the year and are visible each period, and act as an immediate aid to data quality (see attachment 4).

4.2 Outcome of testing

Nine visits were undertaken to a range of local councils and agencies in Victoria and six to local councils in NSW. The respondents were familiar with the spreadsheets and so had not specially looked at them as 'e-forms', and had no 'EDC-instrument' issues, seeing them purely as spreadsheets.

the spreadsheet was usually completed by one person who has a very good knowledge of the data and supporting systems, generally a trained accountant;

there were no significant instructions, definitional or other help requirements needs due to the ABS data items being closely aligned with Grants Commission and Local Government accounting standards and terminology;

there were no significant graphics, navigation or layout issues because the respondents were comfortable with EXCEL;

the main source of data was internal and statutory management reports, systems and statements, with data usually being manually transcribed from other reports to the ABS spreadsheet;

respondents were generally opposed to the collection of other ABS information on the same form.

5. TESTING SUMMARY

The overall conclusions drawn from all three sets of testing were that, in general, respondents expected behaviour and functionality comparable with other usage of the tool that they were familiar with. For instance, for the local government spreadsheets, they were comfortable with worksheets and the visually dense presentation of the instrument; they usually recognised and used hyperlinks and scrolling in a single page form for the HTML Webpage-like instruments, and they expected a very high level of functionality from the 'Windows application' instrument.

Respondent initiated action sequences for editing, verification, saving and data submission and their associated buttons, dialogue boxes and descriptive wording proved particularly difficult to get right. This applied to both the result the users expected and the wording of individual text messages.

Respondents appear to be even more reluctant to read questions, notes and other data-specific instructions on-screen than they are for paper instruments, but want access to help or definitions to be convenient and immediate.

Other themes and definite support for issues already identified from interface design, useability and emerging electronic form literature are that:

5.1 Design

the presentation and implementation of help and navigation are significant problems once there is any move away from the most basic sequential completion of a scrolling page format;

the wording of pop-up dialogue windows respondents interact with for edits, validation, saving etc require care and testing;

basic and accepted screen design and interface principles also hold for EDC instruments. For visible design areas these include text and background colours, fonts and font sizes, user dialogue boxes and wording, grouping and spacing.

5.2 Functionality

most respondents said that a print function is important, both to keep a hard copy record and to view the form off-screen before completion;

nearly all providers commented that for more complicated ABS surveys, the form would either need to be resumable or have the ability to be saved locally.

5.3 Testing procedure

All respondents said that they would prefer to receive access to the form via e-mail with a hyperlink to the form. Many commented that they bank on-line and are comfortable

with the level of security offered by on-line banking;

it can prove quite difficult to get respondents to focus on the instrument (rather than the survey content) in a controlled way, and a large part of several tests was taken up with comments on, and criticisms or questions about, form wording and definitions; and

testers are easily distracted by trivial errors, bugs and wording they dislike, and expect all instrument functionality to be present even during testing.

6. OTHER EDR PROJECTS AND ISSUES

6.1 The Electronic Transactions Act

From 1 July 2001 the Electronic Transaction Act (ETA) will apply to all Australian Commonwealth entities unless specifically exempted by regulation. Under the Act, if a person or business is required or permitted to give information in writing to a Commonwealth entity the requirement is taken to have been met if the person gives the information by means of an electronic communication. However, the Act also provides that the requesting Commonwealth entity can specify the particular information technology requirements used.

Confidentiality and security are major concerns in both policy and practical terms. The federal Government has developed a Public Key Infrastructure (PKI) known as Gatekeeper to facilitate the take up of online delivery of government services in Australia. PKI is a technology and trust framework, which involves the use of digital signature certificates for assuring the identity of certificate holders and the integrity of the online messages they exchange. This is managed by The National Office for the Information Economy (NOIE). NOIE has also invited the Privacy Commissioner to consider issuing guidelines on the privacy implications and good practices for Commonwealth agencies using PKI for individuals.

While the ABS fully supports the government directions in respect of electronic delivery of services, the ABS will only implement the capturing of data over the Internet if and when we can assure our providers that their data are secure, data quality is acceptable, and collection can be undertaken in a cost effective manner.

In light of the above, the ABS will continue to pursue the development of appropriate strategies for the collection of data electronically. In practice this has meant developing 'ETA-fallback' instruments and a small-scale secure lodgement facility policy to accommodate businesses and people who insist or need to report to us electronically.

Drawing on the testing above as well as our other experience, a range of short-term options were considered. EXCEL spreadsheets are the most cost effective and viable option and will be used as the 'ETA-fallback' electronic data collection vehicle for businesses. To minimise costs for what is expected to be an interim solution with a limited-life, these electronic forms will not include any validation, links to databases, or initial information about a provider such as contact information as is the case with many paper forms.

6.2 The Business ETA Instrument

The expectation is that an uncertain but limited number of respondents across the entire range of ABS business surveys and form types will elect to report electronically. Producing upwards of 150 electronic versions of our forms 'just in case' is clearly not a sensible option, so, after a centrally controlled filtering and explanatory process, instruments will be generated on a just in time basis for respondents who wish to report electronically.

These will be based on two or three templates or shells created in EXCEL. These shell forms will be used to create reasonably faithful on-screen representations of the paper form, and will be modified for particular collections where necessary.

They consist of the standard front page, usage and reporting instructions, and the core of the standards question wording (SQW) used in many ABS business collections. Collection-specific industry and activity-specfic questions will be added to complete the electronic instruments for individual surveys.

The template forms are in single and double column formats with no colour (using only shades of grey, black and white), and, when printed, appear similar to the paper equivalent, including question and instruction wording, font sizes and page breaks.

The use of tabs to navigate to worksheets is a familiar feature of spreadsheets. It simplifies instrument production when the options for the vertical alignment of text and data entry boxes etc are determined by the columns set at the top of a page or worksheet. However, because our implementation looks more like a form than a standard spreadsheet, standard spreadsheet navigation may not be appropriate so several navigation options will be usability tested internally. These are: standard spreadsheet tabs; a single scrolling page; previous/next buttons; and hypertext. We expect that even basic qualitative testing will show some clear preferences among users between the different forms.

No edits, field formatting or other checks will be implemented, and instruments will not contain any respondent identifiers or other structural information. Identification of the source of a return will be tracked through the access identifiers used to lodge the instrument.

Immediate concerns with any significant use of the interim instrument follow from the absence of any additional statistical advantages with the spreadsheet approach, which potentially combines the worst aspects of both paper and electronic reporting. In particular:

> assuming reporting is no less accurate on screen than on paper, the lack of edits implies at least the same level of clerical contact to resolve edit failures;
>
> well-developed sample and frame maintenance procedures for the treatment of frame deficiencies, non-response and changes to statistical units based on name, address and contact details is lost; and
>
> we have no indication of the likely data quality and modal bias effects of the instrument, the more so because the self-selected nature of respondents will confound any quantitative comparative analysis with paper reporting.

We will be concerned if use - ie the contribution of the instrument to estimates - in any given collection rises beyond a certain (low) level. It is seen as a temporary solution, has not been field tested, and offers no clear longer-term design development path.

Considering ABS and overseas experience to date, the expectation is that short to medium term demand for electronic reporting will be minimal. Together with the likely significant changes to the systems and organisation that support our business surveys, this approach is seen as acceptable in the short term and will give us time to articulate and work through all EDR issues with a more strategic perspective.

6.3 The 2001 Census ETA Instrument

The date of effect of the ETA (1 July 2001) and the legal advice received on its application to ABS collections came late in the cycle for the 2001 Australian Census of Population and Housing. It was therefore not possible to develop and test field procedures and to put IT systems and infrastructure in place in time to facilitate electronic lodgement on a wide scale for 2001. The 2001 Census electronic lodgement strategy, therefore, aims to ensure electronic lodgement facilities are available but to minimise their use and any additional costs to ABS.

6.4 The Instrument

A single session (ie not resumable) Internet 'Web-style' form has been developed. Because of the size of the Census form, it is in multiple sections which cannot be returned to once they are completed and saved/submitted. An introductory section is followed by the section for entry of dwelling details, then a section for brief details of the persons in that dwelling , and then individual person sections are generated based on the details in the previous section. The householder will need access to an IBM compatible computer with Microsoft Internet Explorer version 4.0 or above.

The instrument includes field completion checks for mandatory fields only, and before a section is saved, respondents are asked to check that the information they have entered is correct. Help is available through hyperlinks to the 'Census Guide'.

6.5 Testing

The instrument has been (externally) focus group tested and (internally) usability tested. Two otherwise identical versions with different levels of automatic editing were tested. The main outcomes of this testing were that:

most testers had experience with other on-line forms and were generally positive about the form;

there was widespread acceptance of a relatively high level of validation editing - however the timing and clarity of the edit messages requires further development;

auto-scrolling the display to position the next question at the top of the screen on exit from a field was confusing and disorienting;

the sudden appearance of list ('drop-down') boxes in response to certain answers was also unpopular;

the lack of question visibility, especially of previous responses, was not popular;

(see attachment 1) testers did not realise that they could not return to a section of the form once they had completed it (this is related to not being able to save a partially completed form); and

little or no use was made of the detailed completion instructions that were accessed through periodic links to the 'Census Guide';

the option of saving and returning to the form was regarded as necessary by users..

7. FUTURE DIRECTIONS - ABS MULTI-MODAL DATA CAPTURE STRATEGY AND CENTRALISED DATA COLLECTION

Apart from the small-scale and essentially empirical, immediate, electronic instrument design and testing discussed above, EDR is increasingly being considered strategically and in light of related developments in the ABS.

The first overarching influence is seeing electronic instruments inside an integrated collection and respondent management environment which envisages:

- more cost effectively using administrative by-product data substituting for directly collected data where possible;
- using 'e-forms' for direct collection of data as part of a multi-modal, fully integrated approach that includes telephone technologies for data collection and/or respondent contacts where appropriate - such as Computer Assisted Telephone Interviewing (CATI), Touch Tone Data Entry (TTDE) and Intelligent Voice Recognition systems (IVR);
- developing the capability to enable extraction of standardised data files from provider systems;
- focusing some initial effort on large volume data providers already familiar with electronic data provision; and
- paper forms continuing to be the most important collection vehicle in the immediate future.

The second major influence is that the ABS recently initiated an organisational review to consider options for improving the effectiveness and efficiency of our business survey operations. One likely outcome of the project is a move to a functional rather than subject matter oriented organisational form, including the establishment of an Integrated Business Data Centre (IBDC) that will centralise despatch, collection and respondent contact aspects for most ABS business collections.

Fitting well with (but not necessarily dependant on) these two influences is the desire to automate, standardise and integrate the production of business survey instruments as far as possible.

Instrument generation is seen as, eventually, being based on common metadata relating to data items and instrument design and functionality. Respondents will have some choice over their means of interaction with the ABS, within parameters set by the ABS and based on the suitability of different collection types and methods.

8. FURTHER WORK

Further methodological work on EDR will be undertaken inside a strategic research and development map we are developing that will inform systems work and accommodate likely organisational changes. It will allow us to identify and prioritise areas where a pilot test or development is necessary to investigate priority methodological aspects that will contribute to the development of the longer term EDR production environment.

At the individual instrument design level several major issues still need to be covered. These include interface design standards such as establishing:

- navigation consistent with a combination of the application or tool and the data sought;
- an appropriate level of and approach to validation, and measuring its effects;
- application or tool suitability - essentially open Internet tools versus proprietary applications; and
- development, testing and evaluation procedures.

From a collection or statistical perspective, the issues are:

identifying and quantifying the extent and effect of modal bias;

which forms of EDR are suitable for different type of providers and/or data items - including telephone technologies as well as screen instruments.

take-up rates and the effect on response rates;

provider management in a mixed mode collection system, particularly security and encryption arrangements and their scalablitiy;

links to collection control and other back end systems;

the effect of necessary security arrangements; and

provider value enhancement and 'incentives', for instance in terms of the 'value proposition' for respondents implied through the balance between the ease of or desire for electronic reporting in different formats and the necessary investment to implement it.

Information of interest will emerge from monitoring EDR flowing from the ETA and existing *ad hoc* collections, particularly on the real demand for EDR and the potential for different forms of it. Specifically 'e-forms' and the most acceptable tools, and the degree of support for various automated data extraction and transfer options.

However, controlled live deployments of various forms of EDR in collections that deal with different types of data and providers are necessary to properly examine and address most of these issues.

9. SUMMARY

ABS' current EDR efforts are focused on ensuring the availability of a Census electronic instrument for 2006 and having a workable ETA fallback instrument for businesses who insist on providing data electronically after July 1 2001.

Future methodological work on EDR will be undertaken inside a strategic research and development 'map' currently being developed by the ABS.

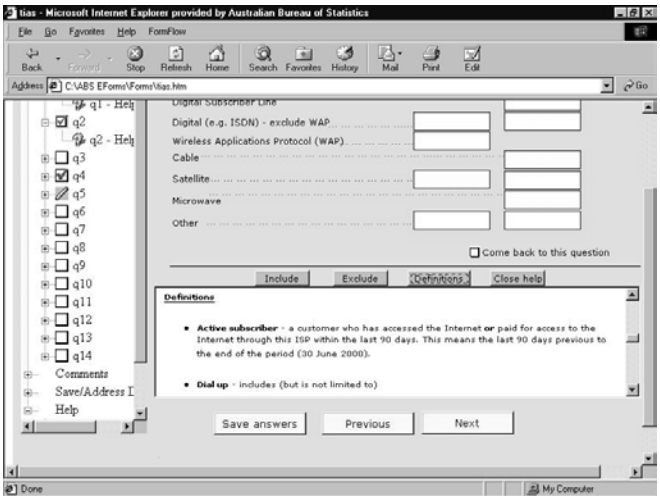
Testing to date has allowed us to specify design standards that we believe will be adequate for the conduct of the live trials of instruments necessary to examine statistical issues.

It appears likely that different electronic collection models, as well as instrument standards and functionality inside the 'e-form' direct collection component, will be better suited to certain data and respondent types. These could range from automatic extraction of standardised data items to Web-based single session or non-resumable (non-savable) forms, through spreadsheets to purpose built (programmed) instruments or those constructed using dedicated electronic form design tools.

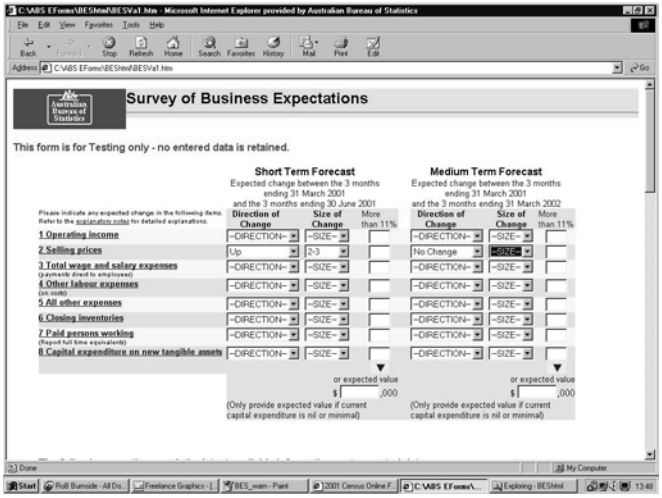
REFERENCES

(1) www. abs.gov.au ->
'Statistics' ->
'Working papers, Research papers and Other ABS Papers' ->
Methodological Advisory Committee Papers ->
'Towards Best Practice for Design of Electronic Data Capture Instruments'

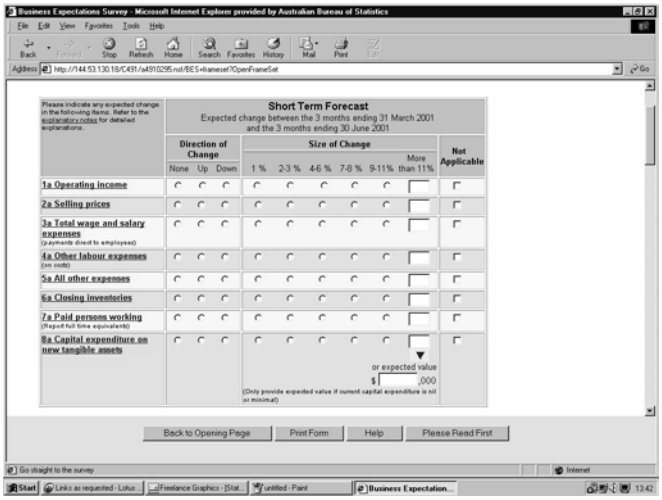
Attachment 1:



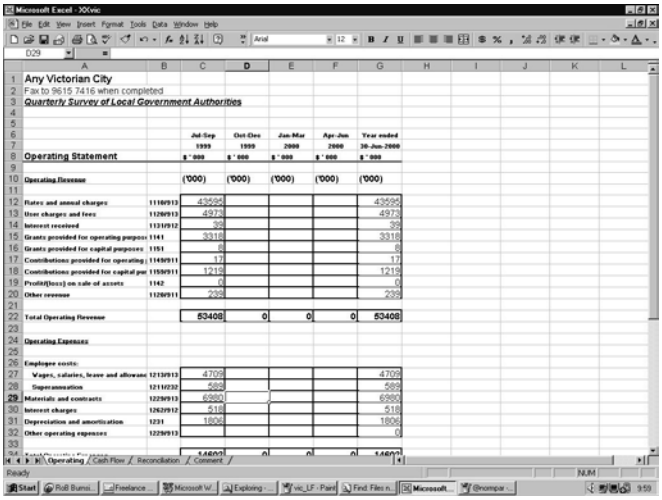
Attachment 2:



Attachment 3:



Attachment 4:



Microsoft Excel - XXvic

	A	B	C	D	E	F	G
1	Any Victorian City						
2	Fax to 9615 7416 when completed						
3	***Quarterly Survey of Local Government Authorities***						
4							
5							
6			**Jul-Sep**	**Oct-Dec**	**Jan-Mar**	**Apr-Jun**	**Year ended**
7			**1999**	**1999**	**2000**	**2000**	**30-Jun-2000**
8	**Operating Statement**		**$ '000**	**$ '000**	**$ '000**	**$ '000**	**$ '000**
9							
10	**Operating Revenue**		**('000)**	**('000)**	**('000)**	**('000)**	**('000)**
11							
12	**Rates and annual charges**	**1110/913**	43695				43695
13	**User charges and fees**	**1120/913**	4973				4973
14	**Interest received**	**1131/912**	39				39
15	**Grants provided for operating purpos**	**1141**	3318				3318
16	**Grants provided for capital purposes**	**1151**	8				8
17	**Contributions provided for operating**	**1140/911**	17				17
18	**Contributions provided for capital pur**	**1150/911**	1219				1219
19	**Profit/(loss) on sale of assets**	**1162**	0				0
20	**Other revenue**	**1120/911**	239				239
21							
22	**Total Operating Revenue**		**53408**	**0**	**0**	**0**	**53408**
23							
24	**Operating Expenses**						
25							
26	**Employee costs:**						
27	**Wages, salaries, leave and allowanc**	**1213/913**	4709				4709
28	**Superannuation**	**1211/232**	589				589
29	**Materials and contracts**	**1229/913**	6980				6980
30	**Interest charges**	**1262/912**	518				518
31	**Depreciation and amortisation**	**1231**	1806				1806
32	**Other operating expenses**	**1229/913**					0
33							

Operating / Cash Flow / Reconciliation / Comment /

Ready

NUM

Start | RoB Burns... | Freelance ... | Microsoft W... | Exploring - ... | vic_LF - Paint | Find: Files n... | Microsoft... | @rompar... | 9:59