

The roles of tax administration data in the production of official statistics in South Africa

Elizabeth Gavin^{*}, Deon Breytenbach, Randall Carolissen, Mamiky Leolo
South African Revenue Service, Pretoria, South Africa,

Tax administration data has been utilised in South Africa for over a decade in determining the sample frame for surveys of businesses used in the production of economic statistics and ultimately for the compilation of National Accounts. Formal agreements govern data flows between the South African Revenue Service (SARS) and South Africa's agency for official statistics, Statistics South Africa (Stats SA), enabled by changes in legislation. SARS and Stats SA discuss data requirements regularly. The modernisation of SARS's systems over the past seven years has led to significant improvements in the quality of tax administration data, as measured in terms of accuracy, completeness and the timeliness of its availability in structured, digital form, suitable for analysis. This has underpinned the annual publication of a re-established *Tax Statistics Bulletin* since 2008, which is being used increasingly by researchers and policymakers. In collaboration with Stats SA, SARS has begun exploring the use of tax administration data to derive statistical information to complement labour market information and income statistics generated from sample surveys. This research area is of significance to policymaking in South Africa, given persistently high unemployment and income inequality.

Keywords: business sample frame, data flows, labour market statistics, tax statistics

1. Introduction

Tax authorities interact regularly with companies, individuals and trusts on a legislated basis and collect information on these taxpayers as part of the revenue collection process. Data collection spans the initial registration details of taxpayers, information provided on tax returns and in the process of deregistration for tax. In addition, the *Tax Administration Act 28 of 2011* (South Africa, 2011) obliges taxpayers or even third parties to provide South Africa's tax authority with additional taxation-relevant information. A statistical agency is also authorised to obtain information directly from companies and households (including individuals making up those households), through the undertaking of surveys to produce official statistics. In South Africa, legislation requires survey respondents to provide Stats SA with the information requested, and failure to do so is punishable by a fine or imprisonment, or both (South Africa, 1999).

Even though statistical agencies have the power to collect information as required from both natural and legal persons, there is a potential benefit to these agencies to utilise data collected in the course of public administration, including tax administration. The utilisation of administrative data can contribute both to incorporating the most complete and accurate information available in the statistical production process, and to minimising respondent burden. The advantages of utilising administrative records in producing statistics for these reasons are reflected in Principle 5 of the Fundamental Principles of Official Statistics (United Nations Statistical Commission, 1999)

It is thus not surprising that there are standing relationships between the tax and statistics authorities in many countries. South Africa is no exception.

^{*} Corresponding author. E-mail: eadam@sars.gov.za

This paper outlines the current and potential future roles of tax administrative data in the production of official statistics in South Africa. Section 2 describes the use of company tax records to construct a sampling frame for the production of economic surveys over the past decade. The publication of statistics on tax revenue collection and administration derived directly from administrative records is the focus of Section 3. In section 4 the ongoing collaboration between SARS and Stats SA is described, regarding the use of administrative records to complement data collected in surveys in areas such as labour market statistics and income statistics. Section 5 contains the main conclusions.

2. Tax administration data as frame for economic statistics surveys

A programme initiated in 1996 by Statistics South Africa to review the production of economic statistics included the exploration of using information from SARS to construct a new business register (Statistics South Africa, 2004). The initial idea of utilising tax administration records for the purposes of designing a business sample frame was rooted in concerns about the completeness and dated nature of the business register in use at that time.

The research on utilising tax administration data in statistical production occurred against a backdrop of a more ambitious proposal to streamline the collection of information required by government (Ertzaas, 2002). Aside from the advantages to statistical production of the availability of a common body of information already supplied to government in terms of a range of statutory obligations, elements of registration information were to become available to other participating government departments and public sector agencies too through a central database. This would, to a large extent, obviate the need to make *ad hoc* (and/or duplicate) requests for data from either other public entities or directly from companies themselves. The project proposal compiled by a team that included members of SARS and Stats SA, outlined the changes required to legislation to realise this freer flow of information, which would nonetheless protect confidentiality of information supplied by companies and individuals.

Stats SA published a new methodology for drawing samples for economic surveys in 2003, which utilised a business sample frame drawing on entities registered for Value Added Tax (VAT) (Statistics South Africa, 2003). Information provided monthly by SARS to Stats SA was to be used in the classification of enterprises, determining births and deaths of enterprises and to provide a measure of size for stratified sampling. The following year (2004) saw the first release of statistical series based on the new sample frame. The use of the new sample frame showed marked increases in the level of economic activity, particularly in some sectors such as manufacturing (Statistics South Africa 2004).

The above changes to the generation of economic statistics were not without technical challenges. For instance, a business enterprise, registered for income tax, might correspond to more than one VAT registered entity. Furthermore, the industry classifications used by SARS and Stats SA do not correspond, and this remains a challenge to date. The transition to a new national industry classification aligned to ISIC 4 is an additional challenge for the future.

Legislation, designed to protect the confidentiality of taxpayer information provided to SARS, had to be amended in order for Stats SA to use SARS information for statistical purposes. The special conditional provision of information by SARS to Stats SA, originally included in 2008 and amended in 2009, was incorporated in the Tax Administration Act in 2011 (South Africa, 2011) as follows:

*‘... A senior SARS official may disclose to—
(a) the Statistician-General the taxpayer information as may be required for the purpose of carrying out the Statistician-General’s duties to publish statistics in an anonymous form;...’*

A Memorandum of Understanding (MOU) between SARS and Stats SA signed in 2005 detailed arrangements for data transfers from SARS to Stats SA that still form the basis for the production of official economic statistics. This MOU is currently being revised, empowering SARS to become a

partner in the National Statistical System (NSS) co-ordinated by Stats SA. The NSS envisages the production of official statistics by organs of state other than Stats SA, provided that they meet prescribed quality criteria as outlined in the South African Statistical Quality Assurance Framework (SASQAF) (Statistics South Africa, 2010).

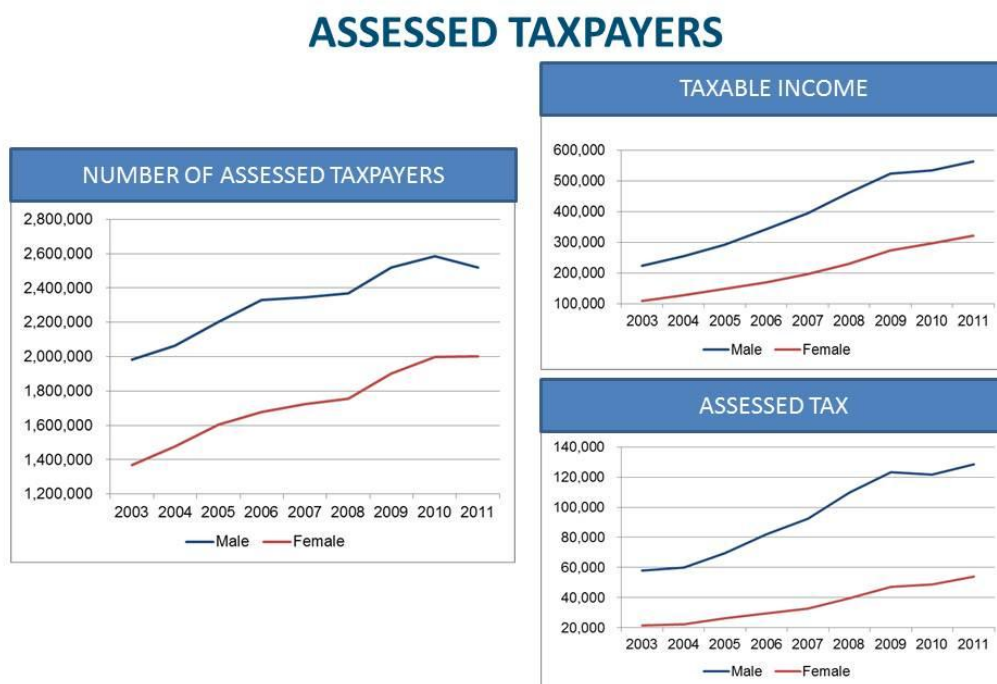
3. Tax statistics from tax administration data

Tax administrative records enable the compilation of tax statistics. Many countries publish statistics on tax revenue nowadays, and internationally comparable country-level tax revenue statistics are collated by international organisations such as the Organisation for Economic Co-operation and Development (OECD).

Printed books of statistical tables based on South African tax records were produced until 1991 and released to a broad range of recipients, including local and international universities, financial institutions, government departments and the general public. The publication of the *Statistics Bulletin* was suspended primarily as a result of changes to administrative systems.

The implementation of an extensive modernisation programme by SARS, which comprises a suite of interlinked technical innovations and process improvements, has laid the foundation for making tax administration data available, while steadily improving its quality. Electronic channels for the submission of returns by taxpayers and the use of third party data sources have led to the improved accuracy and completeness of tax records. Using electronic channels has also helped to improve timeliness of data availability due to the elimination of most post-submission scanning and/or key capture, and the marked improvements in pre-deadline submissions by taxpayers. For example, over 99% of personal income tax (PIT) returns were lodged electronically in 2011/12 and over a period of two years the proportion of individual taxpayers filing their tax returns on time increased by 44% (South African Revenue Service, 2012). An example of the kind of data available is shown in Figure 1 below.

Figure 1: Tax Bulletin data on assessed taxpayers



An annual *Tax Statistics Bulletin* was re-established in 2008, published jointly by SARS and the National Treasury. The *Tax Bulletin* is described in more detail in the contribution to the ISI's 58th World Statistical Congress by Breytenbach and Leolo (Breytenbach, 2011). Internal policy and standard operating procedures for this publication cover quality facets included in SASQAF. The *Tax Bulletin* with accompanying digital tables is now published online on the SARS website (South African Revenue Service, 2012) in October of each year.

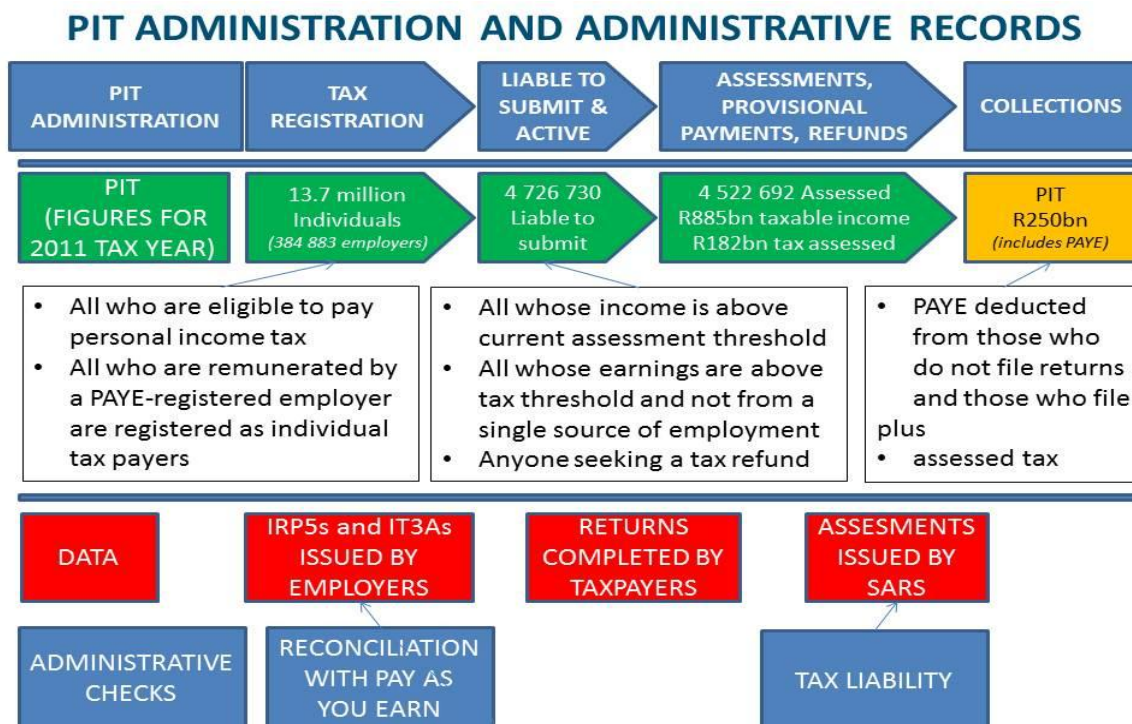
With five editions of the *Tax Statistics Bulletin* currently available, providing tax data time series spanning ten years, interest in the data continues to grow. SARS constantly receives requests from economists, policymakers and researchers for modifications and/or additions to the published data in order to meet user requirements. In addition to electronic channels for communication with users, workshops are convened to provide a forum for tax statistics users to share insights gained from analysing tax statistics and provide input regarding enhancements to future editions. Some user requests are based on potential uses of tax administration records that go beyond studying the dynamics of tax revenue. For example the breakdown of different income sources by taxable income groups has been requested, as well as the nature of deductions that predominate in the tax declarations of companies that show negative or zero taxable income.

4. Other statistics from tax administration data

The revision of the MOU between SARS and Stats SA acknowledges the potential application of data held by SARS in the direct production of statistics. This is over and above the role of tax administration data in providing continuously updated information on active taxpayers in South Africa, which underpins sample surveys. A joint SARS-Stats SA working group has been established to optimise the use of tax and customs administration data in the production of official statistics.

The potential application of PIT related records (see Figure 2 below) to the analysis of employment and income dynamics sparked renewed exploration of the additional uses of tax administration data in collaboration with Stats SA.

Figure 2: Administrative data available through PIT processing



Recent changes in the administration of PIT have significantly improved the quality of information on the remuneration of employees available through tax certificates (IRP5s and IT3(a)s), which employers issue to employees and provide to SARS. This quality improvement underpins an investigation of the picture of employment and income dynamics that can be provided by tax administration records. In principle, tax certificates provide demographic information about the recipient, together with the exact duration of employment of an individual with a particular employer and the associated remuneration. Sample surveys of households and companies are currently used by Stats SA to provide labour market statistics quarterly. The use of administrative data may thus provide a complementary picture to that of sample surveys. According to the National Development Plan the first major challenge facing South Africa is that *'too few people work'* (National Planning Commission, 2012). This is an indication of how important labour market measurements are to the South African government.

Given that sample surveys provide aggregate statistics pertaining to a particular reference period, administrative data could contribute to the understanding of patterns of employment of individuals in particular income brackets, as well as providing employment levels and job numbers associated with tax registered employers for various periods. Most of the process-embedded consistency checks on tax certificate data supplied by employers to SARS relate to remuneration and tax withheld. In addition, the quality of demographic and temporal information needs to be evaluated thoroughly to understand fully the possible limitations of such analyses.

Tax data also hold the potential to contribute to our understanding of income levels and dynamics, given the well-known limitations of obtaining individual and household income figures through household surveys. SARS has been asked to participate in a research programme on employment, income distribution and exclusive growth aimed at informing public policy. One of the projects proposed involves the blending of tax administration derived income information with that of household surveys, in order to build a micro-simulation model (Woollard, 2013).

A further example of the potential use of administrative records in generating statistics relates to spatial information provided as part of tax returns. A recently introduced requirement on employers to include annually-updated physical addresses of their employees on tax certificates adds a new geographical dimension to the data. The difficulty of defining where economic activity takes place on the basis of company-related information alone limits the level of geographical disaggregation to which economic data can be provided, given that companies would generally be associated with a single physical address. Given the spatial component to development in South Africa, data users frequently express the need for economic data at a sub-provincial level.

5. Conclusion

This paper has sketched the current role played by tax administration data in the production of official statistics in South Africa, as well as the initiatives that are underway to expand its usage. As the use of administrative data for generating statistics to be available in the public domain is extended, challenges arise relating to the adoption of standards and classifications that facilitate comparison with other data sources. The need to standardise industry classifications used within the tax administration is an example of this. However, the rigorous application of standards to administrative data also benefits the originating organisation by enabling analyses across previously non-comparable data sources.

As the quality of administrative records improves, driven by improvements in technologies and processes applied, further uses of these records will emerge. The potential use of the spatial component of tax records remains to be explored. SARS, together with Stats SA, will continue to explore such emerging opportunities.

The growing demand for statistics that facilitate development planning, coupled with statistical production resource constraints, guarantee increased reliance on administrative data in general. This is borne out by the past and current usage of tax administration data in South Africa.

Bibliography

Breytenbach, Deon and Leolo, Mamiky. 2011. *Tax Statistics in South Africa*. Proceedings of 58th World Statistical Congress.

Ertzaas, Ulf. 2002. *South African Register Centre*. Feasibility Study.

National Planning Commission. 2012. *National Development Plan*.
<http://www.npconline.co.za/pebble.asp?relid=25> . Last checked 15 April 2013.

South Africa. 1999. *Statistics Act 6 of 1999*. http://www.statssa.gov.za/about_statssa/statistics_act.asp. Last checked 15 April 2013.

South Africa. 2011. *Tax Administration Act 28 of 2011*. <http://www.sars.gov.za/home.asp?pid=5923>. Last checked 15 April 2013.

South African Revenue Service. 2012. *Annual Report 2011/12*.
<http://www.sars.gov.za/home.asp?pid=286> . Last checked 15 April 2013.

South African Revenue Service, 2012. *Tax Statistics Bulletin 2012*.
<http://www.sars.gov.za/taxstatistics> . Last checked 15 April 2013.

Statistics South Africa. 2004. *Gross Domestic Product 1st quarter 2004*. Statistical release.
<http://www.statssa.gov.za/publications/statsdownload.asp?PPN=P0441&SCH=2914> . Last checked 15 April 2013.

Statistics South Africa. 2004. *Manufacturing January, February, March 2004*. Statistical Release.
<http://www.statssa.gov.za/publications/statsdownload.asp?PPN=P3041.2&SCH=3097> . Last checked 15 April 2013. Statistics South Africa. 2003. *Sampling Methodology*. Discussion document.

<http://www.statssa.gov.za/Publications/statsdownload.asp?PPN=DiscussSamplingMeth> .Last checked 15 April 2013.

Statistics South Africa. 2010. *South African Statistical Quality Assessment Framework*.
www.statssa.gov.za/inside_statssa/standardisation/SASQAF_Edition_2.pdf. Last checked 15 April 2013.

United Nations Statistical Commission. 1999. *Fundamental Principles of Official Statistics*.
<http://unstats.un.org/unsd/methods/statorg/FP-English.htm> .Last checked 15 April 2013.

Woollard, Ingrid. 2013. *Development of a micro-simulation model*. Workshop presentation. (Slides in possession of author. Contact: eadam@sars.gov.za)