TECHNICAL ASSISTANCE REPORT

MALAWI

Report on the Improving Estimates of Gross Domestic Product Mission (June 24–July 5, 2024)

JULY 2024

Prepared By
Robin Youll

AFRITAC East

MEMBERS
Eritrea, Ethiopia, Kenya, Malawi, Rwanda, South Sudan, Tanzania, Uganda

PARTNERS FOR PHASE V
The Netherlands, United Kingdom, Norway, China, Switzerland, European Union, Germany

DISCLAIMER

The contents of this report constitute technical advice provided by the staff of the International Monetary Fund (IMF) to the National Statistical Office of Malawi (the "CD recipient") in response to their request for technical assistance. This report (in whole or in part) or summaries thereof may be disclosed by the IMF to IMF Executive Directors and members of their staff, as well as to other agencies or instrumentalities of the CD recipient, and upon their request, to World Bank staff, and other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of AFRITAC East, unless the CD recipient specifically objects to such disclosure (see Operational Guidance for the Dissemination of Capacity Development Information). Publication or Disclosure of this report (in whole or in part) or summaries thereof to parties outside the IMF other than agencies or instrumentalities of the CD recipient, World Bank staff, other technical assistance providers and donors with legitimate interest, including members of the Steering Committee of AFRITAC East, shall require the explicit consent of the CD recipient and the IMF's STA Department.

MEMBERS


Eritrea


Ethiopia


Kenya


Malawi


Rwanda


South Sudan


Tanzania


Uganda

PARTNERS FOR PHASE V

Ministry of Foreign Affairs of the Netherlands

The Netherlands

UKaid from the British people

United Kingdom

Norwegian Ministry of Foreign Affairs

Norway


China

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Economic Affairs SECO

Switzerland


European Union

Germany

Table of Contents

Acronyms and Abbreviations

BOP	Balance of Payments
BPD	Business Processing Documentation
CPI	Consumer Price Index
FCE	Final Consumption Expenditure
GDP	Gross Domestic Product
GDP-E	Gross Domestic Product by expenditure category
GDP-P	Gross Domestic Product by production activities
GFCE	General government Final Consumption Expenditure
GFCF	Gross Fixed Capital Formation
GG	General government
GVA	Gross value added
HFCE	Household Final Consumption Expenditure
IC	Intermediate consumption
ISIC	International Standard Industrial Classification of All Economic Activities
NPISH	Non-profit institutions serving households
NSO	National Statistical Office of Malawi
QGDP-P	Quarterly Gross Domestic Product by production activities
SNA	System of National Accounts
SSB	Statistics Norway
SUT	Supply and use table
TA	Technical assistance
VAT	Value Added Tax

Summary of Mission Outcomes and Priority Recommendations

1. A mission from AFRITAC East (AFE) visited the National Statistical Office (NSO) of Malawi between June 24–July 5, 2024 to assist with improvements to the quality of the published annual estimates of Gross Domestic Product based on the production approach (GDP-P), progress the use of Value Added Tax (VAT) data as a basis for developing quarterly series, and support development of current price estimates of GDP based on the expenditure approach (GDP-E). This mission builds on the work prior AFE missions which took place during 2001–2022.

2. The mission reviewed the basis for estimation of GDP-P, including the data sources, methodology, and compilation system used. Regarding data, very limited information is available from individual companies, with the main source being data captured by the NSO in the head office in Zomba from scanned copies of financial statements maintained by the Malawi Revenue Authority (MRA). For example, estimates for the manufacturing sector are based on financial returns for around 65 companies, compared with around 800 businesses registered with the MRA in this sector. For Wholesale and Retail trade, the sample covers just 10 businesses, compared to over 2,700 registered with MRA. The main methodological shortcoming is the estimation of value added by extrapolation of the base year (2017) values using output indicators (both nominal and quantity). This method is used because of lack of regular data on the costs incurred by businesses to produce output. That said, the compilation system used is well structured, logical, and well-maintained by the NSO team.

3. Beyond this initial assessment of the quality of the GDP series, the mission focused on three developments: finalizing Supply and Use Tables (SUT) for 2017; developing current price estimates of GDP-E; and initiating estimation of quarterly current price GDP-P for some activities based on data for company sales from the VAT system.

4. SUT for 2017 had been developed by the NSO but were never completed and were not consistent with the published estimates of value added by economic activity. Part of the issue was that there were some concerns with the published estimates and AFE recommended to fix the issues before proceeding. However, in order to progress on the work on national accounts it is best to move forward and address any issues in the historical period during the next GDP rebasing schedule which is being supported by the Norwegian statistics office (SSB). On this basis, the mission supported the compilation of fully balanced SUT for 2017, consistent with the published numbers and using the latest data from the Customs' department for external trade, government accounts, and balance of payments (BOP). After manual balancing, small imbalances were resolved using the IMF's automatic balancing tool (r.SUTB). The tool was installed on the laptops of the NSO staff. The balanced SUT for 2017 was then used as the basis for compiling annual SUT for 2017–22. These were based on annual data for external trade, BOP, government accounts, and published estimates for Non- Profit Institutions Serving Households (consistent with the series for GDP-P). The results in each year were constrained to the published estimates of value added, with Initial estimates of Household Final Consumption Expenditure (HFCE) and Gross Capital Formation (GCF) for each product being based on ratios from the 2017 SUT. The SUTs for each year were balanced using the automatic procedure.

5. Finally, the mission initiated the development of quarterly current price estimates of output based on data for companies reporting to the MRA VAT system. Monthly data for 'total supply' (broadly equivalent to sales) for around 6,000 companies had been acquired from the MRA.

These were integrated into a procedure for producing indexes of sales values by economic activity (based on the *International Standard Industrial Classification* Revision 4, ISIC). A method of pruned matched pairs was used for imputation of missing data, and a basis for identifying and correcting reporting errors established. Indexes of current price sales for 4-digit, 2-digit, and ISIC sections (A, B C, and so on) were prepared. These were benchmarked to the published value-added series using the Denton method. Analysis of the results was undertaken. As part of this development, a comprehensive business list was established, comprising around 8,000 business. Around 80 percent of these had been coded to ISIC activities before the mission (accounting for around 88 percent of reported sales over the period 2017–2022).

6. **Detailed Business Processing Documentation (BPD), together with video recordings describing the compilation procedure and methodology, were developed and provided to the national accounts team.** The data workbooks and the documentation were shared via OneDrive with the team. This will remain as shared resource with AFE.

7. **The NSO also receives Technical Assistance (TA) for the development of economic statistics from SSB.** This includes the deployment of a long-term advisor on national accounts. The mission met with advisor and agreed the basis for co-ordination of TA in the coming period.

8. **The NSO national accounts team comprises four members of staff, all of whom showed considerable dedication to improving the quality of the estimates of GDP and expanding the range of outputs produced.** The leadership in the office is also committed to supporting improvements. The technical assistance which will be provided by SBB should enable considerable improvements to be made in the coming period years. However, the IT infrastructure in the office is limited: there is no Wi-Fi and no central server, and most staff use their personal laptops, some of which are using very outdated software. The NSO plans to relocate staff to new premises in Lilongwe later in 2024, at which time some of these shortcomings should be resolved.

9. **To support progress in the above work areas, the mission proposed the following priority recommendations needed to improve estimates of GDP.**

10. **To support progress in the above work areas, the mission recommended a detailed one-year action plan with the following priority recommendations carrying particular weight to make headway in improving National Accounts quality.**

TABLE 1. Priority Recommendations

Target Date	Priority Recommendation	Responsible Institutions
September 2024	*Agree a Memorandum of Understanding (MoU) with the MRA for the provision of the VAT-based data. This should specify format of the data, and an agreed timetable for delivery.*	NSO
October 2024	*Finalize the classification of business in the register to ISIC 4-digit for all businesses which account for more than 0.1 percent of the totals sales in the register.*	NSO

Target Date	Priority Recommendation	Responsible Institutions
December 2024	*Publish the series for GDP by Expenditure components on an experimental basis, and seek feedback from users on its usefulness.*	NSO

11. Further details on the priority recommendations and the related actions/milestones can be found in the action plan under Detailed Technical Assessment and Recommendations.

Section I. Introduction

12. The NSO publish annual nominal and constant price (base year 2017) estimates of GDP-P in line with the ISIC Revision 4 classification. Estimates of GDP-E were last published in 2020 and were based on 2009 and are not consistent with the 2017 base GDP-P. Quarterly GDP estimates are not compiled in Malawi.

13. Remote missions in January and in July 2021 provided support for the development of annual estimates of GDP-E. However, as a result of continued lack of relevant data, the series were not complied. A further remote mission in December 2022 supported the NSO with the development of methods and tools for compiling higher frequency series based on VAT data. Key to this was the establishment of Memorandum of Understanding (MoU) with the MRA for sharing the VAT data and Income Tax data on a regular basis. To date, the MoU has not been signed by both institutions.

14. The current mission reviewed the basis for estimation of GDP-P, including the data sources, methodology, and compilation system used. Regarding data, very limited information is available from individual companies, with the main source being data captured by the NSO in the head office in Zomba from scanned copies of financial statements maintained by the Malawi Revenue Authority (MRA). For example, estimates for the manufacturing sector are based on financial returns for around 65 companies, compared with around 800 businesses registered with the MRA in this sector. For Wholesale and Retail trade, the sample covers just 10 businesses, compared to over 2,700 registered with MRA.

15. The main methodological shortcoming is the estimation of value added by extrapolation of the base year (2017) values using output indicators (both nominal and quantity). This method is used because of lack of regular data on the costs incurred by businesses in producing output. That said, the compilation system used is well structured, logical, and well-maintained by the team.

16. The mission focused on three developments: finalizing Supply and Use Tables (SUT) for 2017; developing current price estimates of GDP-E; and initiating estimation of quarterly current price GDP for some activities based on data for company sales from the VAT system. These developments are considered in the following sections.

Section II. Developing Supply and Use Tables for 2017

17. To facilitate the development of estimates of current price GDP-E (see next section) it was useful to first develop SUT for 2017. In fact, SUT for 2017 had been developed by the NSO with support from the International Labor Organization and AFE. However, the tables contained significant data anomalies and were never fully balanced. The results were also inconsistent with the published estimates of value added by economic activity.

18. The mission therefore supported the compilation of fully balanced SUT for 2017, consistent with the published estimates of GDP-P, despite the concerns with the published estimates noted earlier. This was achieved using the latest data from the Customs' department for external trade, government accounts data for taxes, and data from the BOP.

19. Some manual balancing of the resulting tables was undertaken, to deal with clear anomalies. For example, the original tables included an estimate of intermediate demand by the Construction activity which accounted for over half of the output of Construction services. This would normally be expected to be a relatively small value since it represents the repair and maintenance of buildings used by the construction activity. There were similar issues with the initial estimates of intermediate demand by the Education and Health activities for the services produced by these activities. Some training in the principles used to balance SUTs was undertaken as part of the data cleaning needed to produce near balanced tables for 2017.

20. Once the manual balancing was complete, the remaining small imbalances were resolved using the IMF's automatic balancing tool (r.SUTB[1]). To ensure that the NSO can replicate the results, the tool was installed on the laptops of the NSO staff.

21. Initial estimates of Household Final Consumption Expenditure (HFCE) and Gross Fixed Capital Formation (GCF) for each product were based on ratios from the 2017 SUT. Furthermore, Initial estimates of VAT were made based on the HFCE and the know rates for VAT on each product (either zero or 16.5 percent). Because the estimated VAT is dependent on the estimated HFCE, which itself is dependent on the estimated GFCF, it was necessary to develop a method to ensure that the resulting estimates for these three vectors in the SUT (VAT, HFCE and GFCF) produced balanced estimates of supply and use for each product. The basis of the calculation is provided in Appendix I. The resulting estimates of VAT were then scaled to the actual know level of non-deductible VAT from government revenue statements. This led to small imbalances on each product row of the SUT.

22. To remove these remaining imbalances the automatic procedure developed by the IMF Statistics Department referred to earlier (that is, r.SUTB) was used.

[1] This utilizes the Cholette-Dagum regression-based reconciliation method, and is undertaken in MS Excel using a procedure developed by the IMF Statistics Department which is incorporated into the Basic Excel R Tools (BERT). BERT provides a basis for incorporating applications developed using the 'R' statistical program into MS Excel.

Section III. Developing Annual Current Price Estimates of GDP-E

23. Once the SUT for 2017 was developed this was used as the basis of estimation of SUT for each year from 2017–2021. The procedure used was similar to that used for the 2017 SUT described in the previous section in that, for each year, relevant data for merchandise trade in goods, BOP, government taxes, GFCE, NPSIH FCE, and the known estimates of value added by economic activity were incorporated.

24. Near balanced tables were then complied using the methods set out in Appendix II.

25. Remaining imbalances were of the order of 1 percent or less. Final balancing to produce fully balanced tables was undertaken using the r.SUTB procedure. The results are shown in Table 2.

TABLE 2. Estimates of GDP (billions of Kwacha)

GDP by production activities	2017	**2018**	2019	**2020**	2021
Output	16.25	17.88	20.25	22.40	24.87
Intermediate consumption	10.12	11.11	12.51	14.07	15.52
Gross Value Added	6.13	6.77	7.74	8.33	9.35
Taxes on Products	0.41	0.47	0.50	0.49	0.63
Gross Domestic Product	**6.53**	**7.23**	**8.24**	**8.82**	**9.98**
GDP by Expenditure category					
Household Final Consumption Expenditure	4.64	5.81	6.45	7.74	8.75
General government Final Consumption Expenditure	1.47	1.60	1.83	1.83	2.06
Non-profit institutions serving households	0.31	0.35	0.41	0.45	0.52
Gross Capital Formation	1.33	0.82	0.91	0.26	0.30
Exports of Goods (FOB)	0.66	0.70	0.80	0.62	0.87
Exports of Services	0.16	0.16	0.27	0.32	0.36
Imports of Goods (FOB)	1.82	1.92	2.09	1.97	2.41
Imports of Services	0.21	0.29	0.34	0.44	0.48
Gross Domestic Product	**6.53**	**7.23**	**8.24**	**8.82**	**9.98**
Difference (GDP-P less GDP-E)	-	-	-	-	-

26. Detailed Business Process Documentation is provided in Appendix VI.

Recommendations:

- The NSO team should familiarize themselves with the details of the methodology for estimation of GDP-E and ensure that the relevant data required to produce estimates each year are acquired.
- Once done, the NSO should publish the series on an experimental basis, and seek feedback from users on its usefulness.

Section IV. Developing Quarterly Current Price Estimates of GDP-E

27. The mission also initiated the development of quarterly current price estimates of output based on data for companies reporting to the MRA VAT system.

28. Monthly data for 'total supply' (broadly equivalent to sales) for around 8,000 companies have been acquired from the MRA for the period January 2018 to September 2022. The mission supported the development of an integrated framework to take these data and produce indexes of current price sales values by economic activity (based on the *International Standard Industrial Classification* Revision 4, ISIC).

29. A previous AFE mission in December 2022 reviewed the quality of the VAT data and assessed their suitability for estimating quarterly indexes of economic output which could be used to produce estimates of quarterly GDP in current prices. One key output of that mission was to develop a basis for coding the companies in the VAT dataset to ISIC revision 4. Since the mission, the NSO team has coded around 80 percent of companies, accounting for some 88 percent of the total reported sales over the period 2018–2022.

30. The dataset from the MRA contains around 226,000 data points (covering the period January 2018–September 2022). On average, around 86 percent of months for active business contain data. That is, there are 'gaps' for around 14 percent of observations. Inspection of many of these suggests that business simply do not report in all periods but, when they do, their response relates to the sales in the reported month The dataset might therefore be considered as a large sample of responses with some unit non-response for some months.

31. To deal with this 'non-response,' a method of pruned matched pairs was used for imputation of missing data. The method is described in Appendix III. The 'pruning' here refers to the treatment of responses as outliers if the ratio of their sales in successive months is greater than some value, x, or less than $1/x$. A value of x=3 was found to be sufficient to remove the vast majority of extreme values (values between 3 and 10 were tested, and values above 3 made only very minor differences to the estimates). To be clear, these 'extreme' values are included in the estimates of total sales in each ISIC 2-digit heading in each period, but are not used in the calculation of the matched-pairs.

32. However, on inspection it was found that some of these extreme values were more likely to be reporting (or data capture) errors than actual levels of sales. For example, some responses were 100 or more time the reported values for the same businesses in any other period. A procedure for identifying potential errors of this kind was therefore established together with a basis for replacing them with estimates based on the judgment of the compilers (or removing them so that the responses would be imputed using the pruned matched-pairs method). Over the 226,000 data points, just 20 were considered to be extreme in this sense.

33. It is important to note that imputation of missing values based on any method should only be used as a last resort where it has not been possible to establish an actual value from other sources. With regard to the VAT dataset, there is generally no mechanism for the NSO to query the reported data with the companies directly. However, where there are missing or questionable responses for the largest companies in those activities where the VAT data are used for estimation (see Appendix IV), it may be possible to request clarification on the data from MRA.

34. Indexes of current price sales for 4-digit, 2-digit, and ISIC sections (A, B C, and so on) were prepared. These were benchmarked to the published annual value-added series using the Denton method.

35. As part of this development, a comprehensive business list was established, comprising all businesses which have reported to the VAT system between January 2018 and September 2022. As noted above, the business which have so far been assigned ISIC codes accounts for around 88 percent of reported sales over the period 2017–2022.

36. Detailed BPD for the compilation and updating of these estimates is provided in Appendix III. Examples of the quarterly indexes are provided in Appendix V.

37. With regard to how the VAT-based estimates might be used in the compilation of GDP-P there are a number of issues to consider.

38. First, it is important to note that VAT-based estimates are not appropriate to estimate output for some economic activities, for example the output of the Agricultural activities, Financial activities, Government and so on. Appendix IV identifies the activities where the VAT-based series might be considered as 'candidates' for estimation of changes in value added. In total this comprises 54 of the 88 ISIC 2-digit headings, and accounts for around 31 percent of total value added.

39. Second, it may be that the data content of the VAT-based series is considerably greater than that used to compile the annual estimates. For example, as noted above, for Wholesale and Retail trade, the sample used to estimate the annual changes in value added is based on just 10 businesses, while the VAT dataset contains over 2,700 in this activity. In such cases, it is legitimate to consider estimation of the annual series based on the 'sum of quarters' of the VAT-based estimates. This method has the advantage that there will be no revision when the annual estimates are used to benchmark the quarterly series (because they are the same). However, this method will still need to be benchmarked occasionally, normally whenever there are comprehensive rebasings based on SUT.

40. Third, to derive constant price series, the VAT-based series will need to be deflated using suitable price indexes. The existing GDP-P compilation system includes many such indexes, but only annually, so that equivalent quarterly series will need to be developed. For most activities the Consumer Price Index (CPI), or some component of that series is used to deflate the annual series. The series are available each month, so quarterly averages of the relevant indexes could be used.

Recommendations:

The NSO should:

- Agree a Memorandum of Understanding (MoU) with the MRA for the provision of the VAT-based data. This should specify format of the data, and an agreed timetable for delivery.
- Acquire the VAT data for all months of 2017. This will enable the series to be taken back to the current base period.
- Allocate ISIC 4-digit codes to all businesses which account for more than 0.1 percent of the totals sales in the business register in the compilation workbook (see Appendix III for details).
- Review the results and decide which series might be suitable as indicators of changes in value added.

- Review of the data content and quality of the existing GDP series to establish if the sum-of-quarters method based on the VAT sales data may produce more robust estimates than the existing annual methodology.

41. For series are to be estimated using the VAT data, identify quarterly deflators based on the annual deflators used in the current annual compilation.

Section V. Detailed Technical Assessment and Recommendations

Priority	Action/Milestone	Target Completion Date
Outcome*:** ***Milestone: Develop estimates of GDP by expenditure components		
H	The NSO team should familiarize themselves with the details of the methodology for estimation of GDP-E and ensure that the relevant data required to produce estimates each year are acquired.	September 2024
H	Once done, the NSO should publish the series on an experimental basis, and seek feedback from users on its usefulness.	December 2024
Outcome*:** ***Milestone: Develop estimates of GDP by quarterly GDP		
H	Agree a Memorandum of Understanding (MoU) with the MRA for the provision of the VAT-based data. This should specify format of the data, and an agreed timetable for delivery.	September 2024
H	Acquire the VAT data for all months of 2017. This will enable the series to be taken back to the current base period.	September 2024
H	Allocate ISIC 4-digit codes to all businesses which account for more than 0.1 percent of the totals sales in the business register in the compilation workbook (see Appendix III for details)	October 2024
M	Review the results and decide which series might be suitable as indicators of changes in value added.	January 2025
M	Review of the data content and quality of the existing GDP series to establish if the sum-of-quarters method based on the VAT sales data may produce more robust estimates than the existing annual methodology.	January 2025
M	For series are to be estimated using the VAT data, identify quarterly deflators based on the annual deflators used in the current annual compilation.	March 2025

A. OFFICIALS MET DURING THE MISSION

Name	Institution	E-mail Address
Mr. Shelton Kanyanda	Commissioner of Statistics, NSO	skanyanda@hotmail.com

Name	Institution	E-mail Address
Mr. Hector Kankuwe	Deputy Commissioner, NSO	cyrilyosefe@gmail.com
Mr. Simeon Yosefe	Deputy Commissioner, NSO	cyrilyosefe@gmail.com
Mr. Titus Nkhoma	Chief Statistician	titusnkhoma@gmail.com
Mrs. Hannah Chabwera	Principle Statistician	chabwerahu@yahoo.com
Mr. Godfrey Bvutani	Statistician	gbvutani@gmail.com
Mr. Aliko Mwenelupembe	Statistician	alimwene@gmail.com
Mr. Håvard Sjølie	SSB, Long-term national accounts adviser to NSO	Havard.Sjolie@ssb.no

Appendix I. Method for Estimation of GCF and HFCE in an Unbalanced SUT

Define (for each product):

'other' supply as

$$\acute{\boldsymbol{S}} = \boldsymbol{Output} + \boldsymbol{Margins} + \boldsymbol{Imports} + \boldsymbol{Other\ taxes}$$

and 'other' uses (excluding HFCE and GCF)

$$\acute{\boldsymbol{U}} = \boldsymbol{ID} + \boldsymbol{GFCE} + \boldsymbol{NPISH} + \boldsymbol{X}$$

where $\boldsymbol{ID}$ is the intermediate demand for the product, and $\boldsymbol{X}$ is the exports.

and $\boldsymbol{r}$ as the rate of VAT on a product

We have

$$\acute{\boldsymbol{S}} + V = \acute{\boldsymbol{U}} + GCF + H \qquad ..(1)$$

and

$$V = rH$$

where H is the value of HFCE to be estimated, and V is the non-deductible VAT.

We wish to estimate H, GCF, and V to satisfy equation (1)

Let $P_B = \frac{GFC_B}{GFC_B + H_B}$, that is the proportion of GCF in GCF + HFCE in the base year, and estimate GCF as

$$GCF = P_B(GFC + H)$$

It can be shown that:

$$GCF = \frac{P_B\left[\frac{(\acute{\boldsymbol{S}} - \acute{\boldsymbol{U}})}{1-r}\right]}{1 + \frac{rP_B}{1-r}}$$

and

$$H = \frac{\acute{\boldsymbol{S}} - \acute{\boldsymbol{U}} - GCF}{1-r}$$

This ensures that

$$Output + Margin + Imports + Other\ taxes + V = ID + H + GFCE + NPISH + GCF + X$$

i.e., each product is fully balanced.

Appendix II. Method for Initial Estimation of SUT Vectors before Final Automatic Balancing

Vector/Matrix	Data /Method
SUPPLY	
Domestic Supply (by product and activity)	Estimates of value added from the GDP compilation system classified by product divided by (1 - the 2017 IC/O ratio)
Trade and transport margins	Based on the margin rates from the 2017 SUT
VAT	Initial estimates based on the estimated HFCE multiplied by the known VAT rates on each product and then scaled to the known total for non-deductible VAT from the government accounts.
Trade taxes	Based on reported taxes in the detailed Customs' trade data, aggregated to SUT product, and scaled to the known taxes for the government revenue statements
Imports of Goods (CIF)	Based on reported FOB value of imports from the detailed Customs' trade data aggregated to SUT product and scaled to the imports (FOB) from the BOP. Insurance and freight charges were added for each product based on the known total and the detailed reported values in the Customs' data.
Imports of Services	Based on the BOP.
CIF/FOB adjustment	The total estimated Freight charges within the value of Imports CIF was allocated to the various mode of transport services (land, sea and air) according to the shares of output for these products in each year. Insurance costs within the CIF cost were allocated to the product 'non-life insurance'.
USE	
Intermediate Demand (by product and activity)	Initial estimates of total Intermediate Demand (ID) across all activities based on the share of ID within (ID + Exports) from the base year. Estimates of ID by product and economic activity achieved by constraining estimates based on the pattern of ID for each product from the base year to the estimated total ID for each activity (that is, the intermediate consumption) estimated from the published value added for each activity in the relevant year.
GFCE	Based on the estimated government output from the GDP-P compilation system
NPISH FCE	Based on the estimated NPISH output from the GDP-P compilation system
HFCE and GFCF	Initial estimates based on the shares of the residual uses (not accounted for elsewhere) according to the method set out in Appendix I
Change in inventories	Assumed to be zero in each year
Exports of goods	Using the same method as for imports of goods.
Exports of services	Based on the BOP

Appendix III. Business Processing Documentation for Estimation of Quarterly GDP based on VAT Data

Introduction

This workbook creates quarterly current price indexes of sales by economic activity. It produces monthly and quarterly index for 2-digit and Section level ISIC headings.

The input data are based on VAT returns for sales. Missing values are imputed using a pruned matched pairs methodology. See Box 1. The 'pruning' here refers to treating responses as outliers if the ratio of the sales in successive months was greater than some value, x, or less than 1/x. A value of x=3 was found to be sufficient to remove the vast majority of extreme values (values between 3 and 10 were tested, and values above 3 made only very minor differences to the estimates). To be clear, these 'extreme' values are included in the estimates of total sales in each SIIC 2-digit heading in each period, but are not used in the calculation of the matched-pairs imputation factors described in Box 1.

In addition, a procedure for identifying potential errors is included in the workbook together with a basis for replacing them with estimates based on the judgment of the compilers (or removing them so that the responses would be imputed using the matched-pairs method). See worksheets in group QA below.

Finally, the workbook benchmarks the indexes to published series (see worksheet BM_Qtly). It is not proposed to use the VAT data to estimate quarterly series for all activities as sales is not appropriate as a measure of activity for some. The activities in Appendix I may be suitable to estimation based on VAT sales:

This document describes each of the worksheet in the workbook and provides a guide to the updating of the estimates.

Box 1. A Note on the Method of Matched Pairs

The matched-pairs method involves identifying businesses which reported in each of two successive periods (for example, in both January and February) and dividing the total reported turnover for these businesses in the second period (in this case February) by the equivalent total for the same businesses in the previous period (in this case January).

This creates aa 'forward' matched pairs factor between January and February. This is then used to impute missing values for February for business which reported in January but not in February, with the estimate for February being calculated as the product of the January value multiplied by the matched-pairs factor between January-February.

In addition, factors are calculated 'backwards', based on the inverse of the matched-pairs ratios. This was needed because if a business reported in January and March, say, the 'forwards' matched pairs estimate based on the matched-pair factor between January and February would result in one estimate of the February value, while an estimate based on the reported March figure divided by the matched-pairs factor between February and March.

The final imputed estimate for February, in this example, was based on the average of the imputations 'forward' from January and 'backwards' from March.

The procedure was carried out for each ISIC 2-digt heading where the total number of matched responses was greater than a value selected by the compiler– a value of 10 was used. So that, If there were fewer than 10 matched responses in any 2-fdigit ISIC division the matched-pairs growth factors were based on the same procedure but at the level of ISIC sections rather than 2-digit headings.

The worksheets

The workbook comprises eight groups of sheets, as presented below:

Registry	Analysis	Series	QA	Imputation	Time_Series	Raw Data	Meta
Reg	BM_Qtly	Ind_Sect_Q	Adju	Imputed_Avg	Pvt_Missing	Make_DB	ISIC_Full
Reg_New	Comp	Ind_Sect_M	Ind_ISIC_2	Imputed_Bck	Pvt_O	Data	Sheets
		Sect_Summary		Imputed_Fwd		2018	
		Ind_2_Dig_A		Last_Obs		2019	
		Ind_2_Dig_Q		First_Obs		2020	
		Ind_2_Dig_M		MP_2Dig		2021	
		2_Dig_Summary		MP_Section			

These are described from right to left, below

Group: Meta

Comprises 2 worksheets

ISIC_Full
Sheets

ISIC_Full is the complete list of ISIC Rev 4 headings for sections, divisions (2-digits) and 4-digits.

Sheets: is just the list of all worksheets in the workbook

Group: Raw Data

Comprises 6 worksheets

Make_DB
Data
2018
2019
2020
2021

The data in 2018–2021 are based on a previous format of the VAT file. They do not need to be updated.

Data contains the data in the latest format. This needs to be updated as more data becomes available. The data needs to be formatted like this:

TPIN	Year	Month	Sales	Trading name
20101101	2022	1	1,313,131	AGASON MOTORS
20101101	2022	4	35,141,550	AGASON MOTORS
20101101	2022	7	1,809,600	AGASON MOTORS

Important note: it is critical to ensure the data contain no duplicate records for the same TPIN in the same period. The preparation of the data in this form is outside of the scope of this document. Ideally, the MRA should provide the data precisely in this format.

Make_Data takes the data in all years and makes a single 'database' arrangement of the data (as per the **Data** worksheet), see columns AV:AY. This is used as the basis of the pivot tables in the workbook

Group: Time_Series
Comprises 2 worksheets

Pvt_Missing
Pvt_O

These are pivot tables based on the **Make_Data** worksheet.
Pvt_O contains only those businesses in the registry worksheet (**Reg**, see later). It presents a time series for each business that adds in the appropriate ISIC codes based on the **Reg** worksheet. The 'O' in Pvt_O stands for Output.
Pvt_Missing contains the same data, but for those business not currently in the **Reg** worksheet.

Group: Imputation
Comprises 8 worksheets

Imputed_Avg
Imputed_Bck
Imputed_Fwd
Last_Obs
First_Obs
MP_2Dig
MP_Section

These undertake the pruned matched pairs imputation, as described in Appendix I.

MP_Section

Calculates the matched-pairs growth factors for each ISIC Section.

Observations of the ratio of sales for any individual business between two successive periods which outside of the range (x, 1/x) are lift out of the factors. This is the 'pruning'. The value of x is determined on the **Menu** worksheet. The factors are calculated from row 52. The worksheet also estimates the growth factors for any section only is the sample of matched pairs observations is greater than the value in cell I1 (determined on the **Menu** worksheet). The sample, sizes are shown from row 78.

MP_2Dig

Calculates the matched-pairs growth factors for each ISIC 2-digit division. The procedure is the same as for **MP_Section**. If the sample size is less than the prescribed value, the factor from the appropriate Section level (from **MP_Section** is used).

First_Obs

Puts the value '1' in every period after the first observation for each business. This is used in the **Imputed_Avg** worksheet, see below.

Last _Obs

Puts the value '1' in every period before the last observation for each business. This is used in the **Imputed_Avg** worksheet, see below.

Imputed_Fwd

Imputes missing values 'forwards' by multiplying the previous value by the matched-pairs growth factors from **MP_2Dig.** See Appendix 1.

Imputed_Bck

Imputes missing values 'backwards' by dividing the next value by the matched-pairs growth factors from **MP_2Dig.** See Appendix 1.

Imputed_Avg

Takes the average of the forward and backwards imputed values. Note that imputations before the first observation and after the last one are not included (this is where the worksheets **Last _Obs** and **Last _Obs** are used).

Group: QA
Comprises 2 worksheets

Adju
Ind_ISIC_2

Ind_ISIC_2: This allows the user to select a specific 2-digit ISIC and review the data, including the imputed values. Its purpose is to allow the user to quality assure each activity. Note: The chart will need to be updated to include the additional periods as they appear.

If an odd value is found, it can be adjusted using the **Adju** worksheet.
Adju this worksheet is used to store any manual adjustments made to the data. The suer needs to enter the TPIN, Year and Month and the estimated value of sales.

Note: To incorporate the adjusted values into the final estimates the user should click on the

Replace outliers

button. This may take a few moments to update (it's updating the Pvt_O workwise pivot table and then recalculating the workbook based on the updated numbers. It may therefore be best to make all adjustments needed for all activities and then update.

Group: Series
Comprises 8 worksheets

Ind_Sect_Q
Ind_Sect_M
Sect_Summary
Ind_2_Dig_A
Ind_2_Dig_Q
Ind_2_Dig_M
2_Dig_Summary

These are the results at various levels both in MKW and presented as indexes.
2_Dig_Summary This is a 2-digit aggregation of the **Imputed_Avg** series

Ind_2-Dig_M This presents the same thing as indexes (the reference period is set on the **Menu** worksheet)
Ind_2-Dig_Q This presents the **Ind_2-Dig_M** as average quarterly indexes
Ind_2-Dig_A This presents the **Ind_2-Dig_A** as average annual indexes
Sect_Summary This is a Section level aggregation of the **Imputed_Avg** series
Ind_Sect_M This presents the same thing as indexes
Ind_Sect_Q This presents the **Ind_2-Sect_M** as average quarterly indexes

Group: Analysis

Comprises 2 worksheets

BM_Qtly
Comp

This group compares the published series for value added with the matched-pairs series.

Comp compares the published series for value added with the matched-pairs series. **Note that it is linked to the workbook:**
[GDP Rebased Series_2017 Base.xls]Summary Internal1

Note: Make sure that the link is working OK, and relink if necessary (using Data-Edit links ->Change Source.) if necessary.

BM_Qtly use the XLPBM procedure to benchmark the quarterly series to the published annual value added series. This is just to update the degree of 'fit' of the published and VAT-based series. This will need to be developed once it is agreed how the VAT series will be used in the compilation of GDP.

Group: Registry
Comprises 2 worksheets
Comprises 2 worksheets

Reg
Reg_New

Reg is the 'registry' or business list used in the system. Businesses have ben populated based on their appearing in the VAT dataset. They need to be coded to ISIC Rev 4.

Reg_New When new data are loaded, if they do not appear in the **Reg** worksheet they will be shown here. They are ordered by their size. The largest ones should be copied to the **Reg** worksheet (just add them at the end) and given appropriate ISIC codes.

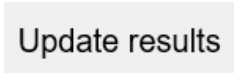
Updating the workbook

To update the workbook follow these steps:

1. Acquire data for the latest months from the MRA The data need to be prepared so that the format corresponds with that in the **Data** worksheet (see above).

   Critical note: You must ensure that there are no duplicate records in the data, that is duplicates of the TPIN + Year + Month

   Note: Also, do not delete any 'old' data. The new data should just be added at below the previous data, in exactly the same format.

2. Click on the 

button (which will update the pivot tables in **Pvt_O** and **Pvt_Missing)** and recalculate the results, including new data.

3. Quality assures the results and make any corrections for extreme outliers using the worksheets in the QA group, above. This is best undertaken by selecting each activity in the Ind_ISIC_2 worksheet and looking for oddities in the charts. If there are 'spikes' for example, identifying on the chart for which period these occur, and then review the data in the relevant column to see which business has provided an odd response (there may be more than one, of course). Then use the Adju worksheet to replace the reported value for that business(s) with an estimated value (which can be based on the judgment of the compiler)

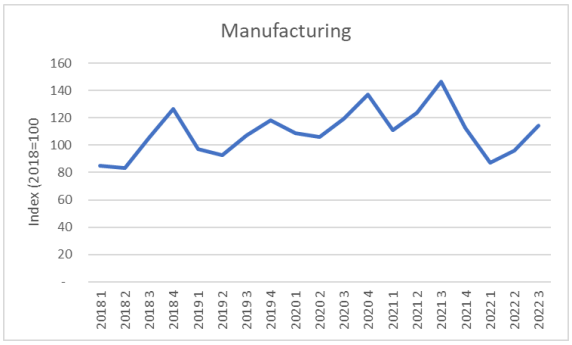
Note: it may be necessary to copy the formulas in many of the worksheets 'to the right', especially in the groups **Imputation** and **Series**.

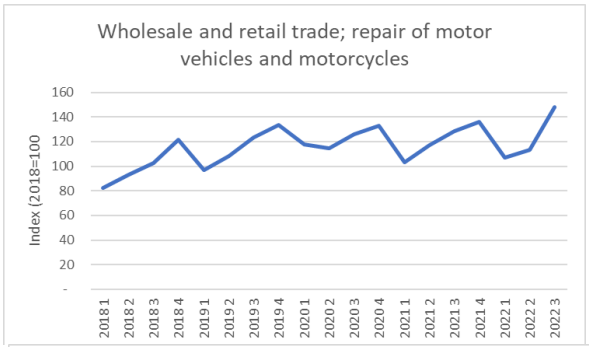
Appendix IV. Economic Activities which May Be Suitable to Estimation Using VAT Data

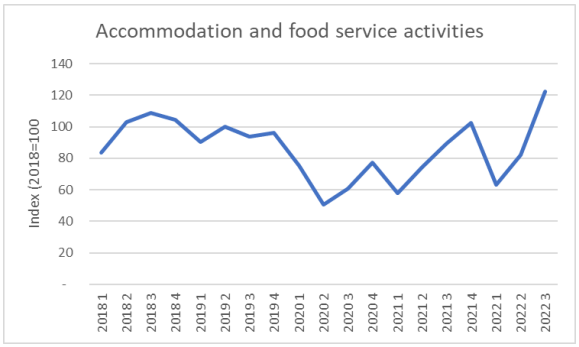
SIC	Description
10	Manufacture of food products
11	Manufacture of beverages
12	Manufacture of tobacco products
13	Manufacture of textiles
14	Manufacture of wearing apparel
15	Manufacture of leather and related products
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
17	Manufacture of paper and paper products
18	Printing and reproduction of recorded media
19	Manufacture of coke and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22	Manufacture of rubber and plastics products
23	Manufacture of other non-metallic mineral products
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29	Manufacture of motor vehicles, trailers, and semi-trailers
30	Manufacture of other transport equipment
31	Manufacture of furniture
32	Other manufacturing
33	Repair and installation of machinery and equipment
45	Wholesale and retail trade and repair of motor vehicles and motorcycles
46	Wholesale trade, except of motor vehicles and motorcycles
47	Retail trade, except of motor vehicles and motorcycles
55	Accommodation
56	Food and beverage service activities
58	Publishing activities
59	Motion picture, video and television programme production, sound recording and music publishing activities
60	Programming and broadcasting activities
62	Computer programming, consultancy, and related activities
63	Information service activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
75	Veterinary activities

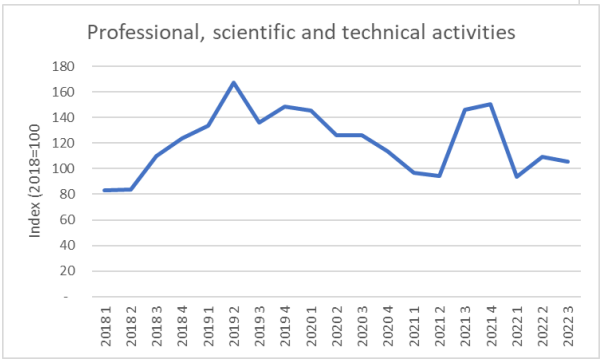
SIC	Description
77	Rental and leasing activities
78	Employment activities
79	Travel agency, tour operator, reservation service and related activities
80	Security and investigation activities
81	Services to buildings and landscape activities
82	Office administrative, office support and other business support activities
90	Creative, arts and entertainment activities
91	Libraries, archives, museums and other cultural activities
92	Gambling and betting activities
93	Sports activities and amusement and recreation activities
94	Activities of membership organizations
95	Repair of computers and personal and household goods
96	Other personal service activities

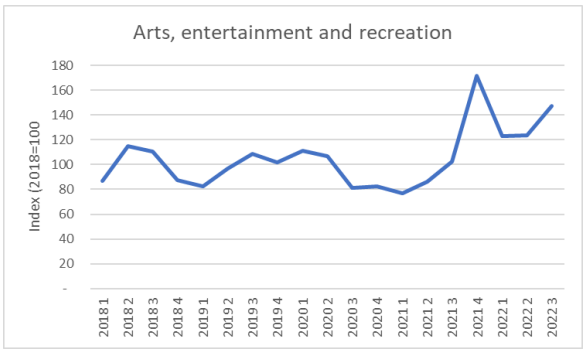
Appendix V. Example Quarterly Indexes of VAT Sales

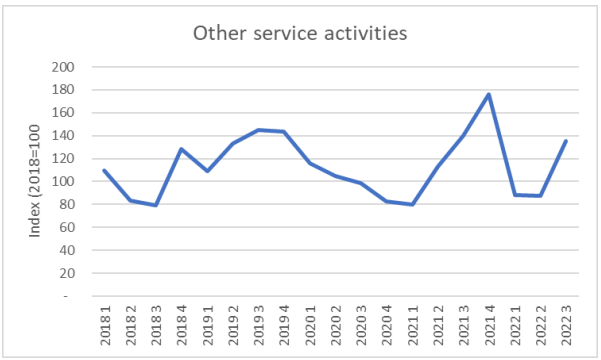












Appendix VI. Business Processing Documentation for Estimation of GDP Based on Expenditure Components

Introduction

This workbook calculates estimates of GDP by expenditure components in current prices. The workbook consists of the following groups of worksheets

Results	Data	Make_SUT	Trdae Vecto	Trade Data	Lookups	SUT
Results	BoP	SUT	CIF_FOB	X_SUT_TS	HS_Desc	Supply
	Merc_Trade	ID_SUT	TiS	X_HS_TS	HS_SUT	Use
	Pub_VA	O_SUT	TiG	CDTax_SUT_TS	Act_Prod	
		CF		CDTax_SUT_TS	SUT_P	
		VAT		CDTax_HS_TS		
		Excise_Dom		ExTax_SUT_TS		
		Imp_Tax		EXTax_HS_TS		
		Margins		M_Ins_SUT_TS		
		O_Non_Mkt		M_Ins_HS_TS		
		O_Prod		M_Fr_SUT_TS		
		VA_Prod		M_Fr_HS_TS		
		VA_SUT		M_SUT_TS		
		U_other		M_HS_TS		
		ID_Prod				

The purpose of these are explained in the following section

The worksheet groups

Group: **SUT**

This links to the 2017 Supply and Use tables and is used as a base for the calculations, for example the margin rates from the SUT are assumed fixed in each year.

Group: **Lookups**

These are the codes used in the calculations, including the relationship between the codes, for example, the links between HS 6-digit codes and SUT product codes.

Group: **Trade Data**

This group calculates time series for the various elements of the trade in goods series for imports and exports in the form of pivot tables. The pivot tables are then aggregated to SUT product level. In this group 'X' is exports, 'M' Imports, HS_TS implies a time series at the HS level and SUT_TS is aggregated to SUT products.

For example, M_HS_TS is the pivot table for Exports based on the HS code, and M_SUT_TS is the aggregated series for Imports up to SUT Product level

The pivot tables can be updated automatically using the Update trade data button on the Instructions worksheet.

Group: **Trade Vectors**

This group calculates the final trade in goods vectors for the SUT.

The worksheet **TiG** scales the mercantile trade and taxes from the group **Trade Data** to the totals from the Balance of Payments. Note: the Balance of Payments data are stored in the worksheet **BOP** (see later).

The worksheet **TiS** allocates the total for Trade in Services from the BOP based on the base year pattern for the group **SUT**

Group: **Make SUT**
Is used to create the matrixes and vectors for the year selected in the worksheet **Results**.

Taking the worksheets in turn

ID_Prod estimates Intermediate Demand (ID) by SUT Product based on the proportion of the ID in the base year compared to the total supply in the base year less exports, multiplied by the estimated total Supply in the currently selected year less the ID and Exports in the current year.

U_other estimates the components of the commodity floe for HFCE and GFCF, based on the algorithm in Appendix .

VA_SUT this just holds the data form the published series for Value Added (VA) in the base year. It is used as the basis of the estimates of Output and IC by activity in the next 2 worksheets estimates

VA_Prod golds the estimates the VA in each year based on the published estimates of VA in each year

O_Prod estimates the Output in each year based on the fixed IC/O ratios and the published estimates of VA in each year

O_Non_Mkt stores the non-market Output for each year for each sector.

Margins estimates margins based on the margin rates in the base year (multiplied by the sum of Output + Imports + import taxes)

Imp_Tax estimates import taxes by product in each year based on the trade data, constrained to the reported taxes in the government accounts

Excise_Dom estimates domestic excise based on the government accounts

VAT holds the reported VAT from the government accounts

CF is a commodity flow which brings together all the vectors for Supply and use for the year selected in the **Results** worksheet.

O_SUT is a complete matrix of the Output by activity and product for the year selected in the **Results** worksheet.

ID_SUT is a complete matrix of the Int. Demand by activity and product for the year selected in the **Results** worksheet.

SUT is a complete SUT by activity and product for the year selected in the **Results** worksheet.

Group: **Data**
This group is where the user need sot put the data to estimate the GDP-E series. There are three worksheets:

BOP are the data from the balance of payments

Merc_Trade are the data from the trade in good series

Pub_VA links to the workbook [GDP Rebased Series_2017 Base.xls] which contains the published VA series.

Group: **Results**

Consists of one worksheet:

Results This allows the user to select year and the result will be updated automatically.

Updating the workbook

To update the workbook

1. Acquire the data needed in the worksheets in the **Data** group from the BOP, merchandise trade statistics and update those worksheets and make sure the link to live (and current) GDP compilation workbook [GDP Rebased Series_2017 Base.xls] (use 'Data-Edit Links->Check status. The workbook should appear as 'OK').

2. Click on the button on the Instructions worksheet to update the pivot tables needed for the trade in good data.

3. Select the year you require in the Results worksheet.

4. Value copy the results from column B to the required column in the time series.