

The Malawi 2014 Enterprise Surveys Data Set

I. Introduction

1. This document provides additional information on the data collected in Malawi between April 2014 and February 2015 under, an initiative of the World Bank. As part of its strategic goal of building a climate for investment, job creation, and sustainable growth, the World Bank has promoted improving business environments as a key strategy for development, which has led to a systematic effort in collecting enterprise data across countries. The Enterprise Surveys (ES) are an ongoing World Bank project in collecting both objective data based on firms' experiences and enterprises' perception of the environment in which they operate.

The Enterprise Surveys currently cover over 130,000 firms in 135 countries, of which 121 have been surveyed following a standard methodology. This allows for better comparisons across countries and across time. Data are used to create statistically significant business environment indicators that are comparable across countries. The Enterprise Surveys are also used to build a panel of enterprise data that will make it possible to track changes in the business environment over time and allow, for example, impact assessments of reforms.

The report outlines and describes the sampling design of the data, the data set structure as well as additional information that may be useful when using the data, such as information on non-response cases and the appropriate use of the weights.

II. Sampling Structure

2. The sample for Malawi was selected using stratified random sampling, following the methodology explained in the *Sampling Manual*¹. Stratified random sampling² was preferred over simple random sampling for several reasons³:

a. To obtain unbiased estimates for different subdivisions of the population with some known level of precision.

b. To obtain unbiased estimates for the whole population. The whole population, or universe of the study, is the non-agricultural economy. It comprises: all manufacturing sectors according to the group classification of ISIC Revision 3.1: (group D), construction sector (group F), services sector (groups G and H), and transport, storage, and communications sector (group I). Note that this definition excludes the following sectors: financial intermediation (group J), real estate and renting activities (group K, except sub-sector 72, IT, which was added to the population under study), and all public or utilities-sectors.

c. To make sure that the final total sample includes establishments from all different sectors and that it is not concentrated in one or two of industries/sizes/regions.

d. To exploit the benefits of stratified sampling where population estimates, in most cases, will be more precise than using a simple random sampling method (i.e., lower standard errors, other things being equal.)

e. Stratification may produce a smaller bound on the error of estimation than would be produced by a simple random sample of the same size. This result is particularly true if measurements within strata are homogeneous.

f. The cost per observation in the survey may be reduced by stratification of the population elements into convenient groupings.

¹ The complete text can be found at http://www.enterprisesurveys.org/documents/Implementation_note.pdf

² A stratified random sample is one obtained by separating the population elements into non-overlapping groups, called strata, and then selecting a simple random sample from each stratum. (Richard L. Scheaffer; Mendenhall, W.; Lyman, R., "Elementary Survey Sampling", Fifth Edition).

³ Cochran, W., 1977, pp. 89; Lohr, Sharon, 1999, pp. 95

3. Three levels of stratification were used in this country: industry, region, and size was not available in the sampling frame for most contacts. The original sample design with specific information of the industries and regions chosen is described in Appendix E.

4. Industry stratification was designed in the way that follows: For panel firms, the universe was stratified into manufacturing industries and a services sector. For fresh firms, the universe was stratified into manufacturing industries and two service sectors (retail and other services).

5. Regional stratification for the Malawi ES was defined by 6 metro areas:

- Blantyre
- Kasungu
- Lilongwe
- Mangochi
- Mzimba
- Zombe

III. Sampling implementation

6. Given the stratified design, sample frames containing a complete and updated list of establishments as well as information on all stratification variables (number of employees, industry, and region) are required to draw the sample.

7. The international firm of Ipsos was hired to conduct the survey and they partnered with local agency Knowledge Research and Consulting in Malawi. The study began with a different subcontractor that was terminated midway through the project. Subsequently, another subcontractor was hired but failed to complete the project before quitting (see appendix for details). Knowledge Research & Consulting was brought in to finalize the survey as they had recently completed the same project in Namibia successfully.

8. For the Malawi ES, multiple sample frames were used: a sample frame was built using data compiled from local and municipal business registries. Due to the fact that the previous round of surveys utilized different stratification criteria in the 2009 survey sample, the following convention was used. The presence of panel firms was limited to a maximum of 50% of the achieved interviews in each cell. That sample is referred to as the Panel.

Malawi, Sample Frame

Panel

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	6	8
	Medium	7	7
	Large	6	9
Northern (Mzimba)	Small	4	3
	Medium	2	3
	Large	1	1
Southern (Blantyre, Zomba, Mangochi)	Small	11	15
	Medium	20	15
	Large	17	15
	Total	74	76

Fresh

	Size	Manufacturing	Retail	Other services
Blantyre	Small	149	55	576
	Medium	173	135	489
	Large	208	232	417
Kasungu	Small	4	0	3
	Medium	3	0	2
	Large	2	0	2
Lilongwe	Small	141	191	518
	Medium	74	88	305
	Large	90	122	332
Mangochi	Small	8	4	15
	Medium	5	6	8
	Large	1	0	7
Mzimba	Small	12	2	43
	Medium	5	0	24
	Large	9	7	12
Zomba	Small	15	22	19
	Medium	6	1	19
	Large	5	5	10
	Total	910	870	2801

10. The sample design for the Malawi Enterprise Survey was generated with the aim of obtaining interviews at 460 establishments.

11. The quality of the frame was assessed at the onset of the project through visits to a random subset of firms and local contractor knowledge. The sample frame was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc. The local contractor had to screen the contacts by visiting them which resulted in slow fieldwork in many cases.

12. Given the impact that non-eligible units included in the sample universe may have on the results, adjustments may be needed when computing the appropriate weights for individual observations. Breaking down by stratified industries, the following sample targets were achieved:

Achieved panel sample

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	4	5
	Medium	2	8
	Large	4	7
Northern (Mzimba)	Small	1	
	Medium		
	Large		
Southern (Blantyre, Zomba, Mangochi)	Small	7	6
	Medium	19	6
	Large	9	9
Total		46	41

Achieved fresh sample

	Size	Manufacturing	Retail	Other services
Blantyre	Small	20	17	34
	Medium	18	23	26
	Large	25	24	16
Kasungu	Small			1
	Medium	2		
	Large			1
Lilongwe	Small	13	17	18
	Medium	24	20	16
	Large	20	13	12
Mangochi	Small	1	1	2
	Medium		3	3
	Large	1		2
Mzimba	Small	1	1	15
	Medium	2		2
	Large	3	2	3
Zomba	Small	2	5	5
	Medium	3	1	10
	Large	2	2	4
	Total	137	129	170

IV. Data Base Structure:

13. The structure of the data base reflects the fact that 2 different versions of the survey instrument were used for all registered establishments. Questionnaires have common questions and respectfully additional manufacturing and services specific questions. The eligible manufacturing industries have been surveyed using the *Manufacturing* questionnaire (includes a common set of core variables, plus manufacturing specific questions). Eligible services have been covered using the *Services* questionnaire. Each variation of the questionnaire is identified by the index variable, *a0*.

14. All variables are named using, first, the letter of each section and, second, the number of the variable within the section, i.e. *a1* denotes section A, question 1 (some exceptions apply due to comparability reasons). Variable names preceded by a prefix "SL" indicate questions specific to Malawi, therefore, they may not be found in the implementation of the rollout in other countries. All other suffixed variables are global and are present in all country surveys over the world. All variables are numeric with the exception of those variables with an "x" at the end of their names. The suffix "x" denotes that the variable is alpha-numeric.

15. There are 2 establishment identifiers, *idstd* and *id*. The first is a global unique identifier. The second is a country unique identifier. The variables *a2* (sampling region), *a6a* (sampling establishment's size), and *a4a* (sampling sector) contain the establishment's classification into the strata chosen for each country using information from the sample frame. The strata were defined according to the guidelines described above.

16. There are three levels of stratification: industry, size, and region. Different combinations of these variables generate the strata cells for each industry/region/size combination. A distinction should be made between the variable *a4a* and *d1a2* (industry expressed as ISIC rev. 3.1 code). The former gives the establishment's classification into one of the chosen industry-strata, whereas the latter gives the actual establishment's industry classification (four digit code) in the sample frame.

17. All of the following variables contain information from the sampling frame. They may not coincide with the reality of individual establishments as sample frames may contain inaccurate information. The variables containing the sample frame information are included in the data set for researchers who may want to further investigate statistical features of the survey and the effect of the survey design on their results.

-*a2* is the variable describing sampling regions

-*a6a*: coded using the same standard for micro, small, medium, and large establishments as defined above. The code -9 was used to indicate units for which size was undetermined in the sample frame.

-*a4a*: coded using ISIC codes for the chosen industries for stratification. These codes include most manufacturing industries (15 to 37), other manufacturing (2), retail (52), and (45, 50, 51, 55, 60, 63, 72) for other Services.

18. The surveys were implemented following a 2 stage procedure. Typically first a screener questionnaire is applied over the phone to determine eligibility and to make appointments. In the case of Malawi, this screener was administered face-to-face. Then a face-to-face interview takes place with the Manager/Owner/Director of each establishment. However, the phone numbers were unavailable in the sample frame, and thus the enumerators applied the screeners in person. The variables *a4b* and *a6b* contain the industry and size of the establishment from the screener questionnaire. Variables *a8* to *a11* contain additional information and were also collected in the screening phase.

19. Note that there are variables for size (*l1*, *l6* and *l8*) that reflect more accurately the reality of each establishment. Advanced users are advised to use these variables for analytical purposes. Variables *l1*, *l6* and *l8* were designed to obtain a more accurate measure of employment accounting for permanent and temporary employment. Special efforts were made to make sure that this information was not missing for most establishments.

20. Variables *a17x* gives interviewer comments, including problems that occurred during an interview and extraordinary circumstances which could influence results. Please note that sometimes this variable is removed due to privacy issues.

V. Universe Estimates

21. Universe estimates for the number of establishments in each cell in Malawi were produced for the strict, weak and median eligibility definitions. The estimates were the multiple of the relative eligible proportions.

23. For some establishments where contact was not successfully completed during the screening process (because the firm has moved and it is not possible to locate the new location, for example), it is not possible to directly determine eligibility. Thus, different assumptions about the eligibility of establishments result in different adjustments to the universe cells and thus different sampling weights.

24. Three sets of assumptions on establishment eligibility are used to construct sample adjustments using the status code information.

25. Strict assumption: eligible establishments are only those for which it was possible to directly determine eligibility. The resulting weights are included in the variable *wstrict*.

Strict eligibility = (Sum of the firms with codes 1,2,3,4,&16) / Total

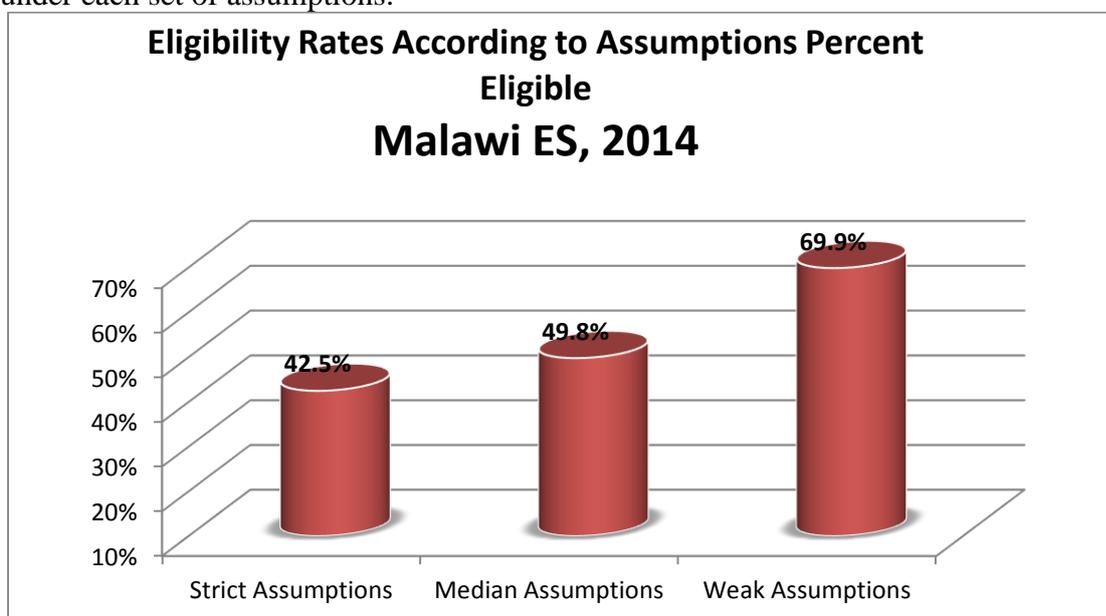
26. Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

$$\text{Median eligibility} = (\text{Sum of the firms with codes } 1,2,3,4,16,10,11, \text{ \& } 13) / \text{Total}$$

27. Weak assumption: in addition to the establishments included in points a and b, all establishments for which it was not possible to contact or that refused the screening questionnaire are assumed eligible. This definition includes as eligible establishments with dead or out of service phone lines, establishments that never answered the phone, and establishments with incorrect addresses for which it was impossible to find a new address. Under the weak assumption only observed non-eligible units are excluded from universe projections. The resulting weights are included in the variable *wweak*.

$$\text{Weak eligibility} = (\text{Sum of the firms with codes } 1,2,3,4,16,91,92,93,10,11,12, \text{ \& } 13) / \text{Total}$$

28. The indicators computed for the Enterprise Survey website use the median weights. The following graph shows the different eligibility rates calculated for firms in the sample frame under each set of assumptions.



29. Universe estimates for the number of establishments in each industry-region-size cell in Malawi were produced for the strict, weak and median eligibility definitions. Appendix D shows the universe estimates of the numbers of registered establishments that fit the criteria of the Enterprise Surveys.

30. Once an accurate estimate of the universe cell projection was made, weights for the probability of selection were computed using the number of completed interviews for each cell.

VI. Weights

31. Since the sampling design was stratified and employed differential sampling, individual observations should be properly weighted when making inferences about the population. Under stratified random sampling, unweighted estimates are biased unless sample sizes are proportional to the size of each stratum. With stratification the probability of

selection of each unit is, in general, not the same. Consequently, individual observations must be weighted by the inverse of their probability of selection (probability weights or *pw* in Stata.)⁴

32. Special care was given to the correct computation of the weights. It was imperative to accurately adjust the totals within each region/industry/size stratum to account for the presence of ineligible units (the firm discontinued businesses or was unattainable, education or government establishments, establishments with less than 5 employees, no reply after having called in different days of the week and in different business hours, no tone in the phone line, answering machine, fax line⁵, wrong address or moved away and could not get the new references) The information required for the adjustment was collected in the first stage of the implementation: the screening process. Using this information, each stratum cell of the universe was scaled down by the observed proportion of ineligible units within the cell. Once an accurate estimate of the universe cell (projections) was available, weights were computed using the number of completed interviews.

VII. Appropriate use of the weights

33. Under stratified random sampling weights should be used when making inferences about the population. Any estimate or indicator that aims at describing some feature of the population should take into account that individual observations may not represent equal shares of the population.

34. However, there is some discussion as to the use of weights in regressions (see Deaton, 1997, pp.67; Lohr, 1999, chapter 11, Cochran, 1953, pp.150). There is not strong large sample econometric argument in favor of using weighted estimation for a common population coefficient if the underlying model varies per stratum (stratum-specific coefficient): both simple OLS and weighted OLS are inconsistent under regular conditions. However, weighted OLS has the advantage of providing an estimate that is independent of the sample design. This latter point may be quite relevant for the Enterprise Surveys as in most cases the objective is not only to obtain model-unbiased estimates but also design-unbiased estimates (see also Cochran, 1977, pp 200 who favors the used of weighted OLS for a common population coefficient.)⁶

35. From a more general approach, if the regressions are descriptive of the population then weights should be used. The estimated model can be thought of as the relationship that would be expected if the whole population were observed.⁷ If the models are developed as structural relationships or behavioral models that may vary for different parts of the population, then, there is no reason to use weights.

VIII. Non-response

36. Survey non-response must be differentiated from item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the refusals to answer some specific questions. Enterprise Surveys suffer from both problems and different strategies were used to address these issues.

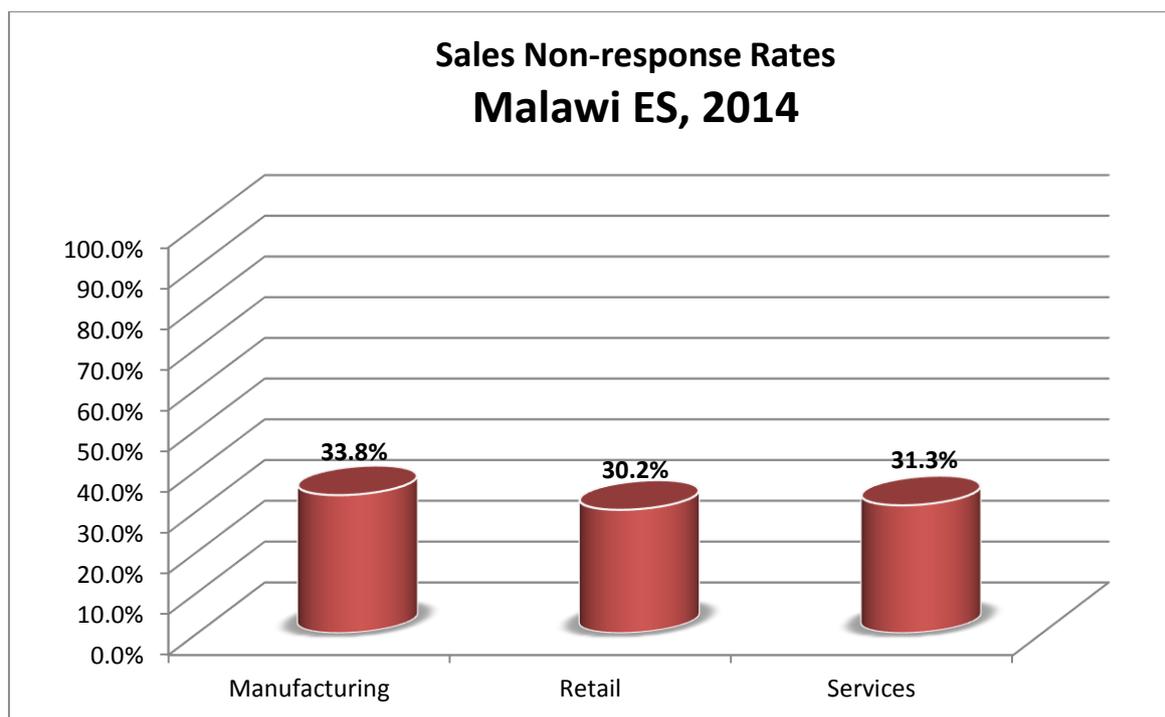
⁴ This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

⁵ For the surveys that implemented a screener over the phone.

⁶ Note that weighted OLS in Stata using the command `regress` with the option of weights will estimate wrong standard errors. Using the Stata survey specific commands `svy` will provide appropriate standard errors.

⁷ The use weights in most model-assisted estimations using survey data is strongly recommended by the statisticians specialized on survey methodology of the JPSM of the University of Michigan and the University of Maryland.

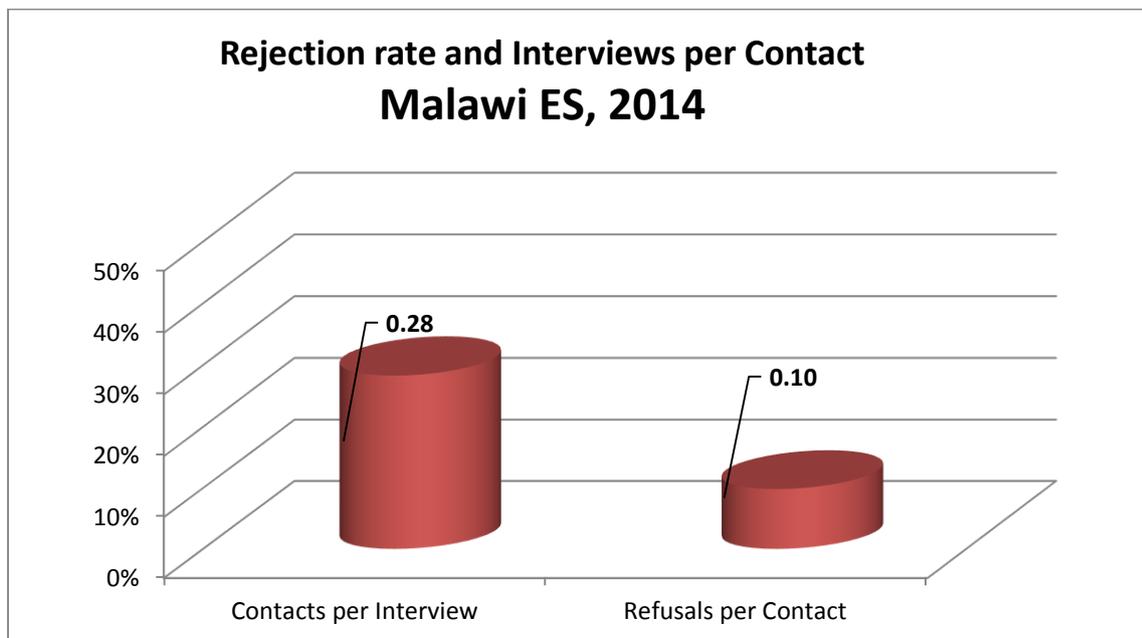
37. Item non-response was addressed by two strategies:
- a- For sensitive questions that may generate negative reactions from the respondent, such as corruption or tax evasion, enumerators were instructed to collect the refusal to respond as a different option from don't know (-7).
 - b- Establishments with incomplete information were re-contacted in order to complete this information, whenever necessary. However, there were clear cases of low response. The following graph shows non-response rates for the sales variable, *d2*, by sector. Please, note that the coding utilized in this dataset does not allow us to differentiate between "Don't know" and "refuse to answer", thus the non-response in the chart below for both enterprise surveys (ES) reflect both categories (DKs and NAs).



38. Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) was suggested for interview. Survey non-response did occur but substitutions were made in order to potentially achieve strata-specific goals. Further research is needed on survey non-response in the Enterprise Surveys regarding potential introduction of bias.

39. As the following graph shows, the number of interviews per contacted establishments was 0.28⁸. This number is the result of two factors: explicit refusals to participate in the survey, as reflected by the rate of rejection (which includes rejections of the screener and the main survey) and the quality of the sample frame, as represented by the presence of ineligible units. The number of rejections per contact was 0.10.

⁸ The estimate is based on the total no. of firms contacted including ineligible establishments.



40. Details on the rejection rate, eligibility rate, and item non-response are available at the level strata. This report summarizes these numbers to alert researchers of these issues when using the data and when making inferences. Item non-response, selection bias, and faulty sampling frames are not unique to Malawi. All enterprise surveys suffer from these shortcomings, but in very few cases they have been made explicit.

References:

Cochran, William G., *Sampling Techniques*, 1977.

Deaton, Angus, *The Analysis of Household Surveys*, 1998.

Levy, Paul S. and Stanley Lemeshow, *Sampling of Populations: Methods and Applications*, 1999.

Lohr, Sharon L. *Sampling: Design and Techniques*, 1999.

Scheaffer, Richard L.; Mendenhall, W.; Lyman, R., *Elementary Survey Sampling*, Fifth Edition, 1996.

Appendix A

Status Codes:

	Malawi
Sample Target	460
Complete interviews (Total)	524
Incomplete interviews	75
Elegible in process	416
Refusals	164
Out of target	158
Impossible to contact	373
Ineligible - coop.	23
Refusal to the Screener	134
Total	1867

Response rate	64%
Out of target + impossible to contact	28%
Impossible to contact	20%

Eligibles	1. Eligible establishment (Correct name and address)	755
	2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	4
	3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	8
	4. Eligible establishment (Wrong address - the firm/establishment has changed address and the address could be found)	19
Ineligibles	5. The establishment has less than 5 permanent full time employees	63
	6. The firm discontinued businesses	64
	7. Not a business: private household	2
	8. Ineligible activity: education, agriculture, finances, governments...	29
Unobtainable	91. No reply (<i>after having called in different days of the week and in different business hours</i>)	114
	92. Line out of order	10
	93. No tone	60
	94. Phone number does not exist	65
	10. Answering machine	0
	11. Fax line - data line	1
	12. Wrong address/ moved away and could not get the new references	123
	13. Refuses to answer the screener	134
	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	353
	151. Out of target - outside the covered regions, firm moved abroad	23
	152. Out of target - firm moved abroad	0
	153. Out of target - Not registered with SAT	0
	Total	1827

Appendix C Weights

Strict Panel Weights

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	1.1	1.1
	Medium	2.9	1.0
	Large	1.3	1.0
Northern (Mzimba)	Small	1.4	
	Medium		
	Large		
Southern (Blantyre, Zomba, Mangochi)	Small	1.3	1.9
	Medium	1.0	2.1
	Large	1.7	1.4

Median Panel Weights

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	1.1	1.1
	Medium	3.0	1.0
	Large	1.4	1.1
Northern (Mzimba)	Small	1.3	
	Medium		
	Large		
Southern (Blantyre, Zomba, Mangochi)	Small	1.3	1.9
	Medium	1.0	2.2
	Large	1.9	1.5

Weak Panel Weights

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	1.5	1.5
	Medium	3.4	1.0
	Large	1.4	1.2
Northern (Mzimba)	Small	3.7	
	Medium		
	Large		
Southern (Blantyre, Zomba, Mangochi)	Small	1.5	2.4
	Medium	1.0	2.3
	Large	1.8	1.5

Strict Fresh Weights

	Size	Manufacturing	Retail	Other services
Blantyre	Small	4.3	1.9	10.6
	Medium	5.8	3.7	12.4
	Large	4.8	5.9	16.5
Kasungu	Small			1.4
	Medium	1.0		
	Large			1.0
Lilongwe	Small	5.2	5.7	15.2
	Medium	1.6	2.4	10.7
	Large	2.2	4.9	14.8
Mangochi	Small	3.1	1.6	3.2
	Medium		1.0	1.2
	Large	1.0		1.5
Mzimba	Small	6.5	1.1	1.7
	Medium	1.4		7.6
	Large	1.7	2.0	2.4
Zomba	Small	4.2	2.6	2.3
	Medium	1.2	1.0	1.2
	Large	1.4	1.5	1.6

Median Fresh Weights

	Size	Manufacturing	Retail	Other services
Blantyre	Small	5.1	2.3	12.2
	Medium	6.4	4.1	13.4
	Large	5.6	6.7	18.7
Kasungu	Small			1.9
	Medium	1.0		
	Large			1.2
Lilongwe	Small	6.3	6.7	17.8
	Medium	1.8	2.6	11.6
	Large	2.6	5.6	17.0
Mangochi	Small	3.9	2.0	3.9
	Medium		1.0	1.4
	Large	1.0		1.8
Mzimba	Small	7.9	1.4	2.0
	Medium	1.6		8.3
	Large	2.0	2.4	2.8
Zomba	Small	4.5	2.7	2.4
	Medium	1.2	1.0	1.2
	Large	1.5	1.5	1.6

Weak Fresh Weights

	Size	Manufacturing	Retail	Other services
Blantyre	Small	6.4	2.9	15.2
	Medium	8.3	5.3	16.9
	Large	7.4	9.1	24.3
Kasungu	Small			2.2
	Medium	1.0		
	Large			1.5
Lilongwe	Small	9.5	10.4	26.3
	Medium	2.7	4.1	17.4
	Large	4.1	9.0	26.1
Mangochi	Small	5.2	2.8	5.1
	Medium		1.4	1.8
	Large	1.0		2.5
Mzimba	Small	9.0	1.6	2.2
	Medium	1.9		9.4
	Large	2.3	2.9	3.2
Zomba	Small	5.4	3.4	2.9
	Medium	1.4	1.0	1.4
	Large	1.9	2.0	2.0

Appendix D

Strict Universe Estimates Malawi

	Size	Manufacturing	Retail	Other services
Blantyre	Small	85	33	359
	Medium	105	86	323
	Large	121	142	265
Kasungu	Small	4	0	1
	Medium	2	0	2
	Large	2	0	1
Lilongwe	Small	68	97	273
	Medium	38	47	170
	Large	44	63	178
Mangochi	Small	3	2	6
	Medium	5	3	4
	Large	1	0	3
Mzimba	Small	7	1	26
	Medium	3	0	15
	Large	5	4	7
Zomba	Small	8	13	12
	Medium	4	1	12
	Large	3	3	6
	Total	507	495	1664

Median Universe Estimates Malawi

	Size	Manufacturing	Retail	Other services
Blantyre	Small	101	39	416
	Medium	116	93	348
	Large	140	162	299
Kasungu	Small	4	0	2
	Medium	2	0	2
	Large	2	0	1
Lilongwe	Small	82	115	320
	Medium	42	52	185
	Large	52	73	203
Mangochi	Small	4	2	8
	Medium	5	3	4
	Large	1	0	4
Mzimba	Small	8	1	30
	Medium	3	0	17
	Large	6	5	8
Zomba	Small	9	14	12
	Medium	4	1	12
	Large	3	3	6
	Total	584	562	1878

Weak Universe Estimates Malawi

	Size	Manufacturing	Retail	Other services
Blantyre	Small	128	50	518
	Medium	149	123	440
	Large	185	218	388
Kasungu	Small	4	0	2
	Medium	2	0	2
	Large	2	0	1
Lilongwe	Small	123	176	473
	Medium	65	81	278
	Large	81	117	314
Mangochi	Small	5	3	10
	Medium	5	4	5
	Large	1	0	5
Mzimba	Small	9	2	34
	Medium	4	0	19
	Large	7	6	10
Zomba	Small	11	17	14
	Medium	4	1	14
	Large	4	4	8
	Total	790	801	2536

Appendix E

Original Sample Design, Malawi:

Panel

		Manufacturing	Services
Central (Lilongwe, Kasungu)	Small	3	4
	Medium	3	3
	Large	3	4
Northern (Mzimba)	Small	2	1
	Medium	1	1
	Large	1	1
Southern (Blantyre, Zomba, Mangochi)	Small	6	7
	Medium	10	7
	Large	8	7
	Total	37	35

Fresh

	Size	Manufacturing	Retail	Other services
Blantyre	Small	28	18	14
	Medium	8	8	18
	Large	14	22	18
Kasungu	Small	2	0	2
	Medium	0	0	0
	Large	0	0	0
Lilongwe	Small	12	8	20
	Medium	18	14	16
	Large	20	22	22
Mangochi	Small	2	2	0
	Medium	2	4	0
	Large	0	0	2
Mzimba	Small	8	2	2
	Medium	2	0	14
	Large	6	4	8
Zomba	Small	0	2	0
	Medium	4	0	0
	Large	4	2	4
	Total	130	108	140

Appendix F

Local Agency team involved in the study:

Local Agency	1st contractor: Research Options Malawi 2nd contractor: Precision Consulting Group 3rd contractor: Knowledge Research & Consulting
Name of Project Manager	1st contractor: Ernest Poya 2nd contractor: Mathias Funsanani 3rd contractor: Knowledge Kaitano
Name and position of other key persons of the project: Local Survey Implementation Team and corresponding supervisor and enumerator codes:	For convenience, only personnel from the third contractor are listed. Project Coordinator: Tafadzwa Mhende 21 enumerators (including 6 recruiters), 4 supervisors

Sample Frame:

Characteristics of sample frame used	Variables: name of establishment, address, sector, region, size, telephone number (for most but not all records)
Year:	2011, 2012, 2013
Comments on the quality of sample frame:	The frame used for the ES was collected from various sources of varying quality. The frame from the National Statistics Office was the best quality of the Malawi frames and comprised detailed regional stratification and ISIC codes. Other sources included local business directories and registries from city Chambers of Commerce - all of which were of variable quality. In order to maintain stratification coherence, sample quality often had to revert to the 'lowest common denominator'.
Year and organism who conducted the last economic census	N/A
Other sources for companies statistics	Global Manufacturers.net, Yellow Pages

Sample:

Comments/ problems on sectors and regions selected in the sample	Region stratification was broadly ok. Regions based on the smaller, provincial cities did suffer to an extent from lower levels of sample availability - leading to interview
--	---

	redistribution within some cells and over-sampling in Lilongwe and Blantyre to off-set shortfalls elsewhere.
Comments on the response rate	Response rates were lower when the first two contractors were responsible for conducting fieldwork. After the third contractor was brought on board in late 2014 - response rates did improve. There was some limited inter-agency crossover in terms of the allocation of contacts - which did lead to a limited number of firms refusing to take part when they were approached twice.
Comments on the sample design:	Panel and Fresh region stratification was harmonized during fieldwork which led to a single sample design being followed for the country.
Other comments:	None

Fieldwork and country situation:

Date of Fieldwork	April 2014- February 2015
Locations	Blantyre, Kasungu, Lilongwe, Mangochi, Mzimba, Zomba
Interview number	524
Problems found during fieldwork	Lower response rates due to lower-than-expected sample quality. One contractor was fired, whilst the second resigned. This had a substantial impact on timings, especially in the ES' latter stages as data collection was nearly complete at the time of resignation.
Other observations:	First two contractors used PAPI, third used CAPI.