

The Rwanda 2011 Enterprise Surveys Data Set

I. Introduction

1. This document provides additional information on the data collected in Rwanda between June 2011 and February 2012 as part of the Africa Enterprise Survey 2011, an initiative of the World Bank. The document also provides information on the Rwanda micro survey carried out, in parallel to the Enterprise Survey, between June 2011 and February 2012.

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 125 countries, of which 113 have been surveyed following the 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 Rwanda 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.

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

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.

3. Three levels of stratification were used in this country: industry, establishment size, and region. 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: the universe was stratified into one manufacturing industry and one service as defined in the sampling manual. For the Rwanda Enterprise Survey (ES), the manufacturing industry and service industry had a target each of 120 interviews. The sample design for the Rwanda micro survey targeted 170 establishments, 18 in manufacturing and 152 in services.

5. For the Rwanda Enterprise Survey (ES), size stratification was defined following the standardized definition for the rollout: small (5 to 19 employees), medium (20 to 99 employees), and large (more than 99 employees). The micro sample consists of firms with 1 to 4 employees. For stratification purposes, the number of employees was defined on the basis of reported permanent full-time workers. This seems to be an appropriate definition of the labor force since seasonal/casual/part-time employment is not a common practice, except in the sectors of construction and agriculture.

6. Regional stratification for the Rwanda ES as well as the Rwanda micro surveys was defined in two regions (City and the surrounding business area): Butare and Kigali City.

III. Sampling implementation

7. 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. Great efforts were made to obtain the best source for these listings. However, the quality of the sample frames was not optimal and, therefore, some adjustments were needed to correct for the presence of ineligible units. These adjustments are reflected in the weights computation (*see below*).

8. TNS Opinion was hired to implement the Africa 2011 enterprise surveys roll out. In Rwanda the local subcontractor was OutReach Development Solutions.

9. For the Rwanda ES, two sample frames were used. The first was supplied by the World Bank and consists of enterprises interviewed in Rwanda 2006. The World Bank required that attempts should be made to re-interview establishments responding to the Rwanda 2006 survey where they were within the selected geographical regions and met eligibility criteria. Due to the fact that the previous round of surveys seemed to have utilized different stratification criteria (or no stratification at all) and due to the prevalence of small firm in the 2006 sample the following convention was used. The presence of panel firms was limited to a maximum of 50% of the achieved interviews. That sample is referred to as the Panel.

The second frame was produced by Rwanda Revenue Authority. A copy of that frame was sent to the TNS statistical team in London to select the establishments for interview. Each database contained the following information

- Coverage;
- Up to datedness;
- Availability of detailed stratification variables;
- Location identifiers- address, phone number, email;
- Electronic format availability;
- Contact name(s).

The Rwanda Revenue Authority sample frame was used also for the Rwanda micro survey.

Counts from sample frames are shown below.

Panel sample counts

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		16	16
	20 to 99	4		4
	100+			
Butare Total		4	16	20
Kigali City	5 to 19	24	111	135
	20 to 99	20	24	44
	100+	11	2	13
Kigali City Total		55	137	192
Grand Total		59	153	212

Fresh Sample Frame

Source: Rwanda Revenue Authority

Enterprise Survey

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19	2	4	6
	20 to 99	1	1	2
	100+		1	1
Butare Total		3	6	9
Kigali City	5 to 19	35	473	508
	20 to 99	37	135	172
	100+	5	43	48
Kigali City Total		77	651	728
Grand Total		80	657	737

Micro

Region name	Employees	Manufacturing	Services	Grand Total
Butare	1 to 4		5	5
Kigali City	1 to 4	58	1210	1268
Grand Total		58	1215	1273

10. The enumerated establishments with 5 employees or more (fresh and panel) were then used as the sample frame for the Rwanda Enterprise Survey with the aim of obtaining interviews at 220 establishments. The enumerated establishments with less than five employees (micro establishments) were used as sample frame for the Rwanda micro survey with the aim of obtaining interviews at 170 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. In addition, the sample frame contains no telephone/fax numbers so the local contractor had to screen the contacts by visiting them. Due to response rate and ineligibility issues, additional sample had to be extracted by the World Bank in order to obtain enough eligible contacts and meet the sample targets.

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. The percentage of confirmed non-eligible units as a proportion

of the total number of sampled establishments contacted for the survey was 18% (90 out of 506 establishments)⁴ for the ES firms and 15% (66 out of 433) for micro firms.

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 (including micro), i.e. manufacturing and the services questionnaire. Both questionnaires have common questions and respectfully additional manufacturing and services specific questions. 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 proceeded by a prefix “*AF*” indicate questions specific to Africa, 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. In the implementation of the Africa roll out 2011 an experiment was carried in some of the countries to better estimate the effects of the use of show cards in data collection. In some of the sections (i.e. innovation) the enumerators were trained to alternatively implement the section using either show cards or asking only the questions without showing any cards, please see the variable “*cards*”.

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

⁴ Based on out of target contacts and impossible to contact establishments

-*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. 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 Rwanda were produced for the strict, weak and median eligibility definitions. The estimates were the multiple of the relative eligible proportions.

22. Appendix B shows the overall estimates of the numbers of establishments in Rwanda based on the sample frame.

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*.

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

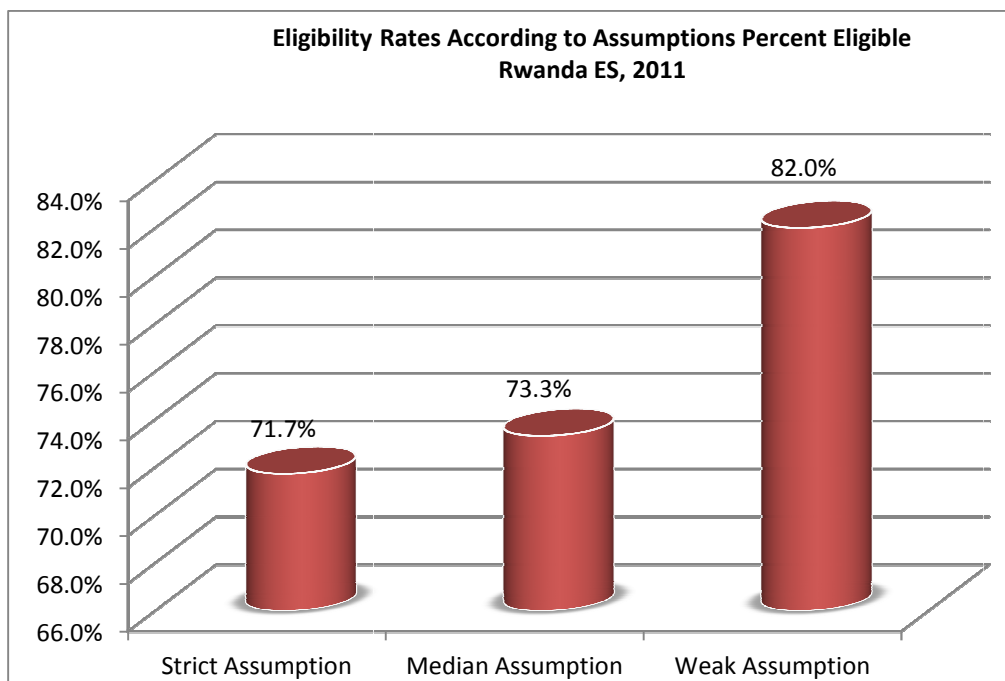
26. Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

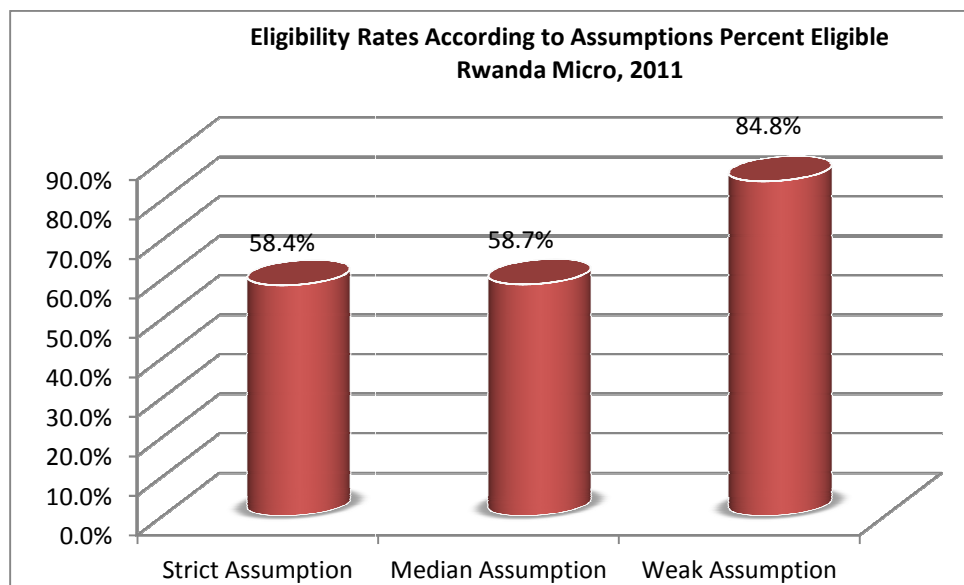
$$\text{Median eligibility} = (\text{Sum of the firms with codes 1,2,3,4,16,10,11, \& 13}) / \text{Total}$$

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*.

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

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. The eligibility rates for micro firms are also provided below.





29. Universe estimates for the number of establishments in each industry-region-size cell in Rwanda 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.)⁵

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 line⁶, wrong address or

⁵ This is equivalent to the weighted average of the estimates for each stratum, with weights equal to the population shares of each stratum.

⁶ For the surveys that implemented a screener over the phone.

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.

33. Appendix C shows the cell weights for registered establishments in Rwanda.

VII. Appropriate use of the weights

34. 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.

35. 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 use of weighted OLS for a common population coefficient.)⁷

36. 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.⁸ 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

37. 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.

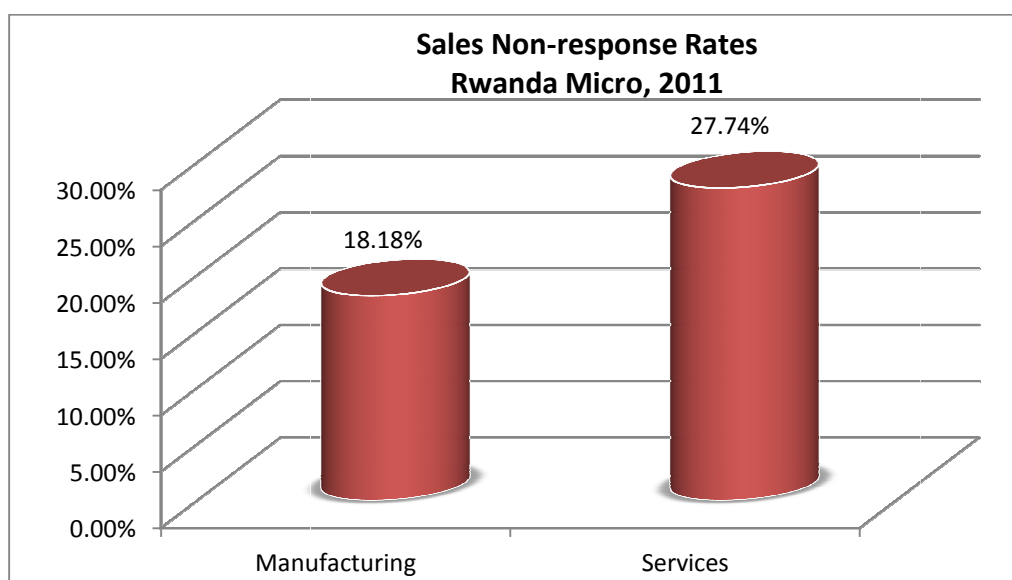
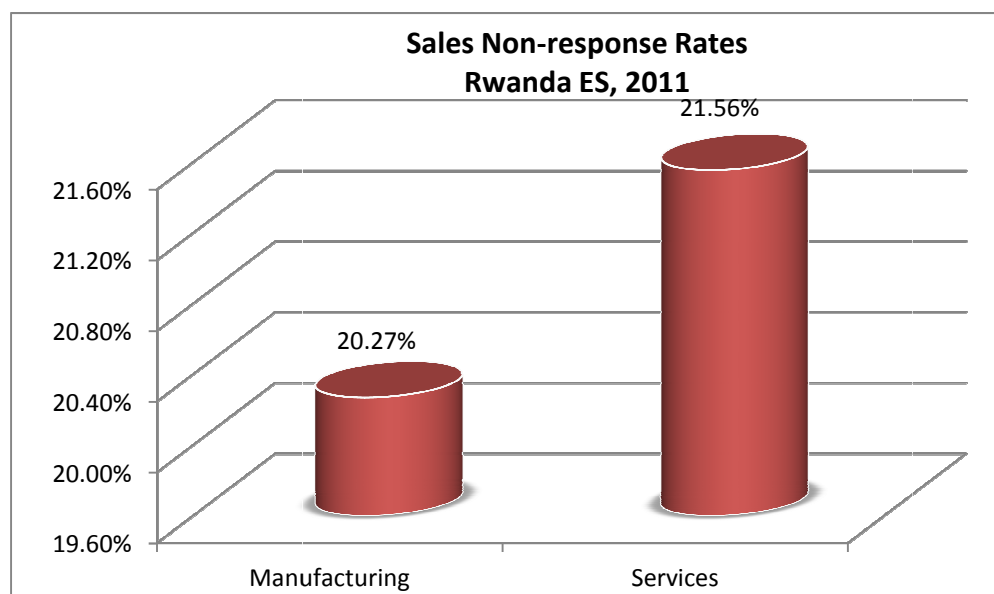
38. Item non-response was addressed by two strategies:

⁷ 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.

⁸ 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.

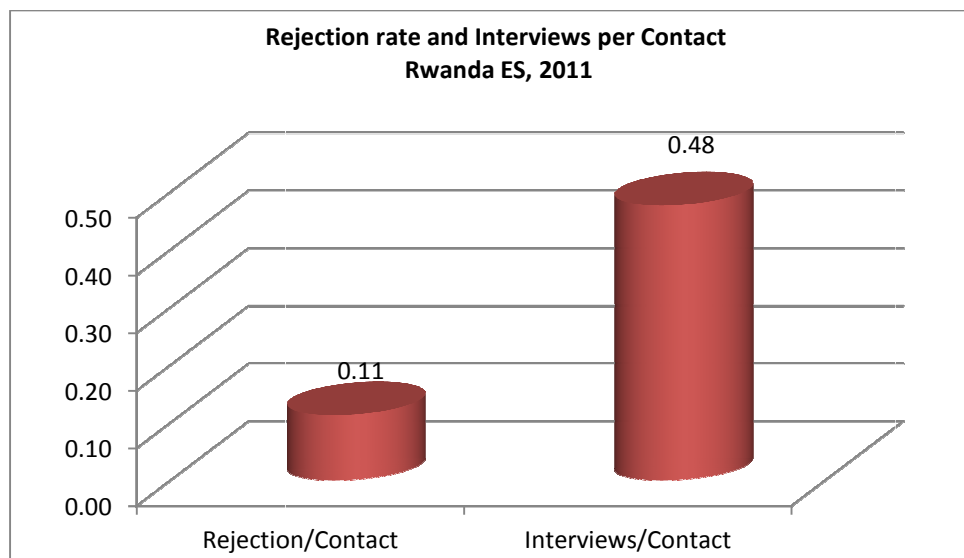
a- For sensitive questions that may generate negative reactions from the respondent, such as corruption or tax evasion, enumerators were instructed to collect the refusal to respond as a different option from don't know (-7).

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 charts below for both enterprise surveys (ES) and micro firms reflect both categories (DKs and NAs).

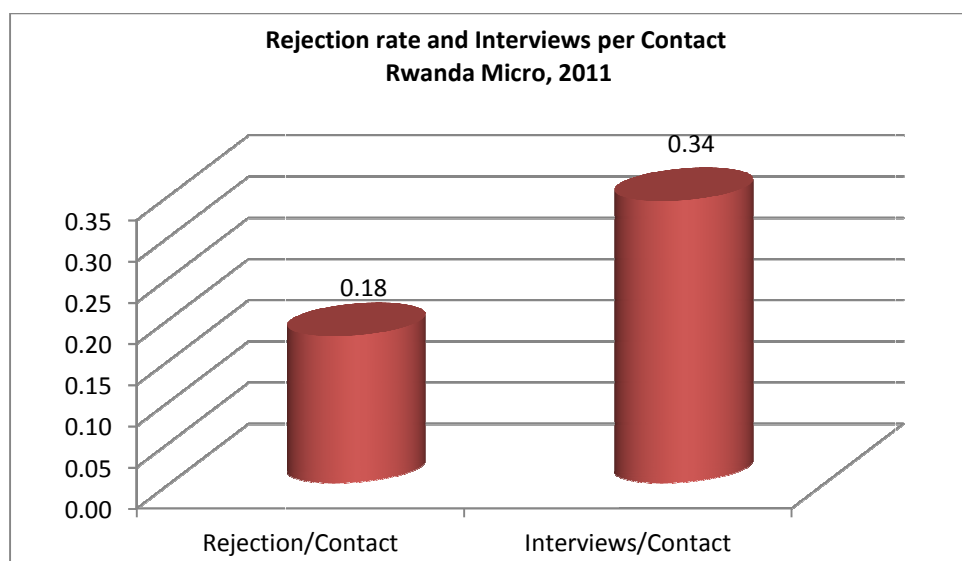


39. Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) 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.

40. As the following graph shows, the number of contacted establishments per realized interview was 0.48 for formal ES firms and 0.34 for micro firms⁹. 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 0.11 for formal ES firms and 0.18 for micro firms.



⁹ The estimate is based on the total no. of firms contacted including ineligible establishments.



41. 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 Rwanda. 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 Enterprise Survey Fresh:

ELIGIBLES		
Eligible	1. Eligible establishment (<i>Correct name and address</i>)	263
	2. Eligible establishment (<i>Different name but same address - the new firm/establishment bought the original firm/establishment</i>)	1
	3. Eligible establishment (<i>Different name but same address - the firm/establishment changed its name</i>)	6
	4. Eligible establishment (<i>Wrong address - the firm/establishment has changed address and the address could be found</i>)	6
	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	0
	6. The firm discontinued businesses	12
	7. Not a business: private household	3
	8. Ineligible activity: education, agriculture, finances, governments...	12
Unobtainable	91. No reply (<i>after having called in different days of the week and in different business hours</i>)	10
	92. Line out of order	11
	93. No tone	7
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	0
	13. Refuses to answer the screener	3
14. In process (<i>the establishment is being called/ is being contacted - previous to ask the screener</i>)		0
151. Out of target - outside the covered regions, firm moved abroad		1
152. Out of target - firm moved abroad		0
153. Impossible to find		16
Total		351

Response Outcomes Enterprise Survey Fresh:

Complete interviews (<i>Total</i>)	171
Incomplete interviews	0
Eligible in process	62
Refusals	43
Out of target	27
Impossible to contact	28
Ineligible - coop.	17
Refusal to the Screener	3
Total	351

Status Codes Enterprise Survey Panel:

ELIGIBLES		
Eligible	1. Eligible establishment (<i>Correct name and address</i>)	80
	2. Eligible establishment (<i>Different name but same address - the new firm/establishment bought the original firm/establishment</i>)	3
	3. Eligible establishment (<i>Different name but same address - the firm/establishment changed its name</i>)	4
	4. Eligible establishment (<i>Wrong address - the firm/establishment has changed address and the address could be found</i>)	0
	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	0
	6. The firm discontinued businesses	17
	7. Not a business: private household	0
	8. Ineligible activity: education, agriculture, finances, governments...	6
Unobtainable	91. No reply (<i>after having called in different days of the week and in different business hours</i>)	1
	92. Line out of order	6
	93. No tone	6
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	3
	13. Refuses to answer the screener	5
	14. In process (<i>the establishment is being called/ is being contacted - previous to ask the screener</i>)	1
	151. Out of target - outside the covered regions, firm moved abroad	3
	152. Out of target - firm moved abroad	0
	153. Impossible to find	20
	Total	155

Response Outcomes Enterprise Survey Panel:

Complete interviews (<i>Total</i>)	70
Incomplete interviews	0
Eligible in process	11
Refusals	6
Out of target	23
Impossible to contact	16
Ineligible - coop.	23
Refusal to the Screener	5
Total	154

Status Codes Micro:

ELIGIBLES

Eligible	1. Eligible establishment (<i>Correct name and address</i>)	243
	2. Eligible establishment (<i>Different name but same address - the new firm/establishment bought the original firm/establishment</i>)	3
	3. Eligible establishment (<i>Different name but same address - the firm/establishment changed its name</i>)	4
	4. Eligible establishment (<i>Wrong address - the firm/establishment has changed address and the address could be found</i>)	3
	16. Panel firm - now less than five employees	0
Ineligible	5. The establishment has less than 5 permanent full time employees	0
	6. The firm discontinued businesses	35
	7. Not a business: private household	1
	8. Ineligible activity: education, agriculture, finances, governments...	5
Unobtainable	91. No reply (<i>after having called in different days of the week and in different business hours</i>)	25
	92. Line out of order	16
	93. No tone	71
	10. Answering machine	0
	11. Fax line - data line	0
	12. Wrong address/ moved away and could not get the new references	1
	13. Refuses to answer the screener	1
	14. In process (<i>the establishment is being called/ is being contacted - previous to ask the screener</i>)	0
	151. Out of target - outside the covered regions	1
	152. Out of target - firm moved abroad	1
	153. Impossible to find	23
	Total	433

Response Outcomes Fresh Micro:

Complete interviews (Total)	148
Incomplete interviews	3
Eligible in process	26
Refusals	76
Out of target	41
Impossible to contact	113
Ineligible - coop.	25
Refusal to the Screener	1
Total	433

Appendix B

Universe Estimates, Rwanda:

Enterprise Survey - Panel

Region name	Sampling Employee size	Manufacturing	Service	Grand Total
Butare	small		16	16
	medium	4		4
	Large			
Butare Total		4	16	20
Kigali	small	24	111	135
	medium	20	24	44
	Large	11	2	13
Kigali Total		55	137	192
Grand Total		59	153	212

Enterprise Survey - Fresh

Source: Source: Rwanda Revenue Authority 2011

Region name	Sampling Employee size	Manufacturing	Service	Grand Total
BUTARE	5 to 19	2	4	6
	20 to 99	1	1	2
	100+		1	1
BUTARE Total		3	6	9
KIGALI CITY	5 to 19	35	473	508
	20 to 99	37	135	172
	100+	5	43	48
KIGALI CITY Total		77	651	728
Grand Total		80	657	737

Micro

Region name	Sampling Employee size	Manufacturing	Service	Grand Total
BUTARE	1 to 4		5	5
KIGALI CITY	1 to 4	58	1210	1268
Grand Total		58	1215	1273

Appendix C

Strict Cell Weights Rwanda:

Panel Enterprise Survey

Region name	Employees	Manufacturing	Services
Butare	5 to 19	1.00	1.11
	20 to 99		
	100+		
Kigali CITY	5 to 19	1.18	2.24
	20 to 99	1.27	2.71
	100+	1.21	

Fresh Enterprise Survey

Region name	Employees	Manufacturing	Services
Butare	5 to 19		1.00
	20 to 99		1.00
	100+		1.00
Kigali CITY	5 to 19	1.26	7.45
	20 to 99	1.19	2.31
	100+	1.00	1.88

Micro

Region name	Employees	Manufacturing	Services
Butare	1 to 4	1.18	1.46
Kigali CITY	1 to 4		5.98

Median Cell Weights Rwanda:

Panel Enterprise Survey

Region Name	Employees	Manufacturing	Services
Butare	5 to 19	1.00	1.11
	20 to 99		
	100+		
Kigali	5 to 19	1.22	2.36
	20 to 99	1.36	2.71
	100+	1.21	

Fresh Enterprise Survey

Region Name	Employees	Manufacturing	Services
Butare	5 to 19		1.00
	20 to 99		1.00
	100+		1.00
Kigali	5 to 19	1.26	7.45
	20 to 99	1.27	2.31
	100+	1.00	1.94

Micro

Region name	Employees	Manufacturing	Services
Butare	1 to 4	1.46	
Kigali CITY	1 to 4	1.20	5.98

Weak Cell Weights Rwanda:

Panel Enterprise Survey

Region name	Employees	Manufacturing	Services
Butare	5 to 19	1.00	1.17
	20 to 99		
	100+		
Kigali CITY	5 to 19	1.46	2.68
	20 to 99	1.55	3.14
	100+	1.21	

Fresh Enterprise Survey

Region name	Employees	Manufacturing	Services
Butare	5 to 19		1.00
	20 to 99		1.00
	100+		1.00
Kigali CITY	5 to 19	1.35	8.20
	20 to 99	1.27	2.64
	100+	1.00	2.12

Micro

Region name	Employees	Manufacturing	Services
Butare	1 to 4	2.08	
Kigali CITY	1 to 4	1.49	8.97

Appendix D

Strict Universe Estimates

Panel Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		7	7
	20 to 99	1		1
	100+			
Butare Total		1	7	8
Kigali	5 to 19	15	56	71
	20 to 99	14	19	33
	100+	8		8
Kigali CITY Total		38	75	113
Grand Total		39	82	120

Fresh Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		2	2
	20 to 99		1	1
	100+		1	1
Butare Total			4	4
Kigali	5 to 19	26	403	429
	20 to 99	31	104	135
	100+	4	32	36
Kigali CITY Total		61	538	600
Grand Total		61	542	604

Micro

Region name	Employees	Manufacturing	Services	Grand Total
Butare	1 to 4		1	1
Kigali CITY	1 to 4	40	675	715
Grand Total		40	676	716

Weak Universe Estimates
Enterprise Survey

Panel Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		7	7
	20 to 99	1		1
	100+			
Butare Total		1	7	8
Kigali CITY	5 to 19	19	67	86
	20 to 99	17	22	39
	100+	8		8
Kigali CITY Total		44	89	133
Grand Total		45	96	141

Fresh Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		2	2
	20 to 99		1	1
	100+		1	1
Butare Total			4	4
Kigali CITY	5 to 19	28	443	471
	20 to 99	33	119	152
	100+	4	36	40
Kigali CITY Total		65	598	663
Grand Total		65	602	667

Micro

Region name	Employees	Manufacturing	Services	Grand Total
Butare	1 to 4		2	2
Kigali CITY	1 to 4	51	1013	1064
Grand Total		173	1692	1865

Median Universe Estimates

Panel Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		7	8
	20 to 99	1		1
	100+			
Butare Total		1	7	8
Kigali CITY	5 to 19	16	59	75
	20 to 99	15	19	34
	100+	8		8
Kigali CITY Total		39	78	117
Grand Total		40	85	125

Fresh Formal ES

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19		2	2
	20 to 99		1	1
	100+		1	1
Butare Total			4	4
Kigali CITY	5 to 19	26	403	429
	20 to 99	33	104	137
	100+	4	33	37
Kigali CITY Total		63	539	603
Grand Total		63	543	607

Micro

Region name	Employees	Manufacturing	Services	Grand Total
Butare	1 to 4		1	1
Kigali CITY	1 to 4	41	675	716
Grand Total		41	676	717

Appendix E

Original Sample Design, Rwanda:

Enterprise Survey

Count of Country		Sample design classification		
Region name	Sampling Employee size	Manufacturing	Service	Grand Total
BUTARE	5 to 19	1	7	8
	20 to 99		1	1
	100+		1	1
BUTARE Total		1	9	10
KIGALI CITY	5 to 19	27	60	87
	20 to 99	36	50	86
	100+	14	23	37
KIGALI CITY Total		77	133	210
Grand Total		78	142	220

Micro Survey

		Sample design classification		
Region name	Sampling Employee size	Manufacturing	Service	Grand Total
BUTARE	1 to 4		1	1
KIGALI CITY	1 to 4	18	151	169
Grand Total		18	152	170

Completed Interviews, Rwanda:

Enterprise Survey

Region name	Employees	Manufacturing	Services	Grand Total
Butare	5 to 19	1	8	9
	20 to 99		1	1
	100+		1	1
Butare Total		1	10	11
Kigali CITY	5 to 19	27	85	112
	20 to 99	36	52	88
	100+	13	17	30
Kigali CITY Total		76	154	230
Grand Total		77	164	241

Micro

Region name	Employees	Manufacturing	Services	Grand Total
Butare	1 to 4		1	1
Kigali CITY	1 to 4	22	125	147
Grand Total		22	126	148

Appendix F

Local Agency team involved in the study:

Local Agency	Name: OutReach Development Solutions Country: Rwanda Activities since: 2007
Enumerators involved:	Enumerators: 15 Recruiters: 10
Other staff involved:	Fieldwork Coordinators: 2 Data Entry: 3 Data Processing: 1

Sample Frame:

Characteristic of sample frame used:	Panel: List from the 2006 Enterprise Survey Fresh: List of registered companies operating in Rwanda
Source:	Rwanda Revenue Authority
Year:	2011

Sectors included in the Sample:

Original Sectors	The manufacturing sector comprises all manufacturing establishments as mentioned in group D The service sector includes Group F (construction), Groups G, Group H (hotels and restaurants), Group I (transport, storage, and communications) and subsector 72 from Group K
Added (top up) Sectors	None

Fieldwork and country situation:

Date of Fieldwork	June 13 th , 2011 to February 22 nd , 2012
Country	Rwanda
Problems found during fieldwork:	▪ In some cases the sample frame did not provide establishment's contact details. In these cases, firms were screened by going to the establishment physical location. ▪ Several interviews were completed in more than one visit. ▪ Respondents were in general inclined to think the study aimed to get tax and social security defaulters. This might have affected the reliability of answers related to revenue and costs.