

## **The Indonesia 2015 Enterprise Surveys Data Set**

### **I. Introduction**

This document provides additional information on the data collected in Indonesia between April 2015 and November 2015. The objective of the Enterprise Survey is to gain an understanding of what firms experience in the private sector.

As part of its strategic goal of building a climate for investment, job creation, and sustainable growth, the World Bank has promoted improving the business environment 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 ES currently cover over 130,000 firms in 135 countries, of which 121 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 ES 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.

This 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**

The sample for 2015 Indonesia ES was selected using stratified random sampling, following the methodology explained in the *Sampling Note*<sup>1</sup>. Stratified random sampling<sup>2</sup> was preferred over simple random sampling for several reasons<sup>3</sup>:

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.

---

<sup>1</sup> The complete text can be found at [http://www.enterprisesurveys.org/~media/GIAWB/EnterpriseSurveys/Documents/Methodology/Sampling\\_Note.pdf](http://www.enterprisesurveys.org/~media/GIAWB/EnterpriseSurveys/Documents/Methodology/Sampling_Note.pdf)

<sup>2</sup> 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).

<sup>3</sup> 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.

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 C.

Industry stratification was designed in the way that follows: the universe was stratified into seven manufacturing industries and two services industries- Food and Beverages (ISIC Rev. 3.1 code 15), Garments (ISIC code 18), Textiles (ISIC code 17), Chemicals (ISIC code 24), Rubber and Plastics (ISIC code 25), Non-metallic mineral products (ISIC code 26), Other Manufacturing (ISIC codes 16, 19-23, 27-37), Retail (ISIC code 52) and Other Services (ISIC codes 45, 50, 51, 55, 60-64, and 72).

For the Indonesia ES, size stratification was defined as follows: small (5 to 19 employees), medium (20 to 99 employees), and large (100 or more employees).

Regional stratification for the Indonesia ES was done across nine regions: Jawa Barat, Jawa Timur, Jawa Tengah, DKI Jakarta, Banten, Sulawesi Selatan, Sumatera Utara, Bali and Lampung.

### **III. Sampling implementation**

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.

Mekong Economics was the main contractor and Kadence International was the subcontractor that implemented the Indonesia 2015 ES.

The sample frame consisted of listings of firms from four sources: First, for panel firms the list of 1444 firms from the Indonesia 2009 ES was used. Second, for fresh firms (i.e., firms not covered in 2009), economic census data from Statistics Indonesia known in Indonesia as Badan Pusat Statistik, henceforth BPS, was used. 2006 BPS data was used for service firms and small manufacturing firms and 2012 BPS data was used for medium and large manufacturing firms.

**Table 1: Indonesia ES Sample Frame (Fresh and Panel Combined)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	650	482	408	84	92	430	1,522	29,749	19,543	<b>63,444</b>
	Medium	883	512	166	408	570	37	1,679	1,550	2,647	
	Large	195	238	66	246	149	28	767	125	218	
<b>Jawa Timur</b>	Small	736	228	155	104	88	601	1,767	28,636	21,561	<b>64,766</b>
	Medium	1,389	305	216	488	272	69	2,118	1,504	2,772	
	Large	314	70	75	215	76	40	641	83	243	
<b>Jawa Tengah</b>	Small	739	582	329	85	102	443	2,066	20,951	16,210	<b>48,728</b>
	Medium	743	461	20	177	133	21	1,365	1,044	2,156	
	Large	137	109	15	111	23	10	495	71	130	
<b>DKI Jakarta</b>	Small	102	430	63	19	31	12	470	16,433	18,326	<b>42,774</b>
	Medium	148	278	48	181	21	15	410	1,280	3,511	
	Large	34	99	15	92	11	10	162	124	449	
<b>Banten</b>	Small	111	74	19	14	36	84	295	6,643	4,322	<b>14,162</b>
	Medium	115	36	39	226	57	12	391	283	545	
	Large	69	62	59	181	49	24	333	26	57	
<b>Sulawesi Selatan</b>	Small	153	35	34	18	4	208	328	5,390	6,803	<b>14,223</b>
	Medium	97	8	4	10	19	3	86	220	650	
	Large	35	1	1	5	15	5	24	16	51	
<b>Sumatera Utara</b>	Small	196	54	38	24	19	173	376	7,501	9,044	<b>19,702</b>
	Medium	303	16	5	99	47	9	194	274	878	
	Large	144	1	3	92	12	5	81	28	86	
<b>Bali</b>	Small	66	42	58	5	9	84	356	5,510	5,591	<b>13,874</b>
	Medium	45	57	6	12	22	7	153	417	1,172	
	Large	25	17	1	2	3	0	27	41	146	
<b>Lampung</b>	Small	165	24	19	5	16	149	174	3,655	3,241	<b>8,189</b>
	Medium	135	9	0	15	8	5	35	135	292	
	Large	53	0	0	5	2	0	17	13	17	
		<b>7,782</b>	<b>4,230</b>	<b>1,862</b>	<b>2,923</b>	<b>1,886</b>	<b>2,484</b>	<b>16,332</b>	<b>131,702</b>	<b>120,661</b>	<b>289,862</b>

Source: World Bank and Statistics Indonesia, known in Indonesia as Badan Pusat Statistik, (Government of Indonesia).

**Table 2: Indonesia Sample Frame (Panel)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	19	26	28	16	12	19	9	16	11	<b>295</b>
	Medium	4	10	10	9	11	9	5	2	3	
	Large	8	13	17	8	14	8	7	1	0	
<b>Jawa Timur</b>	Small	35	8	20	10	8	31	8	23	21	<b>283</b>
	Medium	8	6	11	3	12	13	12	1	3	
	Large	12	1	7	9	6	6	9	0	0	
<b>Jawa Tengah</b>	Small	19	26	29	17	15	19	14	23	12	<b>270</b>
	Medium	8	10	6	8	7	8	6	3	2	
	Large	4	7	8	7	7	0	5	0	0	
<b>DKI Jakarta</b>	Small	5	15	8	4	8	0	4	15	24	<b>161</b>
	Medium	2	8	2	4	9	1	3	5	11	
	Large	1	8	1	7	3	1	5	1	6	
<b>Banten</b>	Small	5	3	3	4	6	6	13	6	3	<b>128</b>
	Medium	4	2	5	8	11	2	6	0	1	
	Large	2	4	4	9	8	6	7	0	0	
<b>Sulawesi Selatan</b>	Small	2	3	2	3	0	15	7	11	8	<b>86</b>
	Medium	4	3	3	0	2	2	8	0	1	
	Large	3	0	1	0	0	4	4	0	0	
<b>Sumatera Utara</b>	Small	13	4	4	5	4	16	3	8	9	<b>81</b>
	Medium	1	1	0	1	2	2	1	2	1	
	Large	2	0	0	0	1	0	1	0	0	
<b>Bali</b>	Small	3	2	6	0	0	5	3	7	7	<b>74</b>
	Medium	3	5	3	0	0	6	3	2	5	
	Large	3	3	0	0	0	0	6	2	0	
<b>Lampung</b>	Small	10	0	5	0	3	10	8	4	5	<b>66</b>
	Medium	2	0	0	1	0	4	2	2	0	
	Large	4	0	0	0	0	0	5	0	1	
		<b>186</b>	<b>168</b>	<b>183</b>	<b>133</b>	<b>149</b>	<b>193</b>	<b>164</b>	<b>134</b>	<b>134</b>	<b>1,444</b>

Data for service firms were updated by cross-checking with lists from several business associations namely Aprindo 2013 for retail, AKI 2013, AKSINDO 2012 and Gapenri 2014 for construction, PHRI 2012 for hotels and restaurants and ALFI/ILFA 2014 for transportation.

The quality of the frame was enhanced by the verification process conducted by Kadence. However, the sample frame was not immune from the typical problems found in establishment surveys: positive rates of non-eligibility, repetition, non-existent units, etc.

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 4.1% (108 out of 2629 establishments)<sup>4</sup>.

Breaking down by industry and size, the following sample targets were achieved (based on the sampling information):

---

<sup>4</sup> Based on out of target and ineligible contacts

**Table 3: Achieved Interviews (Fresh and Panel Combined)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	5	8	8	6	4	9	6	15	7	<b>202</b>
	Medium	7	7	9	8	12	8	6	4	5	
	Large	7	9	11	9	10	7	7	3	5	
<b>Jawa Timur</b>	Small	6	5	6	5	4	9	8	15	13	<b>218</b>
	Medium	11	8	12	8	11	11	11	5	5	
	Large	8	8	11	9	11	7	7	2	2	
<b>Jawa Tengah</b>	Small	6	6	7	4	6	8	8	11	11	<b>191</b>
	Medium	9	8	9	8	9	8	8	5	4	
	Large	7	9	5	9	9	4	6	3	4	
<b>DKI Jakarta</b>	Small	3	6	6	3	6	3	2	13	15	<b>173</b>
	Medium	5	7	6	9	6	7	4	5	7	
	Large	6	9	9	10	4	7	4	5	6	
<b>Banten</b>	Small	3	4	7	3	4	6	2	3	2	<b>150</b>
	Medium	6	6	8	7	8	7	5	1	1	
	Large	7	9	9	9	10	12	7	3	1	
<b>Sulawesi Selatan</b>	Small	3	5	7	6	2	10	2	3	5	<b>95</b>
	Medium	5	4	0	5	7	1	6	1	3	
	Large	6	1	0	2	6	1	2	1	1	
<b>Sumatera Utara</b>	Small	6	3	4	4	4	9	2	6	5	<b>98</b>
	Medium	5	6	2	6	5	3	2	2	1	
	Large	5	0	0	8	2	2	4	1	1	
<b>Bali</b>	Small	2	3	7	2	2	7	2	3	4	<b>97</b>
	Medium	5	9	3	5	3	1	2	4	7	
	Large	5	7	1	1	2	0	4	3	3	
<b>Lampung</b>	Small	4	4	9	3	6	7	5	2	2	<b>96</b>
	Medium	8	5	0	8	4	1	2	3	3	
	Large	8	0	0	2	1	0	5	1	3	
		<b>158</b>	<b>156</b>	<b>156</b>	<b>159</b>	<b>158</b>	<b>155</b>	<b>129</b>	<b>123</b>	<b>126</b>	<b>1,320</b>

**Table 4: Achieved Interviews (Panel)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	2	5	4	4	2	5	3	7	1	<b>80</b>
	Medium	2	2	4	4	5	4	1	0	0	
	Large	3	4	5	4	4	2	3	0	0	
<b>Jawa Timur</b>	Small	3	2	3	2	2	4	4	7	6	<b>93</b>
	Medium	5	4	7	3	7	4	5	1	2	
	Large	4	0	5	4	4	1	4	0	0	
<b>Jawa Tengah</b>	Small	3	3	3	2	2	4	4	5	6	<b>86</b>
	Medium	5	4	4	4	3	4	4	2	2	
	Large	3	4	4	4	4	0	3	0	0	
<b>DKI Jakarta</b>	Small	1	3	3	1	4	0	1	8	9	<b>63</b>
	Medium	2	3	0	3	3	0	2	2	3	
	Large	0	4	0	4	0	1	2	1	3	
<b>Banten</b>	Small	1	2	3	1	2	3	1	1	1	<b>58</b>
	Medium	2	2	4	3	3	2	3	0	0	
	Large	1	4	2	4	4	5	4	0	0	
<b>Sulawesi Selatan</b>	Small	1	2	2	2	0	3	1	1	3	<b>26</b>
	Medium	2	2	0	0	1	0	5	0	0	
	Large	1	0	0	0	0	0	0	0	0	
<b>Sumatera Utara</b>	Small	4	1	0	2	2	3	1	4	2	<b>25</b>
	Medium	0	1	0	1	1	1	0	1	0	
	Large	0	0	0	0	0	0	1	0	0	
<b>Bali</b>	Small	1	0	3	0	0	3	1	1	1	<b>29</b>
	Medium	2	3	2	0	0	0	1	1	3	
	Large	1	2	0	0	0	0	2	2	0	
<b>Lampung</b>	Small	2	0	4	0	3	3	4	1	1	<b>31</b>
	Medium	2	0	0	1	0	0	1	2	0	
	Large	3	0	0	0	0	0	3	0	1	
		<b>56</b>	<b>57</b>	<b>62</b>	<b>53</b>	<b>56</b>	<b>52</b>	<b>64</b>	<b>47</b>	<b>44</b>	<b>491</b>

#### IV. Data Base Structure:

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 (*core* module) and respectfully additional manufacturing- and services-specific questions. The eligible manufacturing industries have been surveyed using the **Manufacturing** questionnaire (includes the *core* module, plus manufacturing specific questions). Retail firms have been interviewed using the **Services** questionnaire (includes the *core* module plus retail specific questions) and the residual eligible services have been covered using the **Services** questionnaire (includes the *core* module). Each variation of the questionnaire is identified by the index variable, *a0*.

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 the prefix “EA” or “MYA” indicate questions specific to Indonesia and other countries in EAP 2015, 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.

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.

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 based on the sample frame, whereas the latter gives the establishment’s actual industry classification (four digit code) based on the main activity at the time of the survey.

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 or outdated 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 small, medium, and large establishments as defined above.

- a4a*: coded following the stratification by sector as defined above.

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, sometimes 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.

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* (number of permanent full-time workers at the end of the last complete fiscal year), *l6* (number of full-time seasonal workers employed during last complete fiscal year) and *l8* (average length of employment of full-time temporary employees during last complete fiscal year) 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.

The firms interviewed had several fiscal years. Most firms had January 2014 to December 2014 as their last complete fiscal year. Variables *eea3a3w* (starting month of last complete fiscal year) and *eea3a3y* (last complete fiscal year) can be used to obtain the last complete fiscal year for each firm.

For questions pertaining to monetary amounts, the unit is the Indonesian rupiah.

## V. Universe Estimates

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

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.

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

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*.

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

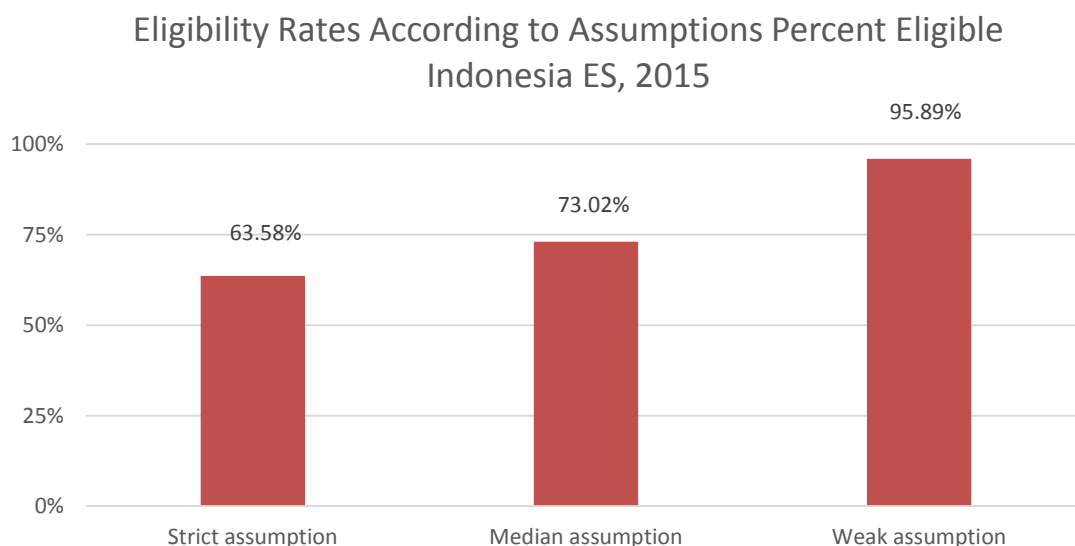
Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

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

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 *wwweak*.

$$\text{Weak eligibility} = (\text{Sum of the firms with codes, 1,2,3,4,16,10,11,13,91,92,93,94,12}) / \text{Total}$$

The indicators computed for the ES 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.



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

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

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.)<sup>5</sup>

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, 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 line<sup>6</sup>, 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

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.

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 have the advantage of providing an estimate that is independent of the sample design. This latter point may be quite relevant for the ES 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.)<sup>7</sup>

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.<sup>8</sup> 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.

---

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

<sup>6</sup> For the surveys that implemented a screener over the phone.

<sup>7</sup> 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.

<sup>8</sup> 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.

## VIII. Non-response

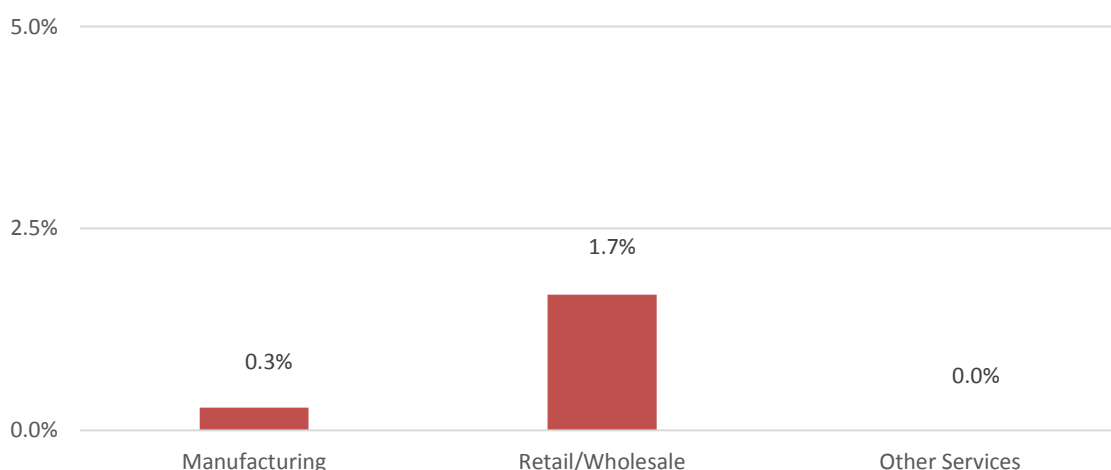
Survey non-response must be differentiated from item non-response. The former refers to refusals to participate in the survey altogether whereas the latter refers to the refusals to answer some specific questions. Enterprise Surveys suffer from both problems and different strategies were used to address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

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 (-8) as a different option from don't know (-9).

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 for this specific question, refusals were not separately identified from "Don't know" responses.

Sales Non-response Rates Indonesia ES, 2015

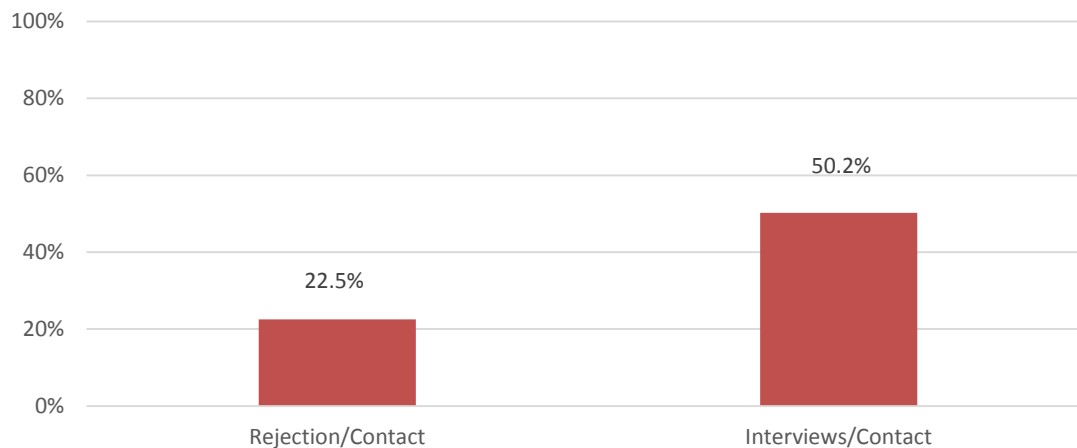


Survey non-response was addressed by maximizing efforts to contact establishments that were initially selected for interview. Attempts were made to contact the establishment for interview at different times/days of the week before a replacement establishment (with similar strata characteristics) was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

As the following graph shows, the number of interviews per contacted establishments was 0.50.<sup>9</sup> 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 share of rejections per contact was 0.23.

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

### Rejection rate and Interviews per Contact Indonesia ES, 2015



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 Indonesia. All enterprise surveys suffer from these shortcomings, but in very few cases they have been made explicit.

#### **References:**

- Cochran, William G., *Sampling Techniques*, New York, New York: John Wiley & Sons, 1977.
- Deaton, Angus, *The Analysis of Household Surveys*, Baltimore, Maryland: Johns Hopkins University Press, 1998.
- Levy, Paul S. and Stanley Lemeshow, *Sampling of Populations: Methods and Applications*, New York, New York: John Wiley & Sons, 1999.
- Lohr, Sharon L. *Sampling: Design and Techniques*, Boston, Massachusetts: Brooks/Cole, 1999.
- Scheaffer, Richard L.; Mendenhall, W.; Lyman, R., *Elementary Survey Sampling*, Fifth Edition, 1996.

## Appendix A

### Status Codes Enterprise Survey (ES):

<b>I</b>	<b>Screening in process</b>	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	<b>I</b>
<b>1671</b>	<b>Eligible</b>	1. Eligible establishment (Correct name and address) 2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment) 3. Eligible establishment (Different name but same address - the firm/establishment changed its name) 4. Eligible establishment (Moved and traced) 16. Eligible establishment (Panel Firm - now less than five employees; this code applies only to panel firms.)	1622 1 12 34 2
<b>246</b>	<b>Screener refusal</b>	13. Refuses to answer the screener	<b>246</b>
<b>70</b>	<b>Ineligible</b>	5. The establishment has less than 5 permanent full time employees 616. The firm discontinued businesses - (Establishment went bankrupt) 617. 618. The firm discontinued businesses - (Original establishment disappeared and is now a different firm) 619. The firm discontinued businesses - (Establishment was bought out by another firm) 620. The firm discontinued businesses - (It was impossible to determine for what reason) 621. The firm discontinued businesses - (Other) 7. Not a business: Private household 8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	11 38 0 4 2 6 5 4 0
<b>38</b>	<b>Out of target</b>	151. Out of target - outside the covered regions 152. Out of target - moved abroad	9 0

		153. Out of target - Not registered with Statistical Authority	27
		154. Out of target - establishment is HQ without production or sales of goods or services	0
		155. Out of target - establishment was not in operation for the entirety of last fiscal year	0
		156. Duplicated firm within the sample	2
<b>603</b>	<b>Unobtainable</b>	91. No reply after having called in different days of the week and in different business hours	60
		92. Line out of order	53
		93. No tone	88
		94. Phone number does not exist	98
		10. Answering machine	1
		11. Fax line- data line	1
		12. Wrong address/ moved away and could not get the new references	302

<b>2629</b>	<b>Total contacted</b>
-------------	------------------------

#### Response Outcomes : Indonesia ES 2015:

<b>Target and totals</b>	Sample target	1320
	Sample target completion rate	100.0%
	Total contacts available in frame	289861
	Total contacts issued	3000
	Total contacts contacted	2629

<b>Screening phase</b>	Screening in process	1
------------------------	----------------------	---

	Eligibles	1671
	Screeners refusal	246
	Ineligible + out of target	108
	Unobtainable	603
<b>Interview phase (only if eligible)</b>	Complete interviews without extra module	1320
	Complete interviews with extra module	0
	Eligible in process + incomplete interviews	0
	Interview refusal	346

<b>Percent breakdown (relative to total contacted)</b>		
	Screening in process rate	0.0%
	Screeners refusal rate	9.4%
	Ineligible + out of target rate	4.1%
	Unobtainable rate	22.9%
	Interview conversion rate	50.2%
	Eligible in process + incomplete interviews rate	0.0%
	Interview refusal rate	13.2%



## Appendix B: Universe Estimate Based on Sampling Weights

### Strict Universe Estimates – Fresh:

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	304	246	178	39	40	209	762	15,365	9,849	<b>32,400</b>
	Medium	433	274	76	198	259	19	881	839	1,397	
	Large	95	126	30	119	67	14	400	67	114	
<b>Jawa Timur</b>	Small	623	211	122	87	69	527	1,601	26,757	19,658	<b>59,814</b>
	Medium	1,232	295	178	429	223	63	2,010	1,472	2,647	
	Large	277	67	62	188	62	37	604	81	230	
<b>Jawa Tengah</b>	Small	365	313	151	42	47	226	1,090	11,402	8,608	<b>26,205</b>
	Medium	384	260	10	91	64	11	754	595	1,199	
	Large	70	61	7	56	11	5	272	40	72	
<b>DKI Jakarta</b>	Small	80	369	46	15	23	10	396	14,279	15,538	<b>36,856</b>
	Medium	122	250	37	148	16	13	362	1,165	3,118	
	Large	28	88	11	75	8	8	142	112	396	
<b>Banten</b>	Small	53	39	8	7	16	42	151	3,510	2,228	<b>7,397</b>
	Medium	58	20	18	112	26	6	210	157	294	
	Large	34	34	27	89	23	12	177	14	31	
<b>Sulawesi Selatan</b>	Small	71	18	15	8	2	101	164	2,778	3,421	<b>7,228</b>
	Medium	47	4	0	5	9	2	45	119	342	
	Large	17	1	0	2	7	3	12	9	27	
<b>Sumatera Utara</b>	Small	96	29	17	12	9	87	196	4,033	4,745	<b>10,443</b>
	Medium	155	9	2	50	22	5	106	154	483	
	Large	73	0	0	46	6	3	44	16	47	
<b>Bali</b>	Small	41	28	33	3	5	53	234	3,737	3,700	<b>9,326</b>
	Medium	29	40	4	8	13	5	105	296	812	
	Large	16	12	1	1	2	0	18	29	101	
<b>Lampung</b>	Small	137	22	15	4	12	128	154	3,341	2,891	<b>7,383</b>
	Medium	117	9	0	13	6	4	32	129	273	
	Large	46	0	0	4	2	0	16	12	16	
		<b>5,004</b>	<b>2,824</b>	<b>1,049</b>	<b>1,851</b>	<b>1,047</b>	<b>1,593</b>	<b>10,939</b>	<b>90,508</b>	<b>82,238</b>	<b>197,053</b>

## Median Universe Estimates – Fresh:

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	310	253	193	43	42	212	839	15,730	10,347	<b>33,625</b>
	Medium	429	273	80	215	267	19	944	835	1,429	
	Large	102	137	34	139	75	15	464	72	127	
<b>Jawa Timur</b>	Small	628	214	131	96	72	530	1,745	27,122	20,447	<b>61,483</b>
	Medium	1,209	292	187	460	228	62	2,132	1,452	2,680	
	Large	294	72	70	218	69	39	694	86	253	
<b>Jawa Tengah</b>	Small	507	439	224	63	68	314	1,640	15,949	12,355	<b>37,194</b>
	Medium	520	354	14	134	90	15	1,105	810	1,675	
	Large	103	90	11	90	17	8	431	59	109	
<b>DKI Jakarta</b>	Small	84	389	51	17	25	10	447	14,996	16,745	<b>39,231</b>
	Medium	124	256	40	164	17	13	398	1,191	3,270	
	Large	31	98	13	90	10	9	169	124	450	
<b>Banten</b>	Small	70	51	12	10	22	55	215	4,633	3,018	<b>9,936</b>
	Medium	74	25	25	157	35	8	290	201	388	
	Large	48	47	40	135	33	17	266	20	44	
<b>Sulawesi Selatan</b>	Small	68	17	15	9	2	96	170	2,675	3,380	<b>7,060</b>
	Medium	44	4	0	5	8	1	45	111	329	
	Large	17	1	0	3	7	3	14	9	28	
<b>Sumatera Utara</b>	Small	104	31	20	14	10	95	230	4,408	5,321	<b>11,587</b>
	Medium	164	9	3	58	24	5	121	164	527	
	Large	84	0	0	58	7	3	54	18	55	
<b>Bali</b>	Small	47	33	41	4	6	62	296	4,389	4,459	<b>11,100</b>
	Medium	33	46	4	10	16	5	130	339	953	
	Large	20	15	1	2	2	0	25	36	128	
<b>Lampung</b>	Small	142	23	16	5	13	133	174	3,498	3,106	<b>7,822</b>
	Medium	119	9	0	14	7	5	36	132	285	
	Large	50	0	0	5	2	0	19	14	18	
		<b>5,424</b>	<b>3,179</b>	<b>1,228</b>	<b>2,216</b>	<b>1,173</b>	<b>1,733</b>	<b>13,091</b>	<b>99,072</b>	<b>91,923</b>	<b>219,038</b>

# Weak Universe Estimates – Fresh:

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	619	448	372	80	86	400	1,414	27,623	18,057	<b>59,297</b>
	Medium	873	494	157	406	554	36	1,619	1,494	2,539	
	Large	198	236	64	251	149	28	760	124	215	
<b>Jawa Timur</b>	Small	708	214	143	101	83	564	1,658	26,858	20,123	<b>61,151</b>
	Medium	1,387	298	207	490	267	67	2,063	1,464	2,685	
	Large	322	70	74	222	77	40	642	83	242	
<b>Jawa Tengah</b>	Small	714	549	304	83	97	417	1,946	19,724	15,185	<b>46,128</b>
	Medium	745	451	19	178	131	21	1,334	1,020	2,097	
	Large	141	110	15	115	23	10	497	71	130	
<b>DKI Jakarta</b>	Small	97	399	57	18	29	11	435	15,203	16,871	<b>39,755</b>
	Medium	146	267	45	179	20	14	394	1,229	3,355	
	Large	34	98	15	94	11	10	160	122	441	
<b>Banten</b>	Small	109	71	18	14	35	81	283	6,367	4,122	<b>13,687</b>
	Medium	117	36	38	232	57	12	389	282	540	
	Large	72	64	59	191	51	25	341	27	58	
<b>Sulawesi Selatan</b>	Small	148	33	32	18	4	197	310	5,099	6,404	<b>13,473</b>
	Medium	98	8	0	10	19	3	84	216	635	
	Large	36	1	0	5	15	5	24	16	51	
<b>Sumatera Utara</b>	Small	159	43	30	20	15	137	297	5,930	7,115	<b>15,631</b>
	Medium	255	13	4	84	39	7	159	225	717	
	Large	125	0	0	80	10	4	68	24	72	
<b>Bali</b>	Small	62	38	52	5	8	76	324	5,008	5,057	<b>12,664</b>
	Medium	44	54	6	12	21	7	144	393	1,100	
	Large	25	17	1	2	3	0	26	40	141	
<b>Lampung</b>	Small	160	23	18	5	15	141	165	3,465	3,057	<b>7,786</b>
	Medium	136	9	0	15	8	5	34	133	286	
	Large	55	0	0	5	2	0	17	13	17	
		<b>7,585</b>	<b>4,044</b>	<b>1,730</b>	<b>2,913</b>	<b>1,831</b>	<b>2,317</b>	<b>15,589</b>	<b>122,253</b>	<b>111,312</b>	<b>269,572</b>

**Appendix C: Original Sample Design**  
**Original Sample Design (Fresh)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	3	3	4	2	2	4	3	8	6	<b>113</b>
	Medium	5	5	5	4	5	4	5	4	3	
	Large	4	5	5	5	5	5	4	3	2	
<b>Jawa Timur</b>	Small	3	3	3	3	2	5	4	8	7	<b>119</b>
	Medium	6	4	5	5	4	4	6	4	3	
	Large	4	8	6	5	5	5	3	2	2	
<b>Jawa Tengah</b>	Small	3	3	4	2	2	4	4	6	5	<b>97</b>
	Medium	4	4	4	4	4	4	4	3	2	
	Large	4	5	3	5	5	4	3	1	1	
<b>DKI Jakarta</b>	Small	2	3	3	2	2	3	1	5	6	<b>106</b>
	Medium	3	4	6	4	3	7	2	3	4	
	Large	6	5	9	5	3	6	2	4	3	
<b>Banten</b>	Small	2	2	3	2	2	3	1	2	1	<b>83</b>
	Medium	3	4	4	4	4	5	2	1	1	
	Large	5	5	7	5	5	5	3	1	1	
<b>Sulawesi Selatan</b>	Small	2	2	4	2	2	4	1	2	2	<b>55</b>
	Medium	3	2	1	3	4	1	1	1	1	
	Large	4	1	0	2	5	1	2	1	1	
<b>Sumatera Utara</b>	Small	2	2	3	2	2	4	1	2	3	<b>76</b>
	Medium	5	5	2	5	4	3	2	1	1	
	Large	5	1	2	5	7	2	3	1	1	
<b>Bali</b>	Small	1	2	3	2	2	3	1	2	2	<b>55</b>
	Medium	3	4	2	3	3	1	1	1	2	
	Large	4	5	1	1	2	0	2	1	1	
<b>Lampung</b>	Small	2	2	4	2	2	4	1	1	1	<b>51</b>
	Medium	4	4	0	6	3	1	1	1	1	
	Large	4	0	0	2	1	0	2	1	1	
		<b>96</b>	<b>93</b>	<b>93</b>	<b>92</b>	<b>90</b>	<b>92</b>	<b>65</b>	<b>70</b>	<b>64</b>	<b>755</b>

**Original Sample Design (Panel)**

		Food	Garments	Textiles	Chemicals	Rubber and Plastics	Non metallic Mineral Products	Other Manufacturing	Retail	Other Services	Grand Total
<b>Jawa Barat</b>	Small	2	3	4	2	2	4	3	7	6	<b>97</b>
	Medium	4	4	4	4	4	4	4	2	2	
	Large	4	4	5	4	4	5	3	1	2	
<b>Jawa Timur</b>	Small	3	2	3	2	2	4	4	7	6	<b>95</b>
	Medium	5	4	5	3	4	4	5	1	3	
	Large	4	1	5	4	4	5	2	1	2	
<b>Jawa Tengah</b>	Small	3	3	3	2	2	4	4	5	4	<b>85</b>
	Medium	4	4	4	3	3	4	4	2	2	
	Large	3	4	2	4	4	3	3	1	1	
<b>DKI Jakarta</b>	Small	1	3	3	1	2	3	1	4	5	<b>64</b>
	Medium	2	3	2	3	3	1	2	2	3	
	Large	1	4	1	4	3	1	2	1	3	
<b>Banten</b>	Small	1	2	3	1	2	3	1	1	1	<b>61</b>
	Medium	2	2	4	3	3	2	1	1	1	
	Large	2	4	4	4	4	5	2	1	1	
<b>Sulawesi Selatan</b>	Small	1	2	2	2	1	3	1	1	2	<b>40</b>
	Medium	2	2	0	3	2	0	1	1	1	
	Large	3	0	0	2	4	0	2	1	1	
<b>Sumatera Utara</b>	Small	1	2	3	2	2	3	1	2	2	<b>43</b>
	Medium	1	1	2	1	2	2	1	1	1	
	Large	2	0	1	4	1	2	1	1	1	
<b>Bali</b>	Small	1	2	3	1	2	3	1	1	1	<b>42</b>
	Medium	2	3	1	3	3	0	1	1	1	
	Large	3	3	0	1	1	0	2	1	1	
<b>Lampung</b>	Small	2	2	3	2	2	3	1	1	1	<b>38</b>
	Medium	2	3	0	1	3	0	1	1	1	
	Large	3	0	0	2	1	0	1	1	1	
		<b>64</b>	<b>67</b>	<b>67</b>	<b>68</b>	<b>70</b>	<b>68</b>	<b>55</b>	<b>50</b>	<b>56</b>	<b>565</b>