

## The Spain 2021 Enterprise Surveys Data Set

### I. Introduction

This document provides additional information on the data collected in Spain between December 2019 and January 2022. 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 210,000 firms in 160 countries, of which 153 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 2021 Spain 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 4: (section C), construction sector (section F), wholesale, retail, transportation, and storage sectors (sections G and H), accommodation and food service sectors (section I), publishing, telecommunications, computer programming (divisions 58, 61, 62 of section J), and repair of household goods (division 95 of section S). Note that this definition excludes the following sectors: electricity, water, steam, waste management supply (sections D and E), financial, insurance and real estate activities (sections K and L), professional, scientific and technical activities

---

<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

(section M), administrative and support activities (section N, except division 79, travel agency which is included in the universe), 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.

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 five manufacturing industries and two services industries- Food and Beverages (ISIC Rev. 4.0 code 10, 11), Fabricated Metal Products (ISIC code 25), Machinery & Equipment (ISIC code 28), Furniture (ISIC code 31), Other Manufacturing (ISIC codes 12-24, 26, 27, 29, 30, 32, 33), Retail (ISIC code 47) and Other Services (ISIC codes 41-43, 45, 46, 49-53, 55, 56, 58, 61, 62, 79 and 95).

For the Spain 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 Spain ES was done across 7 regions: Northwest, Northeast, Community of Madrid, Center, East, South, and Canary Islands.

### **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.

Cámara de Comercio de España, the main contractor, in collaboration with SigmaDos (the sub-contractor), implemented the Spain 2021 ES.

The sample frame consisted of listings of establishments obtained from Cámara de Comercio de España.

**Table 1: Spain ES Sample Frame**

		Food	Fabricated Metal Products	Machinery and Equipment	Furniture	Other Manufacturing	Retail	Other Services	Grand Total
<b>Northwest</b>	Small (5-19)	133	124	111	114	125	140	123	<b>2620</b>
	Medium (20-99)	146	180	85	26	135	169	138	
	Large (100 or more)	153	50	39	10	269	178	172	
<b>Northeast</b>	Small (5-19)	133	124	124	126	125	141	123	<b>2932</b>
	Medium (20-99)	145	129	180	64	135	171	137	
	Large (100 or more)	116	101	116	19	343	204	176	
<b>Community of Madrid</b>	Small (5-19)	133	124	124	161	125	141	222	<b>3172</b>
	Medium (20-99)	145	128	137	57	135	170	137	
	Large (100 or more)	144	42	54	5	322	490	176	
<b>Center</b>	Small (5-19)	133	124	124	220	126	140	148	<b>2751</b>
	Medium (20-99)	144	181	84	61	135	169	138	
	Large (100 or more)	165	34	45	15	277	112	176	
<b>East</b>	Small (5-19)	133	124	124	126	125	167	517	<b>3318</b>
	Medium (20-99)	144	129	129	186	134	171	137	
	Large (100 or more)	165	107	132	33	162	195	178	
<b>South</b>	Small (5-19)	132	123	124	151	125	141	272	<b>3011</b>
	Medium (20-99)	145	129	122	147	135	170	138	
	Large (100 or more)	178	51	32	17	186	315	178	
<b>Canary Islands</b>	Small (5-19)	133	147	19	34	126	166	124	<b>2603</b>
	Medium (20-99)	154	22	1	1	117	508	138	
	Large (100 or more)	47	1	0	0	47	269	549	
		<b>2921</b>	<b>2174</b>	<b>1906</b>	<b>1573</b>	<b>3409</b>	<b>4327</b>	<b>4097</b>	<b>20407</b>

Source: Cámara de Comercio de España

Necessary measures were taken to ensure the quality of the frame; however, the sample frame was not immune to 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 6.6% (1332 out of 20316 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 2: Achieved Interviews**

		Food	Fabricated Metal Products	Machinery and Equipment	Furniture	Other Manufacturing	Retail	Other Services	Other Manuf and Furniture	Grand Total
<b>Northwest</b>	Small (5-19)	7	11	11	18	4	6	6	0	<b>150</b>
	Medium (20-99)	6	15	12	0	7	4	9	0	
	Large (100 or more)	12	3	3	0	8	1	6	0	
	Medium and Large (20+)	0	0	0	1	0	0	0	0	
<b>Northeast</b>	Small (5-19)	4	5	4	4	2	1	3	0	<b>126</b>
	Medium (20-99)	9	13	12	4	9	0	0	0	
	Large (100 or more)	6	9	5	3	18	0	0	0	
	Medium and Large (20+)	0	0	0	0	0	6	9	0	
<b>Community of Madrid</b>	Small (5-19)	6	8	8	24	6	7	12	0	<b>166</b>
	Medium (20-99)	8	14	19	0	10	4	7	0	
	Large (100 or more)	7	3	4	0	7	6	5	0	
	Medium and Large (20+)	0	0	0	1	0	0	0	0	
<b>Center</b>	Small (5-19)	7	10	14	17	12	10	8	0	<b>191</b>
	Medium (20-99)	8	18	14	9	6	10	12	0	
	Large (100 or more)	9	1	3	2	14	2	5	0	
<b>East</b>	Small (5-19)	7	11	5	6	8	5	30	0	<b>162</b>
	Medium (20-99)	5	7	6	12	8	6	7	0	
	Large (100 or more)	6	6	6	1	7	5	8	0	
<b>South</b>	Small (5-19)	8	7	11	7	6	6	17	0	<b>178</b>
	Medium (20-99)	10	13	10	19	9	7	5	0	
	Large (100 or more)	4	7	5	2	10	4	11	0	
<b>Canary Islands</b>	Small (5-19)	7	0	0	0	6	6	0	4	<b>78</b>
	Medium (20-99)	4	0	0	0	0	20	0	6	
	Large (100 or more)	2	0	0	0	0	2	9	0	
	Medium and Large (20+)	0	0	0	0	5	0	0	0	
	Small and Medium (5-99)	0	0	1	0	0	0	0	0	
	Small, Medium, and Large (5+)	0	6	0	0	0	0	0	0	
		<b>142</b>	<b>167</b>	<b>153</b>	<b>130</b>	<b>162</b>	<b>118</b>	<b>169</b>	<b>10</b>	<b>1051</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 “BM” indicate questions specific to Spain other countries in Europe and Central Asia 2018/2019 and Middle East and North Africa 2019, therefore, they may not be found in the implementation of the rollout in other countries. Variable names preceded by the prefix “COV” indicate question specific to the effects induced by the COVID-19 pandemic. 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. Since the outbreak of the COVID-19 pandemic, interviews in Spain were conducted over the phone or with virtual meetings. Variable *method* in the dataset contains information on the mode of interview. The variables *a4b* and *a6b* contain the industry and size of the establishment from the screener questionnaire.

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 last complete fiscal year is December 2018 to December 2021. Variables *a20m* (starting month of last complete fiscal year) and *a20y* (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 Euro.

## V. Universe Estimates

Universe estimates for the number of establishments in each cell in Spain 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*.

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

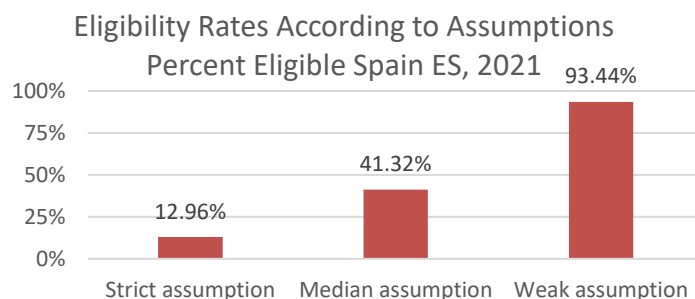
Median assumption: eligible establishments are those for which it was possible to directly determine eligibility and those that rejected the screener questionnaire or an answering machine or fax was the only response. The resulting weights are included in the variable *wmedian*.

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

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 *wweak*.

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

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 Spain 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.

Due to non-response rates, some stratification cells were collapsed for the purposes of weighting, to preserve the representativeness of the sample. The following cells have been transformed:

- (i) Medium and large firms are treated as one cell:

<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.

- In Northwest for Furniture.
  - In Northeast for Retail and Other Services.
  - In Community of Madrid for Furniture.
  - In Canary Islands for Other Manufacturing.
- (ii) Small, Medium, and Large firms are treated as one cell in Canary Islands for Fabricated Metal Products.
- (iii) Small and Medium firms are treated as one cell in Canary Islands for Machinery and Equipment.
- (iv) Furniture and Other Manufacturing are treated as one cell in Canary Islands for Small and Medium Firms.

## **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.

## **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:

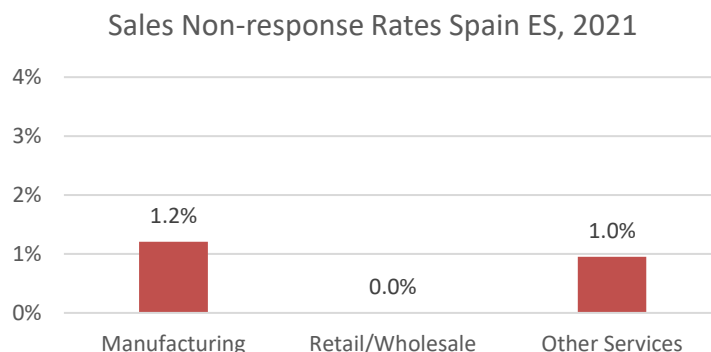
---

<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.

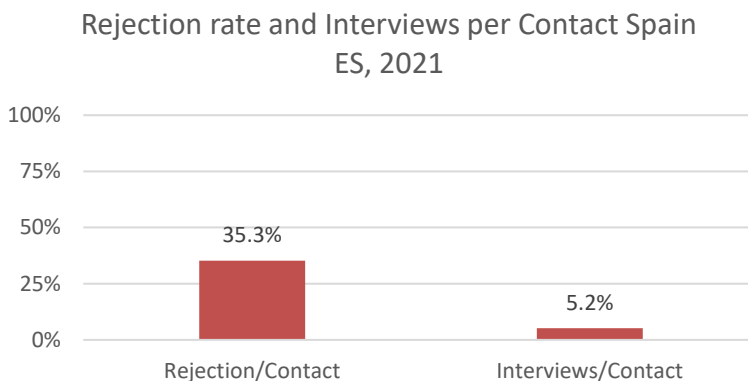
<sup>8</sup> The use weights in most model-assisted estimations using survey data is strongly recommended by the statisticians specialized on survey methodology of the JPSM of the University of Michigan and the University of Maryland.

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.



As the following graph shows, the number of interviews per contacted establishments was 0.05.<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 0.35.



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

## XI. Comments on Implementation

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

Two different contractors were engaged for this survey. The first one acted from 11/2018 until 05/2019 and collected 241 observations. The second one started the survey exercise ex-novo. All the contacts previously contacted (regardless of the outcome of the attempt to contacting) were re-added to the sample frames and re-contacted by the second contractor. The published dataset contains the records obtained from the second contractor.

The fieldwork was interrupted for a few months because of the pandemic outbreak. When the fieldwork resumed after the pandemic, the questionnaire was updated to get insights into the economic effects of the pandemic. In order to keep the questionnaire at a manageable length, the 'green module' was removed from the questionnaire and substituted by a set of covid-related questions (identified by the prefix 'COV' in the questionnaire). All the establishments that were interviewed before the pandemic outbreak were re-contacted and asked to participate in a short Covid-module interview, which contains exactly the same Covid-related questions included in the post-Covid questionnaire.

### **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: Brookes/Cole, 1999.
- Scheaffer, Richard L.; Mendenhall, W.; Lyman, R., *Elementary Survey Sampling*, Fifth Edition, 1996.

## Appendix A

### Status Codes Enterprise Survey (ES):

<b>0</b>	<b>Screening in process</b>	14. In process (the establishment is being called/ is being contacted - previous to ask the screener)	<b>0</b>
<b>2633</b>	<b>Eligible</b>	1. Eligible establishment (Correct name and address)	1977
		2. Eligible establishment (Different name but same address - the new firm/establishment bought the original firm/establishment)	5
		3. Eligible establishment (Different name but same address - the firm/establishment changed its name)	12
		4. Eligible establishment (Moved and traced)	639
		16. Eligible establishment (Panel Firm - now less than five employees; this code applies only to panel firms.)	0
<b>5620</b>	<b>Screener refusal</b>	13. Refuses to answer the screener	<b>5620</b>
<b>640</b>	<b>Ineligible</b>	5. The establishment has less than 5 permanent full time employees	158
		616. The firm discontinued businesses - (Establishment went bankrupt)	21
		618. The firm discontinued businesses - (Original establishment disappeared and is now a different firm)	11
		619. The firm discontinued businesses - (Establishment was bought out by another firm)	2
		620. The firm discontinued businesses - (It was impossible to determine for what reason)	16
		621. The firm discontinued businesses - (Other)	14
		71. Ineligible legal status: not a business, but private household	221
		72. Ineligible legal status: cooperatives, non-profit organizations, etc.	138
		8. Ineligible activity: Education, Agriculture, Finances, Government, etc.	59
<b>692</b>	<b>Out of Target</b>	151. Out of target - outside the covered regions	0
		152. Out of target - moved abroad	1
		153. Out of target - Not registered with Statistical Authority	35
		154. Out of target - establishment is HQ without production or sales of goods or services	54
		155. Out of target - establishment was not in operation for the entirety of last fiscal year	4
		156. Duplicated firm within the sample	68
		157. Out of target - location that is not HQ and does not have financial statements prepared separately	530
<b>10731</b>	<b>Unobtainable</b>	91. No reply after having called in different days of the week and in different business hours	8967
		92. Line out of order	382
		93. No tone	995
		94. Phone number does not exist	7
		10. Answering machine	12
		11. Fax line- data line	129
		12. Wrong address/ moved away and could not get the new references	239
		<b>20316</b>	<b>Total contacted</b>

**Response Outcomes : Spain ES 2021:**

<b>Target and totals</b>	Sample target	1051
	Sample target completion rate	100.0%
	Total contacts available in frame	20407
	Total contacts issued	20407
	Total contacts contacted	20316

<b>Screening phase</b>	Screening in process	0
	Eligibles	2633
	Screener refusal	5620
	Ineligible + out of target	1332
	Unobtainable	10731
<b>Interview phase (only if eligible)</b>	Complete interviews without extra module	861
	Complete interviews with extra module	190
	Eligible in process + incomplete interviews	0
	Interview refusal	1545

<b>Percent breakdown (relative to total contacted)</b>	Screening in process rate	0.0%
	Screener refusal rate	27.7%
	Ineligible + out of target rate	6.6%
	Unobtainable rate	52.8%
	Interview conversion rate	5.2%
	Eligible in process + incomplete interviews rate	0.0%
	Interview refusal rate	7.6%

## Appendix B: Universe Estimate Based on Sampling Weights

### Universe:

		Food	Fabricated Metal Products	Machinery and Equipment	Furniture	Other Manufacturing	Retail	Other Services	Other Manuf and Furniture	Grand Total
<b>Northwest</b>	Small (5-19)	893	677	116	113	1399	2970	11215	0	<b>21853</b>
	Medium (20-99)	286	274	72	0	648	833	1799	0	
	Large (100 or more)	62	31	8	0	159	109	163	0	
	Medium and Large (20+)	0	0	0	25	0	0	0	0	
<b>Northeast</b>	Small (5-19)	1048	1583	405	174	2120	2725	11424	0	<b>25425</b>
	Medium (20-99)	411	653	276	60	1081	0	0	0	
	Large (100 or more)	67	53	56	13	299	0	0	0	
	Medium and Large (20+)	0	0	0	0	0	501	2476	0	
<b>Community of Madrid</b>	Small (5-19)	363	691	175	179	1959	5479	19026	0	<b>36872</b>
	Medium (20-99)	166	144	75	0	621	1664	4867	0	
	Large (100 or more)	28	8	10	0	182	215	978	0	
	Medium and Large (20+)	0	0	0	42	0	0	0	0	
<b>Center</b>	Small (5-19)	1640	691	186	201	1804	2923	13207	0	<b>24969</b>
	Medium (20-99)	554	186	76	51	584	458	1950	0	
	Large (100 or more)	97	27	6	2	157	41	126	0	
<b>East</b>	Small (5-19)	1771	2323	810	525	7724	12643	42410	0	<b>87065</b>
	Medium (20-99)	789	664	373	149	3085	2684	8650	0	
	Large (100 or more)	190	56	49	11	632	373	1154	0	
<b>South</b>	Small (5-19)	1719	768	242	492	2508	7015	24773	0	<b>45346</b>
	Medium (20-99)	571	183	102	119	760	1033	4151	0	
	Large (100 or more)	86	28	10	6	137	103	539	0	
<b>Canary Islands</b>	Small (5-19)	261	0	0	0	373	2750	0	7393	<b>13220</b>
	Medium (20-99)	87	0	0	0	0	399	0	1433	
	Large (100 or more)	13	0	0	0	0	27	253	0	
	Medium and Large (20+)	0	0	0	0	90	0	0	0	
	Small and Medium (5-99)	0	0	14	0	0	0	0	0	
	Small, Medium, and Large (5+)	0	126	0	0	0	0	0	0	
		<b>11103</b>	<b>9166</b>	<b>3062</b>	<b>2163</b>	<b>26322</b>	<b>44946</b>	<b>149161</b>	<b>8826</b>	<b>254749</b>

**Note:** The sampling frame used and the universe are from separate sources. Universe figures for the number of establishments in each cell in Spain were obtained from National Statistics Institute of Spain.

## Universe Estimates:

		Food	Fabricated Metal Products	Machinery and Equipment	Furniture	Other Manufacturing	Retail	Other Services	Other Manuf and Furniture	Grand Total
<b>Northwest</b>	Small (5-19)	893	677	116	113	1399	2970	11215	0	<b>21853</b>
	Medium (20-99)	286	274	72	0	648	833	1799	0	
	Large (100 or more)	62	31	8	0	159	109	163	0	
	Medium and Large (20+)	0	0	0	25	0	0	0	0	
<b>Northeast</b>	Small (5-19)	1048	1583	405	174	2120	2725	11424	0	<b>25425</b>
	Medium (20-99)	411	653	276	60	1081	0	0	0	
	Large (100 or more)	67	53	56	13	299	0	0	0	
	Medium and Large (20+)	0	0	0	0	0	501	2476	0	
<b>Community of Madrid</b>	Small (5-19)	363	691	175	179	1959	5479	19026	0	<b>36872</b>
	Medium (20-99)	166	144	75	0	621	1664	4867	0	
	Large (100 or more)	28	8	10	0	182	215	978	0	
	Medium and Large (20+)	0	0	0	42	0	0	0	0	
<b>Center</b>	Small (5-19)	1640	691	186	201	1804	2923	13207	0	<b>24969</b>
	Medium (20-99)	554	186	76	51	584	458	1950	0	
	Large (100 or more)	97	27	6	2	157	41	126	0	
<b>East</b>	Small (5-19)	1771	2323	810	525	7724	12643	42410	0	<b>87065</b>
	Medium (20-99)	789	664	373	149	3085	2684	8650	0	
	Large (100 or more)	190	56	49	11	632	373	1154	0	
<b>South</b>	Small (5-19)	1719	768	242	492	2508	7015	24773	0	<b>45346</b>
	Medium (20-99)	571	183	102	119	760	1033	4151	0	
	Large (100 or more)	86	28	10	6	137	103	539	0	
<b>Canary Islands</b>	Small (5-19)	261	0	0	0	373	2750	0	7393	<b>13220</b>
	Medium (20-99)	87	0	0	0	0	399	0	1433	
	Large (100 or more)	13	0	0	0	0	27	253	0	
	Medium and Large (20+)	0	0	0	0	90	0	0	0	
	Small and Medium (5-99)	0	0	14	0	0	0	0	0	
	Small, Medium, and Large (5+)	0	126	0	0	0	0	0	0	
		<b>11103</b>	<b>9166</b>	<b>3062</b>	<b>2163</b>	<b>26322</b>	<b>44946</b>	<b>149161</b>	<b>8826</b>	<b>254749</b>

**Note:** Adjustments for strict, median and weak assumptions were not applied to universe estimates because the universe and the frame were from separate sources hence scenarios that apply to the frame cannot be assumed in the universe.

## Appendix C: Original Sample Design

### Original Sample Design

		Food	Fabricated Metal Products	Machinery and Equipment	Furniture	Other Manufacturing	Retail	Other Services	Grand Total
<b>Northwest</b>	Small (5-19)	5	5	5	23	5	5	4	<b>150</b>
	Medium (20-99)	5	15	15	2	8	6	6	
	Large (100 or more)	12	3	4	0	12	5	5	
<b>Northeast</b>	Small (5-19)	3	4	4	6	5	4	5	<b>150</b>
	Medium (20-99)	7	10	9	4	7	6	5	
	Large (100 or more)	8	20	17	2	12	7	5	
<b>Community of Madrid</b>	Small (5-19)	5	5	5	22	5	5	10	<b>150</b>
	Medium (20-99)	5	13	14	2	5	7	5	
	Large (100 or more)	10	4	4	0	11	8	5	
<b>Center</b>	Small (5-19)	5	5	9	23	5	5	6	<b>160</b>
	Medium (20-99)	5	14	13	8	5	5	5	
	Large (100 or more)	10	2	6	5	11	8	5	
<b>East</b>	Small (5-19)	5	5	4	5	12	12	27	<b>151</b>
	Medium (20-99)	4	5	6	12	5	5	7	
	Large (100 or more)	6	6	6	2	5	5	7	
<b>South</b>	Small (5-19)	5	5	5	5	5	9	14	<b>150</b>
	Medium (20-99)	5	6	6	17	5	5	5	
	Large (100 or more)	9	12	5	2	10	7	8	
<b>Canary Islands</b>	Small (5-19)	13	6	3	0	5	8	4	<b>140</b>
	Medium (20-99)	17	5	0	0	17	17	5	
	Large (100 or more)	6	0	0	0	5	21	8	
		<b>150</b>	<b>150</b>	<b>140</b>	<b>140</b>	<b>160</b>	<b>160</b>	<b>151</b>	<b>1051</b>