Department for Business Innovation & Skills

English Business Survey
Technical Annex
February 2012

Content

1. Introduction

The English Business Survey (EBS) is an important new survey commissioned by the Department for Business, Innovation and Skills (BIS) to provide a monthly assessment of business perceptions of current economic and business conditions in each English region. In the summer of 2011 BIS commissioned an independent research company, TNS-BMRB, to conduct the survey on its behalf. It will be conducted continuously between November 2011 and spring 2014 covering workplaces in England.[1]

A detailed understanding of businesses' perceptions and plans across England will inform the Government's economic growth and rebalancing agenda. This information will be valuable to the Government, policy makers, businesses, and local partners in helping to guide policy and support private sector growth in all parts of England. The outputs of the survey should also provided valuable intelligence of the local economic and business conditions to businesses.

This report presents the technical details of the EBS. The report is intended primarily for analysts who wish to make use of the data, who will need to understand the sample design, the details of coding and the actual questions asked. While every care has been taken to produce reliable statistics, this survey is still in its infancy and full quality assurance of the data is not yet possible. Until this fuller quality assurance is possible estimates from the survey have been designated as Experimental Official Statistics, meaning results may be subject to revision if improvements in methodology are identified. All revisions would be pre-announced. Results should be used with this in mind.

This is the first version of this document. If you have any feedback on its content, clarity and readability you may e-mail ebsurvey@bis.gsi.gov.uk. Any suggestions will inform the next update of the document.

2. Sample design

2.1 Overview

The sample for the EBS is drawn from the Inter-departmental Business Register (IDBR) (see Section 2.2).

The main distinction of this survey compared with most other business surveys is that the primary objective is to achieve a sample that is as close as possible to being proportionate to the employment distribution within England. Employment is used as a measure of economic importance for each workplace.[2] Turnover is another metric that could have been used to assess economic importance but this information is not available at the workplace level on the sample frame (see Section 2.2) (only at the enterprise or business level). Employment

[1] The first month of fieldwork was conducted in November and the outputs from that month will be the October 2011 data given October was the key reference period workplaces were asked about in the survey. In the rest of the document we will specify reference or fieldwork period.

[2] A business can be a single local unit (hereafter called workplace) or a workplace that belongs to a larger enterprise, for example a local store of a larger retail chain. The EBS is a workplace survey (see Section 2.2).

data are available at workplace level and therefore is used as a proxy for economic importance.[3]

The sample is drawn using a stratified random sampling method, with probability proportional to size. A random probability sampling approach is used. The fact that the sampling approach fulfils the random sample assumption, enables confidence intervals to be calculated around survey estimates and allows formal statistical tests to be carried out. The size measure is the employment count at a workplace. This approach was taken to maximise the similarity between the drawn sample and the employment distribution in England, as well as providing a high level of precision for estimates that are correlated with employment totals.

The sample is also selected to try to ensure that a minimum of 60 interviews are carried out in each NUTS2 sub region per month.[4] This is so that sample sizes are large enough to enable some sub regional analysis every quarter. This will also enable analysis of some Local Enterprise Partnership areas on a quarterly basis.[5]

2.2 Sample frame

The sample of workplaces is drawn from the IDBR held by the Office for National Statistics (ONS). The EBS is conducted at the level of the workplace rather than at the level of the business or enterprise. The sample is therefore selected at this level as well. The IDBR covers all businesses registered to pay VAT or that have a PAYE scheme in operation. It is widely regarded as the most comprehensive business sampling frame available, although it does not cover unregistered businesses, so excludes the very smallest non-employing businesses that are not registered for VAT/ PAYE. The sample of workplaces is selected on a quarterly basis from across all industry sectors (including public sector organisations). The sample from the IDBR is drawn by BIS, according to a specification set out by TNS-BMRB. The sample is then passed to TNS-BMRB, who have been granted access as a sub-contractor of BIS through the ONS Micro-data release procedure.

2.3 Telephone matching

As the IDBR does not contain telephone numbers for all entries, the selected quarterly sample is put through a telephone number matching process in order to make it usable for telephone interviewing.

The first sample drawn from the IDBR, to be issued in November and December 2011, did not contain any telephone numbers, as BIS's access to the IDBR at this time did not include access to the telephone numbers. Thus all sample records received by TNS-BMRB were sent for matching.

However, the second sample drawn, to be issued between January and March 2012, *did* contain some telephone numbers. Around six in ten (61%) sampled workplaces were supplied with a telephone number. Most of the numbers however, were head office numbers

[3] The use of employment rather than turnover to allocate the sample, is likely to be preferred even if turnover data were available, as employment is generally less volatile than turnover.
[4] The nomenclature of territorial units for statistics (NUTS) classification is a hierarchical system for dividing up the economic territory of the EU.
[5] The coverage of local authorities by LEP can be found on the BIS website at http://www.bis.gov.uk/policies/economic-development/leps/statistics

and the proportion of the sample presumed to include workplace numbers was just 25%.[6] Any records without a unique telephone number were subject to telephone number matching. Future samples drawn will also contain telephone numbers so matching rates are expected to mirror the experience from 2012 Q1 fieldwork sample, rather than November and December 2011 fieldwork sample.

For the November and December 2011 fieldwork sample, the overall matching rate was 73%, and the 2012 Q1 matching rate fieldwork sample reached the same level (see Table 1).

Table 1: Overall telephone matching rates

	Nov/Dec 2011 Fieldwork sample	**2012 Q1 (Jan-Mar 2012) Fieldwork sample**
	Overall	**Overall**
	%	**%**
Overall selected sample with telephone numbers	73	73
- *With existing unique telephone number*	-	*25*
- *Successful look-up*	*73*	*47*
Base	*63,370*	*46,245*

The telephone matching rates represent the success rate after a two stage number look-up process, firstly using an automated searching facility and secondly using a manual search service. The match rates are the proportion of records returned with a number successfully appended, which is then ready to be issued for fieldwork regardless of the quality of that number (although duplicate numbers are removed). The proportion of issued numbers that were not useable due to the telephone number not being valid, for sample issued in November 2011 fieldwork is detailed in Section 4.5. Details for later months are not yet available.

Matching rates were typically higher for workplaces with greater employment (see Table 2). This is desirable since it may result in a lower overall likely error. However the overall match rate for workplaces with an employment of 250 or more in the 2012 Q1 fieldwork sample was lower (72%) compared with November and December (85%). Matching rates for the other size bands between the two quarters were more similar.

Table 2: Telephone matching rates, by employment band

	Nov/Dec 2011 Fieldwork sample				**2012 Q1 (Jan-Mar 2012) Fieldwork sample**			
	IDBR employment band				**IDBR employment band**			
	0-9	10-49	50-249	250+	0-9	10-49	50-249	250+
	%	%	%	%	%	%	%	%
Overall selected sample with telephone numbers	59	78	83	85	61	78	79	72
- *With existing unique telephone number*	-	-	-	-	*9*	*32*	*34*	*31*
- *Successful look-up*	*59*	*78*	*83*	*85*	*52*	*46*	*45*	*40*
Base	*20,787*	*25,594*	*14,528*	*2,461*	*14,862*	*16,636*	*12,372*	*2,377*

[6] Duplicate telephone numbers i.e. the same phone number for different workplaces, arose as the telephone numbers supplied were attached at the enterprise level. For many records, the local units within the same enterprise all had the same telephone number and so could not be used.

2.4 Sample selection

2.4.1 Stratification

The businesses listed on the IDBR can be grouped into statistical units, which themselves are arranged in a hierarchical manner. The top level comprises Enterprise groups. These are groups of legal units under common ownership. Enterprises comprise the layer below that. These are the smallest combination of legal units that have a certain level of autonomy within an Enterprise Group. They are the closest equivalent to a business. The lowest level comprises workplaces (also known as local units) which are individual sites within an Enterprise, for example factories, shops or offices. **The EBS is conducted at this lowest level of workplaces rather than at the enterprise level.** This approach was taken as it is important to pick up conditions at the local level rather than at the location of the enterprise head office.

Prior to selection, all the workplaces on the IDBR are divided into strata based on their NUTS2 sub region and the employment count at the workplace.[7] The employment count is split into three groups: 0-49, 50-249 and 250+ and there are 30 NUTS2 areas in England.[8] So in total, 90 sampling strata are defined at the first stage.

Within each of the 90 strata the workplaces are sorted by their local authority code (LAU1; previously known as NUTS4 areas), industry section (SIC 2007 definition[9]) and the raw employment count of the workplace.

Once a workplace is selected, it is excluded from selection for the next three quarters. This is so that no workplace can be selected for the EBS more than once in a 12 month period; thus reducing the survey burden on businesses.

The number of cases to be drawn from each stratum is determined in advance, based on the target number of interviews and estimates of sample-to-interview conversion rates.

2.4.2 'Annual' sample selection

The selection process is done in two stages with an 'annual' sized sample being selected first. All workplaces in the 250+ strata are selected for the 'annual' sample, while a probability proportional to size (PPS) sampling approach is used within the other 60 strata.[10] From this sample, a quarterly sample is selected. The 'annual' sample is typically around four times the size of the sample that is needed for the quarter.

2.4.3 Sample selection for the quarter

Having selected the 'annual' sample, a quarterly sample is taken using a systematic random sampling method. The order of the file is retained from the 'annual' sample drawn. Within each of the 90 original strata, workplaces are picked systematically from the 'annual' sample

---

[7] Employment is defined as the number of employees plus any working proprietors (owners directly involved in the business).

[8] Some workplaces on the IDBR have an employment count of zero so they are given an employment count of one to ensure they have a chance of selection. Some of these are investment funds, pension funds, holding companies, property trusts etc. However others are workplaces with part time workers. Giving these workplaces a chance of being interviewed, by setting the employment count to one, seems the most pragmatic way of dealing with this.

[9] Standard Industrial Classification (SIC) 2007 is a way of classifying business establishments by the type of economic activity in which they are engaged.

[10] In practice, there are other strata where all workplaces are selected at this first stage due to the small numbers within the stratum.

using a random start and fixed sampling interval. For example, if the objective is to select 1,000 workplaces from 4,000 workplaces within a stratum, then the sampling interval is 4,000/1,000=4 with a random starting position between 1 and 4, so that every 4th workplace is selected until the required quarterly sample size is achieved.

By selecting a quarterly sample from an annual sample using a systematic random sampling method, it ensures that an even spread of different sized workplaces is selected within a stratum. If a PPS method was employed as the only sampling stage, this would select the largest workplaces in earlier quarters, leaving only smaller workplaces available to select in later quarters, since the rule imposed is not to select a workplace for the survey more than once in a 12 month period.

2.5 Issuing sample for fieldwork

The EBS has an achieved sample size of circa 3,000 interviews every calendar month. The sample unit remains active for up to three months to maximise the opportunity to complete an interview with each selected workplace. In any monthly dataset there is therefore a mix of older, harder-to-reach cases carried over from previous sample issues as well as fresh cases issued that month. Over time the expectation is that the mix of old and new sample cases reaches a steady state, so that comparisons between monthly datasets are valid. However for the first month, there were no existing records to call although 3,000 interviews still needed to be achieved. Thus the number of records issued in the first month was substantially higher than would normally be issued, and the mix of easier and harder to contact workplaces will not have the same balance as when the survey is more established. This imbalance may lead to volatility in some of the survey estimates for the first few months of the survey, although this depends on the extent to which opinions and experiences differ on the basis of contact difficulties.

As is good practice when trying to maximise the response rate on a survey, rather than issue the entire drawn sample, the monthly sample was randomly allocated to batches and was issued batch by batch. This maximised the likelihood of achieving the target number of interviews without issuing too many cases (which would have led to a sub-optimal response rate).

Batches of sample are issued weekly for use by the telephone centres.[11]

3. Questionnaire development

3.1 Initial development

It was important that the questionnaire was designed to produce high quality data on the key directional indicators that it was intended to track each month. A rigorous period of questionnaire development was undertaken prior to the survey's launch in November 2011.

The questionnaire length is eleven minutes, consisting of nine minutes of 'core' questions which are asked every month, and two minutes of ad hoc questions, which can be changed prior to the start of each calendar month's interviewing. The questionnaire development focused primarily on finalising the 'core' questions.

[11] However, for the first two months larger batches were issued less frequently, owing to the unusual time of year for interviewing, which necessitated more sample to be front loaded and the lack of existing contacts.

A separate report documents the questionnaire development process which incorporated a questionnaire development workshop, a period of cognitive testing and a large-scale pilot exercise.[12] In this technical annex, a summary of each of the stages is detailed below.

3.2 Questionnaire development workshop

TNS-BMRB, in partnership with BIS, embarked on an intense period of development work in the summer of 2011. The first step involved BIS developing an initial questionnaire based on data needs.

TNS-BMRB then worked with BIS research team members to redraft the questionnaire into a more complete version, before a workshop was held involving key stakeholders. The workshop was held at BIS's London office but was facilitated by TNS-BMRB researchers. There were two aims for the workshop: to review and finalise the content of the draft questionnaire for cognitive testing (in terms of the question coverage) and to think about how the results would be reported on a monthly basis. The workshop was held on 26 July 2011.

3.3 Cognitive testing

The next stage of development was to cognitively test the questionnaire. This was to ensure that respondents understood and answered the questions as anticipated.

A total of 28 cognitive interviews were conducted (10 face-to-face and 18 by telephone) by the TNS-BMRB research team between 22 August 2011 and 2 September 2011 with workplaces from a wide range of industry sectors and different levels of employment.

Overall, respondents found the questions to be generally clear and easy to answer. Some question specific issues were highlighted during this phase and were addressed for the telephone pilot. The full details of this testing can be found in the separate questionnaire development report.

3.4 Telephone pilot

Following the cognitive testing, the questionnaire for the pilot stage was developed. The main purpose of the pilot was to test the contact procedures across a broader range of workplaces, to test the questionnaire further and to get a clear idea of the questionnaire length.

A total of 253 telephone interviews were conducted by TNS-BMRB between 19 September 2011 and 30 September 2011, using sample drawn from the IDBR. The sample was drawn from a wide range of industry sectors and workplace sizes in order to try to capture responses from a greater breadth of workplaces.

As with the cognitive testing, overall, respondents found the questions to be generally clear and easy to answer. The pilot helped to ascertain that the questionnaire was too long in length, as well as clear up final issues picked up at the cognitive testing stage. Further details of the pilot findings can be found in the separate questionnaire development report.

[12] Report on Questionnaire Development: http://www.bis.gov.uk/analysis/statistics/sub-national-statistics/ebsurvey

3.5 Finalising the questionnaire

Following the pilot, further changes were made to the questionnaire by TNS-BMRB to reduce the questionnaire length, in collaboration with BIS. Questions were identified as high or low priority and then identified for deletion by BIS. The core questionnaire is available on the EBS section of the BIS website.[13]

The broad questionnaire structure of the final questionnaire used to start interviewing for the November 2011 fieldwork is summarised in Table 3.

Table 3: EBS Questionnaire structure

Topic areas covered
Contact / introduction
Characteristics
Output
Exports
Domestic output
Stocks
Jobs and hours
Prices
Credit conditions
Investment
Capital investment
Data linkage / re-contact permissions

3.6 CATI system

For interviewing on the survey TNS-BMRB uses the IBM SPSS Data Collection Family Suite of interviewing software, previously marketed as SPSS Dimensions. For the management of sample, TNS-BMRB uses SPSS-MR's Quancept CATI software.

Prior to the survey's launch in November 2011, both the questionnaire program and sample management system were extensively and systematically checked by the TNS-BMRB research team. This included checking the questionnaire content, question wording, routing, internal consistency checks and text substitution. Further checking was carried out by the data processing and telephone project management teams. Before release into the field, the questionnaire was further checked by running automatically generated test results through the program then checking the results against the questionnaire they were based on.

The questionnaire changes each calendar month, due to the rotation of the ad hoc questions, thus the questionnaire program checking procedure is conducted every month to ensure that the previous ad hoc questions have been removed correctly and the new ad hoc questions have been inserted correctly.

[13] English Business Survey Questionnaire: http://www.bis.gov.uk/analysis/statistics/sub-national-statistics/ebsurvey

4. Data collection and fieldwork management

4.1 Interviewer briefings

Prior to starting work on the survey from November 2011, interviewers and telephone centre supervisors were personally briefed by researchers from TNS-BMRB.

This briefing covered:

- the background to the survey and its aims;
- the sample design and methodology;
- introducing the survey to respondents;
- identifying the appropriate respondent to interview
- maximising response; and
- the questionnaire structure.

All interviewers were given a set of project instructions, which provided further information about the survey and key definitions and rules. As the survey progresses and more interviewers are needed to work on the survey, TNS-BMRB researchers will continue to brief them personally covering the same issues as above.

4.2 Reassurance letter

Advance letters are not sent out to businesses prior to interviewing so a reassurance letter is provided to interviewers. This can be sent to those who are contacted and wish to receive notification of the survey in writing. This is included in Appendix A.

The letter can be e-mailed or faxed by interviewers while they are interviewing, as the facility is built into the CATI system. The letters are despatched on BIS-headed paper and explain the purpose and nature of the research.

Although workplaces are encouraged to contact TNS-BMRB in the first instance, the name and direct telephone number of a contact in the project team at BIS is also provided. This is so workplaces can be reassured that the survey is genuine and allows them to ask questions that are more relevant to BIS.

4.3 Telephone interviewing

Telephone fieldwork is carried out by TNS-BMRB's telephone interviewing centres in Ealing and Hull. Each of these centres has a long-standing track record in conducting large-scale random sample surveys among businesses for central government clients. These telephone centres are equipped with around 300 CATI stations and 1,150 interviewers.

4.4 Quality control

Strong emphasis is placed on quality control, in terms of ensuring that the surveys are administered according to the instructions and conducted to the highest standards. At least 7% of interviews are monitored as the interview occurs by the supervisory team using undetected listening facilities.

Prioritising and selecting which interviewers are monitored is done on a systematic basis, taking into account factors such as experience, previous performance and how often they have been monitored in the past. Monitoring results are graded using a standard benchmark and interviewers receive regular feedback on their performance.

4.5 Response rate

A breakdown of fieldwork outcomes is provided in Table 4 for the sample issued in November 2011. The sample was issued at the beginning of November 2011, staying active for three months and was removed from the CATI system at the end of January 2012. A total of 4,848 interviews were achieved over this time period. This represents a response rate of 43% from workplaces where a telephone number was found and was useable. The overall conversion rate, taking into account all cases allocated to November 2011 regardless of whether a telephone number could be matched, was 25%.

The response rate for November sample issued (which stayed live until the end of January 2012) provides an indication of the likely response rates going forward, although it should be borne in mind that the structure of fieldwork was slightly out of the ordinary compared with interviewing at other times of the year (due to the Christmas period and it being the start of the survey). The quoted response rates therefore remain provisional and firmer inferences should only be made once further fieldwork batches are completed.

Table 4: Fieldwork outcomes from cases issued in November 2011

	n	% telephone matched	% in and out of scope	% in scope resolved
TOTAL SAMPLE ALLOCATED TO NOVEMBER FIELDWORK BATCHES	**19,026**	**100.0**		
No telephone number	5,919	31.1		
TOTAL SAMPLE ISSUED	**13,107**	**68.9**	**100.0**	
OUT OF SCOPE (DEADWOOD)	**1,957**		**14.9**	
Fax	189		1.4	
Unobtainable	479		3.7	
Residential number (not a business)	109		0.8	
Business Moved	216		1.6	
Business Closed	190		1.4	
Incorrect Sample details	734		5.6	
Other	40		0.3	
IN SCOPE	**11,150**		**85.1**	**100.0**
Unresolved Live Sample	**3,832**			**34.4**
General Call Back	695			6.2
No reply/engaged/fax/modem/ answer machine (after maximum number of calls made) - Made Contact	2,316			20.8
No reply/engaged/fax/modem/ answer machine - No Contact Made	779			7.0
Other	42			0.4
Resolved Sample (ineffective)	**2,470**			**22.1**
Abandoned interview	228			2.0
Hard refusal	630			5.7
No such job title and no-one responsible	43			0.4
Soft refusal	779			7.0
Proxy refusal	682			6.1
Office refusal	82			0.7
Other	26			0.2
Total Interviews	**4,848**			**43.5**

Of the total sample issued, the level of deadwood was 15%. The largest categories within these deadwood records were those where the sample details from the IDBR were incorrect (6% of total sample issued) and where the number dialled did not connect (4% of total sample issued).

Of the remaining in-scope records, over four in ten (43%) workplaces completed the survey. One in five (22%) workplaces contacted resulted in an unproductive outcome e.g. refused to take part in the interview. However the remaining third (34%) of the in-scope records were not fully resolved before fieldwork on these records ceased.

The majority of these records were where an interviewer had made contact at least once, but subsequent contact attempts received no reply, an engaged tone or reached a fax, modem

or answer machine, and had reached the maximum number of calls (60% of the unresolved records)[14].

Table 5 presents the broad fieldwork outcomes from cases issued in November 2011, broken down by workplaces with different employment levels.

Table 5: Fieldwork outcomes from cases issued in November 2011, by IDBR employment size band

	Overall	IDBR employment		
		0-49	50-249	250+
	%	%	%	%
TOTAL SAMPLE ALLOCATED TO NOVEMBER FIELDWORK BATCHES	100.0	100.0	100.0	100.0
No telephone number	31.1	34.1	23.1	21.8
TOTAL SAMPLE ISSUED	100.0	100.0	100.0	100.0
OUT OF SCOPE (DEADWOOD)	14.9	15.6	12.8	17.4
IN SCOPE	100.0	100.0	100.0	100.0
Unresolved Live Sample	34.4	34.2	37.2	20.0
Resolved Sample (ineffective)	22.1	21.5	22.8	28.3
Total Interviews	43.5	44.3	40.0	51.7

As already noted in Table 2, larger workplaces had a higher telephone matching rate and this is reflected in the proportion of cases issued to the telephone centres.

The deadwood rates were similar across the three size bands (ranging from 13% to 17%). However it was noticeable that a higher response rate of 52% was achieved across workplaces with an employment level of 250 or more. This compared with a response rate of 44% across those with an employment count less than 50 and 40% across workplaces with an employment count of between 50 and 249.

One of the key features of the EBS is the ability to produce regional estimates. Table 6 looks at the fieldwork outcomes achieved in each English region for cases issued in November 2011.

[14] The maximum number of calls if an interviewer has made contact previously is 40.

Table 6: Fieldwork outcomes from cases issued in November 2011, by region

	Overall	Region		
		North East	North West	Yorkshire and Humber
	%	%	%	%
TOTAL SAMPLE ALLOCATED TO NOVEMBER FIELDWORK BATCHES	100.0	100.0	100.0	100.0
No telephone number	31.1	30.7	27.4	29.4
TOTAL SAMPLE ISSUED	100.0	100.0	100.0	100.0
OUT OF SCOPE (DEADWOOD)	14.9	16.4	15.2	14.0
IN SCOPE	100.0	100.0	100.0	100.0
Unresolved Live Sample	34.4	33.8	37.0	33.0
Resolved Sample (ineffective)	22.1	21.5	21.3	22.3
Total Interviews	43.5	44.8	41.7	44.8
		East Midlands	**West Midlands**	**East of England**
	%	%	%	%
TOTAL SAMPLE ALLOCATED TO NOVEMBER FIELDWORK BATCHES	100.0	100.0	100.0	100.0
No telephone number	31.1	30.8	29.4	30.9
TOTAL SAMPLE ISSUED	100.0	100.0	100.0	100.0
OUT OF SCOPE (DEADWOOD)	14.9	13.7	14.6	12.9
IN SCOPE	100.0	100.0	100.0	100.0
Unresolved Live Sample	34.4	33.5	35.8	34.8
Resolved Sample (ineffective)	22.1	20.5	20.1	22.0
Total Interviews	43.5	46.0	44.1	43.2
		London	**South East**	**South West**
	%	%	%	%
TOTAL SAMPLE ALLOCATED TO NOVEMBER FIELDWORK BATCHES	100.0	100.0	100.0	100.0
No telephone number	31.1	35.5	33.1	32.6
TOTAL SAMPLE ISSUED	100.0	100.0	100.0	100.0
OUT OF SCOPE (DEADWOOD)	14.9	19.7	14.4	13.6
IN SCOPE	100.0	100.0	100.0	100.0
Unresolved Live Sample	34.4	37.9	33.3	29.9
Resolved Sample (ineffective)	22.1	23.7	24.7	23.0
Total Interviews	43.5	38.5	42.0	47.1

Telephone matching rates did show a difference between regions although these differences were small. The highest non-matching rate was from London workplaces where over one third (36%) of records could not be matched to a telephone number. The region with the greatest proportion of records matched to a telephone number was the North West (only 27% of records remained unmatched).

London also exhibited the highest deadwood rate from cases issued to the telephone centres. One in five (20%) records were out of scope for interviewing in this region. Coupled with the lower telephone matching rate, this meant that fewer records were available to achieve the target number of interviews in this region.

It should be noted that the number of records initially drawn from the IDBR was based on an assumption of equal telephone matching and sample conversion rates across regions and workplace size bands, as this information was not available prior to starting fieldwork. However as more information becomes available the number of cases drawn in each region and workplace size group will vary according to the actual experience gained from administering the survey.

Finally, there were response rate differences between regions from the 'in scope' sample. London achieved the lowest response rate across all of the English regions (38%) while the South West achieved the highest (47%).

The data were weighted to compensate for this difference in response across different areas (see Section 6).

5. Data preparation

5.1 Coding of open questions and other answers

The EBS questionnaire contains no open ended questions within the core set of questions. Occasionally, the ad hoc questions that can be added to a calendar month's questionnaire may require the need for coding of open ended and 'other (specify)' responses. When this is the case, all 'other' answers will be inspected to check whether they should have been assigned to one of the pre-coded answers. Where there are cases that give similar or identical responses additional codes will be added to the code frames to simplify analysis.

5.2 Tabular output

Monthly data tables are produced in Microsoft Excel format displaying data weighted to the population of employment. The tables include the weighted percentage totals only; all absolute numbers are suppressed with the exception of unweighted base sizes.

Suppression rules have been devised through reviews of other surveys and through discussions between BIS, the ONS and TNS-BMRB. Subsequently, unweighted base counts are rounded to base 5 and column percentages are rounded to the nearest whole number. Figures are suppressed for any columns where the base is less than 50. Cell sizes with an unweighted count of less than 5 but not equal to one are also suppressed.

5.3 Provision of data in SPSS

In producing the SPSS data file, checks are carried out by TNS-BMRB to ensure that all variables are fully-labelled in terms of both variable names and code frame labels. There is also consistent formatting of standard codes (e.g. Don't know, Refused codes) and consistent definitions of missing values.

The SPSS data set is fully checked by research executives for both accuracy and comprehensiveness. The SPSS data set is checked against the whole counts to ensure that the data set fully reflects the raw data from which it has been exported. The SPSS data set is also fully checked to ensure that all filters fully match the specification i.e. checking whether the correct number of respondents answered each question that they were supposed to, according to the paper version of the questionnaire.

5.4 BIS processing and validation

When BIS receive the Excel tables from TNS-BMRB they are run through an automated Quality Assurance Procedure. This involves extracting the same tables from the microdata and then directly comparing results with the TNS-BMRB table sets. This quality assurance procedure ensures that the correct figures are displayed in the tables, the correct suppression rules have been applied and the data have been rounded accordingly.

5.5 Online reporting tool

An online reporting tool is currently being developed to assist with dissemination of the survey results. More details will be available at a later date on the EBS page of the BIS website.

5.6 Dataset release

There are plans to deposit the SPSS data in the ONS Virtual Microdata Laboratory (VML)[15] and an aggregated version of the data in the UK Data Archive[16] to enable other researchers and analysts to use the data. Further details will be available at a later date on the EBS page of the BIS website.

6. Weighting and estimation

6.1 Employment (economic importance) weights

To enable inferences from the sample of workplaces to the entire population of workplaces in England the sample data need to be weighted. This entails assigning each responding workplace a weight based on its employment as listed on the IDBR.

Employment is used as a measure of economic importance for each workplace. Turnover is another metric that could have been used to assess economic importance but this information is not available at the workplace level on the IDBR (only at the enterprise or business level), so employment is therefore used in its place. These weights are then scaled so that they sum to the total number of interviews collected each month. These are not traditional survey weights, because of the inclusion of employment in the weight. This means that they should not be used with survey variables which already report quantities related to the size of the business e.g. totals, and extra care is needed when carrying out statistical inference, including calculating standard errors.

When reading the results, for example, reading a national estimate of 10%, this should be interpreted as workplaces accounting for 10% of employment in England providing this response, rather than 10% of workplaces in England providing this response.

---

[15] The Virtual Microdata Laboratory (VML) is a facility within the Office for National Statistics (ONS) which enables access to restricted microdata for research purposes. Researchers from government and academia use the VML to carry out research on ONS surveys and other confidential datasets. Further details can be found at http://www.ons.gov.uk/ons/about-ons/who-we-are/services/vml/index.html.

[16] The UK Data Archive is curator of the largest collection of digital data in the social sciences and humanities in the United Kingdom. Further details can be found at http://www.data-archive.ac.uk.

As the sample of workplaces for the EBS is chosen with unequal probabilities, compensating weights need to be used in order to produce unbiased estimates of population quantities. The primary weight created for the EBS represents the employment profile across England, rather than the profile of workplaces.

There are a number of stages to the weighting, outlined below:

1. The first stage weight (sampling weight) is calculated as the inverse of the probability of selection. As outlined in Section 2, the probability of selection for an individual workplace is determined by the NUTS2 sub region and employment size band it is assigned to on the IDBR, as well as the raw employment count of the workplace.

2. The sampling weight is then adjusted by multiplying it by the employment total of each interviewed workplace as recorded on the IDBR.

3. This adjustment to the design weight produces a 'base weight' for calibrating to population totals. These totals are based on the latest available IDBR information for (a) total employment band, (b) NUTS2 sub region, and (c) industry sector with no interactions between these variables.

4. To avoid excessive variance in weights both across the sample and within key subgroups, these weights are trimmed within each NUTS1 region. The minimum weight is set at 1/5 of the median untrimmed weight in the relevant region; the maximum weight is set at 5 times the median untrimmed weight in the relevant region. Once this stage is complete, a second calibration to population totals is carried out but this time excluding total employment band.

5. The final weights ensure that the sample is representative of employment by NUTS1 region and by industry sector and as representative as possible of total employment band.

Table 7 illustrates the final unweighted and weighted profile of the November 2011 fieldwork interviews by IDBR employment band, industry sector and region, when the data are weighted to the employment profile of England.

It can be seen that, in practice, workplaces in the '250+' employment band are under-represented by the final weights. Ultimately, the weights represent a pragmatic trade-off between two competing types of error: bias and variance. The weighting scheme will continue to be evaluated to ensure that the trade off continues to meet the requirements of the survey.

Table 7: Profile of November 2011 EBS data, weighted by employment

	Unweighted	Weighted by employment	IDBR population (employment)
	%	%	%
IDBR employment band			
1-2	7.5	7.3	6.7
3-4	8.1	5.6	5.0
5-9	14.0	9.8	8.8
10-19	17.8	11.4	10.2
20-49	24.8	17.5	15.5
50-99	14.4	14.0	12.4
100-249	9.0	16.2	14.4
250+	4.4	18.1	27.0
Industry (SIC 2007 section)			
A: Agriculture, Forestry And Fishing	3.1	1.3	1.3
B: Mining And Quarrying	0.1	0.1	0.1
C: Manufacturing	8.6	9.1	8.5
D: Electricity, Gas, Steam And Air Conditioning Supply	-	-	0.4
E: Water Supply; Sewerage, Waste Management And Remediation Activities	0.4	0.4	0.6
F: Construction	4.8	4.8	4.8
G: Wholesale And Retail Trade; Repair Of Motor Vehicles And Motorcycles	20.2	16.4	16.4
H: Transportation And Storage	3.1	4.6	4.6
I: Accommodation And Food Service Activities	9.2	6.2	6.2
J: Information And Communication	2.2	3.9	3.9
K: Financial And Insurance Activities	1.2	3.8	3.8
L: Real Estate Activities	1.8	1.7	1.7
M: Professional, Scientific And Technical Activities	6.7	7.6	7.6
N: Administrative And Support Service Activities	6.1	8.1	8.1
O: Public Administration And Defence; Compulsory Social Security	2.0	5.3	5.3
P: Education	12.1	9.5	9.5
Q: Human Health And Social Work Activities	12.8	12.5	12.5
R: Arts, Entertainment And Recreation	2.8	2.5	2.5
S: Other Service Activities	3.0	2.2	2.2
T: Activities Of Households As Employers; Undifferentiated Goods-And Services-Producing Activities Of Households For Own Use	-	-	0.0
U: Activities Of Extraterritorial Organisations And Bodies	-	-	0.0
Region			
1 North East	8.5	4.3	4.3
2 North West	13.8	12.8	12.8
3 Yorkshire and The Humber	10.8	9.5	9.5
4 East Midlands	9.9	8.2	8.2
5 West Midlands	12.6	10.0	10.0
6 East of England	9.6	10.6	10.6
7 London	8.7	18.3	18.3
8 South East	12.6	16.2	16.2

	Unweighted	Weighted by employment	IDBR population (employment)
	%	%	%
9 South West	13.6	10.0	10.0
Base	*3,080*	*3,080*	*23,205,600*

Note: The dash (-) means that there were no workplaces in this subgroup interviewed.

6.2 Experimental workplace weights

For analyses at the workplace level a different weight is derived. Since the profile of workplaces by size tends to go in the opposite direction to employment and given that the sampling method was chosen to most closely represent the employment distribution within England, a less efficient sample is expected when weighted to represent workplaces.

The weighting approach is similar to the derivation of the employment (economic importance) weights outlined in Section 6.1. The major difference is that no adjustment is made for a workplace's employment count as listed on the IDBR. The stages are outlined below:

1. The same first stage weight (sampling weight) is used, as when deriving the employment weights. This is calculated as the inverse of the probability of selection.

2. This sampling weight is then calibrated to population totals for workplaces. These totals are based on the latest available IDBR information for (a) total employment band, (b) NUTS2 sub region, and (c) industry sector with no interactions between these variables.

3. As with the employment weighting, to avoid excessive variance in weights both across the sample and within key subgroups, these weights are trimmed within each NUTS1 region. The minimum weight is set at 1/5 of the median untrimmed weight in the relevant region; the maximum weight is set at 5 times the median untrimmed weight in the relevant region. Once this stage is complete, a second calibration to population totals is carried out but this time excluding total employment band.

4. The final weights ensure that the sample is representative of workplaces by NUTS1 region and by industry sector and as representative as possible of total employment band.

Table 8 illustrates the final unweighted and weighted profile of the November 2011 fieldwork interviews by IDBR employment band, industry sector and region, when the data are weighted to the workplace profile of England.

In practice, although the industry and region profiles of the sample match the population profiles, the profile of workplaces in the various employment bands are significantly different. The smallest workplaces (those with employment of 2 or less) are considerably under-represented. However, as with the employment weighting, ultimately, the weights represent a pragmatic trade-off between two competing types of error: bias and variance. It should also be borne in mind that the sampling procedure was designed with a view to maximising the efficiency when weighted to employment, not workplaces. As with many aspects of the EBS, this will be reviewed as the survey progresses. All tables initially published from the EBS will use the employment weights.

Table 8: Profile of November 2011 EBS data, weighted to workplaces

	Unweighted	Weighted to workplaces	IDBR population (workplaces)
	%	%	%
IDBR employment band			
1-2	7.5	22.0	52.9
3-4	8.1	20.5	15.6
5-9	14.0	25.3	14.5
10-19	17.8	16.2	8.1
20-49	24.8	9.8	5.5
50-99	14.4	3.4	2.0
100-249	9.0	1.9	1.0
250+	4.4	0.9	0.5
Industry (SIC 2007 section)			
A: Agriculture, Forestry And Fishing	3.1	4.4	4.4
B: Mining And Quarrying	0.1	0.2	0.1
C: Manufacturing	8.6	5.4	5.3
D: Electricity, Gas, Steam And Air Conditioning Supply	-	-	0.1
E: Water Supply; Sewerage, Waste Management And Remediation Activities	0.4	0.3	0.4
F: Construction	4.8	10.7	10.7
G: Wholesale And Retail Trade; Repair Of Motor Vehicles And Motorcycles	20.2	19.0	19.0
H: Transportation And Storage	3.1	3.2	3.2
I: Accommodation And Food Service Activities	9.2	6.2	6.2
J: Information And Communication	2.2	6.6	6.6
K: Financial And Insurance Activities	1.2	2.7	2.7
L: Real Estate Activities	1.8	3.6	3.6
M: Professional, Scientific And Technical Activities	6.7	14.3	14.3
N: Administrative And Support Service Activities	6.1	7.2	7.2
O: Public Administration And Defence; Compulsory Social Security	2.0	0.9	0.9
P: Education	12.1	2.6	2.6
Q: Human Health And Social Work Activities	12.8	5.6	5.6
R: Arts, Entertainment And Recreation	2.8	3.0	3.0
S: Other Service Activities	3.0	4.0	4.0
T: Activities Of Households As Employers; Undifferentiated Goods-And Services-Producing Activities Of Households For Own Use	-	-	0.0
U: Activities Of Extraterritorial Organisations And Bodies	-	-	0.0
Region			
1 North East	8.5	3.4	3.4
2 North West	13.8	11.6	11.6
3 Yorkshire and The Humber	10.8	8.5	8.5
4 East Midlands	9.9	7.9	7.9
5 West Midlands	12.6	9.6	9.6
6 East of England	9.6	11.6	11.6
7 London	8.7	18.2	18.2
8 South East	12.6	18.1	18.1

	Unweighted	Weighted to workplaces	IDBR population (workplaces)
	%	%	%
9 South West	13.6	11.0	11.0
Base	*3,080*	*3,080*	*2,161,200*

Note: The dash (-) means that there were no workplaces in this subgroup interviewed.

6.3 Balance figures

Some of the individual tables within the monthly data tables set include a 'balance' figure. This is the difference between the positive response and negative response. For example, if at the question for the level of business activity last month compared with 3 months before (Qoutput) 55% said 'Higher', 20% said 'The same' and 25% said 'Lower', the balance would be 30% (55%-25%).

The balance figures are additional measures that are intended to be read in conjunction with the underlying figures for those tables.

7. Sampling errors

7.1 Individual estimates

Sampling errors result from taking a sample of the population rather than a census, since a different sample may produce a different estimate. The extent of this depends on the natural variation in any measure that is collected and the sample size achieved.

It is important to take the sampling design, weighting strategy and stratification into account when calculating estimates of standard errors and confidence intervals. Ignoring the sample design (i.e. assuming simple random sampling and taking into account only sample size and element variance) usually results in estimated standard errors that are too small. An adjustment factor, called a design factor, needs to be calculated to account for these differences. The design factor is the ratio of the standard error of the complex design to the standard error of a simple random sample (SRS) of the same size.

Put another way, the design factor (deft) is the factor by which the standard error of an estimate from a simple random sample has to be multiplied to give the true standard error of the complex design (this is referred to as the complex standard error or the true standard error).

A design factor of less than one arises when a sample is more precise and has smaller standard errors than a SRS, while a design factor greater than one arises when a sample has standard errors larger than those that would be obtained from a SRS of the same size. Each estimate has its own particular design factor.

Standard errors form the basis for calculating confidence intervals associated with particular estimates. A 95% confidence interval for a population estimate is ±1.96 standard errors around the estimate calculated from the sample. Table 9 presents the confidence intervals around survey estimates of various magnitudes if the sample is assumed to be a simple random sample (SRS).

Table 9: 95% Confidence intervals associated with different survey estimates, using a simple random sample

Sample size	10% or 90%	25% or 75%	50%
100	+/- 5.9%	+/- 8.7%	+/- 9.8%
150	+/- 4.9%	+/- 7.1%	+/- 8.2%
200	+/- 4.3%	+/- 6.1%	+/- 7.1%
250	+/- 3.7%	+/- 5.5%	+/- 6.2%
300	+/- 3.5%	+/- 5.0%	+/- 5.8%
400	+/- 3.0%	+/- 4.3%	+/- 5.0%
1,000	+/- 1.9%	+/- 2.8%	+/- 3.1%
1,500	+/- 1.5%	+/- 2.3%	+/- 2.5%
2,000	+/- 1.3%	+/- 1.9%	+/- 2.2%
3,000	+/- 1.1%	+/- 1.6%	+/- 1.8%

However the EBS does not use a simple random sampling approach; it has a complex sample design. More precise design factors can be calculated by using statistical software packages which account for the survey design.

The confidence intervals can be calculated for a complex sample design by multiplying the SRS confidence interval by the relevant design factor.

For example, if the SRS standard error was 0.9% and the design factor was 1.38, the complex standard error would equal 1.2%. If the survey estimate was 50%, the SRS confidence interval would be:

50% +/- (1.96*0.9%) = between 48.2% and 51.8%

The 95% confidence interval for the survey employing the complex sample design would be:

50% +/- (1.96*1.2%) = between 47.6% and 52.4%

Thus in 19 samples out of 20, the population value would be expected to lie within the 95% confidence intervals constructed i.e. between 47.6% and 52.4%. Only one time out of twenty would the population value be expected to lie outside the confidence interval for the survey estimate.

Table 10 shows the estimated design factors associated with a number of estimates based on the interview data from November 2011 fieldwork (October 2011 EBS), weighted to employment. These design factors were calculated using the Complex Samples module within IBM's SPSS software. This software incorporates the sample design when producing statistics and so generates the complex standard errors.

It should be noted that these are provisional estimates of confidence intervals and further exploratory work is being conducted to finalise these, the results of which will be published when available.

Table 10: 95% Complex standard errors for survey estimates

Variable	Estimate (%)	Standard error (%)	95% confidence interval		Design factor (deft)	N (unweighted)
			Lower (%)	Upper (%)		
QOUTPUT Level of business activity or volume of output last month compared with 3 months before						
-7 Don't know	2.2	0.3	1.7	2.9	1.128	85
1 Higher	35.1	1.2	32.7	37.5	1.434	985
2 The same	38.1	1.2	35.8	40.5	1.393	1,150
3 Lower	24.6	1.0	22.7	26.6	1.281	865
QFUTOUT Expectation of level of business activity or volume of output in 3 months time						
-7 Don't know	3.9	0.5	3.1	5.0	1.379	120
1 Higher	26.1	1.2	23.9	28.5	1.489	750
2 The same	39.3	1.2	36.9	41.8	1.410	1175
3 Lower	30.6	1.1	28.5	32.9	1.330	1035
QEXPT Whether site produces or sells goods or services for export directly overseas						
-7 Don't know	0.6	0.4	0.2	2.0	2.290	10
1 Yes	16.8	1.1	14.8	19.1	1.435	330
2 No	82.6	1.1	80.3	84.7	1.460	2,045
QEXPTFT Expected volume of exports in 3 months time compared with this month						
-7 Don't know	4.2	1.3	2.3	7.6	1.174	15
1 Higher	28.8	3.5	22.4	36.0	1.395	90
2 The same	46.1	3.5	39.3	53.1	1.281	155
3 Lower	20.9	3.0	15.7	27.3	1.321	70
QSTOCK Volumes of stocks or inventories this month compared with 3 months ago						
-7 Don't know	2.7	0.6	1.7	4.1	1.390	30
1 Higher	27.3	1.6	24.3	30.6	1.338	375
2 The same	38.9	1.7	35.7	42.3	1.289	560
3 Lower	17.7	1.2	15.5	20.2	1.175	280
5 Do not have any stocks	13.4	1.2	11.2	15.9	1.301	155
QFTSTOCK Expected volume of stocks or inventories in 3 months compared with this month						
-7 Don't know	2.8	0.6	1.9	4.1	1.249	35
1 Higher	11.8	1.1	9.9	14.1	1.236	185
2 The same	40.2	1.7	36.9	43.6	1.302	575
3 Lower	32.4	1.6	29.3	35.6	1.282	460
5 Do not have any stocks	12.8	1.2	10.7	15.2	1.300	150
QPPL Number in employment last month compared with 3 months before						
-7 Don't know	0.5	0.3	0.2	1.5	2.159	10
1 Higher	18.6	1.0	16.6	20.7	1.497	495
2 The same	63.9	1.2	61.4	66.2	1.406	2,090
3 Lower	17.0	1.0	15.1	19.1	1.496	485
QFTPPL Expected number in employment in 3 months compared with this month						
-7 Don't know	0.8	0.3	0.4	1.7	1.954	15
1 Higher	11.3	0.9	9.7	13.2	1.572	295
2 The same	67.8	1.2	65.4	70.2	1.467	2,235
3 Lower	20.1	1.1	18.0	22.3	1.515	535

Variable	Estimate (%)	Standard error (%)	95% confidence interval		Design factor (deft)	N (unweighted)
			Lower (%)	Upper (%)		
QHRS Number of actual hours worked last month compared with 3 months before						
-7 Don't know	1.3	0.3	0.8	2.2	1.668	35
1 Higher	29.3	1.2	27.0	31.6	1.443	765
2 The same	51.8	1.3	49.3	54.2	1.393	1,665
3 Lower	17.6	0.9	16.0	19.5	1.298	615
QPRCBKa Wage and other labour costs last month compared with 3 months before						
-7 Don't know	4.6	0.6	3.5	5.9	1.607	135
1 Higher	23.2	1.0	21.2	25.2	1.351	765
2 The same	65.7	1.2	63.3	68.0	1.392	1,950
3 Lower	6.6	0.6	5.4	7.9	1.384	205
QPRCFT Expected wage and other labour costs in 3 months compared with last month						
-7 Don't know	2.1	0.4	1.5	3.0	1.346	65
1 Higher	21.9	1.0	20.0	24.0	1.358	680
2 The same	67.3	1.2	65.0	69.5	1.360	2,015
3 Lower	8.6	0.7	7.4	10.0	1.321	290
QPCHRD Average prices charged last month compared with 3 months before						
-7 Don't know	3.0	0.5	2.1	4.2	1.480	65
1 Higher	18.0	1.1	15.9	20.2	1.389	440
2 The same	67.3	1.3	64.6	69.8	1.380	1,610
3 Lower	11.8	0.9	10.2	13.7	1.341	270
QPCHRDFT Expected average prices charged in 3 months compared with last month						
-7 Don't know	4.2	0.6	3.2	5.4	1.381	100
1 Higher	20.2	1.1	18.1	22.5	1.361	505
2 The same	64.9	1.3	62.3	67.4	1.356	1,540
3 Lower	10.7	0.8	9.2	12.4	1.288	250
QOUTYR Volume of business activity or output compared with the same time last year						
-7 Don't know	3.2	0.5	2.4	4.3	1.480	105
1 Higher	43.1	1.3	40.6	45.6	1.432	1,225
2 The same	26.9	1.1	24.9	29.0	1.333	870
3 Lower	26.7	1.1	24.6	28.9	1.367	875
region_code Region (NUTS1)						
1 North East	4.3	0.1	4.1	4.6	0.347	260
2 North West	12.8	0.3	12.3	13.4	0.469	425
3 Yorkshire and The Humber	9.5	0.2	9.1	10.0	0.423	335
4 East Midlands	8.2	0.2	7.7	8.6	0.493	305
5 West Midlands	10.0	0.3	9.5	10.6	0.483	385
6 East of England	10.6	0.3	10.1	11.1	0.459	295
7 London	18.3	0.5	17.4	19.2	0.652	270
8 South East	16.2	0.3	15.6	16.8	0.466	385
9 South West	10.0	0.2	9.6	10.5	0.410	420
industry_code Industry section (SIC 2007)						
1 Agriculture, Forestry And Fishing	1.3	0.1	1.0	1.6	0.724	95

Variable	Estimate (%)	Standard error (%)	95% confidence interval		Design factor (deft)	N (unweighted)
			Lower (%)	Upper (%)		
2 Mining And Quarrying	*	*	*	*	*	*
3 Manufacturing	9.1	0.7	7.8	10.5	1.291	265
4 Electricity, Gas, Steam And Air Conditioning Supply	-	-	-	-	-	-
5 Water Supply; Sewerage, Waste Management And Remediation Activities	0.4	0.2	0.2	0.9	1.347	10
6 Construction	4.8	0.5	4.0	5.8	1.214	145
7 Wholesale And Retail Trade; Repair Of Motor Vehicles And Motorcycles	16.4	0.8	14.8	18.0	1.235	620
8 Transportation And Storage	4.6	0.6	3.5	5.8	1.551	95
9 Accommodation And Food Service Activities	6.2	0.4	5.5	7.1	0.980	285
10 Information And Communication	3.9	0.6	2.9	5.2	1.693	70
11 Financial And Insurance Activities	3.8	0.7	2.6	5.5	2.075	35
12 Real Estate Activities	1.7	0.3	1.3	2.3	1.096	55
13 Professional, Scientific And Technical Activities	7.6	0.7	6.4	9.1	1.412	205
14 Administrative And Support Service Activities	8.1	0.7	6.8	9.6	1.476	190
15 Public Administration And Defence; Compulsory Social Security	5.3	0.9	3.9	7.3	2.130	60
16 Education	9.5	0.7	8.3	10.9	1.244	375
17 Human Health And Social Work Activities	12.5	0.8	11.1	14.2	1.316	395
18 Arts, Entertainment And Recreation	2.5	0.3	2.0	3.2	1.088	85
19 Other Service Activities	2.2	0.3	1.7	2.8	1.025	90
employment_cod IDBR employment band						
1 0-4	12.9	0.6	11.8	14.2	1.019	480
2 5-9	9.8	0.5	8.9	10.9	0.953	430
3 10-49	28.9	0.7	27.6	30.3	0.822	1310
4 50-99	14.0	0.5	13.0	15.1	0.879	445
5 100-249	16.2	0.8	14.7	17.8	1.181	275
6 250-499	11.7	0.8	10.2	13.4	1.386	95
7 500-999	3.5	0.6	2.4	4.9	1.905	25
8 1000-4999	3.0	0.8	1.8	4.9	2.506	15
9 5000+	-	-	-	-	-	-

Note: The star (*) means that the row was suppressed because there were too few workplace responses in this particular subgroup to produce reliable statistics. The dash (-) means that there were no workplaces in this subgroup interviewed.

As can be seen in Table10 there are design factors that are larger than one. This is because weights are required to compensate for different sampling rates in different strata. However this is countered by the stratification and the targeting of more variable strata, which leads to design factors less than one for the regional breakdowns where the latter effect is more dominant.

7.2 Comparing estimates to see whether a significant difference exists

One of the main uses of the EBS is to compare estimates across regions. However due to variation between random samples that could be drawn from the same population, differences between survey estimates may arise purely by chance, rather than as a result of genuine differences between the underlying populations. On the data tables accompanying the Statistical Release subgroup percentages have been tested against the equivalent percentage for England and if shaded, the difference was found to be statistically significant at the 95% level. The methodology for testing whether differences in results are significant continues to be developed. Tables from November 2011–January 2012 used a more conservative testing approach than the more precise approach employed for the February 2012 tables onwards. The two approaches are outlined below. The development work is not yet finalised therefore all significance testing marked on the tables is provisional.

The significance test employed is a t test. The t statistic is calculated as the difference in the proportions of the two samples, divided by the standard error of the difference between those proportions. If a t statistic is greater than the critical value of 1.96, or lower than -1.96, then the difference is significant at the 95% level.

Differences in the balance scores have also not been tested as the methodology for doing so is currently in development and will be published at a later date.

7.2.1 Significance testing for November 2011–January 2012 tables

Significance testing results up to and including the January 2012 EBS data had been reported using the following approximation method for estimating the variance of the difference outlined below.

The variance of the difference between the estimated subgroup percentage (y_1) and the estimated England percentage (y_2) is first calculated as:

$$V(y_1 - y_2) = V(y_1) + V(y_2) - 2R\sqrt{V(y_1)*V(y_2)}$$

where R is the correlation between two estimated sample proportions y_1 and y_2 and V represents the sampling variance. [17] With independent samples, R would be 0. With overlapping samples, it is generally positive. However, it is technically difficult to estimate R directly. In this instance, since the subgroup makes up part of the England total, the correlation coefficient for the overlapping proportion is 1 and for the non-overlapping proportion is 0. Thus, we have approximated R using P the proportion of all cases coming from this region as follows:

[17] Taken from Kalton, G. (1983), *Introduction to Survey Sampling*, Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-035, Beverly Hills, CA, and London: Sage Publications.

$$V(y_1 - y_2) = V(y_1) + V(y_2) - 2P\sqrt{V(y_1)*V(y_2)}$$

It should be noted that this method is conservative and may result in fewer instances of significant differences than would be the case with a more precise method, which is being investigated.

7.2.2 Significance testing for February 2012 tables onwards

In the above approach substituting the correlation coefficient with the proportion is not ideal despite the range for the values being the same. Subsequent development work led to a more precise approach being derived from standard statistical identities. This is defined as:

$$V(y_2 - y_1) = V(y_1) + V(y_2) - 2PV(y_1)$$

This more precise variance estimation method has been used for reporting the significance testing results for the February 2012 data and onwards.

As mentioned earlier the estimates of standard errors and confidence intervals continue to be provisional as further exploratory work is being conducted to finalise these.

Appendix A: Reassurance letter

Dear Sir, Madam,

ENGLISH BUSINESS SURVEY

The Department for Business, Innovation and Skills (BIS) has asked TNS-BMRB, an independent research agency, to conduct a survey among businesses across the country. The survey aims to provide BIS with regular and up to date assessments of how businesses perceive current economic and business conditions across England.

The questionnaire will be covering business issues such as employment, costs and investments. It will not be asking for any detailed information about your business performance, just whether things have got better or worse over the last few months and what you expect to happen in the next few months. If your organisation has more than one workplace we would like you to answer the questions based on the specific site that we refer to in the interview.

Each interview will take about 10 minutes. We would be very grateful if you would agree to take part. If you would like to find out any further information about this research please visit the English Business Survey website at www.ebsurvey.co.uk .

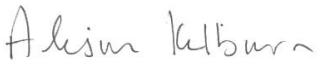
As a result of this survey the government and your local enterprise partnership will have a much better understanding of how businesses in your area are faring in the current economic climate. Findings from the survey will be placed on the BIS website on a regular basis. More information is available at http://www.bis.gov.uk/ebsurvey

TNS-BMRB are bound by the Market Research Society Code of Conduct. This means that all survey responses and data will be treated confidentially. The full survey data, including company demographic information, will be used for research purposes only by researchers within BIS, the National Institute for Economic and Social Research, TNS-BMRB and the ONS data lab. Published data, held on the BIS and the English Business Survey website, will be anonymised. No individual firm will be identifiable in published findings.

If you have any further questions about this research please contact Hannah Kilshaw at TNS-BMRB on 0207656 5748 or Hannah.Kilshaw@tns-bmrb.co.uk quoting the ID number in the subject header of the correspondence.

Thank you in advance for your cooperation.

Yours sincerely,



Alison Kilburn
Deputy Director, Enterprise and Economic Development Directorate
Department for Business, Innovation and Skills
t: +44 (0)20 7215 3107 e: EEDAdeputydirector@bis.gsi.gov.uk

This publication is also available on our website at www.bis.gov.uk
Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

If you require this publication in an alternative format, email enquiries@bis.gsi.gov.uk, or call 020 7215 5000.

URN 12/602X