Standard Report
on
Methods and Quality (v1)

for
Irish Health Survey

This documentation applies to the reporting period:

2015

Last edited: <3 November 2017>

CENTRAL STATISTICS OFFICE

Skehard Road, Cork

021 4535000

www.cso.ie

Table of Contents

1. Overview

The Irish Health Survey (IHS) was collected under Regulation (EC) No 1338/2008 of the European Parliament and of the Council of 16 December 2008 on Community statistics on public health and health and safety at work1. Outputs from this survey fulfil the need for public health policies to obtain reliable data on health status, health care usage and health determinants.

Demographic information was collected in the field by a team of face-to-face interviewers using Computer Assisted Personal Interviewing (CAPI) on laptop computers (using a Blaise application). This was followed by a issuing a paper questionnaire (with a web-based option) to a random respondent aged 15 or over, which was then posted back to the Central Statistics Office.

The survey population is all individuals living in households who are aged 15 years or older. The sample achieved was 10,323 respondents.

2. General Information

2.1Statistical Category

Primary Statistical Survey

2.2 Area of Activity

Health statistics

2.3 Organisational Unit Responsible, Persons to Contact

The relevant persons are part of the Social and Demographic Statistics Directorate.

The work of the IHS is part of the Social Analysis division, which is headed by a Senior Statistician, who also has responsibility for other survey areas.

Irish Health Survey queries:

Damien Lenihan Tel: +353 021 453 5424 Email: ihs@cso.ie

2.4 Objectives and Purpose; History

The IHS was carried out in Ireland for the first time in 2015. The survey is part of a wider European survey called the 'European Health Interview Survey'. This is the second wave of the survey. Ireland did not partake in the first wave.

2.5 Periodicity

The periodicity of the survey is every six years.

2.6 Client

Each Member State of the EU must undertake a European Health interview Survey (EHIS) to provide information on key health indicators. The IHS is the Irish implementation of the EHIS. It provides important inputs to national policy makers.

2.7 Users

- European Union/Eurostat
- Government departments (Department of Health, Department of Social Protection, Department of Children and Youth Affairs, etc.)
- Research centres and universities involved in social research
- National media
- The general public

2.8 Legal basis

The IHS is carried out under EU Council Regulation No 1338/2008. There are a number of implementing regulations which outline various aspects of the survey. These can be found at the following link;

http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_%E2%80%93_main_features_and_legal_basis

While Ireland as a Member State is obliged to undertake the IHS, participation in the survey is voluntary for respondents.

3 Statistical Concepts, Methods

3.1 Subject of the Statistics

Health statistics

3.2 Units of Observation/Units of Presentation

The survey population is individuals living in private households. It therefore excludes individuals living in institutions or communal accommodation and persons of no fixed abode. The collection units are households containing at least one individual aged 15 years or over for whom it is the main residence. Information is then collected from one random person aged 15 year years or over who is selected randomly from those living in the household.

The main units of presentation are:

- Demographic variables such as Sex, Age, Nationality
- Other personal characteristics such as Regional Classification NUTS 3 (Nomenclature of Territorial Units), disability status, level of advantage/disadvantage etc.
- International Labour Office (ILO) employment status

While the above are the primary presentation units for publication, the IHS can present data according to a wide variety of classifications based on the comprehensive range of questions asked in the QNHS, of which the IHS was asked as a sub-module. Such analysis is often provided on an ad hoc basis following user requests.

3.3 Data Sources

Information for the IHS is collected from individuals in households.

3.4 Reporting Unit/Respondents

The respondent to the IHS is a randomly selected person aged 15 years or older, who is a 'usual resident'[1]. One individual is selected to answer per household. Proxy responses are not allowed. A Proxy response refers to data which is collected from other another member of the household due to the unavailability of the specific respondent at the time of interview.

3.5 Type of Survey/Process

Sample survey

[1] A person is defined as a "Usual Resident" of a private household if he or she

(i) Lives regularly at the dwelling in question

And

(ii) Shares the main living accommodation (i.e. kitchen, living room or bathroom) with the other members of the household.

3.6 Characteristics of the Sample/Process

3.6.1 Population and Sampling Frame

The reference population is all individuals aged 15 years or older living in private households in Ireland. It therefore excludes persons with no usual address or those with a usual address in a public institution, such as hospitals, nursing homes etc. All usual residents in each household are included. The sampling frame is all private households in Ireland. The sampling frame is based on the Census of Population 2011.

3.6.2 Sample Design

The sample design is based on the sample design used in the Quarterly National Household Survey (QNHS), which is the national survey carried out to measure the labour force.

With this design, a two-stage sample is used. The sample frame of households is clustered into blocks (small areas) with each block containing a minimum of 60 occupied households on the night of the 2011 Census of Population. In the case of the sample introduced in 2012, the sample frame is stratified using administrative county and population density. In the first stage 1,300 blocks are selected using Probability Proportional to Size (PPS) sampling and in the second stage 20 households are selected using Simple Random Sampling (SRS). This ensures that each household in the ample frame has an equal probability of selection, and results in a total sample of 26,000 households.

Following this, an individual aged 15 years or older is randomly selected to take part in the IHS.

3.7 Survey Technique/Data Transfer

The core health variables on the IHS were collected by way of a self-administered paper questionnaire. These questionnaires were returned to the Central Statistics Office and inputted once a week.

The demographic data, including variables such as sex, age, NACE etc. were collected as part of the QNHS. QNHS interviewers visit each household and ask the residents to voluntarily participate in the survey. Data are collected on encrypted tablets using CAPI (Computer Assisted Personal Interviewing). Every week, the survey interviewer transmits the weekly data collected from households to head office using a secure encrypted data tunnel.

3.8 Questionnaire (including explanations)

The IHS questionnaire contains approximately 56 questions covering the following topics;

- European Health Status Module (EHSM) – This module on health status is the central point of the survey. It allows measurement of the health status of the population in general and not only in relation with specific health problems. It covers different aspects and dimensions of health: physical and mental health, chronic and temporary problems, specific conditions but

also their general impact on the functional status and the limitation in activities of daily living of the respondents.

- European Health Care Module (EHCM) - The ECHM module collects data on the use of health care services and the unmet needs for health care. Information on health care consumption is an essential part of the health information system in order to assign necessary resources to the population. Administrative data may provide more reliable and accurate data on health care services, but not necessarily comparable between countries. The advantage of observing the data via EHIS is that firstly, we can receive comparable data for all countries due to same method of data collection; and secondly, it enables linking the data with characteristics of health status, health determinants and socio-economic characteristics. This allows analysing the relations between health consumption and several determinants such as health status, lifestyles or socio-demographic characteristics as well as the relations between different types of health care use. As such EHIS data permit the comparison of the health needs and health consumption and thus make it also possible to explore the concepts of vertical and horizontal equity in health care.
- European Health Determinants Module - The general focus of the module is to measure some aspects in lifestyles or health-related behaviours having a positive or negative impact on someone's health state. Better lifestyles are probably the main potential source of improvement in the health of the population. For public health actors in health promotion it is essential to measure regularly the prevalence of specific health-related behaviours and their trends at population level and in specific population subgroups. Such measurement is imperative for the evaluation of programmes and policies and for raising awareness of the population.

The IHS questionnaire is available on the CSO website:

http://www.cso.ie/en/methods/surveyforms/irishhealthsurvey/

3.9 Participation in the Survey

Participation in the survey is voluntary for respondents.

3.10 Characteristics of the Survey/Process and its Results

The survey is designed to provide information on various aspects of health for the State.

The data which were published refer to the estimate number of persons aged 15 years or over in the state classified by various characteristics. These characteristics include;

- 10-year age groups
- Sex

- Economic status in accordance with the ILO definition etc.

For a full list of published indicators please see the IHS release on

http://www.cso.ie/en/releasesandpublications/ep/p-ihs/irishhealthsurvey2015/

3.11 Classifications Used

There are a number of different classifications used in the IHS:

- The main classifications of economic activity are the standard ILO definitions of persons in employment, unemployment (which are summed to derive the labour force), and persons not in the labour force.
- Regional data is coded to NUTS3 as described in Section 3.12
- Disability is calculated as following: Questions ten and twelve of the questionnaire identify various difficulties that individuals may experience. These difficulties are both physical and non-physical in nature. These questions have been used to identify an individual's disability status. If a respondent has identified 'a lot of difficulty' or 'cannot do at all' as a response to any of the sub-categories in question ten, the respondent is identified as having a disability. Similarly, if the respondent has identified 'Quite a bit' or 'Extremely' as a response to question twelve, they are identified as having a disability. All other responses to these two questions implied that a respondent does not have a disability.
- The Pobal Haase-Pratschke Deprivation Index is used to analyse Irish Health Survey questionnaire responses experienced by individuals. The Index uses Census data to measure levels of disadvantage or affluence in a particular geographical area. More detailed information on the index can be found here: https://www.pobal.ie/Pages/New-Measures.aspx. The results are presented by quintiles, five equal-sized groups of households, with the first quintile representing the least deprived/most affluent area and the fifth quintile representing the most disadvantaged areas.

3.12 Regional Breakdown of Results

The regional classifications used are based on the NUTS (Nomenclature of Territorial Units) classification used by Eurostat. The NUTS3 regions correspond to the eight Regional Authorities established under the Local Government Act, 1991 (Regional Authorities) (Establishment) Order, 1993, which came into operation on 1 January 1994. The NUTS2 regions, which were proposed by Government and agreed by Eurostat in 1999, are groupings of the NUTS3 regions. The sample itself is designed to be representative at NUTS3 level although some results may be repressed depending on the number of observations in the particular cell. The composition of the regions is set out below:

NUTS2: Border, Midlands, and Western		Southern and Eastern	
NUTS3: Border	Cavan	Dublin	Dublin
	Donegal		Dún Laoghaire-Rathdown
	Leitrim		Fingal
	Louth		South Dublin
	Monaghan		
	Sligo	Mid-East	Kildare
			Meath
Midlands	Laoighis		Wicklow
	Longford		
	Offaly	Mid-West	Clare
	Westmeath		Limerick City
			Limerick County
West	Galway City		North Tipperary
	Galway County		
	Mayo	South-East	Carlow
	Roscommon		Kilkenny
			South Tipperary
			Waterford City
			Waterford County
			Wexford
		South-West	Cork City
			Cork County
			Kerry

4 Production of the Statistics, Data Processing, Quality Assurance

4.1 Data Capture

Demographic data is collected in the field by a tea, of interviewers using tablets (CAPI using a Blaise application) and data are then transmitted to the main processing unit in the CSO.

Data from the questionnaire, once received, was scanned and then converted into a SAS file for analysis.

4.2 Coding

Occupation and Industry text strings are captured in the field to the relevant classifications (see Section 3.11) by the interviewers using the Blaise application. The codes assigned are then subsequently checked for quality purposes.

4.3 Data Editing

In the QNHS, the majority of questions only allow answers to be entered to a limited set of pre-defined categories and therefore the number of edits required is limited.
However, once the IHS paper questionnaire was scanned and inputted, logic edits were carried out. These included reviewing questions that were answered in error by the respondent e.g. a male answering a question supposed to be answered by females only.

4.4 Imputation

No imputation for unit non-response was carried out on the IHS. However, imputation was carried out to address item non-response. For categorical variables, un-weighted sequential hot deck imputation for global variables was used. This method involves dividing the data into different strata. Data is then selected from observations, and substituted into other observations where that datum is missing.

Regression imputation was used for numerical variables. This technique involves setting the missing value as a dependent variable in a regression equation, and then using correlated variables as the corresponding independent variables. A tendency for imputed values to be located directly on the regression line is avoided by allowing an error rate to be added or subtracted from the imputed value. Zero values were also imputed for missing numerical variables where appropriate. There were low levels of imputation carried out.

4.5 Grossing and Weighting

Once a final dataset has been created, the data is weighted to population totals. To derive grossing factors, population estimates are produced by the CSO's Demography Unit each quarter by sex, age (5-year age groups) and region (8 NUTS3 regions). The individual returns are matched to these population estimates and the grossing factor is calculated by dividing the total estimates of population in a given cell by the number of valid responses in that cell – i.e. if there are 1,000 respondents from the sample in a given cell and an estimated population of 40,000 then each respondent in this cell will

have a grossing factor of 40. The grossing factors calculated for each individual cell are a function of the number of valid responses in that cell. The overall number of valid responses was as follows:

Year	Total valid responses - persons aged 15+	Average grossing factor assigned - persons aged 15+
2015	10,323	350.52

4.6 Computation of Outputs, Estimation Methods Used

Output results are aggregated to produce various totals published. These aggregates are produced using key variables such as sex, age group, region, etc. All aggregations are carried out by way of various SAS procedures (e.g. PROC FREQ, PROC TABULATE etc.). The aggregate results produced for any given set of classifications will be the sum of the individual grossing factors of the valid responses which belong to that set of classifications and no estimates are made unless the data itself has been captured within the survey.

4.7 Other Quality Assurance Techniques Used

The quality of the data was regularly checked as it was received in head office. Interviewers receive regular feedback on the quality of the data. Interviewers undergo training when they are first employed and regular meetings are held with field-coordinators to ensure standards are maintained. All interviewers received training on the IHS.

5 Quality

5.1 Relevance

The data is required to be collected under Regulation (EC) No 1338/2008 of the European Parliament and of the Council of 16 December 2008 on Community statistics on public health and health and safety at work.

The data is used by various users. These include Government departments, academics, Eurostat, internal users of data within the CSO, and the general public. The data is need to inform policy decision making, satisfy European regulation, and inform the general public interest.

5.2 Accuracy and Reliability

5.2.1 Sampling Effect & Representativeness

As the IHS is a sample survey, it is subject to sampling error. The following table shows estimated standard errors and confidence intervals for some of the key estimates in the IHS. The 95% confidence intervals indicate the range within which we can be 95% confident the true value of the estimate in question will lie based on measurable sampling error.

2015 Irish Health Survey; Standard Error, 95% Confidence Interval and Design Effect *deff*					
Respondents aged 15 years or over in good or very good health (HS1)	**Number of respondents - n (unweighted)**	**Estimated proportion - *p* *(weighted)***	**Standard error - *SE* *(with respect of sampling plan)***	**95% confidence interval | lower; upper |**	**Design effect deff**
All	7979	81.41	0.7167	79.909935 | lower; upper | 82.9321064	1.625
Women	4462	81.35	1.1028	79.868157 | lower; upper | 82.8613211	1.5127
Men	3517	81.48	0.9246	79.8403346 | lower; upper | 82.892846	1.6985
Respondents aged 15 years or over with a longstanding illness or health problem (HS2)					
All	3791	30.73	0.5432	29.1865357 | lower; upper | 32.3010058	1.4048
Women	2078	32.11	0.7943	30.596891 | lower; upper | 33.6610948	1.4003
Men	1713	29.29	0.7383	27.7265348 | lower; upper | 30.8934817	1.3817
Respondents aged 15 years or over that were severely limited in activities people usually do because of health problems for at least the past 6 months (HS3)					
All	448	3.68	0.5736	1.3288641 | lower; upper | 60.94598	1.4971
Women	248	4.03	0.8661	1.6894661 | lower; upper | 6.4291404	1.4164
Men	200	3.31	0.7557	0.969405 | lower; upper | 5.73888	1.5513
Respondents aged 15 years or over declaring having been hospitalized in the past 12 months (HO1) (men and women)	1177	10.88	0.3880	10.0396001 | lower; upper | 11.7737516	1.5891
Respondents aged 18 years or over who are obese (BMI>=30, men and women)	2570	28.09	0.5633	26.8311944 | lower; upper | 29.3659482	1.6595

5.2.2 Non-Sampling Effects

In addition to known sampling errors, any survey will be subject to other non-sampling errors (for example, measurement errors arising from questions not capturing the desired information accurately). Non-sampling error is far more difficult to measure than sampling error and no formal estimate of non-sampling error is available in the IHS.

5.2.2.1 Quality of the Data Sources Used (other than survey registers)

Not applicable

5.2.2.2 Register Coverage

As the IHS was carried out as a module of the QNHS, the entire stock of private households in the country represents the full sampling frame for the IHS. The sample based on the 2011 Census was first introduced on a wave by wave basis in Q4 2012 and was fully in effect as of Q4 2013.
To avoid response burden, old samples are examined so as to avoid an overlap, i.e. blocks in the old sample were deliberately excluded from the new sample to avoid response burden for included households.

5.2.2.3 Non-response (Unit and Item)

Currently there is no non-response adjustment used in the QNHS. The table below gives a breakdown of the response and non-response for the IHS, based upon figures from the QNHS.

	Survey Total
Target individuals	22,013
of which:	
Uncontactable	351
Refusal	11,076
Actual responses received	
of which:	
Insufficient data provide for analysis	263
Final total number of individuals used to generate outputs	10,323

5.2.2.4 Measurement Errors

No formal evaluation of sources of error is available, although measures are in place to minimise error.

- Members of the field staff were fully trained on the questionnaire before implementation.
- Respondent effects – most of the requested information was readily available to respondents. Proxy responses are not allowed for all questions on the IHS.
- Comprehension errors – An effort is made to ensure that the terms used in the survey are clear and readily understood.

5.2.2.5 Processing Errors

a. Data capture errors: These errors are minimised by logic checks and limits on values implemented by the relevant statistician once the data has been inputted at head office.

5.2.2.6 Model-based Effects

For categorical variables, un-weighted sequential hot deck imputation for global variables was used. This method involves dividing the data into different strata. Data is then selected from observations, and substituted into other observations where that datum is missing.

Regression imputation was used for numerical variables. This technique involves setting the missing value as a dependent variable in a regression equation, and then using correlated variables as the corresponding independent variables. A tendency for imputed values to be located directly on the regression line is avoided by allowing an error rate to be added or subtracted from the imputed value. Zero values were also imputed for missing numerical variables where appropriate. There were low levels of imputation carried out. The following table shows the level of missing data for each variable in the IHS.

Variable Name	Percentage of data missing	Variable Name	Percentage of data missing
HS1	2.16%	HO1	37.78%
HS2	5.25%	HO3	36.45%
HS3	7.79%	AM1	4.23%
CD1A	10.47%	AM2	2.86%
CD1B	11.26%	AM3	18.84%
CD1C	11.82%	AM4	8.33%
CD1D	11.63%	AM5	9.83%
CD1E	8.25%	AM6A	5.23%
CD1F	12.07%	AM6B	7.93%
CD1G	12.71%	AM7	5.54%
CD1H	9.18%	MD1	2.54%
CD1I	10.96%	MD2	5.43%
CD1J	11.27%	PA1	10.15%
CD1K	11.31%	PA2	2.22%
CD1L	12.31%	PA3	3.78%
CD1M	10.76%	PA4	7.58%
CD1N	11.68%	PA5	14.43%
CD1O	10.39%	PA6	10.05%
AC1A	7.17%	PA7	3.45%
AC1B	7.20%	PA8	4.50%
AC1C	7.95%	UN1A	8.33%
AC2	3.87%	UN1B	14.10%
AW1	12.18%	UN2A	8.03%
PL1	2.86%	UN2B	8.01%
PL2	4.47%	UN2C	8.04%
PL3	5.32%	UN2D	12.47%
PL4	8.32%	BM1	10.35%
PL5	8.31%	BM2	5.08%
PL6	6.69%	PE1	3.01%
PL7	6.33%	PE2	5.55%
PC1A	1.88%	PE3	1.57%
PC1B	1.71%	PE4	6.63%
PC1C	1.82%	PE5	0.87%
PC1D	1.95%	PE6	13.10%
PC1E	1.74%	PE7	2.18%
PC2	0.04%	PE8	0.63%
PC3	0.03%	FV1	1.36%
HA1A	1.77%	FV3	1.76%
HA1B	1.88%	SK1	3.05%
HA1C	1.71%	SK2	0.25%
HA1D	1.90%	SK4	16.56%
HA1E	2.03%	AL1	3.42%
HA1F	1.90%	AL2	0.67%
HA1G	1.87%	AL3	0.60%
HA2	0.01%	AL4	0.77%
HA3	0.03%	AL5	0.53%
PN1	3.32%	AL6	3.23%
PN2	1.40%	SS1	1.96%
MH1A	6.97%	SS2	2.76%
MH1B	8.06%	SS3	1.86%
MH1C	6.41%	IC1	5.85%
MH1D	6.22%	IC2	0.18%
MH1E	7.63%	IC3	0.75%
MH1F	7.97%		
MH1G	7.46%		
MH1H	8.55%		

5.3 Timeliness and Punctuality

5.3.1 Provisional Results

No provisional outputs are published for this survey

5.3.2 Final Results

There was release date fixed for this survey. The data collection period ended on the 15th of April 2016. The final results were released by way of an electronic publication on the CSO website on the 16th of November 2016.

The microdata were to be submitted to Eurostat in accordance with Commissions Regulation (EU) No 141/2013 (Annex 1) no later than 9 months after the end of the collection period. This data was submitted on the 15th of January 2017.

Variables which were not published, but were required by Eurostat, include for example the following;

- HATLEVEL – Highest level of education achieved in accordance with ISCED classification
- FT_PT – Whether the respondent was involved in full-time or part-time employment
- HHNBPERS – The number of persons in the household
- DEG_URB – The degree of urbanisation
- MASRTALEGAL – The legal marital status of the respondent

Variable which were included on the national publication which were not provided to Eurostat included;

- DI_5 – The level of advantage or disadvantage of the household of the respondent.
- Dis – This is a derived variable providing the disability status of the respondent.

Unpublished variables are used in the calculation of indicators such as the Oslo Social Support scale. Such variables are provided on an ad-hoc request basis.

5.4 Coherence

Coherence checks can be carried out with auxiliary sources. An example of this is the 'Minimum European Health Module', which is comprised of the first three questions on the IHS. These are self-perceived health, long-standing conditions, and general activity limitation. These questions are also asked in EU-SILC, which ensures cross-domain comparability. They are also asked in the Healthy Ireland questionnaire, which is a Department of Health survey carried out annually.

5.5 Comparability

The IHS is comparable with other European versions of the survey. This is due to the following;

- A common implementing regulation, ensuring that the same concepts are measured across countries
- A sample questionnaire provided by Eurostat

- Standard classification including NACE regional classifications, and ILO employment classifications.

5.6 Accessibility and Clarity

5.6.1 Assistance to Users, Special Analyses

The IHS publication is available on the CSO website. Information on the methodology used is also available on the website, and the background notes provide some detail on the survey.
A press statement was issued with the release of the publication.
Ad-hoc analysis can be produced upon request.
Anonymised microdata is made available to researchers cia the Irish Social Science Data Archive (ISSDA). Such data is accessible to researchers by applying directly to the ISSDA.

For further information see:

https://www.ucd.ie/issda/

Access to Research Microdata Files (RMFs) can be requested from the CSO under the CSO's microdata access policy. Extensive use of this facility is made by the research community.

For further information see:

http://www.cso.ie/en/aboutus/dissemination/accesstomicrodatarulespoliciesandprocedures/researchaccesstomicrodatafiles/

5.6.2 Revisions

Not applicable

5.6.3 Publications

5.6.3.1 Releases, Regular Publications

The release of the IHS can be found here:

http://www.cso.ie/en/releasesandpublications/ep/p-ihs/irishhealthsurvey2015/

5.6.3.2 Statistical Reports

IHS contributes to the Men and Women in Ireland statistical release on a national level.

5.6.3.3 Internet

The IHS is available on the CSO website in publication format. In addition, data is made available via the CSO's main databank dissemination tool (Statbank) and is also hosted on the CSO website in Excel format:

Releases and Publications:

http://www.cso.ie/en/releasesandpublications/ep/p-ihs/irishhealthsurvey2015/

Databank dissemination

http://www.cso.ie/px/pxeirestat/Database/eirestat/Irish%20Health%20Survey/Irish%20Health%20Survey_statbank.asp?SP=Irish%20Health%20Survey&Planguage=0

5.6.4 Confidentiality

The confidentiality of all information provided to the CSO by individual respondents is guaranteed by law under the 1993 Statistics Act. All CSO office and field personnel become "Officers of Statistics" on appointment and are liable to penalties under this Act if they divulge confidential information to any outside person or body. Extreme precautions are taken to ensure that there are no violations of this principle throughout the survey process. The laptops/tablets on which the data was collected are encrypted and contain several layers of password protection. Data are only published in aggregate form and care is taken to ensure that the data are aggregated to avoid the indirect identification of respondents. Confidentiality is also ensured within the anonymised microdata by using coded variables instead of original values for key characteristics. For example, age groupings are provided instead of single year of age.

6 Additional documentation and publications

The main CSO Health homepage can be found at the following link:

http://www.cso.ie/en/statistics/health/

Statcentral – Ireland's portal to official statistics:

http://statcentral.ie/

Eurostat issue many releases which use IHS data and the central repository for such data can be found at the following link:

http://ec.europa.eu/eurostat/publications/all-publications