

Article

# Developing a method for measuring time spent on green tasks: March 2022

A methods article on estimating time spent on green tasks, as part of research into measuring “green jobs”. Describes the use of task-level data from the US Occupational Information Network (O\*NET) database, occupation conversion from US to UK occupation codes, and application to UK labour market data.

Contact:  
Josh Martin and Ellys Monahan  
environment.accounts@ons.gov.  
uk  
+44 1633 455783

Release date:  
7 March 2022

Next release:  
To be confirmed

## Table of contents

1. [Overview of developing a method for measuring time spent on green tasks](#)
2. [Estimating time spent on green tasks in the Occupational Information Network \(O\\*NET\) database](#)
3. [Conversion between occupation classifications](#)
4. [Application to UK labour market data](#)
5. [Data used to develop the method](#)
6. [Related links](#)

# 1 . Overview of developing a method for measuring time spent on green tasks

This methodology contributes to research on defining and measuring “green jobs”, as set out in the [green jobs workplan](#). There are many ways to define a green job, and no internationally agreed approach. The estimates, produced using this new experimental methodology, are based on task-level information from the US. As highlighted in the workplan, the task- and occupation-based approach taken by this research provides one potential approach to measuring green jobs and we plan to develop this further.

While the task- and occupation-based approach in this research offers some advantages over other approaches, it also has some drawbacks. The purpose of this research is to explore these pros and cons, rather than to provide official estimates of “green jobs”.

Estimates based on this research are published in [Research into “green jobs”: time spent doing green tasks, UK: 1997 to 2019](#) with the accompanying datasets providing additional results.

## 2 . Estimating time spent on green tasks in the Occupational Information Network (O\*NET) database

### About the O\*NET database

The Occupational Information Network ([O\\*NET database](#)) is the result of a large data collection in the US, run by the National Center for O\*NET Development with support from the US Department of Labor's Employment and Training Administration. Among other things, it collects very detailed task-level information for almost 1000 occupations. The information includes attitudes, preferences, skills, work activities, and more.

The task-level data are updated for around 100 occupations each year. Data are collected either by interviewing people in the US labour market who fall within the relevant occupation category (incumbents), or by consultation with occupational experts, or both. More dynamic occupations tend to be reviewed more frequently than average.

An updated dataset is released each year, and archive versions of the dataset with historic data are [available to download](#) from O\*NET. The program started in 1998 with information collected from occupational experts only. Databases consistent with the current methods began in 2003, and [O\\*NET encourages researchers to use that as a starting point](#). Minor revisions, between the large annual updates, are also included in the archive, but these do not contain reviewed occupation data. The latest update to the data at the time of this research was for 2021.

### Definition of green tasks in the O\*NET database

There are many ways to define a "green job" or the "green economy". This research is built on the O\*NET database's definition of the green economy introduced in [Greening of the world of work: Implications for O\\*NET® - SOC and new and emerging occupations](#), which is "The economic activity related to reducing the use of fossil fuels, decreasing pollution and greenhouse gas emissions, increasing the efficiency of energy usage, recycling materials, and developing and adopting renewable sources of energy".

## Identifying green tasks in the data

In a [2011 study](#), the O\*NET database was augmented to include new occupations and green tasks. Existing occupations likely to be related to the "green economy" (as defined in the O\*NET database) were reviewed and task lists updated or refreshed to better reflect green tasks. Some new "green occupations" were also separately included.

Markers on green tasks were included in each release of the dataset from 2011 until 2019, at which point they were removed. With each annual update to the dataset, newly updated occupations also had "green task" markers added as appropriate.

For releases of the database after 2019 and before 2011, the "green task" markers are not included in the data. For this research, we assigned "green task markers" using a combination of automated and manual processes.

For years before 2011, we needed to assume that only occupations that have green tasks from 2011 onwards can have green tasks before 2011. This is consistent with an increasing size and prevalence of the green economy over time. We matched the task list from 2011 with each earlier release, and if a green task was present in a "green occupation" we assigned it a green task marker. We also used a list of keywords associated with green tasks, identified using the 2011 data. Any tasks containing one or more of those green keywords done by a "green occupation" was flagged for review, and assigned a green task marker, if appropriate.

For years after 2019, we matched the green task list forward from 2019 into the later releases of data, and if a green task is present we assign it a green task marker. We also used the same set of keywords and approach to identify additional new green tasks. We did not require that the occupation was previously "green", to allow for the emergence of newly green occupations after 2019.

## Estimating time spent on green tasks

We combined two pieces of data on each task to estimate the time spent on that task, these were:

- [Relevance](#) of the task -- the proportion of job incumbents who rated the task as relevant to their job (multiplied by 100)
- [Frequency](#) with which the task is performed) - the distribution of respondents across seven frequency categories, ranging from "Yearly or less" to "Hourly or more"

The distribution of responses to the frequency question came only from respondents who said the task was relevant. As such, there was implicitly an eighth frequency category, which was "Never" This relates to those respondents who said the task was not relevant. As such, we imputed this eighth frequency category as 100 minus the relevance score, and rescaled the original frequency distribution accordingly, such that the adjusted frequency distribution again summed to 100. An example is given in Table 1.

Table 1: Example of original distribution of responses, and the rearrangement of the frequency distribution for a task

<b>Measure</b>	<b>Proportion of respondents (%)</b>	<b>New frequency category</b>	<b>Revised proportion of respondents (%)</b>
Relevance	80	"Never"	20
Frequency: "Yearly or less"	5	"Yearly or less"	4
Frequency: "More than yearly"	10	"More than yearly"	8
Frequency: "More than monthly"	10	"More than monthly"	8
Frequency: "More than weekly"	20	"More than weekly"	16
Frequency: "Daily"	30	"Daily"	24
Frequency: "Several times daily"	15	"Several times daily"	12
Frequency: "Hourly or more"	10	"Hourly or more"	8

Source: Office for National Statistics

We then assigned estimated time weights to each frequency category. That is, for each task done at a given frequency, what fraction of a worker's total time would we expect to be spent on that task? We reviewed the tasks most often reported for each frequency category, and used regression analysis across the occupations and tasks in the data to estimate an appropriate set of parameters. The ones that we used in this research are shown in Table 2.

Table 2: Time weights used in this research

**Frequency category Time weight**

“Never”	0
“Yearly or less”	0.02
“More than yearly”	0.03
“More than monthly”	0.06
“More than weekly”	0.08
“Daily”	0.1
“Several times daily”	0.13
“Hourly or more”	0.16

Source: Office for National Statistics

The time weights are more evenly distributed than the frequency descriptions might suggest. This is because we believe the intensity with which a task is done varies across the frequency categories. For instance, a task that is done “Hourly or more” might only be done briefly each hour (for instance, for five minutes each hour, equating to 8.3% of total time). Whereas a task that is done “Yearly or less” might be infrequent but lengthy, such as compiling an annual report, which may take several days. The factors are fairly arbitrary, and we welcome further research on this method.

The frequency distribution (Table 1) and the time weights (Table 2) were multiplied to give a composite score for the time spent on each task. An example is given in Table 3.

Table 3: Example of the application of the time weights to the revised frequency distribution to produce a composite time score for a task

<b>Frequency category</b>	<b>Revised proportion of respondents (%)</b>	<b>Time weight</b>	<b>Composite score for time spent</b>
“Never”	20	0	0
“Yearly or less”	4	0.02	0.08
“More than yearly”	8	0.03	0.24
“More than monthly”	8	0.06	0.48
“More than weekly”	16	0.08	1.28
“Daily”	24	0.1	2.40
“Several times daily”	12	0.13	1.56
“Hourly or more”	8	0.16	1.28
Total	100		7.32

Source: Office National Statistics

These scores are then rescaled across tasks to represent shares of total time. The green time share for the occupation is calculated by summing the time shares of green tasks. An example is given in Table 4.

Table 4: Example of the calculation of a time share for each task, and the green time share, for an occupation

Task	Green	Composite	Time share (%)
		score for time spent	
A	Y	7.32	44.4
B	N	8.13	49.3
C	Y	1.05	6.4
Total		16.5	100
of which: green			50.7

Source: Office for National Statistics

## Interpolation and extrapolation

The previous steps produce an estimate of the green time share for each of the nearly 1,000 occupations in the O\*NET database, for the years in which the data for that occupation is collected or updated. The task-level data are updated for around 100 occupations each year (around 10% of the total number), so occupations will have only a few observations over the course of the time period available (2003 to 2021). The majority have two or three data points over the period, but some have only one, and some have more.

To avoid discontinuities in the time series, and to better reflect the gradual change of tasks over time, we use linear interpolation (drawing a straight line) between data points. This gives a smoother transition between observations. For occupations with only one data point across the time series, we hold the green time share constant over time.

We also use linear extrapolation (projecting a straight line) for periods after the latest observation, and before the earliest observation. These extrapolations are bounded between 0% and 100%.

This gives a full time series, between 1997 and 2021, of the estimated percentage of time spent doing green tasks (green time share) for each of the nearly 1,000 O\*NET occupations. The next step, is to convert these data onto the UK occupation classification.

## 3 . Conversion between occupation classifications

To apply the green task time shares calculated from the Occupational Information Network (O\*NET) database to UK labour market data, we need to convert from the occupation classification used in O\*NET to the [UK Standard Occupation Classification \(SOC\)](#). Since there is no direct correspondence table between these classifications, we instead use a number of conversions in sequence.

### Conversion between releases of the O\*NET occupation classification

Since we are using many releases of the O\*NET database, we first have to convert them all to a common occupation classification. The occupation classification used in O\*NET changed in 2006, 2009, 2010, and 2020. We use [correspondence tables](#) made available by O\*NET to convert between these classifications.

### O\*NET SOC 2019 to US SOC 2018

The occupation classification used in O\*NET since 2020 is O\*NET SOC 2019, which is based on the US SOC 2018. O\*NET SOC 2019 is an extension of US SOC 2018, with additional detail beneath most US SOC 2018 codes. As such, to convert from O\*NET SOC 2019 to US SOC 2018, we remove the two digits at the end of the O\*NET SOC 2019 code, which abbreviates it to its US SOC 2018 code parent. Equivalently, a [crosswalk](#) is made available by O\*NET.

## US SOC 2018 to ISCO-08

There is again no direct mapping from the US SOC 2018 to the UK SOC 2010 (or any other release of the UK SOC). As such, we convert via the International Standard Classification of Occupations (ISCO) 2008 (known as ISCO-08). We do this using a [publicly available crosswalk](#), which is based on a [crosswalk from the US SOC 2010 to ISCO-08](#) available from the US Bureau of Labor Statistics (BLS).

Since these crosswalks are mostly not one-for-one, the green time shares calculated for O\*NET SOC codes are converted to several US SOC and ISCO codes. This means that, at each point of the conversion, we have several green time share estimates for each occupation, based on different converted codes. Unlike other studies, we chose not to average (collapse) these at each stage. We believe this avoided the structure of the US SOC or ISCO playing too large a role, since our ultimate aim is to convert to the UK SOC. Different choices would lead to slightly different results.

## ISCO-08 to UK SOC

Unlike on the other steps of the conversion, there is no ready-made crosswalk between ISCO-08 and UK SOC. The Office for National Statistics (ONS) published [an exploration of the link between UK SOC 2010 and ISCO-08](#), but this was not a complete crosswalk since many UK SOC codes are unmatched, and it is in the wrong direction for this exercise. That is, it considered a conversion from UK SOC 2010 to ISCO-08, which is the opposite direction to our conversion. Where the relationships between the classifications are not all one-for-one (which is almost always true), conversions in different directions require different crosswalks.

In the absence of a ready-made crosswalk, we developed one based on the latest (at the time of the research) [coding index for UK SOC 2020](#). The coding index is a long list of job titles and their corresponding occupation codes, which can be used by data collectors to assign occupation codes to data collected by job title. For instance, the Annual Survey of Hours and Earnings (ASHE) asks respondents for a job title and job description, rather than an occupation code, since most respondents would not be able to report an occupation code easily; these job titles are then mapped to occupation codes using the coding index.

The UK SOC 2020 coding index contains around 30,000 job titles and the corresponding codes from SOC 1990, SOC 2000, SOC 2010, SOC 2020 and ISCO-08. This can be used to produce a crosswalk between ISCO-08 and UK SOC classifications, in either direction. One option is a modal crosswalk, where each ISCO-08 code is associated with a single SOC code, based on the most common pairing of codes. An alternative is a proportional crosswalk, where each ISCO-08 code is associated with one or more SOC code, based on the proportion of job titles with each pairing. An example is given in Table 5.

Table 5: Example of conversion methods using coding index

Job title	ISCO-08 code	SOC 2010 code	Modal conversion	Proportional conversion	Truncated proportional conversion
Electrician, marine	7412	8232	0	0.33	0.4
Electrician, chief (shipping industry)	7412	8232			
Electrician, auto	7412	5231	1	0.5	0.6
Electrician, automobile	7412	5231			
Electrician, vehicle	7412	5231			
Electrician, aircraft	7412	5234	0	0.17	0

Source: Office for National Statistics

Both modal and proportional conversions have pros and cons. The main drawback of the modal conversion is that it would not reflect that an ISCO-08 code could have more than one associated SOC 2010 code. This can often lead to situations where a SOC 2010 code has no associated ISCO-08 code, and thus no data assigned to it. The drawback of the proportional conversion is that it potentially spreads the data around too much, influenced by niche or rare matches as well as more common ones.

In both cases, a limitation of using the coding index is that we do not know which job titles are most common. We are implicitly assuming that each job title is equally as common in UK labour market, but if one were much more or less common than others, then we would attribute it too little or too much weight. For instance, if "electrician, aircraft" in Table 5 was the most common of all the job titles associated with ISCO-08 code 7412, then it should be the modal conversion and have a much higher share of the proportional conversion. Without that information, treating it as equally common as all the other job titles, it is given a lower weight.

We chose to use a truncated proportional conversion for this work, an example of which is also shown in Table 5. This is somewhere between a modal and proportional conversion. We exclude:



- unique pairs of ISCO-08 and SOC 2010 codes (that is, where that pairing appears only once in the coding index)
- pairs that have a weight in the proportional conversion of less than 2%
- pairs that appear only twice and have a weight in the proportional conversion of less than 5%

These rules help to remove unusual or niche pairings, and help to avoid allocating green time shares to spurious SOC 2010 codes. This results in all SOC 2010 codes having associated ISCO-08 codes, but the conversion is still reasonably parsimonious. We established these rules based on testing and judgement. Other rules could be used, which would produce different results.

After applying all of these conversion steps, there are many data points for each UK SOC 2010 code, based on the one-to-many matches at each step of the conversion process. We then take the arithmetic average (mean) of the data points, to collapse it down to a single estimate for each UK SOC 2010 code per year. We do the same for SOC 2000 codes as well, since the UK labour market data is coded to SOC 2000 in earlier years and SOC 2010 in later years.

## 4 . Application to UK labour market data

We apply the converted green time shares to UK labour market data. The green time shares will partly reflect regulation, economic conditions and other factors in the US, as a result of using the US Occupational Information Network (O\*NET) data to estimate them. However, the application to UK labour market data means that the estimates will also reflect UK employment dynamics and labour market trends (for example, increasing prevalence of jobs with higher green time shares).

We use two main sources in the analysis: the Annual Population Survey (APS), and the Annual Survey of Hours and Earnings (ASHE).

### Annual Population Survey

The Annual Population Survey (APS) is the UK's largest continuous household survey. It is not a standalone survey but uses data combined from Wave 1 and Wave 5 (the first and last wave) of the main [Labour Force Survey](#) (LFS) plus a boost from the Local Level Labour Force Survey for England, Wales and Scotland. Industries and occupations are coded by interviewers based on information given by respondents. It covers employees and the self-employed, but the quality of the industry allocation is generally considered to be less reliable than business surveys. As such, we use APS data for our UK and country breakdowns, but not our industry breakdowns. See the [APS QMI](#) for more details.

We use APS microdata to estimate the number of jobs doing green tasks, and hours worked on green tasks, for the UK and country breakdown, for 2004 to 2019. Use of the microdata avoids issues with rounding or suppression from using publication datasets. We use the occupations and work location of workers in their main job (a worker can have more than one job). The measure of hours is "total actual hours worked in main job in reference week", a derived variable based on respondent answers.

The microdata is coded to Standard Occupation Classification (SOC) 2010 from 2011 onwards, and to SOC 2000 between 2004 and 2011. Estimates based on SOC 2010 and SOC 2000 are not entirely consistent. To avoid a discontinuity we use the level of the SOC 2010 based estimates and extend backwards based on the growth rate of the SOC 2000 based estimates known as splicing.

Data from the APS on the number of workers by occupation on a SOC 2010 basis are available to download from [Nomis](#) for calendar years back to 2006. We use this to quality assure the splicing, which confirms it is robust. We use LFS data from the April to June period, by occupation on a SOC 2000 basis, available in the [Employment by occupation dataset](#), to extend the APS series for the UK back to 2001.

## Annual Survey of Hours and Earnings

The Annual Survey of Hours and Earnings (ASHE) is based on a 1% sample of employee jobs taken from HM Revenue and Customs (HMRC) Pay As You Earn (PAYE) records. Employers supply information on earnings and hours. This information is treated confidentially. Occupations are coded by the Office for National Statistics (ONS) based on written responses from businesses. Industries are allocated based on information held on the Inter-Departmental Business Register (IDBR). ASHE covers employees only. See the [ASHE QMI](#) for more details.

Since the industry information in ASHE comes from the IDBR, it is considered higher quality than that reported by workers on the LFS. As such, we use the ASHE-based estimates for the industry breakdowns. We also show the whole-economy estimates from ASHE to demonstrate the consistency with APS-based estimates. However, since ASHE and APS have a different coverage of the economy, they are not directly comparable, and the APS-based estimates are preferred.

We use ASHE microdata to estimate the number of jobs doing green tasks, and hours worked on green tasks, for the UK and by industry, for 1997 to 2019. Estimates are on a SOC 2000 basis between 1997 and 2011, and on a SOC 2010 basis between 2011 and 2019, with 2011 used as the joining year, as described for the APS. The measure of hours is "total paid hours" (the sum of "basic paid hours" and "paid overtime hours") as reported by the respondent employers.

Industries in the ASHE data since 2008 are coded using the Standard Industrial Classification (SIC) 2007, while those prior are coded using SIC 2003 and SIC 1992. To produce consistent industry estimates over the full time series, we convert the SIC 1992 and SIC 2003 codes to SIC 2007 using a [standard proportional mapping](#) based on employment data on the IDBR.

## 5 . Data used to develop the method

### [Truncated proportional conversion between ISCO-08 and UK SOC classifications](#)

Dataset | Released 07 March 2022

Conversion between International Standard Classification of Occupations (ISCO-08) and the UK Standard Occupation Classification (SOC) 2000 and 2010. Developed as part of research into "green jobs" using an occupation and task-based approach. Used to convert US O\*NET data to UK SOC codes.

### [Estimated time spent on green tasks by occupation code](#)

Dataset | Released 07 March 2022

Estimates of the fraction of time spent doing green tasks, by occupation code using the Standard Occupation Classification (SOC) 2000 and SOC 2010. Part of research into "green jobs", using an occupation and task-based approach.

### [Time spent on green tasks](#)

Dataset | Released 07 March 2022

Experimental estimates of the time spent doing green tasks, over time, by UK country and by industry. It uses a new method, based on task-level data from the O\*NET database in the US.

## 6 . Related links

[Research into "green jobs": time spent doing green tasks, UK: 1997 to 2019](#)

Article | Released 7 March 2022

This article provides new experimental estimates of the time spent doing green tasks, over time, by UK country and by industry. It uses a new method, based on task-level data from the O\*NET database in the US

["Green jobs", current and upcoming work: March 2022](#)

Article | Released 7 March 2022

How the ONS has contributed to understanding "green jobs" through regular estimates and research articles, and what our future work on "green jobs" will include.