

Guidance on Calculating Compliance Costs

The UK Code of Practice for Official Statistics requires all producers of National and Official Statistics to:

Report annually the estimated costs (for example, on businesses, service providers, or the public) of responding to statistical surveys ..."

Principle 6, Practice 1

The following document provides guidance on how to calculate the costs specified in the above practice. The guidance is tailored to two areas:

- businesses and local authorities; and
- households and individuals.

General

The methods outlined below should be used for both voluntary and statutory surveys.

A calculation of compliance is required for all questionnaire or interview types. If different questionnaires or interviews are used, then the calculations should be conducted for each type and then summed for an overall compliance cost. Similarly, if there are different groups of respondents (e.g. large and small businesses) then the calculations should be made separately and summed.

Achieved response should be used to measure actual burden of completed surveys.

Anticipated response should be used to report burden in the future (eg. new surveys, surveys where data collection is on-going and estimating regular surveys for future years).

Some sections below refer to respondent questionnaires or reviews. The UK Code of Practice for Official Statistics requires producers of National and Official Statistics to "*seek to achieve continuous improvement in statistical processes by, for example, undertaking regular reviews ...*". Best practice suggests that for monthly or quarterly surveys a review should be conducted every three years (triennial reviews) and for annual surveys (or surveys run less often than every year, eg. once every two or three years), a review should be conducted every five years (quinquennial reviews).

To establish any change in respondent burden following major changes to the questionnaire or interview, the impact should be evaluated as part of the pilot, or through the cognitive interview process. This should be followed by a full review within two years of these changes taking place.

Surveys to Businesses and Local Authorities

Inclusions

A financial cost is calculated based on the time taken to complete the questionnaire / interview, an appropriate hourly rate and external costs incurred (eg. costs to the business of a bookkeeper or accountant to aid in completion of the survey).

Subsequent contact to validate responses is also included since it contributes to actual burden on respondents. Estimates for this will be made internally and should be department specific and evidence-based.

Exclusions

The following are excluded from the calculations of compliance cost:

- set-up costs (costs involved in setting up any systems);
- overheads.

Non-response is assumed to place no burden on respondents, thus it is not included within the calculations.

Compliance Cost Model:

burden =

$$\left[\{n_{\text{resp,main_surv}} \times \text{med}(t_{\text{main_surv}})\} + \{n_{\text{val,main_surv}} \times \text{med}(t_{\text{val}})\} \right] \times \text{hourly_rate} \\ + \\ [prop_{\text{external costs}} \times n_{\text{resp,main_surv}}] \times \text{med}(\text{external cost})$$

where:

$n_{\text{resp,main_surv}}$ is the number of responses to the survey or interview including full and partial responses, even if some are invalid¹.
Achieved response should be used to measure actual burden of completed surveys.
Anticipated response should be used to report burden in the future (eg. new surveys, surveys where data collection is on-going and estimating regular surveys for future years).

$\text{med}(t_{\text{main_surv}})$ is the median time taken to complete the questionnaire or interview.

This information can be obtained from:

- respondent questionnaires (ie reviews)²;
- pilots;
- historical data³;
- paradata⁴; or
- survey manager expertise.

$n_{\text{val,main_surv}}$ is the number of respondents from the main survey re-contacted for the purpose of validating their responses.

$\text{med}(t_{\text{val}})$ is the median time taken in any additional re-contact of respondents for validation purposes. If this is not known, $\text{med}(t_{\text{main_surv}})$ may be substituted on the assumption that a respondent may need to review the entire questionnaire.

hourly_rate is the (estimated) hourly rate as decided by each department, based on the Annual Survey of Hours and Earnings (ASHE) and updated annually. The decision on the hourly rate should be justifiable⁵ and should be based on:

- evidence from previous respondent questionnaires (ie reviews); or
- survey manager expertise.

prop_{external costs} is the proportion of respondents who incur external costs. This is estimated from the review questionnaires as:

$$n_{\text{resp with external costs, review sample}} / n_{\text{resp, review sample}}$$

med(external cost) is the median cost incurred (eg. accountant's fees) by those businesses and local authorities that say they do incur external costs. This information can be obtained from:

- respondent questionnaires (ie reviews);
- pilots.

An example calculation is included in Annex A.

Surveys to Individuals and Households

For surveys of individuals and households, respondent burden is calculated on the basis of time taken. The financial element (seen above) is excluded from these calculations.

There may be scenarios where it is sensible to include focus groups and cognitive interviews but generally such activities should be excluded from these calculations.

Compliance Cost Model:

$$\text{burden} = n_{\text{resp,main_surv}} \times \text{med}(t_{\text{main_surv}})$$

However, if validation of the responses is required, the following model should be used:

$$\text{burden} = \left[\{n_{\text{resp,main_surv}} \times \text{med}(t_{\text{main_surv}})\} + \{n_{\text{val,main_surv}} \times \text{med}(t_{\text{val}})\} \right]$$

where:

n_{resp,main_surv} is the number of responses to the survey or interview including full and partial responses, even if some are invalid¹.
Achieved response should be used to measure actual burden of completed surveys.
Anticipated response should be used to report burden in the future (eg. new surveys, surveys where data collection is on-going and estimating regular surveys for future years).

$med(t_{\text{main_surv}})$ is the median value of time spent completing the questionnaire, interview, travel diaries, etc. Time taken to establish eligibility should also be included (eg. time spent by the interviewer on the doorstep determining household eligibility).

This should not be an additional piece of information required of respondents as this can be either estimated via:

- paradata ; or
- data collected at a pilot.

If a respondent is deemed out of scope, the time taken to establish eligibility should still be included.

$n_{\text{val,main_surv}}$ is the number of respondents from the main survey re-contacted for the purpose of validating their responses.

$med(t_{\text{val}})$ is the median time taken in any additional re-contact of a respondent for validation purposes. If this is not known, $med(t_{\text{main_surv}})$ may be substituted on the assumption that a respondent may need to review the entire questionnaire.

An example calculation is included in Sub-Annex 2.

¹ Questionnaires returned blank should to be treated as a non-response.

² Questions on time taken to complete the survey may also be included as part of the questionnaire itself.

³ Surveys/studies within the past 5 years.

⁴ Paradata are data about the process by which the survey data were collected. This would usually include information about the length of interview.

⁵ Justifications for hourly rates used will be subject to scrutiny.

Example calculation for surveys of Businesses and Local Authorities

The following example is based on the Annual Business Survey (ABS) conducted by the Office for National Statistics (ONS). The ABS is the ONS financial information survey and samples businesses and other related establishments across the UK.

Compliance cost formula

burden =

$$\left[\{n_{\text{resp,main_surv}} \times \text{med}(t_{\text{main_surv}})\} + \{n_{\text{val,main_surv}} \times \text{med}(t_{\text{val}})\} \right] \times \text{hourly_rate} \\ + \\ [prop_{\text{external costs}} \times n_{\text{resp,main_surv}}] \times \text{med}(\text{external cost})$$

where:

$$n_{\text{resp,main_surv}} = 49,896 \text{ (as this is a completed survey, } \textit{achieved} \text{ response is used)}$$

$$\text{med}(t_{\text{main_surv}}) = 1.5 \text{ (hours)}$$

$$n_{\text{val,main_surv}} = 2,294$$

$$\text{med}(t_{\text{val}}) = 1.5 \text{ (hours) information on time taken for validation not available,} \\ \text{thus used: } \text{med}(t_{\text{main_surv}})$$

$$\text{hourly_rate} = \text{£}19.20$$

$$prop_{\text{external costs}} = n_{\text{resp with external costs, review sample}} / n_{\text{esp, review sample}} = 13 / 167 = 7.8\%$$

$$\text{med}(\text{external cost}) = \text{£}62.50$$

$$\begin{aligned} \text{Burden} &= [(49,896 \times 1.5) + (2,294 \times 1.5)] \times 19.20 \\ &\quad + (0.078 \times 49,896) \times 62.50 \\ &= 1,503,072 + 243,243 \\ &= \text{£}1,746,315.00 \end{aligned}$$

Example calculation for surveys of Individuals and Households

burden =

$$n_{\text{resp,main_surv}} \times \text{med}(t_{\text{main_surv}})$$

The following example shows how this calculation can be made for various stages of the survey process and summed to provide an overall compliance cost. Please refer to the table below the example for the calculation.

Example survey process:

Stage 1:

Households (e.g. 1,000 households) are screened, by phone or on the doorstep, to establish their eligibility to take part in the survey. This screening process may take around 5 minutes.

Stage 2:

Some of the households screened may not be eligible to take part (e.g. 300 may be ineligible, leaving 700 potential households) and others may refuse to take part (e.g. 100 may refuse, leaving 600 households).

Stage 3:

The main respondent in each of the eligible and willing households is then interviewed (e.g. for 90 minutes).

Stage 4:

In addition to the main respondent, some households (e.g. 360 of them) may also include partners, who are available for interview (e.g. for 35 minutes). Partners in other households may not be available and thus proxy interviews may be conducted with the main respondent to collect information on the unavailable partner (e.g. for 15 minutes). Other households may not include partners and thus no additional interviews are conducted.

Calculation:

	$n_{\text{resp,main_surv}}$	$\text{med}(t_{\text{main_surv}})$ (minutes)	Compliance cost (minutes)
Total households screened for eligibility (stage 1)	1,000	5	5,000
Total eligible households (stage 2)	700		
Total responding households	600		
Total main respondent interviews (stage 3)	600	90	54,000
Total partner interviews (stage 4)	360	35	12,600
Total proxy partner interviews	50	15	750
Total			72,350