

Aadne Aasland and Guri Tyldum

A note on sampling procedures for the hotel survey

Title: **A note on sampling procedures for the hotel survey**

Author: Aadne Aasland and Guri Tyldum

NIBR Working Paper: 2011:105

ISSN: 0801-1702
ISBN: 978-82-7071-875-7

Project number: O-2806
Project name:: Industrial relations under global stress

Financial supporter The Research Council of Norway

Head of project: David Jordhus-Lier

Abstract: The paper describes the sample for the survey of hotel employees in Oslo and Akershus. It gives details of challenges met and measures made to obtain a representative sample of the target population. The paper also presents the calculation of sampling weights to adjust for the sample not being self-weighting.

Date: February 2011

Pages: 12

Publisher: Norwegian Institute for Urban and Regional Research
Guastadalléen 21,
N-0349 OSLO

Telephone (+47) 22 95 88 00
Telefax (+47) 22 60 77 74
E-mail: nibr@nibr.no
<http://www.nibr.no>

Org. nr. NO 970205284 MVA

© NIBR 2011

Preface

This is the second working paper presenting methodological background information related to the hotel worker survey conducted as a part of the research project *Industrial relations under global stress: fragmentation and the potential for representation of workers in the Norwegian hospitality sector*, funded by the Research Council of Norway under the VAM programme.

The project explores the politics of work in a part of the Norwegian labour market which is characterised by high levels of labour migrants, low-wage and low-skilled jobs, relatively low unionisation levels and an increasing outsourcing of services. In many ways, the hotel sector provides a contrast to the “Norwegian model” of labour relations. The project will focus on the possibilities of representation of workers in the workplace and in industrial relations. The data collected through this survey is an important first stage in this process, followed by workplace case studies and interviews with key informants in the hotel sector.

Research on the role of trade unions in society is in line with the general focus on the politics of civil society in NIBR’s Department of International Studies, although the team encompasses other parts of NIBR as well as other research institutions. While published as a NIBR Working Paper, this publication is co-authored by researchers from NIBR and Fafo.

Oslo, February 2011.

Marit Haug
Research director

Table of Contents

Preface	1
Tables.....	3
Figures	3
1 Introduction.....	4
2 Sampling.....	5
2.1 First stage sample	5
2.2 Second stage sample.....	5
2.3 Third stage sample	7
2.4 Non-response.....	7
3 Subsequent stages of research.....	8
3.1 The follow-up survey.....	8
3.2 Inclusion probabilities and weights	10
3.3 Implications for analysis	11
References.....	12

Tables

Table 3.1....Distribution of respondents in hotel mapping and in follow-up survey and percentage recapture.....	9
--	---

Figures

Figure 2.1... Estimated distribution of hotel workers in time units during 24 hours....	6
---	---

1 Introduction

This document describes the sample for a survey of hotel sector employees in Oslo and Akershus. Creating a probability sample of hotel sector employees introduces a number of challenges, as no sampling frame (that is a complete list of employees in hotels in Oslo and Akershus) is available. In a sector characterised by frequent use of outsourced and supplement labour, official registers is likely severely to underestimate the true number of employees in a given hotel. Furthermore, these biases are believed to be systematic and to correlate with the size and type of the hotel.

Our sample is a three stage sample; first hotels were selected, after which two time-periods of two hours were selected in each hotel, and finally all employees working within those two hours were selected for interviews. In the following we give a more detailed description of each of the stages involved.

2 Sampling

2.1 First stage sample

It would be quite costly (in terms of transportation) and inefficient to select respondents at all hotels in Oslo and Akershus. Therefore, a selection of hotels had to be made. We needed to ensure that different types of hotels according to their key characteristics: location, size, ownership, outsourcing of labour and degree of unionisation were covered. As one of the key concerns of the survey was to map variation in working conditions, a factor that is likely to vary with size of hotel, we wanted to secure that a sufficient number of small hotels were included in the sample. We therefore decided not to draw the hotels with probability proportionate to size. The sample is therefore not self-weighting (see section on weights below).

We also found that the best approximation of hotel workers in Oslo and Akershus based on available information is number of rooms reported in each hotel. Although not a completely accurate measure of number of hotel employees, due to differences in hotel types (conference hotels, hotels with/without restaurant, hotel quality, etc.) it was the best approximation that we could obtain. We checked the number of rooms against the number of employees as reported to the company register (*Bronnøysundregisteret*), which showed some discrepancy but was considered to be less accurate due to the large number of hotels using outsourced and supplement labour (see also Jordhus-Lier et al. 2010).

Thus, to secure a representation of hotels of different sizes, the sample was implicitly stratified by size (estimated based on number of rooms), through sorting of the hotel list from the largest to the smallest. It was decided to make interviews at 40 of the 110 hotels listed (36.3 per cent). The hotels were drawn systematically with a random start from the presorted hotel list.

2.2 Second stage sample

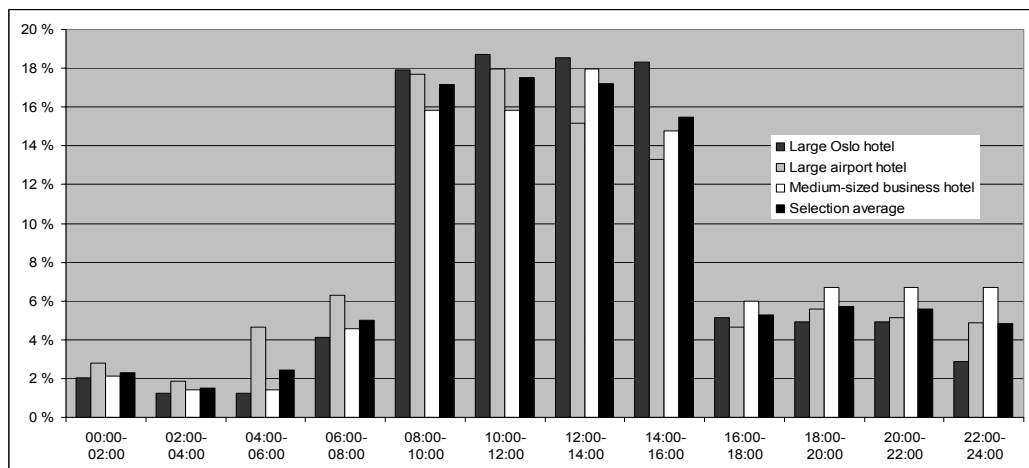
The second stage concerned how to reach employees working at the selected hotels. After selecting the 40 hotels we aimed to draw a sufficient number of employees to be able to draw firm conclusions about the surveyed population. A target of 600+ hotel workers was agreed for the survey based on feasibility and cost considerations. Also, we needed to ensure that people with different types of employment in the hotels had a known probability of being selected (different job categories are not randomly scattered throughout the day). However, to reach people in a hotel setting is a challenge, since it is unrealistic to obtain full lists of hotel employees from the

employers (in particular since our aim was to reach also employees with more informal employment relationship to the hotel). Thus, it was decided that the best way to reach them would be to encounter them in their work environment.

To reach people of different job categories it was decided that we would select two 2-hour time units in hotels and interview all hotel employees present at the hotel during such randomly selected time units. In this way we would get a sufficiently broad selection of workers at different times of the day (altogether 80 time units). Since the size of the staff at hotels tends to vary between days of the week - especially the weekend tends to differ from the rest of the week - weekdays of interviewing were also drawn randomly.

Time units were systematically selected to reflect the proportion of employees working there, so that day shifts would have a probability of being selected according to the number of employees expected to work there at the time. The average percentage of workers present at various hours was calculated based on information received from three different hotels of different types in telephone interviews with hotel administration. Based on averages of the three we could estimate the number of workers present in the workplace at different types of the day. The estimated distribution of workers in these time units is shown in Figure 1.1.

Figure 2.1. *Estimated distribution of hotel workers in time units during 24 hours*



Source: Telephone interviews with hotel administration in different types of hotels.

The 2-hour slots were selected with probability proportionate to size, i.e. proportionate to the number of employees present. In other words, people working at different shifts should have an equal probability of selection.

A drawback with the chosen sampling procedure was the risk of respondents falling into two of the worker units drawn. We therefore decided not to draw the two time periods independently of each other, but rather to make sure that the two time units were as far as possible apart. This was done based on a hypothetical list of employees ordered from the first to come to work that day to the last. The first 2-hour time unit was drawn by random non-zero positive number x_1 , where $x_1 \leq N$ (N = number of employees), while the second slot x_2 was defined as $x_1 + N/2$. In this way each time

unit was selected as a function of a hypothetical employee working in this time period being selected – the first based on a random number, the other halfway-around-the-clock away from the first person. Still, since many hotels have few employees and many work long hours, one could not eliminate the risk of one person selected twice (i.e. working during both the time periods selected).

2.3 Third stage sample

The number of employees working during the selected two-hour time period was registered, and everybody were asked to participate in the survey. Those present but refusing to participate were recorded according to their sex, age group, type of work (when known) and region of origin (according to rough categories). This enables us to find any systematic discrepancies between the sampled population and those responding and, if necessary, to adjust for non-response.

All employees working at the selected hotel during the selected time unit were regarded as eligible respondents, including temp agency staff, part-time and short-term workers. Senior managers were the only group not included in the sample. Kitchen, waiters, cleaning, administration/booking and reception staff formed the main population.

The success of this approach, of course, hinges on the willingness of chain and hotel management to allow researchers to enter the workplace during work hours. A carefully drafted letter of inquiry was distributed to make sure that management was duly informed about the research. In this way we would get more accurate estimates of hotel employees in the Oslo and Akershus region from which relative weights to adjust for sampling errors would be calculated.

2.4 Non-response

The survey organisers and interviewers were able to conduct interviews in all together 35 of the 40 hotels selected (in two hotels we were denied access, while three hotels were considered irrelevant for various reasons, see Underthun et al. 2011). A total of 1135 records of employees were registered by us during the two shifts of these 35 hotels. Of these 268 persons were recorded at both the selected shifts. In other words, we recorded 867 individuals present at the hotels. This is substantially higher than the 600 + respondents that we had aimed for.

Of these 867 individuals only 36, or 4.2 percent, did not participate in the brief interview at the workplace (questions about working hours, country of birth, work tasks, etc.). The reasons for non-response were “no time to answer” (the respondent was busy with work task he/she could not leave), interviewers unable to find the worker, or refusal to answer (in a few cases due to language problems). We have no accurate information about the nationality, work task or other composition of persons that were present at the hotel but that we for various reasons were not able to reach, but the interviewers’ impression is that it is non-systematic and random (for more on the fieldwork organisation, see Underthun et al. (2011)).

3 Subsequent stages of research

3.1 The follow-up survey

All respondents were asked to participate in a follow-up survey with more in-depth questions. Therefore e-mail addresses were recorded for those who were willing to answer the questionnaire on a web-based survey, but they could optionally receive a printed version of the questionnaire and send it in a prepaid envelope to NIBR. A total of 87% of the recorded respondents (including those not answering in person) said that they were willing to participate in the follow-up survey.

It turned out, however, that 44 percent (381 persons) participated in the follow-up survey, of whom 342 persons used the web-based questionnaire, the remaining 39 filled out a paper questionnaire and sent it to NIBR. However, not all respondents who started filling out the web-based questionnaire filled it out to the end, so item response tends to be slightly lower. In fact 21 persons were removed from the data file altogether due to low item response.

Since we have recorded key information about each of the respondents and are able to identify which of the respondents who answered, we are able to identify the discrepancies between the characteristics of respondents answering and the non-response of the follow-up survey. Key characteristics are recorded in Table 2.1.

Table 3.1 *Distribution of respondents in hotel mapping and in follow-up survey and percentage recapture.*

		Hotel mapping (N)	Follow-up survey (N)	Recapture (%)
Sex	Female	517	222	42.9
	Male	350	126	36.0
Age	-24	150	67	44.7
	25-34	288	156	54.2
	35-44	209	80	38.3
	45-54	138	39	28.3
	55+	70	18	25.7
	Missing – refused	12	-	
Country of birth	Norway	383	203	53.0
	Other Nordic	64	29	45.3
	(Other) EU	77	36	46.8
	Outside EU	331	92	27.8
	Missing – refused	12	-	
Main task	Cleaning /dishwashing	264	59	22.3
	Reception	102	48	47.1
	Waiter / bar	119	46	38.7
	Cook	84	35	41.7
	Administration and booking	88	69	78.4
	Conference	61	38	62.3
	Middle management	85	50	58.8
	All others	59	15	25.4
	Missing, refused	5	-	
Employment	Employed by the hotel	742	331	44.6
	Unsure	10	-	
	Outsourced, temp staff agency, others	111	29	26.1
	Missing, refused	4	-	

Note: In a few cases, and most notably so for Employment, there are discrepancies between the responses given by the same respondents in the mapping at the hotels and the follow-up survey. In these cases we stick to the original data conducted at the hotels, since the purpose of this table is to document the recapture of the same respondents.

Although, as can be seen from the table, there are some differences between the share of the respondents according to sex, age, country of birth and main task performed at the hotel, these differences are not considered to be so large that we need to compensate by computing weights for under-represented groups. Moreover,

we are aware that there may be other systematic discrepancies between those answering and not answering the follow-up survey, in terms of integration in the labour market (people less integrated are likely to be less prone to answer), language barriers, etc. Thus, analysis of the data requires a certain amount of caution and the uncertainty of the estimates should be pointed out when reporting survey data.

3.2 Inclusion probabilities and weights

As the sample as described above is not self-weighting (as employees at different hotels have had different probabilities for being included in the sample) sampling weights were made to adjust for this in the analysis.

In this section the following notation will be used:

$P_{\frac{h}{s}}$	is the inclusion probability for the hotel, the time-period or the employee
N_h	is the number of hotels in Oslo and Akershus
n_h	is our target sample of hotels
N_s	is number of employees in hotel (estimated).
M_t	is the percentage of staff at work during a given two hour period (estimated).

The sample was selected in three stages. In stage one and three the inclusion probability was the same for all.

The inclusion probability for each hotel was:

$$P_h = \frac{N_h}{n_h} = \frac{110}{40} = 2.75$$

The variation in inclusion probabilities and weights is introduced with the second stage sample.

$$P_t = M_t \times N_s$$

For the last stage sample all eligible employees working a selected time period were included with probability 1, which does not introduce additional variation in weights.

Inclusion probabilities for each employee are consequently:

$$P_s = P_h \times P_{t_s}$$

Weights are calculated as the inverse of the selection probabilities:

$$W_s = \frac{1}{P_s}$$

This produced a weight with mean 5.75, with a standard deviation of 4.91.

As we expected much of the variation between individual experiences to be related to the working place, it was an explicit decision to make a sample where small hotels were over-represented (hotels were selected with equal probability independent of size). This has also led to relatively large variation in weights; however this was done in order to allow for proper coverage of factors that we expected to be relevant.

If the weights introduced by the second stage sample are analysed separately we find that the weights have a mean of 2.07 and standard deviation of 1.77, ranging from 0.86 to 13.

The relatively strong expansion from the second stage sample was expected. It should be attributed to the fact that the number of officially registered employees was used to draw the sample, while the widespread use of temp agencies and part time employees led to more people being at work than our sampling frame had predicted.

3.3 Implications for analysis

The weights are used in the data analysis to adjust for the sample not being self-weighting: different hotels had different probabilities for being included in the sample. However, it was decided that weights would not be used to adjust for different groups of respondents being somewhat over- or underrepresented in the survey. Thus, when analysing the results, the authors should always stress that frequency distributions represent rough estimates and that results observed in the sample may differ somewhat from those of the whole population. The data are better suited to discern statistically significant differences between groups for which there is a substantial number of respondents than to give exact descriptive data on frequency distributions for the population as a whole.

References

Jordhus-Lier, D. C., A. C. Bergene, H. M. Knutsen and A. Underthun (2010). Hotel workplaces in Oslo and Akershus. NIBR Working Paper 2010:106. Oslo, Norwegian Institute for Urban and Regional Research: 1-30.

Underthun, A., H. Hellvik, J. Loge and D. C. Jordhus-Lier (2011). On-site workplace survey and internet-based questionnaire for hotel workers in the Oslo region: Some methodological reflections. NIBR Working Paper 2011:106. Oslo: 1-19.