

## Redesigning the German job vacancy survey - assessing the impact of high nonresponse rates

Hans Kiesl, Susanne Rässler  
Institute for Employment Research, Regensburger Str. 104, 90478 Nuremberg, Germany

### Abstract

Information about job vacancies in Germany is provided by the Federal Employment Agency, which collects reported job openings from business units. They cover less than 50% of all vacancies, as many businesses prefer not to cooperate with the Agency. Therefore, the Institute for Employment Research conducts a yearly job vacancy mail survey to estimate the total number of vacancies. Response rates have dramatically declined over the years. In 2006, less than 20% of the sampled units participated in the survey. We tried to gather information about the potential nonresponse bias by varying the length of the questionnaire. During the latest wave of the survey with the usual questionnaire of eight pages in length, we conducted an additional survey with a one-page questionnaire. Though this one-pager resulted in a higher response rate, the estimates from the different surveys did not differ significantly. We conclude that nonresponse due to higher response burden does not necessarily lead to biased estimates.

**Keywords:** job vacancy survey, nonresponse, response burden, length of questionnaire

### 1. Background

Official information about the number of job vacancies in Germany is provided by the Federal Employment Agency. Business units are encouraged to report their job openings to the Agency, which publishes monthly job vacancy statistics by region and industry sector. This official numbers cover less than 50% of all vacancies, however, as many business units prefer not to cooperate with the Agency. Therefore, the Institute for Employment Research conducts a yearly job vacancy mail survey among all business units to get an estimate of the total number of job vacancies (reported to the Federal Employment Agency or not). The questionnaire is sent out during the 4<sup>th</sup> quarter of each year.

#### 1.1 Basic Sampling Design

The sampling units are chosen by stratified random sampling from the business register of the Institute for

Employment Research. As is usually the case with business surveys, strata are defined by region (West and East Germany), by NACE sector (16 sectors) and by size, i.e. number of employees covered by the social insurance system (7 size classes). Smaller units get a smaller chance of being selected. Table 1 shows the sampling rates and sample sizes for the different size classes in the fall 2006 survey.

| Size (no. of employees) | West Germany  |             | East Germany  |             |
|-------------------------|---------------|-------------|---------------|-------------|
|                         | sampling rate | sample size | sampling rate | sample size |
| < 10                    | 0.9%          | 11,067      | 3.3%          | 10,963      |
| 10 - 19                 | 5.9%          | 8,533       | 26.1%         | 9,546       |
| 20 - 49                 | 7.3%          | 6,745       | 33.6%         | 8,282       |
| 50 - 199                | 8.6%          | 4,541       | 13.3%         | 1,815       |
| 200 - 499               | 32.9%         | 3,249       | 50.3%         | 1,111       |
| 500 - 999               | 68.3%         | 1,806       | 73.0%         | 395         |
| >= 1000                 | 75.1%         | 984         | 84.8%         | 193         |
| total                   | 2.4%          | 36,925      | 8.0%          | 32,305      |

Table 1: Sampling plan for job vacancy survey in 2006

The survey questionnaire is roughly 8 pages in length. Apart from questions concerning the number and types of vacancies, the survey contains more detailed questions on recruiting strategies. There are questions on main characteristics of the business unit and its economic environment and on specific distinctions of the job openings. Openings are divided into openings to be filled immediately (vacancies) and openings to be filled later. Additionally every year specific questions on recent labor market issues are asked, for instance on the effects of labor market reforms from the firms' perspectives. The last part of the questionnaire is related to the very last case of an engagement in the firm. Questions concern the job requirements, characteristics of the hired person, information on recruitment channels, reasons why positions are difficult to fill etc.

The job vacancy survey is a voluntary survey, and low response rates have always been a problem. During the recent years, the nonresponse rates got even worse. The response rate declined from 32.1% in 1995 and 26.8% in 2000 to 16.8% in 2005 and 19.6% in 2006 (see Figure 1).

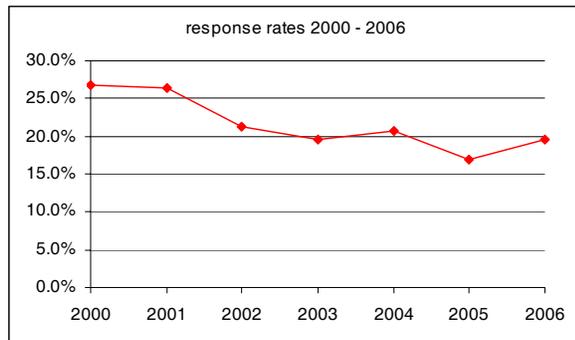


Figure 1: Response rates between 2000 and 2006

Due to these high nonresponse rates, we use auxiliary information in the estimation stage to calibrate the weighting factors to known totals (see section 3 for more details).

## 2. Reasons for nonresponse

In the first quarter of 2005 we conducted a nonresponse follow-up study to learn about the reasons for the large nonresponse rates. For that purpose, a stratified sample of size 1,700 was drawn from the approx. 45,000 nonrespondents of the 2004 job vacancy survey. We tried to conduct a CATI survey among these units. 26% of them could not be reached, as we had no or a wrong telephone number in the frame. 16% could be contacted, but were not willing to respond. The other 58% agreed in responding to a short interview, in which they were asked to give reasons for not having participated in the job vacancy survey (more than one answer was possible).

The following reasons were given most frequently (percentages weighted with sampling rates in different strata):

- 75%: did not have time; had too much work to respond (88% among largest units)
- 27%: did not have job vacancies (41% among smallest units)
- 25%: no relevant topic for the business unit (44% among small units)
- 17%: business unit only takes part in a survey if it is mandatory
- 9%: business unit never takes part in surveys.

Obviously, for the smallest units (with less than 10 employees), our detailed questionnaire on vacancies and recruiting strategies was irrelevant. Most of them probably had no vacancies for years and did not realize that this information would indeed be valuable for estimating the total number of vacancies. Thus, the sample will probably be biased in the sense that units with job vacancies are overrepresented.

The other lesson to be learned from this study was that the top reason for not taking part in the survey is having too much work to do. Since the questionnaire is 8 pages in length, we concluded that shortening the questionnaire and thus reducing response time and burden should result in higher response rates. On the other hand, shortening the questionnaire results in fewer questions to be asked and thus in less information for our Institute.

## 3. Impact of length of questionnaire

### 3.1 Study design

To assess the impact of the length of our questionnaire on response rates and final estimates, we conducted a split survey during fall 2006. We drew two different samples from our business unit frame. We sent the usual 8 page questionnaire to the main sample (for sample sizes see section 1.1), while the additional sample received a condensed 1 page questionnaire which was almost identical to the first page of the long form questionnaire. Only those questions needed to estimate the number of job vacancies were included in this short form.

Stratification for the additional sample was similar to the one used for the long form. Due to the reduced sample sizes (for budget reasons), fewer strata were defined: region (West/East Germany), NACE sector (5 sector groups) and size class (5 classes).

| Size (no. of employees) | sample size |        |
|-------------------------|-------------|--------|
|                         | 8 pages     | 1 page |
| < 10                    | 22,030      | 3,898  |
| 10 - 19                 | 18,079      | 2,631  |
| 20 - 49                 | 15,027      | 1,943  |
| 50 - 199                | 6,356       | 2,266  |
| 200 - 499               | 4,360       | 770    |
| 500 - 999               | 2,201       | -      |
| >= 1000                 | 1,177       | -      |
| total                   | 69,230      | 11,508 |

Table 2: Sample sizes for main and additional survey

Units with 500 employees or more were excluded from the additional sample due to the very high sampling rates in the main survey. Also, business units from the public sector were excluded from the additional survey, since they usually have a comparatively high response rate in the main survey. Table 2 compares the sample sizes (total Germany) for the two surveys.

### 3.2 Impact on response rates

| Size (no. of employees) | 8 pages | 1 page | abs. diff. | rel. diff. |
|-------------------------|---------|--------|------------|------------|
| < 10                    | 21.0%   | 27.3%  | + 6.3%     | + 29.9%    |
| 10 - 19                 | 17.0%   | 30.7%  | + 13.7%    | + 80.1%    |
| 20 - 49                 | 18.9%   | 28.2%  | + 9.3%     | + 49.1%    |
| 50 - 199                | 23.2%   | 28.6%  | + 5.4%     | + 23.2%    |
| 200 - 499               | 18.1%   | 27.1%  | + 9.0%     | + 49.9%    |
| 500 - 999               | 18.4%   | -      | -          | -          |
| >= 1000                 | 26.8%   | -      | -          | -          |

Table 3: Response rates by size of business unit

Since the short survey requires much less time to complete, we expected higher response rates for the 1-pager, which turned out to be true. In every size class up to 499 employees, response rates were higher for the 1 page survey (28% on average compared to less than 20% for the 8 page questionnaire). Although this is only an absolute increase of 8 percentage points on average, the relative increase in the different size classes ranges from 23% to 80% (see Table 3).

|   | 8 pages | 1 page | abs. diff. | rel. diff. |
|---|---------|--------|------------|------------|
| Agriculture, farming, forestry, fishing   | 24.8%   | 36.5%  | + 11.6%    | + 46.9%    |
| Manufacturing   | 19.6%   | 29.8%  | + 10.2%    | + 52.3%    |
| Construction  | 21.5%   | 28.8%  | + 7.3%     | + 34.1%    |
| Sale, trade, hotels and restaurants, transport, communications                  | 14.9%   | 23.3%  | + 8.4%     | + 56.1%    |
| Financial intermediation, insurance, renting and business activities            | 17.3%   | 24.9%  | + 7.6%     | + 44.0%    |
| Public administration, education, health, social work, other service activities | 23.0%   | -      | -          | -          |

Table 4: Response rates by industry sector

If we look at the different industry sectors, we see that response rates for the additional survey were uniformly higher in all strata. The relative increase ranges from 34% (construction sector) to 56% (hotels, transport, communication; see Table 4).

### 3.3 Estimation

Due to the high nonresponse rates, using the pure Horvitz-Thompson estimator for stratified sampling without replacement is not the best choice. We incorporate auxiliary information to calibrate the Horvitz-Thompson weights to known totals. The following variables are used as calibration variables:

- number of business units by region, size classes and industry sectors;
- number of employees (covered by the German social insurance system) by region, size classes and industry sectors; information from business register (= sampling frame);
- number of registered job vacancies by region and industry sectors; information from the official statistics of the Federal Employment Agency.

Note specifically that we calibrate on the *registered* number of job vacancies. In the questionnaire, the business units are asked for the number of their job openings and how many of them they already reported to the Federal Employment Agency. Calibration is with regard to the number of these reported openings.

We calculate the calibrated weights by means of the generalized regression estimator (GREG); for the mathematical details see e.g. Särndal, Swensson and Wretman (1992). We use the SAS macro CLAN from Statistics Sweden to compute point and variance estimates (c.f. Andersson and Nordberg, 1998).

### 3.4 Impact on estimates

The generalized regression estimator was used to estimate the total number of job vacancies for West Germany, East Germany and total Germany. Since the largest size classes and the public sector were not included in the additional survey, we copied all responding units from the main survey that belong to one of the corresponding strata (public sector or size 500+) into the additional survey (i.e. we made duplicates of these cases). This had to be done in order to get estimates from both surveys for the whole population.

The estimates from the additional survey (1 page, higher response rate) are slightly higher than the estimates from the main survey (8 pages, lower

response rates); c.f. Table 5. The absolute differences are 73,000 for West Germany and 17,000 for East Germany; the relative differences are 9.4% for West Germany and 8.0% for East Germany.

|         | <b>Estimate</b> | <b>standard error</b> | <b>coeff. of variation</b> |
|---------|-----------------|-----------------------|----------------------------|
|         | 8 pages         | 8 pages               | 8 pages                    |
|         | 1 page          | 1 page                | 1 page                     |
| West    | 775,000         | 29,000                | 3.7%                       |
|         | 848,000         | 51,000                | 6.0%                       |
| East    | 212,000         | 10,000                | 4.7%                       |
|         | 229,000         | 32,000                | 14.0%                      |
| Germany | 987,000         | 31,000                | 3.1%                       |
|         | 1,077,000       | 60,000                | 5.6%                       |

Table 5: Estimates for total number of job vacancies from the two different surveys

At first glance, this result seems to indicate that the nonresponding units tend to have more job vacancies than the responding units (if we are willing to accept the hypothesis that the respondents to the additional survey include some potential nonrespondents to the main survey). Note, however, that the estimator is calibrated to the registered number of job vacancies. Therefore the right conclusion is that the nonresponding units tend to have a lower ratio of registered vs. total vacancies (since calibration leads to the same estimate for the number of registered vacancies in both surveys). This is a very plausible result: business units that are not willing to report their job openings to the Federal Employment Agency are less inclined to take part in a job vacancy survey.

If we look at the estimates for different size classes, the picture is not that clear. The short form leads to higher estimates for job vacancies among small and large units, whereas the long form leads to higher estimate for units with 20 to 199 employees (see Figure 2).

An important question remains: are these differences significant? The (estimated) standard errors of the estimated number of job vacancies (in West, East and total Germany) are shown in Table 5. The difference between the estimates from the two surveys is smaller than the sum of the estimated standard errors for the two estimates, which shows that the differences in the estimates are not significant (for simplicity, we disregard the fact that the two surveys are not completely independent). This is a quite surprising result, as we had expected that higher response rates would lead to significantly different estimates. To be precise: we expected that higher nonresponse results in

more biased estimates. However, our findings seem to indicate that the nonresponse bias in our job vacancy survey might not be as large as expected.

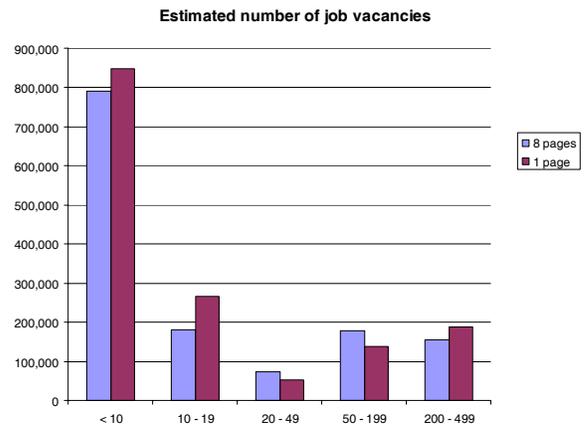


Figure 2: Estimates for total number of job vacancies within different size classes

#### 4 Conclusions

Sending out a 1 page questionnaire instead of an 8 page form considerably increased the response rate from approx. 20% to approx. 28%. The length of the questionnaire has a substantial effect on the response rates. Nevertheless, we had initially hoped to get even better results. On the other hand, estimates from the two surveys were surprisingly close, i.e. changes in response rates seem to have only little effect on the main survey estimates. Since the 8 page form collects much more information than the 1 pager, we will not consider the 1 pager as a real alternative.

However, we will adopt a split questionnaire design (c.f. Raghunathan and Grizzle 1995) in the future to find a compromise between acceptable response rates and a maximum amount of information. Apart from this, other strategies like a non-respondent CATI follow-up survey have to be considered to increase the response rates. Shortening the questionnaire alone obviously does not suffice.

#### References

Andersson, C. and Nordberg, L. (1998): A User's Guide to CLAN 97 - a SAS-program for computation of point- and standard error estimates in sample surveys. Statistics Sweden.

Raghunathan, T. E. and Grizzle, J. E. (1995): "A Split Questionnaire Survey Design." *Journal of American Statistical Association*, 90:55-63.

Särndal, C.E., Swensson, B. and Wretman, J. (1992): *Model Assisted Survey Sampling*. Springer, New York.