Are Japanese Wage Statistics Representative?

By

Daiji Kawaguchi (The University of Tokyo)
Takahiro Toriyabe (The University of Tokyo)

January 2022

CREPE DISCUSSION PAPER NO. 115

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

Are Japanese Wage Statistics Representative?*

Daiji Kawaguchi Takahiro Toriyabe

January 21, 2022

Abstract

This paper assesses the representativeness of two major government wage surveys of Japan: the Basic Survey on Wage Structure (BSWS) and the Statistical Survey of Actual Status for Salary in the Private Sector (SSPS). We examine whether the two-step procedure that involves random sampling of establishments and random sampling of workers within selected establishments ensures the representativeness of sampled workers in the population. We find lower response rates for establishments operating in the service industry, with fewer employees, and located in urban prefectures, but no relationship between the response rate and past wages. The mean wages calculated from the sampled individual payroll records coincide with establishment-level aggregate records, indicating that workers are randomly selected within establishments. Overall, we only find evidence of non-random sample selection based on observed characteristics of establishments, thus the selection is considered to be *ignorable* and the unbiased mean wage can be estimated with an appropriate weighting. We also discuss coverage issues due to the sampling design.

Keywords: Wages, Random Sampling, Comparison, Coverage, Non-response, Bias
JEL codes: C81, C83, J31

*This research is based on the research report "Research on Comparative Verification of Wage-Related Statistics" by Metrics Work Consultants, Inc. commissioned by the Director-General for Policy Planning (Statistical Policy) of the Japanese Ministry of Internal Affairs and Communications. We thank Junichi Suzuki of the Statistics Commission Office and Hiroshi Nishino of Metrics Work Consultants, Inc. for their efficient management of the project and substantive input on the research report. We also thank Satoru Hagino, Hitoshi Konno, Masaaki Sato, and Masahiro Haramiishi of the Statistics Commission Office, Shinichi Kakui of the Ministry of Health, Labour and Welfare, and Atsuhiro Miyamoto, Yasutaka Yoneta of the National Tax Agency for their comments and administrative assistance. We also thank participants of the interim and final report meetings of the Bureau of Statistics for their valuable comments. This research is financially supported by JSPS KAKENHI Grant Number 20KK0032. Editorial assistance was provided by Philip MacLellan. All remaining errors and omissions are our own.

1 Introduction

Wage statistics plays crucial roles for policy making through providing one of the most important signals for monitoring the current state of the economy. For example, changes in nominal wages are regularly cited in Bank of Japan quarterly publications, and these figures play a crucial role in determining its monetary policy stance.[1] The government also uses wage statistics for administrative purposes such as wage inflation adjustments of unemployment benefits and worker compensation for occupational injury. Researchers also use wage statistics to analyze a variety of topics ranging from the impact of minimum wages to macroeconomic wage dynamics (Kawaguchi and Mori, 2021; Hoshi and Kashyap, 2020). Given the importance of wage statistics for both policy makers and researchers, the long-standing practice of the statistics section of the Ministry of Health, Labor and Welfare which does not follow approved statistical sampling procedures attracts much public attention and casts serious doubt on the credibility of wage statistics in Japan.[2] Independent from this particular scandal, the long-term declining trend in response rates for government surveys which is common across developed countries raises questions about the representativeness of wage statistics due to possible sampling bias. Yet despite heightened attention to this issue, systematic assessment of the method of collecting wage statistics in Japan has been scarce.

Previous studies of sampling bias in earnings surveys, based mainly on the US Current Population Survey (CPS), have set an important path for investigating the issue. The literature demonstrates that sample selection bias is *ignorable* if the non-response behavior is characterized by observable characteristics, as the population distribution can be recovered by applying an appropriate weight. On the other hand, sample selection bias is *non-ignorable* if the non-response behavior depends on unobserved determinants of the outcome variable. In this latter case, recovering the population distribution from the sample becomes difficult unless there exists a credible instrumental variable that affects the response behavior but does not affect the outcome variable. In this study, we analyze the dependence of response behavior on both observed and unobserved characteristics, with the latter approximated by past responses obtained by exploiting a panel feature of the wage surveys.

For our investigation of the credibility of wage statistics in Japan, we examine two large-scale surveys that are conducted independently by two different national government ministries. These are the Basic Survey on Wage Structure (BSWS, *Chingin Kouzou Kihon*

[1] Quarterly Bank of Japan publications include the Outlook for Economic Activity and Prices of the Bank of Japan, which provide background information for monetary policy meetings.

[2] See Higo et al. (2020) for an overview of the procedural flaw in the implementation of the Monthly Labor Survey and its effect on the published statistics.

Tokei Chosa) by the Ministry of Health, Welfare, and Labor and the Statistical Survey of Actual Status for Salary in the Private Sector (SSPS, *Minkan Kyuyo Jittai Tokei Chosa*) by the National Tax Agency. We assess their quality in terms of their representativeness of the population and the quality of their wage measurements.

Both the BSWS and SSPS collect the wage information of individual workers through establishment surveys drawing on a similar two-step sampling design. In the first step, the surveyor randomly selects establishments according to the stratified sampling method and in the second step, the surveyor instructs personnel managers of establishments to randomly select employees from its payroll records. Sampling bias can occur at each of these two steps. To examine the representativeness of the sampled establishments, we first match the list of establishments selected for the survey with those that actually responded. Then, we examine how the probability of unit non-response differs by observed establishment characteristics such as size, industry, city size and prefecture in which the establishment is located. In addition to selection on observables, we also attempt to quantify the degree of selection on unobservables by exploiting the short-panel feature of the BSWS and SSPS created by its oversampling of large establishments. In particular, we choose past average wages of the establishment as a proxy for unobserved establishment characteristics and examine whether they impact survey responses.

We first document the characteristics of the establishments that actually respond to the survey as compared to the full list of establishments selected for the survey. For the BSWS, we find that establishments operating in the service industry, with fewer employees, and located in urban prefectures are less likely to respond to the survey. We also find a lower response rate for service sector establishments with fewer employees in the SSPS. Together, these results suggest a possible sample selection bias based on observed characteristics.

We further examine the possibility of sample selection bias based on unobserved characteristics by drawing on the panel nature of the BSWS and SSPS. Although both surveys are principally cross sectional, to attain oversampling of large establishments, the same large establishments are repeatedly selected up to a certain number of times. Using this feature, we construct short-panel datasets to examine the dependence of response behavior on the past average wages of each establishment. In both surveys, we find that establishments with higher average wages in the previous survey are more likely to respond to the current survey in statistically significant ways, but the sizes of the estimated impacts are quite limited.

In the end, the establishment level analysis suggests sample selection bias based on observed characteristics such as establishment size, industry and the location, but no meaningful bias based on unobserved characteristics, at least to the extent that these

unobserved characteristics are captured by past average wages. Thus, any potential bias due to sample selection of establishments can be corrected by proper weighting based on establishment observed characteristics.

The idiosyncratic nature of the sampling methodology for Japanese wage statistics occurs in the second step. Both the BSWS and SSPS ask selected establishments to randomly pick workers from the establishment's payroll record, which obviously raises the concern that the random sampling may not be properly implemented by administrators who are not experts in sampling methodology. To address this concern, we compare the mean wages calculated based on individual workers in the sample and the mean wages calculated based on the establishment-level aggregate wage bill and total hours worked. If instructions were properly followed, these two numbers should match. Since the BSWS does not collect information on each establishment's wage bill and total hours worked, we first match the establishment in the BSWS with the same establishment in the Monthly Labor Survey (MLS, *Maitsuki Kinro Toukei*) which includes establishment-level number of workers, aggregate wage bill and total hours worked. Both the BSWS and MLS are conducted by the Ministry of Health, Labor and Welfare and about 40% of BSWS establishments with 500 or more employees can be matched to MLS establishments because of the commonality of the sampling structure. The analysis shows that while the difference is statistically significant due to the large sample size, the difference in the two numbers is less than one percent of the BSWS establishment means. Unlike the BSWS, the SSPS does collect information on number and total wages of salaried workers, so this analysis is more straightforward, and we find virtually no difference in the means of annual wages calculated based on individual payroll records and the aggregate figures of establishments. In sum, as individual and establishment-level data largely coincide, the random sampling of workers appears to have been conducted properly in both the BSWS and SSPS.

In addition to the assessment of the sampling bias of the wage statistics, we further examine the limitation on the coverage of workers because of the sampling design of the wage statistics. We first focus on the fact that both BSWS and SSPS do not cover freelancers. To compliment this uncovered workers, we draw on the multiple waves of the Employment Status Survey, which is the household survey that covers population including self-employed workers and those who do not work. We find that about 6 percent of the all workers are the freelance workers, and they earn substantially less than those who are employed. In the end, the wage inequality captured by the two representative wage statistics of Japan underestimate the overall earnings inequality.

We next compare the coverage of workers by BSWS and SSPS. While the BSWS is the wage statistics cited more frequently than and the SSPS, the SSPS has wider coverage of

workers at the both bottom and top tails of the wage distribution. Regarding the bottom end of the wage distribution, the SSPS covers workers who work for micro establishments that hire 4 or fewer workers. On the other hand, regarding the top end of the wage distribution, the SSPS asks establishments to report all individuals who earn 20 million Yen (approximately US$ 200,000) or more annually. Reflecting the differences in the sample design, the BSWS underestimate the wage gap between the median and the 10th percentile and the 90th percentile and the median.

The rest of the paper is organized as follows. Section 3 introduces the two representative wage surveys of Japan: the Basic Survey of Wage Structure (BSWS) and the Statistical Survey of Actual Status for Salary in the Private Sector (SSPS). Section 4 examines the survey response behavior of establishments and Section 5 assesses whether the establishments randomly select workers from their payroll records. Following a discussion in Section 6 about workers who are not covered by the two wage surveys, Section 7 examines the evolution of wage inequality based on two wage statistics. Section 8 concludes the paper.

2 Literature on Non-response Bias

We articulate the central issues in the sample selection bias problem arose from survey non-response by drawing on an extensive literature on the the Current Population Survey (CPS) and Census of the US.[3]

A strand of literature pays attention to item non-response of earnings questions in the CPS. For instance, Hirsch and Schumacher (2004) find that about 30% of the survey respondents do not report their earnings. When examining non-response bias, an important distinction is whether or not the non-response is *ignorable*, or missing at random conditional on observed characteristics, for if the bias is ignorable, then statistical inference is properly implemented by simply ignoring the sample selection process (Rubin, 1976). Assuming that the sample selection bias is ignorable, statistics agencies often impute the missing earnings based on observed characteristics, for instance, by the hot deck imputation procedure. Hirsch and Schumacher (2004) and Bollinger and Hirsch (2006) assess the performance of the hot deck imputation procedure by comparing estimates obtained with non-imputed and imputed samples, and Bollinger and Hirsch (2006) propose a re-weighting procedure based on the difference in the response rate by observed

[3]There are few studies of non-response behavior outside of the US available in English. Barnes et al. (2008) analyzes the unit non-response of the UK Labor Force Survey, and in their discussion of rotation group bias in the US CPS, Krueger et al. (2017) also refer to bias in the UK and Canadian counterparts. The European Commission (2020), which tabulates the time series of unit non-response rates from 2014 to 2018 among 35 European countries, finds that the evolution of the non-response rate is different across countries.

characteristics.

Statistical inference becomes complicated when the sample selection bias is *non-ignorable*, or conditional on unobserved characteristics, and there are two general approaches. The first is a Heckman-style sample selection correction using excluded variables that affect response behavior but not wages (Lillard et al., 1986; Vella, 1998). The second approach applies a Manski-style bound to the sample selection to infer the distribution of wages among non-respondents (Manski, 2016), with the bound substantially tightened by the presence of an excluded variable that affects sample selection but not earnings. For either method to be practical, the presence of an excluded variable is crucial, but this is generally difficult to obtain *ex post*. Dinardo et al. (2021) propose constructing a credible excluded variable by randomizing the subjects among non-respondents who receive intensive reminders.

In addition to analyzing item non-response regarding earnings or wages, another strand of literature has examined unit non-response as well, but the challenge here is in obtaining information on the survey non-respondents. Korinek et al. (2006) show that the CPS unit non-response rate increases as local income increases, suggesting that high earners are less likely to respond to the CPS. Further, a series of studies examining the survey response behavior of households by matching survey respondents with administrative data that presumably represents the population (Kline and Santos, 2013; Bollinger et al., 2019). Kline and Santos (2013) developed a statistic that characterizes the type of sample selection bias on a continuum ranging from missing randomly to missing non-randomly. Bollinger et al. (2019) concludes that the sample selection bias of the CPS is non-ignorable in the tails of the earnings distribution but that the impact on the estimation of *mean* earnings is minimal.

3 Overview of Japanese wage statistics

In this section, we introduce the two major surveys of Japanese wage statistics analyzed in this study as well as an additional survey used for establishment matching. These are the Basic Survey of Wage Structure (BSWS), the Statistical Survey of Actual Status for Salary in the Private Sector (SSPS), and the Monthly Labour Survey (MLS, *Maitsuki Kinrou Tokei Chosa*). [4] The BSWS and SSPS collect individual worker-level records on wages while the MLS collects establishment-level wage expenses.

The BSWS is an annual survey conducted by the Ministry of Health, Labour and

[4]In addition to these three surveys, the National Personnel Authority conducts a wage survey of the private sector (*Shokushubetsu Minkan Kyuyo Jittai Chosa*) to benchmark public sector salaries, but this paper does not cover that survey as its micro-data was not available. We refer the reader to Kawaguchi (2013) for an analysis using aggregated data.

Welfare that samples private establishments with five or more employees and public establishments with ten or more employees.[5] The BSWS is a repeated cross-sectional survey that over-samples large establishments with 500 or more employees, so it is not rare for these establishments to be sampled in multiple (but not necessarily consecutive) years. Between 2012 and 2017, 21 percent of establishments in the BSWS were sampled twice or more, and 14 percent in the SSPS were sampled at least twice between 2012 and 2019. This sampling structure enables us to construct short-panel data for these large establishments. The BSWS survey, which is conducted in July, asks for the monthly salary and hours worked including overtime in June and the worker's annual bonus payment in the previous year for each individual worker based on payroll records.

The SSPS, which is conducted by the National Tax Agency, is similar to the BSWS in that it is an annual worker-level survey, but it covers a wider range of establishments including those with only 1–4 employees that are not covered by the BSWS. It does not cover some self-employed individuals who do not have any employees if they are not liable for withholding tax. Another difference between the BSWS and the SSPS is their coverage of the high income population, with the SSPS covering executives but the BSWS not. Furthermore, the SSPS asks establishments to report information on all employees whose annual salary exceeds 20 million JPY (approximately US$ 200,000). Therefore, the SSPS is expected to capture the lower and upper tails of the wage distribution better than the BSWS.

A notable feature of both the BSWS and SSPS is that the random sampling of establishments is implemented by the surveyors but the random sampling of workers is left to the surveyed establishments (see Table 1 for more details on the sampling design). Although a guideline is provided to the establishments, it is conceivable that the sampling procedure might not be conducted appropriately because establishments are not experts in surveying. We thus examine whether random sampling of employees is properly implemented in Section 5.

The MLS is a monthly survey of establishments with five or more employees, and collects the wage bill of employees at the establishment level. The rotation sampling design of MLS with over-sampling of large establishments enables to construct establishment panel data. The BSWS and MLS can be matched by using the common establishment identifier because the Ministry of Health, Labour and Welfare implement both surveys, select establishments based on a common population database, and assign the same identifier. In Sections 4 and 5, we exploit this feature to validate whether the within-establishment random sampling of workers is implemented properly.

[5]The definition of employees excludes temporary workers with contract periods of less than 1 month. Workers whose contract periods extend one month or more are known as "permanent workers" (*Joyo Rodo Sha*).

Table 1: Summary of Japanese wage statistics

	Basic Survey of Wage Structure	Monthly Labour Survey	Statistical Survey of Actual Status for Salary in the Private Sector
Population	Private est. with 5+ workers Public est. with 10+ workers	Private est. with 5+ workers Public est. with 5+ workers	Persons liable for withholding tax
Survey timing	July	Every month	January or February
Surveyed wage	June	Previous month	Previous year
Sample size	78,000 est. 1,650,000 workers	33,000 est.	29,000 est. 320,000 workers
Surveyor	Ministry of Health, Labour and Welfare	Ministry of Health, Labour and Welfare	National Tax Agency
Sampling prob. of est.	5–9: 3% 10–29: 5% 30–99: 9.2% 100–499: 19.9% 500–999:56.2% 1,000–4,999: 65.1% 5000–14,999: 90.4% 15,000+: 100%	NA (Depend on region and industry) 30–99: 0.4%–50% 100–499: 0.7%–100% 500+: 100%	1–9: 0.25% 10–29: 0.5% 30–99: 1.7% 100–499: 6.7% 500–999: 33.3% 1000+: 100% Headquarters: 100%
Sampling prob. of workers	5–29: 100% 30–99: 50% 100+: Depend on est. size and industry		1–9: 100% 10–29: 50% 30–99: 16.7% 100–499: 5% 500–999: 1% 1,000–4,999: 0.5% 5,000+: 0.5% (up to 100 employees) Headquarters: 5%
Stratification	Prefecture, est. size and industry	5–29: Region and industry 30+: Est. size and industry	Est. size etc. (Details NA)

Note: The definition of employees in the BSWS and MLS excludes temporary workers whose contract period lasts less than 1 month. The workers whose contract period extend one month or more are called *Joyo Rodo Sha*. See `https://www.mhlw.go.jp/toukei/list/dl/30-1d-02.pdf` for the detailed sampling probability of establishments in the MLS. See `https://www.mhlw.go.jp/toukei/itiran/roudou/chingin/kouzou/detail/dl/20210622-detail-06.pdf` for the detailed sampling probability of workers in the BSWS. In the SSPS, "headquarters" means headquarters of joint stock company with employees less than 500, and capital of more than 1 billion JPY.

4 Establishment-level Survey response

Over the last half-century, Japan has experienced a continual decline in the response rate for government surveys, raising questions about potential selection bias in measures of the mean wage and mean hours worked. For instance, the response rate of the BSWS was 87 percent in 1982 but has declined to around 70 percent in recent years.[6] The BSWS employs a stratified sampling design in which the strata are defined by prefecture, industry and establishment size. As the sampling probability of establishments differs according to the characteristics of the strata, the appropriate weight for recovering the population means is the inverse of the ratio of the number of valid responses and the number of establishments in the population. The SSPS, on the other hand, stratifies establishments by size, and so the weight is the inverse of the valid responses and the number of establishments within the strata.

These weights, which are designed to address the heterogeneity of response rates by strata, are appropriate for recovering the population means as long as the non-response behavior is *ignorable* within the strata because it is not correlated with the unobserved determinants of wages or hours worked. However, if the survey response probability is *non-ignorable*, meaning that the response probability depends on unobserved characteristics of establishments, then the published weights will not accurately recover the population means. Thus the source of the unit non-response plays a key role in determining whether bias correction is possible via sampling weights and/or regression techniques.

In this light, it is important to check if the selection is ignorable or non-ignorable by analyzing how much the survey response behavior of an establishment depends on observed and "unobserved" establishment characteristics. As a way to approximate the "unobserved chracteristics" we will use the past average wages of the establishment as a proxy variable.

In this section, we assess whether the sample selection is ignorable in the BSWS or the SSPS by constructing establishment-level panel data from establishments that have been surveyed multiple times. For this analysis, we examine whether the current response behavior depends on past wages. While the ignorability assumption cannot be tested directly because we do not observe the wages of non-responding establishments, if we find that the response probability depends on past wages for the responding establishments, we can infer that the ignorability assumption is likely to have been violated.

[6]Source: `https://www.mhlw.go.jp/toukei/list/dl/chinginkouzou_01.pdf`

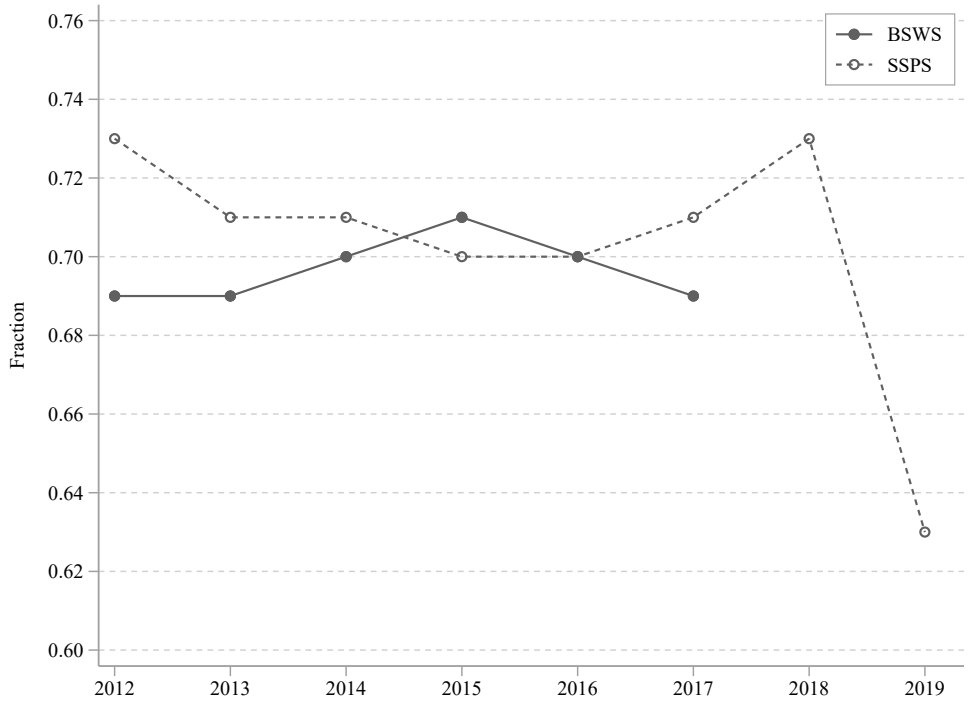


Figure 1: Response rate of the BSWS and SSPS

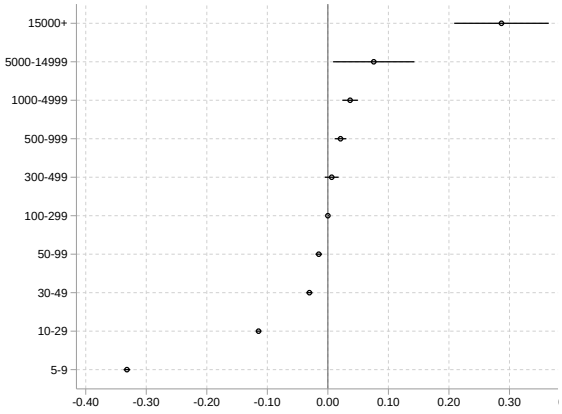
4.1 Relationship between survey responses and establishment characteristics

The Basic Survey on Wage Structure

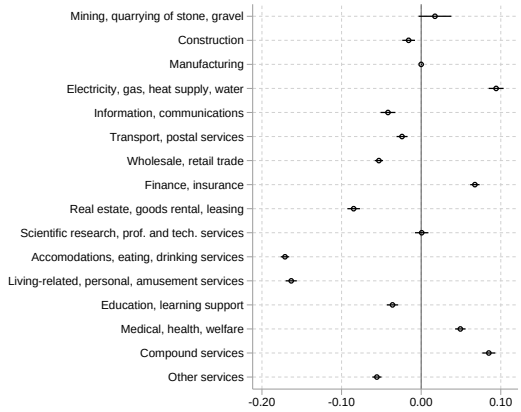
To characterize the survey response behavior in terms of observed characteristics of establishments, we analyzed the responses by establishment size, industry, prefecture, and city size using the BSWS from 2012 to 2017. Drawing from the list of sampled establishments, we matched the list with the micro data of the BSWS valid responses to generate the indicator for the valid response at the establishment level. The resulting matched data includes the indicator variable for the valid response and the variables used for stratification including the category variables of establishment size, industry, prefecture, and city size.[7] Table A1 indicates the apparent relationship that response rate increases with establishment size. In terms of industry, the response rate is relatively low in wholesale and retail trade, real estate and goods leasing, accommodation and food services, lifestyle-related services and entertainment, and education and learning support services (Table A2). We also investigated the response rate by city size, but did not find substantial heterogeneity.

Since the establishment characteristics examined above seem highly correlated with

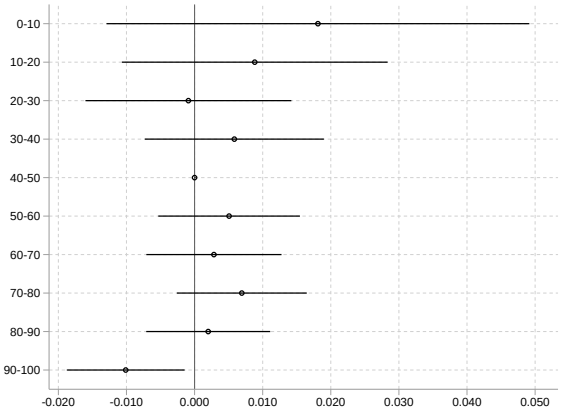
[7]The city size category variable was obtained from the population size known from the municipality code.



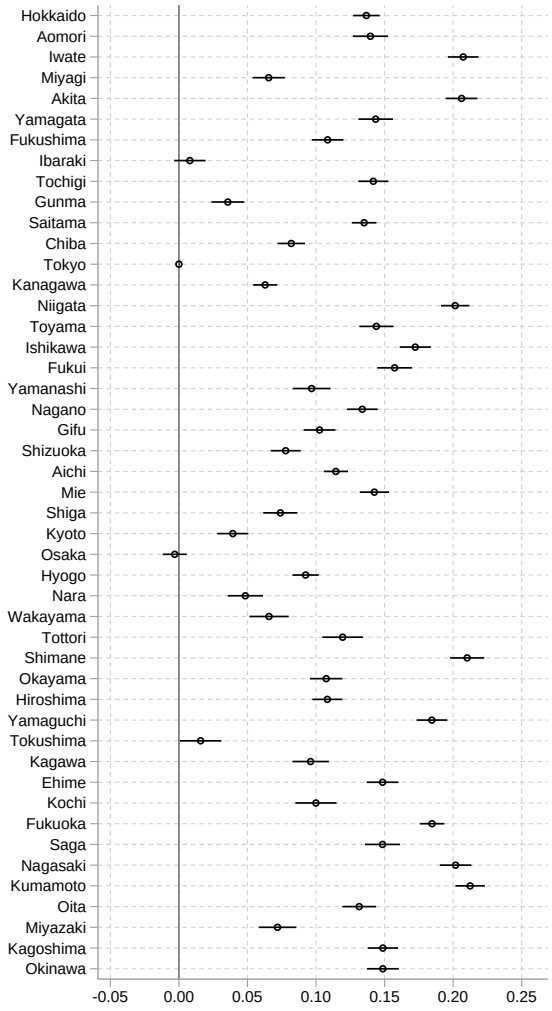
(a) Establishment size



(b) Industry



(c) City size deciles



(d) Prefecture

Figure 2: Survey response and establishment characteristics (BSWS)

Note: This figure shows the estimation results of equation (1). Points and bars indicate the estimates and 95 percent confidence intervals, respectively. Standard errors are clustered by each establishment.

each other, we disentangle the correlation by the following regression:

$$y_{itp} = x'_{it}\beta + \tau_t + \pi_p + u_{itp}, \tag{1}$$

where i, t, p indicate establishment, year and prefecture, respectively, and τ_t and π_p are year and prefecture fixed effects. The vector of explanatory variables, x_{it}, includes establishment size, industry and city population. The dependent variable is a dummy variable for the survey response. Figure 2 shows the results of the regression analysis, which basically confirms the previous findings that establishments with smaller number of employees and in service industries are less likely to respond. However, after conditioning on the establishment size and the industry, this regression analysis additionally indicates that the response rate is low in the largest city group, which mainly consists of cities designated by government ordinance (*seirei shitei toshi*) and core cities (*chukakushi*). In terms of prefecture, the response rate is lowest in Osaka, followed by Tokyo. Thus, this regression analysis suggests that the survey non-response is particularly severe in urban areas.

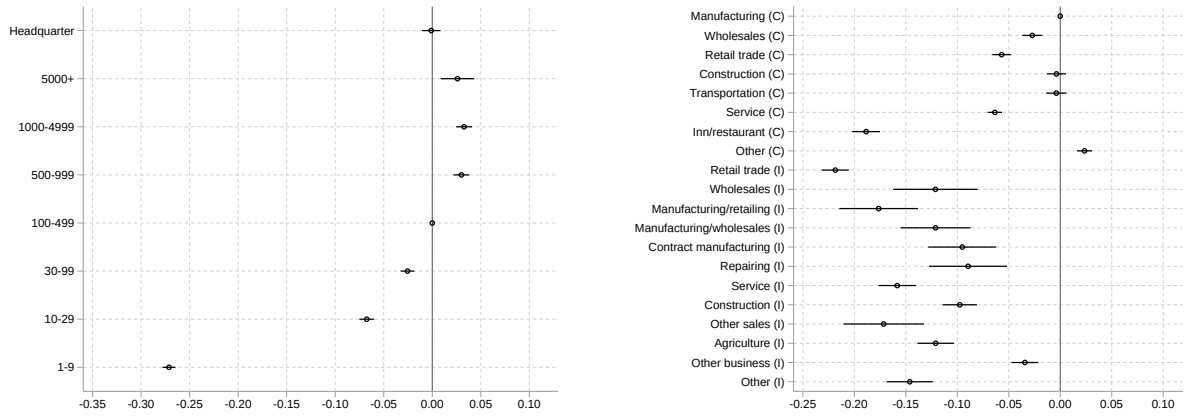
The Statistical Survey of Actual Status for Salary in the Private Sector

As with the BSWS, we calculated the response rate of the SSPS for 2012–2019 by matching the list of sampled establishments with the micro data of the SSPS. Since regional information was not available, we focus on establishment size and industry in this analysis. Although the response rate was generally stable across years, there was a decrease from 73 percent in 2018 to 63 percent in 2019, which is enumerated in 2020, due to the novel coronovirus pandemic.[8] In terms of establishment size, as in the BSWS, the response rate of small establishments is substantially lower (Table A4)), but for establishments with 100 or more employees, we did not observe a clear relationship between the establishment size and the response rate. The reason why this relationship is different from the BSWS is unknown. By industry, the response rates for retail trade, inn/restaurant, and service are low (Table A5), similar to the BSWS. In addition, the response rate for sole proprietors is significantly lower than that for corporations. These tendencies are confirmed by the regression analysis (Figure 3).

4.2 Analysis using panel structure

To examine whether the survey non-response is ignorable conditional on observed characteristics of the establishments, we investigate whether the response behavior depends on

[8]Reporting for 2019 occurred in early 2020, and requests and inquiries to business establishments were suspended at that time due to the declaration of a state of emergency in April 2020.



(a) Establishment size (b) Industry

Figure 3: Survey response and establishment characteristics (SSPS)

Note: This figure shows the estimation results of equation (1). Points and bars indicate the estimates and 95 percent confidence intervals, respectively. Standard errors are clustered by each establishment. In the SSPS, "headquarters" means headquarters of joint stock company with employees less than 500, and capital of more than 1 billion JPY. "(C)" indicates corporation, and "(I)" indicates Individual business.

the past average characteristics of workers, including wages, at the establishment level. To implement the analysis, we constructed panel data for establishments using the BSWS and SSPS. Since the sampling probability of large establishments is relatively high in both the BSWS and SSPS, it is not rare for the same establishment to be sampled in multiple years. Table A1 of Appendix B reports the sampling and response distributions for the BSWS from 2012–2017, and we see that 21 percent of the sampled establishments were chosen multiple times. For the SSPS from 2012–2019, the result was 14 percent (Table A2). This over-sampling allowed us to construct panel data of establishments in order to analyze the response status of the BSWS and the SSPS.

Regarding the response behavior of establishments conditional on the number of occasions selected for the survey, irrespective of the number of sampled years, the mode of the distribution of the number of responses is equal to the number of sampled years (Tables A1 and A2). Thus, many establishments respond to the survey whether or not they have been selected for multiple years but, on the other hand, some establishments do not respond to the survey at all even though they have been selected for multiple years. This raw data suggests a strong serial correlation in the survey response behavior.

In order to shed further light on this apparent serial correlation of survey response behavior, we estimated a linear probability model of the survey response that included past response behavior as an explanatory variable. In particular, we estimated the following model:

$$y_{itp} = z'_{it}\delta + x'_{it}\beta + \tau_t + \pi_p + u_{itp},$$

Table 2: Correlation between survey response and previous survey response(s)

Dep.Var. Survey resp.	BSWS				SSPS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NR	-0.530	-0.476			-0.475	-0.461		
	(0.004)	(0.004)			(0.006)	(0.006)		
NR →R			-0.205	-0.174			-0.188	-0.180
			(0.012)	(0.011)			(0.008)	(0.008)
R →NR			-0.396	-0.360			-0.348	-0.344
			(0.013)	(0.012)			(0.009)	(0.009)
NR →NR			-0.696	-0.643			-0.632	-0.622
			(0.010)	(0.010)			(0.008)	(0.008)
Constant	0.876		0.920		0.898		0.920	
	(0.001)		(0.002)		(0.001)		(0.001)	
Controls		X		X		X		X
Observations	107010	107010	29378	29378	70989	70989	48461	48461

Note: Standard errors clustered by each establishment are reported in parentheses.

where z_{it} is the survey response status in the previous survey(s), and the remaining part of the model is identical to equation (1). The parameter δ indicates the serial correlation of the survey response behavior. The first column of Table 2 reports the estimation results for the BSWS, and we found a substantial difference in the response rate depending on the previous survey response. If an establishment responded to the previous survey, the response probability for the next survey is 0.88, but if not, the probability drops to 0.35 (Column 1 of Table 2). This relationship remains robust even after controlling for observed establishment characteristics. Columns 3 and 4 reveal that the response pattern is not explained by a simple Markov process. The variable $NR \to R$ indicates that the establishment did not respond to the survey at $t-2$ but responded at $t-1$, where t is the timings of the survey in which the establishment was sampled. Even among those who responded to the previous survey, the response rate differs by 20 percentage points depending on the response status two surveys ago. Similarly, among those who did not respond to the previous survey, the response status two surveys ago raises the response rate by 30 percentage points. These results indicate that establishments with a low tendency to respond are less likely to respond conditional on the response behavior of the immediate past. Columns 5–8 report the estimation results for the SSPS. The magnitude of the estimates is almost identical to that of the BSWS, suggesting that the underlying non-response structure is common across the surveys. In sum, the survey response behavior of establishments in the current period is heavily dependent on past response behavior.

There are at least two possible explanations for this observed strong serial correlation

in response behavior; namely, state dependence of response behavior and selection of establishments based on unobserved characteristics, and these two explanations have very different consequences in terms of estimating the means of wages or hours worked. Regarding state dependency, such as a Markov process driven by k previous survey responses, the sample still represents the population as far as the sample selection is ignorable conditional on the past responses. Thus, the wage statistics are unbiased as long as the Markov process is mean reverting, for this type of serial correlation only affects the standard error of the estimated wage means. However, in the case of selection on unobserved characteristics, the wage statistics are likely to be biased. For instance, if highly productive establishments are more likely to respond, the average wages estimated from the responding establishments would overestimate the population mean of the wages. Thus, decomposing the serial correlation to the state dependence and the unobserved heterogeneity is critically important in this exercise. However, it is well known in the literature that in fixed effects estimations with a lagged dependent variable, disentangling state dependence conditional on unobserved heterogeneity is impossible without either imposing an assumption on the time series structure of the idiosyncratic determinants of survey response or obtaining a credible instrumental variable that affects past response behavior but not current response behavior (Bond, 2002). As we were neither confident in assuming a specific time series structure of the unobserved determinants of the survey response or had a credible instrumental variable, we pursued a different approach.

As a way to characterize the importance of serial correlation in the survey response behavior, we used lagged worker characteristics as a proxy for unobserved establishment characteristics. Specifically, we assigned each establishment-year record with the average of workers' wages, age, female proportion, and hours worked in the previous survey. We then regressed the survey response status on those worker characteristics and the other control variables used in equation (1).

Table 3 reports the regression results for the current response indicator variable on the past characteristics of establishments. Overall, the results show a statistically significant but not economically meaningful correlation between the lagged establishment characteristics and response probability.[9] For example, in the BSWS, a 20 percent increase in the mean hourly wage increases the response rate only by 1.1 percentage points, which is negligible compared with the overall response rate of 87.5 percent in this sample (Column 1 of Panel A).[10] Similarly, the estimated coefficients for workers' age, female proportion and work hours are minute.

[9]We also estimated the model using earnings instead of the hourly wage and found quantitatively similar results (Table F1).

[10]The response rate in this sample is higher than that of the full sample because the analysis sample here is those establishments that responded to the survey at least once.

Table 3: Response rate and lagged establishment characteristics

Dep.Var. Survey resp.	Panel A: BSWS				
Establishment size	All	500+	100–499	30–99	5–29
	(1)	(2)	(3)	(4)	(5)
L.$ln(wage)$	0.056	0.026	0.063	0.045	0.073
	(0.004)	(0.013)	(0.007)	(0.006)	(0.006)
L.Age	-0.000	-0.001	0.001	-0.001	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
L.Female	0.021	0.003	0.038	0.011	0.025
	(0.006)	(0.026)	(0.014)	(0.011)	(0.010)
L.$ln(WorkHours)$	0.065	0.035	0.085	0.034	0.075
	(0.006)	(0.023)	(0.015)	(0.011)	(0.009)
Mean of Dep.Var.	0.875	0.916	0.910	0.896	0.825
Observations	91282	8369	21560	27532	33821
Dep.Var. Survey resp.	Panel B: SSPS				
Establishment size	All	500+	100–499	30–99	1–29
	(1)	(2)	(3)	(4)	(5)
L.$ln(earnings)$	0.022	0.018	0.047	0.020	0.104
	(0.003)	(0.003)	(0.011)	(0.023)	(0.025)
L.Age	0.001	0.001	0.001	0.001	-0.002
	(0.000)	(0.000)	(0.001)	(0.002)	(0.002)
L.Female	0.021	0.017	0.027	0.057	0.056
	(0.006)	(0.007)	(0.019)	(0.045)	(0.061)
Mean of Dep.Var.	0.898	0.903	0.874	0.859	0.784
Observations	59966	53392	4940	975	653

Note: Standard errors clustered by each establishment are reported in parentheses.

Since the response rate varies substantially by establishment size, it is of particular interest whether the selection on unobservables is substantial among small establishments. However, the results of the subsample analysis, reported in Columns (2)–(5) of Table 3, indicate little systematic heterogeneity. If anything, the selection in terms of the lagged hourly wage is relatively distinct for the smallest establishment group, but the size of its estimate is at most moderate, and the estimation results show that a 10 percent increase in the lagged mean wage increases the unit response rate only by 0.7 percentage points, or 0.9 percent. Panel B of the same Table reports the estimation results of the same analysis using the SSPS 2012–2019, in which we used annual earnings instead of hourly wage because work hours are not surveyed. The SSPS estimates are comparable with those from the BSWS, suggesting again a common underlying selection structure. Overall, as the correlation between the survey response and lagged establishment charac-

teristics is negligible, the results argue in favor of state dependency rather than selection on unobservable.

One caveat to the analysis in this section is that it does not speak to sample selection bias based on unobserved characteristics among establishments that never responded to the survey. In particular, our analysis cannot say anything about establishments that were sampled twice or more but never responded; however, the proportion of those establishments is relatively low, at 12 percent in the BSWS and 7 percent in the SSPS. While it is beyond the scope of this study, administrative data would be helpful in characterizing establishments that have never responded to the survey in order to gauge their potential impact on wage statistics.[11]

4.3 Supplementary establishments in the Basic Survey on Wage Structure

In the BSWS, when a selected establishment does not respond to the survey, an additional establishment is sampled from a pool of *supplementary establishments*. We found that about 6 percent of the establishments that responded to the survey did not appear in the original list of sampled establishments, and so these were presumably selected from the supplementary establishment pool. Supplementary establishments are typically smaller (see Appendix C for more detailed characteristics) and so if only establishments that are likely to respond to the survey are used as a supplementary sample, the establishment distribution in the sample would differ from that of the population. In particular, if the pool tends to have more establishments with high wages, the bias could be serious. In this subsection, we address this situation by comparing the wage payments of establishments listed and not listed in the supplementary establishment pool.

Although any difference in the mean wages of establishments on the original survey list and the supplementary list would bias the estimate of mean wages, if the difference is negligible within a prefecture-firm size-industry cell, using a proper weighting procedure would correct the potential bias. To examine if this is the case, we regressed the individual workers' wage on the indicator of the supplementary establishment, along with the survey year, establishment size, industry and prefecture fixed effects. Table 4 reports the estimation results, and Column 1 shows that the average wage of supplementary establishments is about 13% lower than that of listed establishments when we do not condition on indus-

[11] In addition, since we use lagged variables in the analysis, our sample did not include establishments that were selected only once regardless of their unit response status. Due to random sampling, however, in principle, excluding those establishments from the analysis should not affect the results, given an appropriate sampling weight. This point was indeed confirmed by our sub-sample analysis by establishment size: Although small establishments are more likely to be sampled only once, the estimation result was comparable with that of larger establishments.

Table 4: Regression analysis: Listed vs unlisted establishments (BSWS 2012–2017)

Dep.Var. ln(*earnings*)	(1)	(2)	(3)	(4)
Supplementary	-0.128	-0.042	-0.014	-0.011
	(0.011)	(0.009)	(0.008)	(0.009)
Year	X	X	X	X
Industry		X	X	X
Establishment size			X	X
Prefecture				X
Observations	7833796	7833796	7833796	7833796

Note: Standard errors clustered by each establishment are reported in parentheses.

try, establishment size and prefecture fixed effects. However, this difference is reduced to only 1.1% and becomes statistically insignificant after conditioning on these fixed effects, as reported in Column 4. Note that, given the sample size, this lack of significance is not due to an imprecise estimate, and indeed the 95 percent confidence interval ranges from -2.8 percent to 0.7 percent, suggesting an economically small difference. Therefore, even if supplementary establishments are used, the resulting bias would not be serious after appropriately adjusting the distribution for the observed establishment characteristics.

5 Sampling of workers within establishments

The peculiar feature of the BSWS and SSPS is that both surveys delegate the sampling of workers within the establishments to the survey respondents. Although both surveys provide instructions on how to conduct the random sampling, it is not inconceivable that an inexpert respondent might implement the random sampling inappropriately. In this section, we investigate whether the random sampling was conducted properly by comparing the mean wages calculated based on establishment-level aggregate data and individual worker level data. If the random sampling was properly implemented, these two statistics should match up.

5.1 The Basic Survey on Wage Structure

We first examine the random sampling of workers for the BSWS. Since the BSWS establishment survey does not ask for the aggregate wage bill and total hours worked, we cannot directly calculate the average wages based on establishment level aggregate statistics. To overcome this data limitation, we first matched the establishment in the BSWS with the same establishment in the Monthly Labor Survey (MLS, *Maitsuki Kinro Toukei*). As

Table 5: Sampling of workers within an establishment: BSWS vs MLS

	BSWS			MLS	Diff.	
	All	500+	Matched		(3) − (4)	(3) − (2)
	(1)	(2)	(3)	(4)	(5)	(6)
Number of employees	33.94	962.34	1147.16	1153.22	-6.06	184.82
	[109.80]	[1101.64]	[1209.31]	[1167.76]	(16.65)	(21.19)
Work days	20.00	20.05	20.13	19.59	0.54	0.08
	[5.25]	[3.58]	[3.43]	[2.22]	(0.06)	(0.04)
Predetermined work hours	140.14	149.46	151.10	147.26	3.84	1.64
	[50.27]	[33.02]	[31.50]	[20.52]	(0.48)	(0.40)
Overtime hours	7.55	12.88	14.66	15.57	-0.91	1.78
	[15.57]	[17.95]	[18.34]	[9.81]	(0.18)	(0.18)
Total work hours	147.70	162.34	165.76	162.83	2.93	3.42
	[56.46]	[40.56]	[39.57]	[25.05]	(0.53)	(0.48)
Salary paid on a fixed basis	22.89	36.65	36.84	36.63	0.21	0.19
	[14.74]	[21.17]	[19.60]	[10.26]	(0.13)	(0.23)
Bonus	3.57	9.93	10.80	11.41	-0.61	0.87
	[5.91]	[10.74]	[9.94]	[6.24]	(0.07)	(0.14)
Total salary	26.47	46.58	47.64	48.04	-0.40	1.06
	[18.84]	[29.16]	[27.36]	[15.43]	(0.18)	(0.34)
Observations	114410	5037	2193			

Note: Unit of earnings is 10 thousand JPY/month. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

explained in the data section, the MLS is a monthly establishment survey asking about the monthly aggregate wage bill as well as aggregate hours worked. Since the BSWS and the MLS can be matched at the establishment level by using a common establishment identifier, we were able to check whether the average wage payment per worker calculated from each of these two surveys was consistent. The MLS data includes the total number of full-time workers, aggregate actual working hours, and aggregate paid salaries, and from this information, we calculated the hours worked per worker by dividing the total number of hours worked by the number of permanent workers at each establishment. Similarly, we calculated the salary per worker by dividing the total salaries paid by the number of permanent workers. If the workers surveyed individually in the BSWS are randomly sampled within an establishment, then the hours worked and salary per full-time worker calculated from the individual surveys should, on average, match those calculated from the MLS. Therefore, by calculating the difference between the values obtained from the BSWS and the MLS, it is possible to infer the randomness of the sampling of workers and, if not random, which type of workers are more likely to be sampled.

The process of constructing the analysis sample is shown in Table 5. Column 1 reports the statistics of all observations in the BSWS. We then first restrict our sample to establishments with 500 or more workers from the 2016–2017 BSWS (Column 2), because

the MLS is in principle a survey of all establishments with 500 or more employees but this rule does not apply to smaller establishments. Compared with the grand mean reported in Column 1, employees working for larger establishments tend to work longer and receive higher wages and bonuses, but limiting the establishment size does not essentially alter our conclusion because we are interested in the random sampling of workers within an establishment. Next, we matched the 5037 establishments in the BSWS from 2016–2017 with those in the MLS from 2016–2017 using the common establishment identifier. Although all establishments in the BSWS are supposed to appear in the MLS as well, the number of establishments that were successfully matched in this way was only 2193, or 44 percent of the 5037 establishments in the BSWS.[12] See Appendix B for characteristics of BSWS establishments that we were not able to match with the MLS.

Next, we compare the statistics of the BSWS establishments matched with the MLS (Column 3) and those of the MLS (Column 4). Since these statistics are from the same (matched) establishments, the means should be identical,[13] and, indeed, there is no statistically or economically significant difference in the number of workers. As this variable is establishment-level information (and not relying on the sampling of workers), it underpins our confidence that the two data sets are matched correctly.

Comparing other statistics, we note that the number of working days and working hours is about 2-3 percent higher for the BSWS and there is no difference in regular salaries, but bonuses are slightly higher in the MLS. This leads to total salaries, defined as the sum of the salary paid on a fixed basis and 1/12 of the annual bonus, being about 4,000 JPY higher in the MLS, which is statistically significant but less than 1 percent of the average value. In the end, we conclude that the average hours worked and wages calculated based on the randomly selected individual employees in the BSWS and the aggregate statistics in the MLS are not substantially different. This suggests that the random sampling of employees delegated to the employer was conducted properly, at least to the extent that the mean wages are unbiased estimates.[14]

[12] One partial explanation for the low match rate is the MHLW non-compliance with the MLS sampling design, which targets all establishments that hire 500 or more employees. However, in Tokyo prefecture, for example, only one third were surveyed.

[13] Since the BSWS asks the amount of bonus paid in the last year and since there is substantial seasonality in the bonus payment, we calculated the bonus amount in the MLS by summing up the bonus payment from January to December in the previous year. Due to this procedure, the sample of the MLS was restricted to establishments that had completed the survey for consecutive 12 months. This sample restriction does not substantially change the sample construction because the attrition rate is relatively low at about 15 percent. Furthermore, our estimation results other than the bonus payment are robust to this sample restriction (Table F2).

[14] Additionally, when we investigated the sampling of workers by gender, we similarly found a bias in earnings across gender, with the magnitude statistically significant but not economically meaningful. Specifically, the earnings of male (female) sampled workers in the BSWS tend to be lower (higher) than those in the MLS. See Appendix D for details.

5.2 The Statistical Survey of Actual Status for Salary in the Private Sector

We now move on to examining the random sampling of employees within an establishment based on SSPS. As the SSPS requires establishments (withholding agents) to report the number of salaried workers and the wage bill (the total amount of salaries paid), we can directly calculate the mean annual earnings per worker without matching to other statistics. At the same time, we can calculate the corresponding figure using the sampled employees. If the sampling of employees is random, as instructed, then the mean annual salary based on the establishment form (the withholding agent form) should coincide with that of the worker form (the payroll income form).

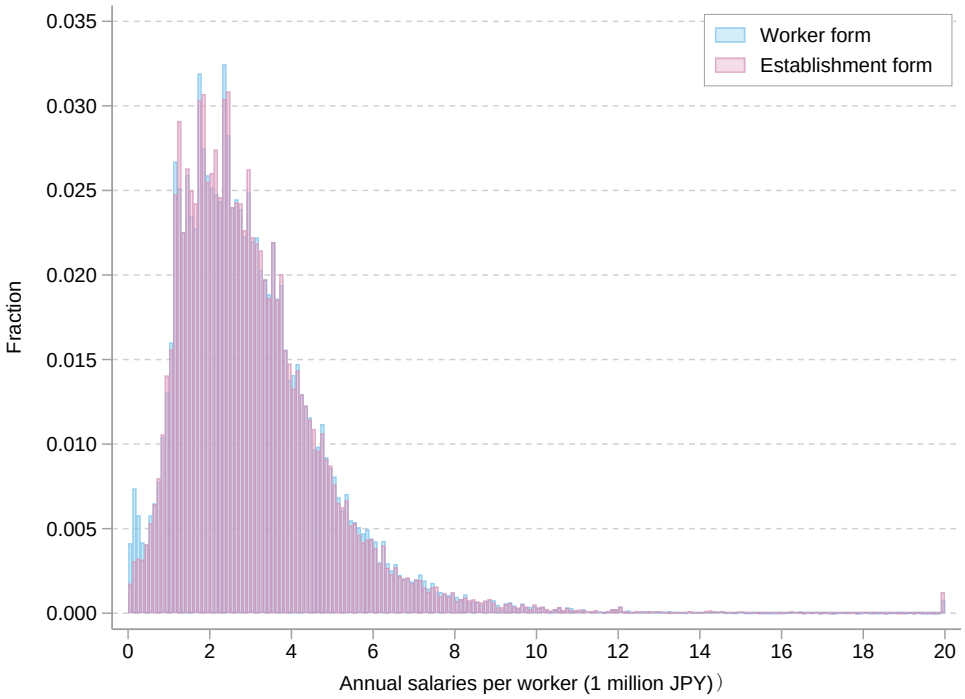


Figure 4: SSPS earnings distribution from 2012–2019 (1 million JPY/year)

Note: For visualization purposes, the amount of earnings is capped at 20 million JPY in this figure.

As a first cut, Figure 4 plots the distribution of annual earnings from each questionnaire, pooling data from 2012 to 2019. Here, the values based on the establishment questionnaire are calculated by dividing the total payroll by the average number of employees in March, June, September, and December. As the two distributions overlap, this provides reassurance that the random sampling of employees within an establishment conducted by employers is implemented properly. There are, however, several notable irregularities. First, the figure shows that a non-negligible proportion of workers earn less

Table 6: Sampling of workers within an establishment: SSPS

	All			Non outliers		
Questionnaire	Worker	Establishment	Diff. (1) − (2)	Worker	Establishment	Diff. (4) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	3.067	3.220	-0.153	3.039	3.030	0.009
	[2.000]	[19.988]	(0.090)	[1.897]	[1.866]	(0.002)
p10	1.181	1.201	-0.020	1.188	1.201	-0.013
			(0.003)			(0.004)
p25	1.775	1.787	-0.011	1.777	1.781	-0.004
			(0.004)			(0.004)
p50	2.703	2.697	0.006	2.692	2.681	0.011
			(0.004)			(0.004)
p75	3.903	3.872	0.030	3.870	3.840	0.030
			(0.006)			(0.005)
p90	5.316	5.253	0.063	5.251	5.169	0.081
			(0.010)			(0.009)
Observations	163993	163993	163993	157661	157661	157661

Note: Unit of earnings is 1 million JPY/year. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses. The standard error of each percentile difference was obtained via establishment-level clustering bootstrap with 1000 replication. In the non-outlier sample, we excluded the top and bottom 1 percent of the distribution of a difference between the worker questionnaire and establishment questionnaire.

than 50 thousand JPY in the distribution calculated from the worker questionnaire. This might have occurred because salaried employees who worked only a few months of the year were sampled, and annual earnings of them are lower reflecting shorter annual hours worked. Second, the distribution calculated from the data on the establishment form shows a slightly larger number of people earning 20 million JPY or more. The data from the establishment form sometimes show extremely high values, and one reason for this may be that the total payroll is divided by the "average" number of employees measured quarterly in order to calculate the payroll per employee. In particular, the measurement error in the number of employees would not be negligible in establishments whose labor demand has substantial seasonality, and in some cases, such measurement error leads to artificially high earnings per worker. With these caveats, the distributions of annual earnings based on the establishment survey and the worker survey are reasonably similar.

We now compare the distribution of mean annual wages calculated based on establishment level aggregates and randomly selected individual payroll records focusing on the means and other percentiles (Table 6). We see that the mean calculated from the worker form is 150,000 JPY lower than the mean from the establishment form. Looking at the percentiles, the low percentile value is lower and the high percentile value is higher in the worker form, which indicates that the establishment level mean annual earnings based on

the worker form has a wider distribution than the one based on the establishment form. However, compared to the difference in means, the differences in percentile values are limited. The reason for the large difference in the mean but the small difference in the distribution is that, as mentioned above, the earnings per worker calculated from the establishment questionnaire may be affected by a small number of outliers. In fact, we found that only a few records have a substantial impact. When the top and bottom 1 percent of records were trimmed, the difference in the mean shrank substantially (Columns 4-6 of Table 6). Although the value from the worker questionnaire is statistically significantly higher, the difference is negligible at 9 thousand JPY, or 0.3 percent of the mean. As the percentile values are not affected by outliers, the results are almost identical, irrespective of outlier treatment.

All in all, both the BSWS and SSPS appear to perform the random sampling of workers adequately. While we found some statistically significant biases in the sampling of workers, the magnitude is not economically meaningful. Thus, the impact of the non-random sampling of workers within an establishment seems very limited.

6 Uncovered population

The analysis thus far has assessed the extent to which the random sampling of the BSWS and SSPS has been implemented properly. We now move on to the coverage of the two surveys, focusing on the fact that neither of which cover freelance workers. Given the heightened attention to the gig economy in general and freelance workers for such services as Uber Eats[15], we attempt to document them using a household survey that covers all workers regardless of the form of employment.

6.1 Freelance workers

Since the BSWS and the SSPS do not cover freelance workers, we also evaluate the impact of this limitation in coverage on wage statistics. To that end, we draw on the Employment Status Survey (ESS, *Shugyo Kouzou Kihon Chosa*) and regard the self-employed without employees as freelance workers in this analysis.[16]

Although, as a long-term trend, the proportion of freelance workers is decreasing, from 13 percent of the working population in 1977 to 6 percent in 2017 (Figure 5), the share

[15]Uber's ride-hailing service is heavily regulated and is not prevalent in Japan.

[16]We note that the definition of freelance workers might vary. For example, the Cabinet Office excludes individual shopkeepers with stores, and agricultural, forestry, and fishery workers from its definition of freelance workers. However, we emphasize that the objective of this analysis is not to reveal the characteristics of freelance workers defined in a certain way but to reveal those of workers not covered by the main Japanese wage statistics, and freelance workers as defined here correspond to that population.

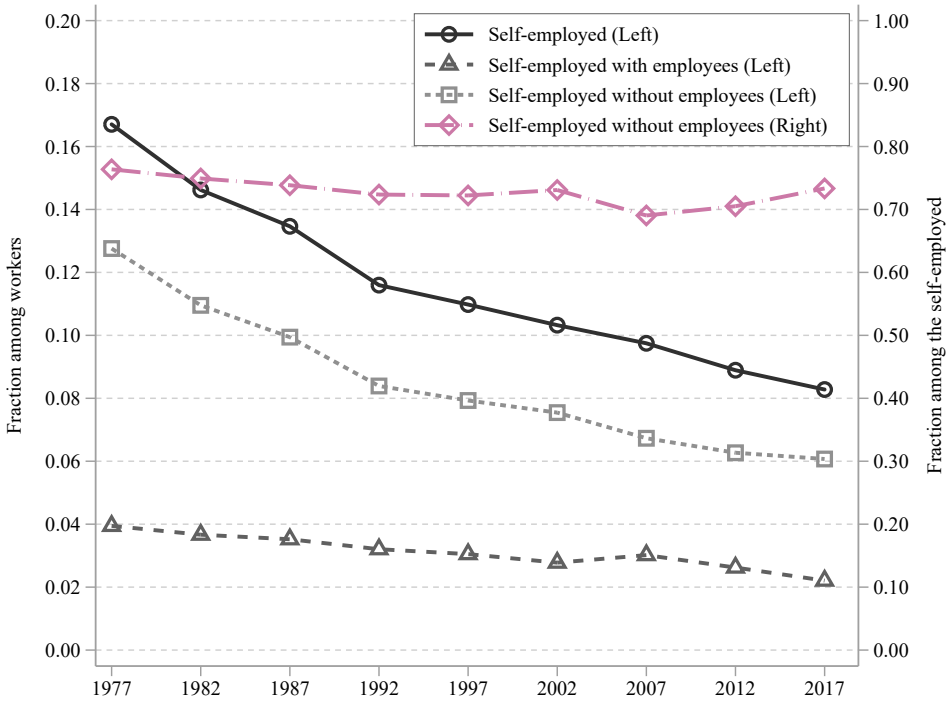


Figure 5: Proportion and composition of self-employed workers

Source: The ESS 1977–2017.

of freelance workers among the self-employed is high at around 70 percent and has been increasing in recent years. According to the 2017 ESS, freelance workers are older, with a higher percentage of males, higher rate of marriage, and lower education, and they chose the current job in order to utilize their knowledge and skills (Tables E1 and E2). Table 7 shows that the earnings of freelance workers tend to be substantially lower than other workers, with 24 percent of freelance workers earning less than 0.5 million JPY/year but only a small fraction (4.7 percent) of other workers earning that amount. Furthermore, the income distribution of freelance workers is stochastically dominated by that of non-freelance workers, with the mean annual income, calculated using the mid-value of each income category, 2.1 million JPY for freelance workers and 3.5 million JPY for other workers. As a result, the Japanese mean income statistics are upwardly biased by 2–3 percent when freelance workers are not counted.

6.2 Small establishments with 1–4 employees

We now move on to the limited coverage of small establishments in the BSWS. This is acknowledged by policy makers and researchers as a limitation of the survey, who find the SSPS more suitable for characterizing the earnings of workers of micro establishments

Table 7: Income distribution (ESS 2017)

	All	Non freelance	Employees	Self-employed	
				non freelance	freelance
	(1)	(2)	(3)	(4)	(5)
Composition	1.000	0.939	0.917	0.022	0.061
Income distribution					
0–0.5 million JPY	0.059	0.047	0.046	0.083	0.244
0.5–0.99 million JPY	0.114	0.112	0.113	0.066	0.147
1–1.49 million JPY	0.104	0.102	0.103	0.084	0.127
1.5–1.99 million JPY	0.075	0.074	0.074	0.080	0.087
2–2.49 million JPY	0.107	0.108	0.108	0.102	0.092
2.5–2.99 million JPY	0.080	0.081	0.081	0.078	0.067
3–3.99 million JPY	0.139	0.141	0.142	0.133	0.097
4–4.99 million JPY	0.100	0.103	0.103	0.092	0.052
5–5.99 million JPY	0.070	0.073	0.073	0.069	0.029
6–6.99 million JPY	0.048	0.050	0.050	0.039	0.017
7–7.99 million JPY	0.035	0.037	0.037	0.027	0.012
8–8.99 million JPY	0.022	0.023	0.023	0.023	0.007
9–9.99 million JPY	0.014	0.014	0.014	0.020	0.005
10–12.49 million JPY	0.020	0.020	0.020	0.039	0.007
12.5–14.99 million JPY	0.005	0.006	0.005	0.015	0.002
15 million JPY or more	0.009	0.009	0.008	0.051	0.004
Mean income (1 million JPY)	3.456	3.545	3.528	4.255	2.097

Note: The mean annual income was calculated using the mid-value of each income category. A *freelance worker* is defined as a self-employed worker without employees.

because it covers the population of withholding agents with one or more employees. Another, perhaps less known, limitation in the coverage of the BSWS is the exclusion of corporate executives. By contrast, the SSPS not only covers corporate executives but also instructs establishments to include every worker (i.e. executive) who earns 20 million yen (about USD200,000) or more annually. Thus, overall, the SSPS covers a wider range of workers than the BSWS. We next use the SSPS to describe the characteristics of workers not represented by the BSWS.

Here, we characterize employees of micro scale firms. First, withholding agents with 1–4 employees account for 83 percent of all withholding agents, but their employee share is only 19 percent. The the characteristics of workers in those micro establishments are distinct from those of other workers (Table 8). In particular, workers in micro-establishments are more likely to be female (4.2 percentage points higher), older (by 7.5 years), and with longer tenure (3.7 years). However, despite their higher age and longer tenure, their earn-

Table 8: Characteristics of workers in establishments with 4 or less employees (SSPS)

Establishment size	1–4 (1)	5 or more (2)	Diff. (1) − (2) (3)
Female	0.470	0.427	0.042
	[0.499]	[0.495]	(0.003)
Age	50.902	43.439	7.463
	[15.201]	[13.815]	(0.076)
Tenure	13.212	9.524	3.688
	[13.175]	[10.295]	(0.077)
Salary excl. bonus	2.970	3.279	-0.309
	[3.033]	[3.200]	(0.019)
Bonus	0.202	0.655	-0.453
	[0.541]	[1.142]	(0.008)
Total salary	3.172	3.933	-0.762
	[3.185]	[3.859]	(0.026)
Observations	163051	2239078	

Note: Unit of earnings is 1 million JPY/year. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

ings are lower than workers in establishment with 5 or more employees, a difference of 9.5 percent in regular salary, 69.2 percent in bonuses, and 19.4 percent in total earnings. When controlling for worker attributes (gender, age, and tenure), the difference in earnings expands to 17.7 percent and 25.1 percent in terms of regular salary and total earnings, respectively (Table 9).

The above analysis suggests that the BSWS, which is often used to describe the wage distribution of Japan, fails to adequately capture workers with low earnings. According to the figures from the SSPS, the BSWS does not cover 19 percent of employees whose earnings are, on average, 19 percent lower than the BSWS sample. In the end, the bias associated with this uncovered population amounts to about 4 percent. This undercoverage of low wage workers is, of course, of particular importance in the assessment of policies such as the minimum wage that targets low wage earners. Neglecting these low wage earners underestimates the proportion of workers who are affected by a minimum wage hike.

7 Application: The Evolution of Wage Inequality

As an example of how this difference in the coverage of the BSWS and SSPS may affect real-world policy-making, we investigated its impact on the depiction of the evolution of

Table 9: Regression analysis: Earnings in establishment with 4 or less employees

Dep.Var.	ln (Salary excl. bonus)		ln (Total salary)	
	(1)	(2)	(3)	(4)
Establishment size ≤ 4	-0.110	-0.177	-0.187	-0.251
	(0.008)	(0.005)	(0.009)	(0.005)
Controls		X		X
Observations	2402114	2402114	2402114	2402114

Note: Standard errors clustered by each establishment are reported in parentheses. We excluded 15 records whose earnings are 0. Covariates includes workers' gender, age and tenure.

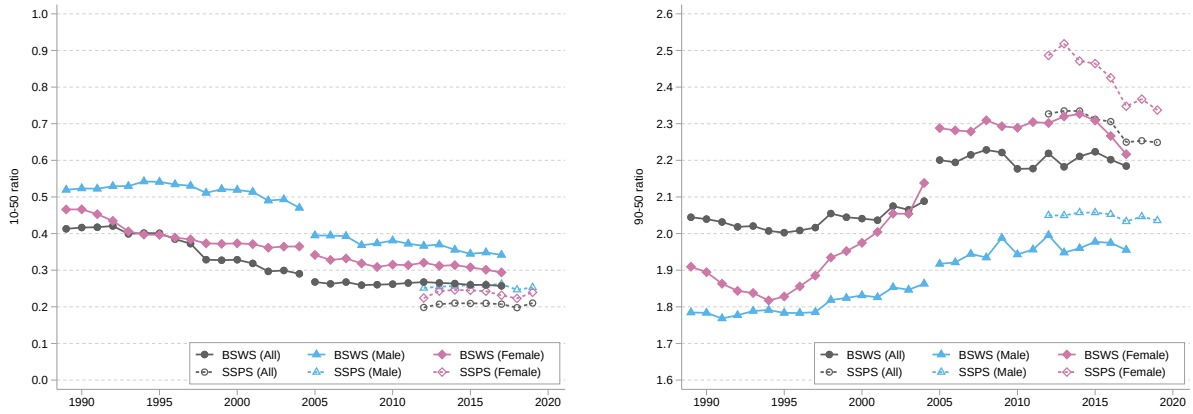
wage inequality. We first examine the evolution of the lower end of the income distribution, and Figure 6a shows the 10-50 percentile ratio using the BSWS and the SSPS. According to the long-run trend from the BSWS, the ratio decreased in the late 1990s and early 2000s, particularly among females, but has been stable in recent years. Since the SSPS covers small establishments not covered by the BSWS and since small establishments tend to pay less, the gap becomes larger when using the SSPS, but the time-series trends are similar to the BSWS.

We now move our focus from the lower tail to the upper tail of the wage distribution. While the BSWS covers only workers and does not include corporate executives,[17] the SSPS covers all salaried workers, including corporate executives. Furthermore, the sampled establishments are supposed to report salary information for all employees whose annual salary exceeds 20 million JPY. As a result, the SSPS captures the right tail of the income distribution more accurately. In this section, we investigate to what extent inequality measures differ between the BSWS and the SSPS.

Figure 6b shows the 90-50 percentile ratio of earnings distribution. According to the BSWS, the 90-50 ratio has been relatively stable over the last few decades, though the ratio made a discrete jump in 2005 due to a change in the sampling design,[18] with the 90 percentile about twice as large as the median. However, the ratio increased from the 1990s and early 2000s within each gender, but particularly striking among females, before becoming relatively stable after 2005 for both genders. Since the SSPS has better coverage of the upper tail of the income distribution, the 90-50 percentile ratio is larger than that of the BSWS, but the difference is at most less than 10 percent. Furthermore, in terms of the yearly trend, the BSWS appears to track the SPSS trend well. The small difference

[17] In the BSWS, those who are paid salaries based on the same standards as ordinary employees are covered by the survey, even if they are executives. However, even in this case, the definition of earnings in the survey excludes executive compensation.

[18] See Shinozaki (2008) for a discussion of this change.



(a) 10-50 percentile ratio

(b) 90-50 percentile ratio

Figure 6: Wage Inequality

Note: The time series of the BSWS is not comparable before and after 2005 due to a change in the sampling design.

in the 90-50 percent ratio between these two surveys is probably because the uncovered population in the BSWS; that is, executives, lie at the very top of the income distribution and so they have limited impact on the 90th percentile.

In order to focus on the very top end of the income distribution, we next calculate the top 1 percent and 0.1 percent earnings shares (Figure 7). The BSWS shows relatively stable trends, though the top 1 percent share slightly increased during the 2000s. In contrast to the BSWS, the SSPS indicates that the top income share increased during the 2010s, with the top 1 percent share growing from 5.8 percent to 6.3 percent, and the top 0.1 percent share growing from from 1.2 percent to 1.6 percent. Moriguchi and Saez (2008) and Alvaredo et al. (2012), who estimate the right tail of the income distribution by fitting the Pareto distribution to the binned tax data and the estimated Pareto coefficient, calculate the top income share in Japan between 1886 and 2010. According to their estimation results, the top 1 percent share was 9.5 percent in 2010 and the top 0.1 percent share was 2.5 percent. These estimates are larger than our estimates, and we suspect that the difference is due to the definition of income, for they include non-labor income such as rents, interest and dividends while our income is limited to labor income.[19]

As performed in Moriguchi and Saez (2008) and Alvaredo et al. (2012), another method to illustrate the top income distribution is to estimate the Pareto distribution:

$$F(y) = 1 - \left(\frac{y_m}{y}\right) \quad (y \geq y_m), \tag{2}$$

[19]Moriguchi and Saez (2008) disentangle total income into these detailed income categories, and they find the top 1 percent share of employment income to be about 7 percent in 2005, which is more in line with our estimates.

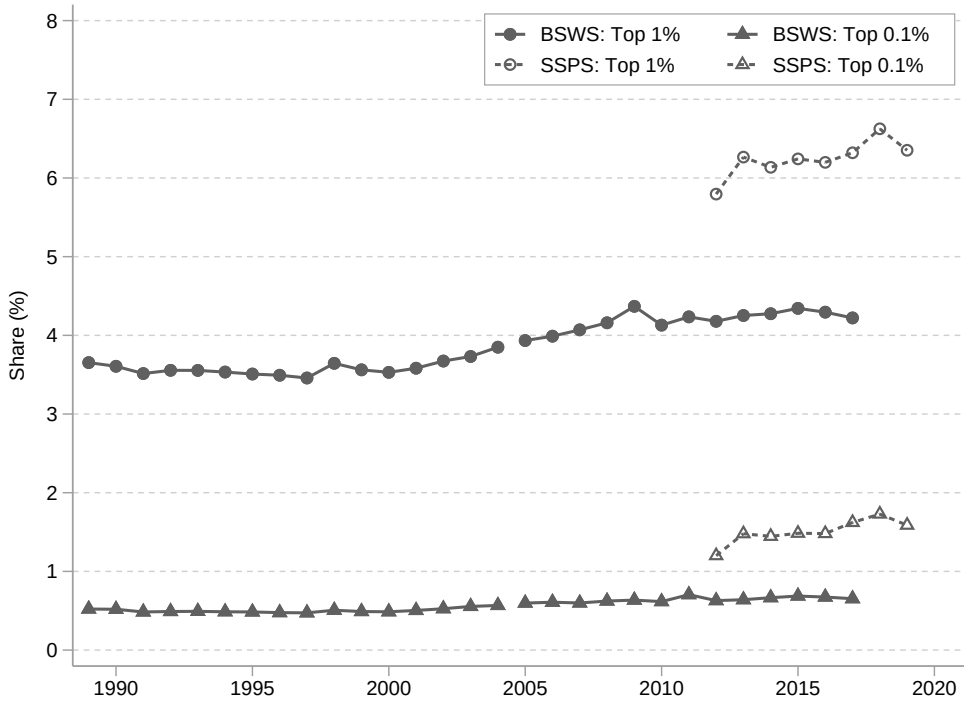


Figure 7: Top income share

Note: The time series of the BSWS is not comparable before and after 2005 due to a change in the sampling design.

where y_m is some threshold parameter and α is a parameter that characterizes the tail of income distribution; the smaller the value of α, the fatter the right tail. In other words, a small value of α indicates substantial inequality among high-income individuals. In particular, when $\alpha \leq 1$, the distribution does not have a mean, and when $\alpha \leq 2$, the distribution does not have a variance. Following Atkinson et al. (2018), equation (2) is rewritten as

$$\ln y = C + \frac{1}{\alpha} \ln \frac{1}{1 - F(y)}, \tag{3}$$

where C is some constant consisting of α and y_m. We estimated this equation using those whose annual income was 9 million JPY or more, which corresponds to the 95 percentile of the income distribution of both the BSWS and SSPS. Figure 8 shows that the Pareto coefficient obtained from the BSWS is larger than that of the SSPS, which implies that the SSPS captures upper tail inequality better than the BSWS. In addition, the SSPS estimates have smaller standard errors, indicating that the SSPS estimate is more accurate. On the other hand, the difference between men and women is limited in both surveys.[20]

[20]The World Inequality Database provides the Pareto coefficient estimated from the binned income tax data, and its estimate for the 2010s is 2, which is smaller than ours; that is, more income inequality. As

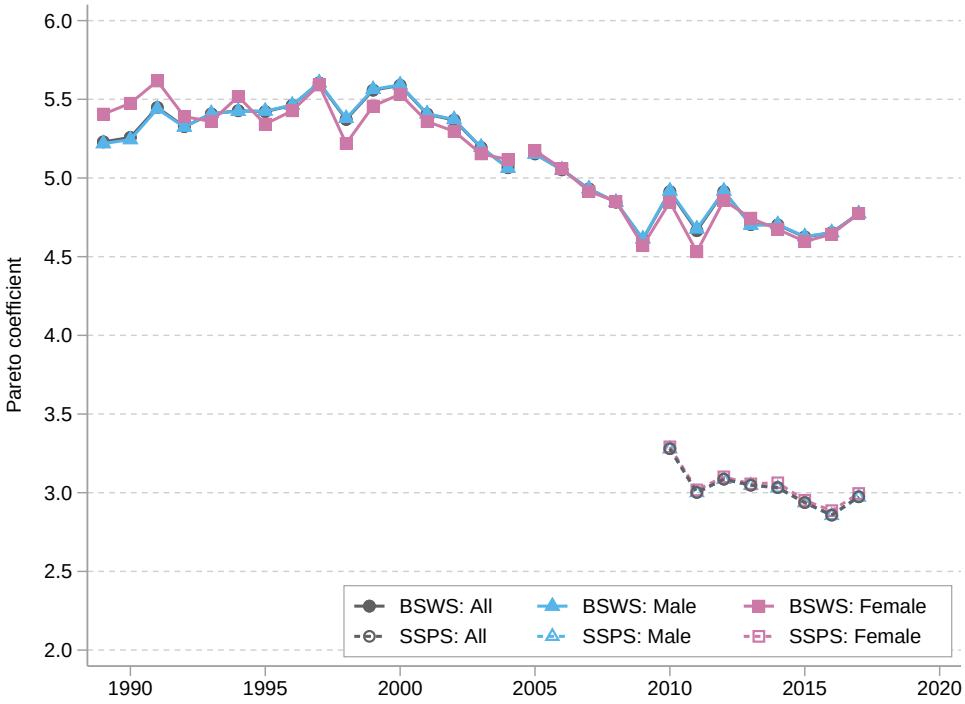


Figure 8: Pareto coefficient

Note: This figure shows the Pareto coefficient estimated using equation (3). The threshold is set to 9 million JPY/year. The time series of the BSWS is not comparable before and after 2005 due to a change in the sampling design.

To summarize, the limited coverage of the BSWS leads to an under-estimation of wage inequality but yet it still picks up the overall trend, so the BSWS seems still useful as long as one is not interested in the absolute level but the trend in income inequality. However, the bias of the BSWS could be serious when the very top of the income distribution and the very bottom of it are concerned.

8 Conclusion

We assess the non-sampling error caused by the unit non-response of two representative large-scale national wage statistics. We find evidence of systematic heterogeneity in the unit non-response rates across industries, establishment sizes and regions. The evidence of the sample selection based on observed characteristics suggests the importance of reflecting the response rate heterogeneity in the calculation of the weight. One option is to use the raking method to iteratively calculate the sampling weight using the population

discussed above, this seems reasonable because the World Inequality Database includes capital income while we do not.

distribution of workers by industry, establishment size and prefecture available from the annual business frame.

In contrast to the sample selection bias by observed characteristics, we did not find clear evidence of sample selection bias based on unobserved characteristics from our analysis of short panel data constructed from the wage statistics at the establishment level. In particular, we found that the mean of current wages does not depend on the past average of wages in any quantitatively significant way. We must acknowledge, however, that this suggestive evidence for the absence of sample selection bias due to unobserved characteristics is not definitive because the panel units are biased toward large establishments. To fully characterize the selection on unobserved characteristics, we need to consider the assessment of sample selection bias based on a credible instrumental variable that affects response behavior but not wages. Randomization of the intensity of follow up among non respondents to calibrate the degree of non-response bias as suggested by Dinardo et al. (2021) would be a promising avenue.

Our findings also have implications for the measurement of wage inequality based on the Basic Survey of Wage Structure (BSWS) that has been widely used in the literature (Shinozaki, 2006; Kambayashi et al., 2008). We have shown that the BSWS under-samples both the lower and upper tails of the wage distribution compared to the Statistical Survey of Actual Status for Salary in the Private Sector (SSPS). This systematic sampling bias implies that wage inequality is underestimated when using the BSWS, and this underestimation is particularly substantial when estimating the top of the income distribution. Despite the SSPS being less popular than the BSWS among researchers, our analysis demonstrates that the SSPS better serves the purpose of examining the tails of the wage distribution than the BSWS. This finding also suggests that administrative tax records are indispensable for examining the behavior of the tails of the wage distribution.

References

Alvaredo, Facundo, Chiaki Moriguchi, and Emmanuel Saez, "Japan Estimates of Top Income Shares: Update up to 2010," *WID.world Technical Note Series*, 2012.

Atkinson, Anthony B., Alessandra Casarico, and Sarah Voitchovsky, "Top incomes and the gender divide," *The Journal of Economic Inequality*, 2018, *16*, 225–256.

Barnes, William, Geoff Bright, and Colin Hewat, "Making sense of Labour Force Survey response rates," *Economic and Labour Market Review*, 2008, *2* (12), 32–42.

Bollinger, Christopher R. and Barry T. Hirsch, "Match bias from earnings imputation in the current population survey: The case of imperfect matching," *Journal of Labor Economics*, 2006, *24* (3), 483–519.

_ , _ , Charles M. Hokayem, and James P. Ziliak, "Trouble in the tails? What we know about earnings nonresponse 30 years after Lillard, Smith, and Welch," *Journal of Political Economy*, 2019, *127* (5), 2143–2185.

Bond, Stephen R, "Dynamic panel data models: a guide to micro data methods and practice," *Portuguese Economic Journal*, 2002, *1* (2), 141–162.

Dinardo, John, Jordan Matsudaira, Justin McCrary, and Lisa Sanbonmatsu, "A practical proactive proposal for dealing with attrition: Alternative approaches and an empirical example," *Journal of Labor Economics*, 2021, *39* (S2), S507–S541.

European Commission, *Quality report of the European Union Labour Force Survey 2018* 2020.

Higo, Masahiro, Kiyohiko G. Nishimura, and Yukie Sakuragawa, "Reforms and crises in government statistics: The case of Japan," *Asian Economic Papers*, Summer 2020, *19* (2), 21–37.

Hirsch, Barry T. and Edward J. Schumacher, "Match bias in wage gap estimates due to earnings imputation," *Journal of Labor Economics*, 2004, *22* (3), 689–722.

Hoshi, Takeo and Anil K Kashyap, "The Great Disconnect: The Decoupling of Wage and Price Inflation in Japan," Technical Report 2020.

Kambayashi, Ryo, Daiji Kawaguchi, and Izumi Yokoyama, "Wage distribution in Japan, 1989–2003," *Canadian Journal of Economics/Revue canadienne d'économique*, 2008, *41* (4), 1329–1350.

Kawaguchi, Daiji, "Wage," *The Japanese Journal of Labour Studies*, 2013, *55* (4), 14–17.

_ and Yuko Mori, "Estimating the effects of the minimum wage using the introduction of indexation," *Journal of Economic Behavior & Organization*, 2021, *184*, 388–408.

Kline, Patrick and Andres Santos, "Sensitivity to missing data assumptions: Theory and an evaluation of the U.S. wage structure," *Quantitative Economics*, 2013, *4* (2), 231–267.

Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion, "Survey nonresponse and the distribution of income," *The Journal of Economic Inequality*, April 2006, *4* (1), 33–55.

Krueger, Alan B., Alexandre Mas, and Xiaotong Niu, "The evolution of rotation group bias: Will the real unemployment rate please stand up?," *The Review of Economics and Statistics*, May 2017, *99* (2), 258–264.

Lillard, Lee, James P Smith, and Finis Welch, "What do we really know about wages? The importance of nonreporting and Census imputation," *Journal of Political Economy*, 1986, *94* (3), 489–506.

Manski, Charles F., "Credible interval estimates for official statistics with survey nonresponse," *Journal of Econometrics*, 2016, *191* (2), 293–301.

Moriguchi, Chiaki and Emmanuel Saez, "The evolution of income concentration in Japan, 1886–2005: Evidence from income tax statistics," *The Review of Economics and Statistics*, 2008, *90* (4), 713–734.

Rubin, Donald B., "Inference and missing data," *Biometrika*, 1976, *63* (3), 581–592.

Shinozaki, Takehisa, "Wage inequality in Japan, 1979–2005," *Japan Labor Review*, 2006, *3* (4), 4–22.

_ , "Changes in survey methodology and wage disparity in the Basic Survey on Wage Structure (*Chingin Kouzou Kihon Toukei Chosa no Chosa Houhou Henkou to Chingin Kakusa no Suii*)," *Humanities and Social Sciences Research (Jinbun Shakai Kagaku Kenkyu)*, 2008, *48*, 131–144. *(in Japanese).*

Vella, Francis, "Estimating models with sample selection bias : A survey," *Journal of Human Resources*, 1998, *33* (1), 127–169.

A Survey response rate of the BSWS and SSPS by establishment characteristics

This appendix section reports the unit response rates of the BSWS and SSPS.

Table A1: Survey response rate of the BSWS by establishment size

Year	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)
All	0.69	0.69	0.70	0.71	0.70	0.69
15000+	1.00	1.00	1.00	1.00	1.00	1.00
5000-14999	0.92	0.81	0.84	0.82	0.94	0.86
1000-4999	0.88	0.85	0.87	0.85	0.84	0.81
500-999	0.87	0.84	0.84	0.84	0.83	0.81
300-499	0.85	0.82	0.81	0.79	0.83	0.83
100-299	0.83	0.81	0.82	0.82	0.80	0.80
50-99	0.79	0.79	0.80	0.81	0.79	0.77
30-49	0.77	0.76	0.77	0.77	0.77	0.75
10-29	0.67	0.67	0.68	0.69	0.69	0.67
5-9	0.45	0.48	0.49	0.50	0.51	0.48

sectionAnalysis using panel structure: Supplementary tables

The rows of Tables A1 and A2 show the number of times a specific establishment was selected in the sample and the columns show the number of responses. The decimals in each cell of the table show the distribution of the number of responses conditional on the times selected in the sample, and the sum of each row is one.

Table A2: Survey response rate of the BSWS by industry

Year	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)
All	0.69	0.69	0.70	0.71	0.70	0.69
Mining, quarrying of stone, gravel	0.70	0.76	0.74	0.74	0.72	0.71
Construction	0.71	0.71	0.72	0.75	0.74	0.71
Manufacturing	0.76	0.75	0.76	0.76	0.75	0.74
Electricity, gas, heat supply, water	0.87	0.85	0.90	0.88	0.87	0.87
Information, communications	0.69	0.70	0.69	0.70	0.70	0.68
Transport, postal services	0.73	0.75	0.73	0.75	0.72	0.71
Wholesale, retail trade	0.69	0.69	0.72	0.71	0.68	0.67
Finance, insurance	0.80	0.81	0.82	0.83	0.83	0.81
Real estate, goods rental, leasing	0.61	0.62	0.62	0.63	0.66	0.67
Scientific research, prof. and tech. services	0.70	0.72	0.74	0.73	0.72	0.72
Accomodations, eating, drinking services	0.59	0.57	0.58	0.59	0.54	0.50
Living-related, personal, amusement services	0.56	0.57	0.59	0.59	0.58	0.56
Education, learning support	0.61	0.63	0.62	0.67	0.70	0.69
Medical, health, welfare	0.80	0.78	0.81	0.81	0.79	0.75
Compound services	0.85	0.87	0.84	0.87	0.85	0.87
Other services	0.70	0.70	0.72	0.71	0.70	0.69

Table A3: Survey response rate of the BSWS by city size

Year	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)
All	0.69	0.69	0.70	0.71	0.70	0.69
0-10	0.69	0.74	0.71	0.70	0.76	0.65
10-20	0.72	0.71	0.74	0.74	0.70	0.70
20-30	0.73	0.70	0.73	0.70	0.73	0.70
30-40	0.71	0.71	0.72	0.71	0.73	0.72
40-50	0.68	0.69	0.70	0.73	0.70	0.68
50-60	0.71	0.73	0.73	0.73	0.72	0.71
60-70	0.69	0.72	0.71	0.72	0.72	0.70
70-80	0.69	0.70	0.72	0.72	0.72	0.70
80-90	0.70	0.70	0.72	0.72	0.72	0.70
90-100	0.68	0.68	0.69	0.70	0.69	0.68

Table A4: Survey response rate of the SSPS by establishment size

Year	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)	2018 (7)	2019 (8)
All	0.73	0.71	0.71	0.70	0.70	0.71	0.73	0.63
Headquarters	0.84	0.83	0.82	0.83	0.83	0.82	0.84	0.77
5000+	0.84	0.82	0.82	0.82	0.82	0.84	0.83	0.73
1000–4999	0.85	0.84	0.86	0.84	0.84	0.84	0.85	0.72
500–999	0.85	0.85	0.85	0.85	0.84	0.85	0.85	0.68
100–499	0.82	0.81	0.82	0.81	0.81	0.81	0.82	0.70
30–99	0.78	0.79	0.79	0.79	0.79	0.80	0.79	0.66
10–29	0.76	0.73	0.73	0.73	0.72	0.74	0.76	0.61
1–9	0.54	0.51	0.48	0.48	0.48	0.49	0.52	0.49

Note: In the SSPS, "headquarters" means headquarters of joint stock company with employees less than 500, and capital of more than 1 billion JPY.

Table A5: Survey response rate of the SSPS by industry

Year	2012 (1)	2013 (2)	2014 (3)	2015 (4)	2016 (5)	2017 (6)	2018 (7)	2019 (8)
All	0.73	0.71	0.71	0.70	0.70	0.71	0.73	0.63
Corporation								
Manufacturing	0.82	0.80	0.80	0.81	0.80	0.82	0.83	0.73
Wholesales	0.76	0.76	0.76	0.74	0.74	0.75	0.77	0.64
Retail trade	0.73	0.73	0.71	0.69	0.72	0.70	0.74	0.63
Construction	0.73	0.70	0.69	0.69	0.68	0.68	0.71	0.63
Transportation	0.84	0.82	0.82	0.82	0.83	0.83	0.84	0.73
Service	0.75	0.73	0.73	0.73	0.72	0.73	0.74	0.62
Inn and restaurant	0.67	0.61	0.63	0.60	0.62	0.62	0.63	0.46
Other	0.80	0.78	0.78	0.77	0.78	0.78	0.80	0.72
Individual business								
Retail trade	0.40	0.38	0.37	0.39	0.36	0.37	0.41	0.39
Wholesales	0.52	0.52	0.44	0.38	0.34	0.52	0.43	0.50
Manufacturing and retailing	0.48	0.38	0.44	0.37	0.41	0.44	0.42	0.40
Manufacturing and wholesale trade	0.44	0.48	0.42	0.51	0.42	0.47	0.36	0.59
Contract manufacturing	0.40	0.52	0.46	0.54	0.49	0.42	0.53	0.49
Repairing	0.52	0.49	0.49	0.46	0.45	0.52	0.47	0.43
Service	0.46	0.47	0.43	0.36	0.38	0.43	0.40	0.43
Construction	0.50	0.48	0.48	0.46	0.45	0.47	0.46	0.47
Other sales	0.44	0.37	0.39	0.40	0.39	0.45	0.48	0.38
Agriculture	0.44	0.46	0.44	0.47	0.43	0.45	0.43	0.47
Other business	0.63	0.58	0.56	0.54	0.53	0.53	0.58	0.53
Other	0.47	0.47	0.38	0.43	0.38	0.39	0.45	0.43

Table A1: Distribution of the number of sampled years and survey responses: The BSWS 2012–2017

Number of sampled years	Number of response 0	1	2	3	4	5	6	Observations	Share
1	0.372	0.628						274422	78.8%
2	0.141	0.193	0.666					58387	16.8%
3	0.070	0.080	0.175	0.675				12293	3.5%
4	0.041	0.033	0.062	0.146	0.718			4064	0.1%
5	0.021	0.019	0.019	0.035	0.109	0.797		2595	0.7%
6	0.031	0.020	0.024	0.041	0.020	0.126	0.737	293	0.1%
Observations	111425	184772	41334	8996	3206	2105	216		
Share	31.6%	52.5%	11.7%	2.6%	0.9%	0.6%	0.1%		

Table A2: Distribution of the number of sampled years and survey responses: The SSPS 2012–2019

Number of sampled years	Number of response 0	1	2	3	4	5	6	7	8	Observations	Share
1	0.377	0.623								139530	86.1%
2	0.117	0.210	0.673							9694	6.0%
3	0.074	0.095	0.202	0.629						3087	1.9%
4	0.045	0.059	0.098	0.217	0.581					1775	1.1%
5	0.035	0.055	0.071	0.095	0.233	0.511				1187	0.7%
6	0.041	0.026	0.041	0.087	0.091	0.219	0.495			849	0.5%
7	0.018	0.027	0.022	0.044	0.061	0.104	0.194	0.530		819	0.5%
8	0.019	0.014	0.020	0.025	0.034	0.045	0.083	0.175	0.584	5130	3.2%
Observations	54269	89504	7563	2679	1611	1111	1007	1331	2996		
Share	33.5%	55.2%	4.7%	1.7%	1.0%	0.7%	0.6%	0.8%	1.8%		

B Establishments unmatched with the Monthly Labour Survey

Since all establishments with 500 or more employees are covered by the MLS, all establishments with 500 or more employees in the BSWS should match completely, but as mentioned in the main text, the match rate was only 44 percent. Possible reasons why large establishments in the BSWS do not match the MLS include (1) those establishments were not surveyed by the MLS due to mis-steps by the Ministry of Health, Labour and Welfare or (2) those establishments declined to respond to the survey.

In order to check the first possibility, we matched the lists of the sampling population of the BSWS and the MLS for 2015–2017, and calculated the match rate by establishment characteristics (Table B1–B3). About 60 percent of establishments with 500 or more employees in the BSWS are successfully matched with the MLS, but the match rate declines across years. Although the match rate is around 60–80 percent in most prefectures, it is only 31 percent in Tokyo. In terms of establishment size, there is no apparent relationship, though the match rate is slightly higher in establishments with 1000–4999 employees. In addition, we found some heterogeneity across industries, but did not find a stable relationship.

Regarding the second point, according to the Ministry of Health, Labor and Welfare, the submission rate for the MLS is 85 percent.[21] Given that, the response rate is generally high in large establishments and so the impact of survey non-response on our data matching would not be serious.

[21] https://www.mhlw.go.jp/toukei/list/dl/maikin-20180927-01.pdf

Table B1: Match rate of the MLS and BSWS by prefecture

Year	2015–2017 (1)	2015 (2)	2016 (3)	2017 (4)
All	0.61	0.75	0.57	0.50
Hokkaido	0.58	0.73	0.57	0.44
Aomori	0.82	0.83	0.84	0.79
Iwate	0.63	0.70	0.61	0.56
Miyagi	0.69	0.81	0.61	0.64
Akita	0.58	0.62	0.53	0.59
Yamagata	0.70	0.73	0.74	0.64
Fukushima	0.80	0.93	0.75	0.69
Ibaraki	0.65	0.82	0.57	0.59
Tochigi	0.71	0.81	0.66	0.65
Gunma	0.74	0.86	0.71	0.66
Saitama	0.60	0.77	0.55	0.48
Chiba	0.63	0.78	0.59	0.51
Tokyo	0.31	0.52	0.25	0.17
Kanagawa	0.59	0.76	0.60	0.41
Niigata	0.77	0.78	0.81	0.73
Toyama	0.87	0.91	0.88	0.82
Ishikawa	0.76	0.89	0.71	0.69
Fukui	0.75	0.83	0.70	0.71
Yamanashi	0.80	0.84	0.81	0.75
Nagano	0.66	0.80	0.60	0.60
Gifu	0.64	0.62	0.68	0.63
Shizuoka	0.75	0.85	0.74	0.63
Aichi	0.72	0.90	0.69	0.58
Mie	0.76	0.89	0.70	0.71
Shiga	0.82	0.91	0.78	0.78
Kyoto	0.68	0.75	0.64	0.64
Osaka	0.60	0.79	0.55	0.45
Hyogo	0.66	0.79	0.64	0.55
Nara	0.47	0.70	0.36	0.39
Wakayama	0.76	0.86	0.72	0.72
Tottori	0.69	0.80	0.64	0.64
Shimane	0.64	0.67	0.69	0.56
Okayama	0.69	0.81	0.67	0.60
Hiroshima	0.56	0.71	0.48	0.48
Yamaguchi	0.75	0.88	0.70	0.66
Tokushima	0.75	0.88	0.76	0.65
Kagawa	0.79	0.85	0.69	0.83
Ehime	0.67	0.79	0.67	0.54
Kochi	0.64	0.78	0.56	0.60
Fukuoka	0.64	0.76	0.58	0.55
Saga	0.73	0.79	0.73	0.64
Nagasaki	0.75	0.79	0.78	0.70
Kumamoto	0.65	0.77	0.55	0.62
Oita	0.60	0.65	0.61	0.54
Miyazaki	0.74	0.82	0.62	0.73
Kagoshima	0.61	0.53	0.65	0.65
Okinawa	0.64	0.79	0.64	0.46

Table B2: Match rate of the MLS and BSWS by establishment size

Year	2015–2017 (1)	2015 (2)	2016 (3)	2017 (4)
All	0.61	0.75	0.57	0.50
15,000+	1.00	1.00	1.00	1.00
5,000–14,999	0.58	0.63	0.57	0.53
1,000–4,999	0.68	0.80	0.65	0.60
500–999	0.57	0.73	0.54	0.46

Table B3: Match rate of the MLS and BSWS by industry

Year	2015–2017 (1)	2015 (2)	2016 (3)	2017 (4)
All	0.61	0.75	0.57	0.50
Mining, quarrying of stone, gravel	0.00	0.00	0.00	0.00
Construction	0.51	0.64	0.44	0.43
Manufacturing	0.77	0.85	0.75	0.70
Electricity, gas, heat supply, water	0.76	0.73	0.77	0.80
Information, communications	0.39	0.57	0.37	0.25
Transport, postal services	0.54	0.70	0.47	0.46
Wholesale, retail trade	0.43	0.58	0.36	0.33
Finance, insurance	0.55	0.68	0.51	0.45
Real estate, goods rental, leasing	0.53	0.75	0.50	0.36
Scientific research, prof. and tech. services	0.57	0.77	0.56	0.41
Accommodations, eating, drinking services	0.42	0.57	0.28	0.37
Living-related, personal, amusement services	0.47	0.58	0.33	0.49
Education, learning support	0.62	0.73	0.56	0.57
Medical, health, welfare	0.60	0.89	0.59	0.38
Compound services	0.23	0.67	0.17	0.13
Other services	0.46	0.69	0.44	0.28

C The characteristics of supplementary establishments

In this section of the appendix, we review the characteristics of the establishments in the original list (listed establishments) and in the supplementary list (unlisted establishments). The first column of Table C1 shows the proportion of supplementary establishments included in our sample, which is 5–8 percent. The remaining columns show the average wage payment of listed and unlisted establishments, and we see that the average wage of supplementary establishments is 10 percent lower than the originally listed establishments.

Table C1: Amount of salary paid on a fixed basis by survey years

Survey year	Frac. unlisted	Monthly salary per worker		
		Unlisted	Listed	Diff: (2) − (3)
	(1)	(2)	(3)	(4)
All	0.06	24.29	27.01	-2.72
		[16.49]	[17.80]	(0.24)
2012	0.06	24.22	27.53	-3.31
		[15.51]	[18.06]	(0.58)
2013	0.08	24.95	27.34	-2.38
		[16.88]	[17.98]	(0.40)
2014	0.07	25.48	26.67	-1.19
		[17.89]	[17.66]	(0.75)
2015	0.06	23.04	26.76	-3.72
		[15.66]	[17.86]	(0.44)
2016	0.05	23.08	26.94	-3.86
		[15.48]	[17.74]	(0.45)
2017	0.05	23.94	26.83	-2.89
		[16.04]	[17.46]	(0.47)

Note: Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

The frequency of the usage of the supplementary list depends on the establishment size, as seen in Column 1 of Table C2. The distribution shows that supplementary establishments tend to be substantially smaller than non-supplementary establishments, for the response rate of small establishments is lower than large establishments. After disentangling by establishment size, the difference in the average wage payment is relatively small particularly among middle and small-sized establishments, ranging from around 2 percent among establishments with fewer than 300 employees to about 6 percent for establishments with 50–99 employees. This difference is still relatively small, at about half of the overall difference in Table C1. While the monthly salary per worker tends to be

Table C2: Amount of salary paid on a fixed basis by establishment size

Establishment size	Frac. unlisted	Monthly salary per worker		
		Unlisted	Listed	Diff: (2) − (3)
	(1)	(2)	(3)	(4)
15000+	0.00		24.13	
			[12.91]	
5000–14999	0.02	29.57	45.12	-15.55
		[22.33]	[19.74]	(3.11)
1000–4999	0.02	40.21	40.48	-0.27
		[23.41]	[21.60]	(1.62)
500–999	0.02	29.63	34.10	-4.47
		[17.55]	[21.09]	(1.83)
300–499	0.02	27.53	30.91	-3.38
		[17.18]	[19.23]	(1.45)
100–299	0.03	27.32	27.92	-0.60
		[16.33]	[17.43]	(0.61)
50–99	0.03	23.56	25.01	-1.45
		[15.68]	[15.66]	(0.52)
30–49	0.04	23.19	23.85	-0.65
		[16.20]	[15.36]	(0.39)
10–29	0.07	22.12	22.51	-0.39
		[15.37]	[15.10]	(0.22)
5–9	0.13	21.89	21.49	0.41
		[13.84]	[13.38]	(0.27)

Note: Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

lower in unlisted establishments than in the listed establishments within each establishment size category, the smallest group is an exception, with the monthly salary of the unlisted establishments 2 percent higher than that of the listed establishments.

The frequency in using the supplementary list does not depend on the industry, as seen in Table C3, where the proportion of unlisted establishments differs only slightly by industry. In contrast, the difference in the average wage payment varies widely among industries and is particularly large in the information and communication industry and the academic research, professional and technical services industry. On the other hand, the wage payment is higher for supplementary establishments in the lifestyle-related services and entertainment, combined services, and unclassified service industries. In sum, the differences in the average wages between the original sample establishments and the supplementary establishments are dependent on establishment size and industry.

Table C3: Amount of salary paid on a fixed basis by establishment size by industry

Industry	Frac. unlisted	Monthly salary per worker		
		Unlisted	Listed	Diff: (2) − (3)
	(1)	(2)	(3)	(4)
Mining, quarrying of stone, gravel	0.07	30.86	32.87	-2.01
		[13.95]	[15.95]	(1.64)
Construction	0.05	30.61	33.03	-2.42
		[13.86]	[15.04]	(0.82)
Manufacturing	0.05	27.77	30.71	-2.94
		[15.64]	[15.37]	(0.54)
Electricity, gas, heat supply, water	0.04	42.26	44.53	-2.27
		[16.85]	[17.89]	(1.10)
Information, communications	0.06	35.30	40.49	-5.18
		[16.88]	[18.96]	(2.25)
Transport, postal services	0.05	28.19	28.73	-0.53
		[13.90]	[14.21]	(0.70)
Wholesale, retail trade	0.06	21.06	22.65	-1.59
		[15.27]	[16.53]	(0.41)
Finance, insurance	0.05	34.61	36.50	-1.89
		[22.53]	[23.60]	(0.69)
Real estate, goods rental, leasing	0.09	24.94	27.82	-2.89
		[16.82]	[18.76]	(0.70)
Scientific research, prof. and tech. services	0.07	32.21	37.28	-5.07
		[16.71]	[19.29]	(0.76)
Accommodations, eating, drinking services	0.07	12.31	12.64	-0.33
		[10.60]	[10.99]	(0.18)
Living-related, personal, amusement services	0.08	19.66	18.98	0.68
		[13.51]	[13.16]	(0.66)
Education, learning support	0.07	24.63	27.23	-2.60
		[20.24]	[21.01]	(0.70)
Medical, health, welfare	0.05	21.82	24.97	-3.15
		[16.95]	[19.28]	(0.45)
Compound services	0.04	28.82	28.16	0.66
		[13.16]	[13.67]	(0.43)
Other services	0.07	23.23	22.09	1.14
		[14.25]	[13.62]	(0.56)

Note: Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

Table D1: Sampling of workers within an establishment: BSWS vs MLS (Male)

	BSWS			MLS	Diff.	
	All	500+	Matched		(3) − (4)	(3) − (2)
	(1)	(2)	(3)	(4)	(5)	(6)
Number of employees	18.77	610.56	798.86	803.39	-4.52	188.30
	[76.14]	[821.34]	[964.70]	[1004.17]	(12.44)	(16.34)
Work days	20.63	20.24	20.29	19.83	0.46	0.05
	[5.05]	[3.62]	[3.45]	[2.36]	(0.06)	(0.04)
Predetermined work hours	150.26	152.92	154.16	150.85	3.30	1.23
	[47.15]	[31.67]	[29.99]	[20.78]	(0.52)	(0.40)
Overtime hours	9.51	14.66	16.58	17.66	-1.08	1.92
	[17.58]	[19.66]	[19.68]	[10.62]	(0.22)	(0.20)
Total work hours	159.76	167.58	170.74	168.51	2.21	3.15
	[53.83]	[39.87]	[38.50]	[25.25]	(0.59)	(0.49)
Salary paid on a fixed basis	27.92	42.88	42.28	42.71	-0.43	-0.60
	[17.38]	[25.31]	[22.92]	[12.79]	(0.22)	(0.29)
Bonus	4.57	11.51	12.21	12.96	-0.76	0.69
	[6.90]	[11.83]	[10.49]	[6.71]	(0.09)	(0.15)
Total salary	32.49	54.40	54.49	55.67	-1.19	0.10
	[21.84]	[32.98]	[30.05]	[17.32]	(0.27)	(0.38)
Observations	114410	5037	2193			

Note: Unit of earnings is 10 thousand JPY/month. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

D Sampling of workers in the BSWS: Subsample analysis by gender

Tables D1 and D2 show the results of the analysis by gender. Regardless of gender, the sampled workers tend to work long hours. On the other hand, we found that the sampling bias in earnings is heterogeneous across gender, though the size of the bias is not economically substantial. For male workers, all salary items are higher in the MLS than in the BSWS, with the difference in total salary about 11,900 JPY or 2.1 percent. For women, regular salaries are higher in the BSWS and bonuses are higher in the MLS, but total salaries are 1.4 percent higher in the BSWS. Combining men and women, the biases cancel each other out and there is not a large difference.

All in all, we conclude that the sampling of workers in the BSWS is well performed, though not perfectly. Althogh the sampling of workers of the BSWS is biased in terms of work hours and earnings, the magnitude of the bias is limited to around 2 percent. The data does not allow us to examine the cause of this bias but, in general, the difference is larger for women than for men. One possibility is that core workers are more likely to be selected, and since the share of core workers is high among men but not among women, the selected male workers may still represent the population of male workers relatively well

Table D2: Sampling of workers within an establishment: BSWS vs MLS (Female)

		BSWS		MLS	Diff.	
	All	500+	Matched		(3) − (4)	(3) − (2)
	(1)	(2)	(3)	(4)	(5)	(6)
Number of employees	15.17	351.78	348.29	349.83	-1.54	-3.48
	[52.87]	[588.32]	[550.82]	[433.77]	(6.37)	(11.13)
Work days	19.36	19.80	19.87	19.07	0.80	0.07
	[5.32]	[3.61]	[3.49]	[2.28]	(0.06)	(0.04)
Predetermined work hours	131.09	145.40	147.19	140.97	6.22	1.79
	[50.88]	[34.33]	[33.02]	[21.87]	(0.52)	(0.46)
Overtime hours	4.36	9.66	9.92	10.18	-0.25	0.26
	[10.39]	[14.27]	[14.19]	[6.86]	(0.19)	(0.17)
Total work hours	135.45	155.06	157.11	151.15	5.97	2.05
	[54.50]	[39.96]	[38.85]	[24.40]	(0.57)	(0.51)
Salary paid on a fixed basis	17.55	28.31	27.65	26.72	0.93	-0.66
	[10.74]	[15.28]	[13.61]	[7.57]	(0.13)	(0.20)
Bonus	2.57	6.94	7.37	7.81	-0.45	0.43
	[4.36]	[7.28]	[6.76]	[4.35]	(0.08)	(0.10)
Total salary	20.12	35.25	35.02	34.55	0.48	-0.23
	[13.77]	[20.78]	[19.02]	[11.12]	(0.18)	(0.28)
Observations	114410	5037	2193			

Note: Unit of earnings is 10 thousand JPY/month. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

whereas the selected female workers may deviate from the population of female workers.

E The characteristics of freelance workers

Table E1: Characteristics of (non-) freelance workers

	Freelance (1)	Other (2)	Diff: (1) (3)
Age	57.321	45.369	11.95
	[15.132]	[14.623]	(0.13
Female	0.294	0.446	-0.15
	[0.456]	[0.497]	(0.00
Never married	0.196	0.300	-0.10
	[0.397]	[0.458]	(0.00
Attending school	0.003	0.027	-0.02
	[0.050]	[0.162]	(0.00
Primary school or junior high school	0.177	0.064	0.11
	[0.382]	[0.245]	(0.00
Senior high school	0.403	0.370	0.03
	[0.491]	[0.483]	(0.00
Professional training college (1 year or more but less than 2 years)	0.076	0.053	0.02
	[0.265]	[0.224]	(0.00
Professional training college (2 years or more but less than 4 years)	0.069	0.086	-0.01
	[0.253]	[0.281]	(0.00
Professional training college (4 years or more)	0.000	0.002	-0.00
	[0.017]	[0.048]	(0.00
Junior college	0.050	0.083	-0.03
	[0.218]	[0.277]	(0.00
College of technology	0.010	0.011	-0.00
	[0.097]	[0.105]	(0.00
College or university	0.198	0.296	-0.09
	[0.399]	[0.456]	(0.00
Graduate school	0.016	0.034	-0.01
	[0.126]	[0.182]	(0.00
Observations	28649	476678	

Source: The ESS 2017.
Note: The freelance worker is defined as the self-employed without employees. Standard deviations are reported in the brackets and standard errors are reported in parentheses.

Table E2: Reason why one chose the current job

	Freelance (1)	Other (2)	Diff: (1) − (2) (3)
Being unemployed	0.065	0.103	-0.038
	[0.247]	[0.304]	(0.002)
Having graduated from school	0.093	0.219	-0.126
	[0.290]	[0.414]	(0.002)
Need to earn income	0.156	0.163	-0.007
	[0.363]	[0.369]	(0.003)
Wanted to make the best use of my knowledge and skills	0.276	0.123	0.153
	[0.447]	[0.328]	(0.004)
Wanted to make a start in working life	0.021	0.030	-0.009
	[0.143]	[0.171]	(0.001)
Had sufficient time to take up a job	0.032	0.037	-0.005
	[0.176]	[0.188]	(0.001)
Wanted to maintain health	0.032	0.009	0.022
	[0.175]	[0.096]	(0.001)
This job has better conditions	0.040	0.129	-0.089
	[0.196]	[0.336]	(0.002)
Other	0.286	0.186	0.100
	[0.452]	[0.389]	(0.004)
Observations	28649	476678	

Source: The ESS 2017.
Note: The freelance worker is defined as the self-employed without employees. Standard deviations are reported in the brackets and standard errors are reported in parentheses.

F Other supplementary tables

Table F1: Response rate and lagged establishment characteristics: The BSWS

Establishment size	All	500+	100–499	30–99	5–29
	(1)	(2)	(3)	(4)	(5)
L.$ln(earnings)$	0.063	0.025	0.071	0.050	0.082
	(0.004)	(0.014)	(0.008)	(0.007)	(0.007)
L.Age	-0.000	-0.001	0.001	-0.001	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
L.Female	0.024	0.001	0.043	0.013	0.027
	(0.006)	(0.026)	(0.014)	(0.011)	(0.010)
L.$ln(WorkHours)$	0.003	0.010	0.018	-0.014	-0.006
	(0.007)	(0.025)	(0.015)	(0.013)	(0.012)
Observations	91282	8369	21560	27532	33821

Note: Standard errors clustered by each establishment are reported in parentheses. Control variables include year and prefecture fixed effects, establishment size, industry and city population.

Table F2: Sampling of workers within establishments: BSWS vs MLS (Non-bonus adjustment sample)

		BSWS		MLS	Diff.	
	All	500+	Matched		(3) − (4)	(3) − (2)
	(1)	(2)	(3)	(4)	(5)	(6)
Number of regular employees	33.91	961.81	1073.18	1071.89	1.29	111.37
	[108.08]	[1078.31]	[1213.24]	[1067.06]	(13.46)	(16.25)
Work days	20.09	20.05	20.09	19.49	0.60	0.04
	[5.24]	[3.61]	[3.51]	[2.42]	(0.06)	(0.03)
Predetermined work hours	140.66	148.88	149.86	145.85	4.01	0.98
	[50.25]	[33.30]	[32.32]	[22.12]	(0.44)	(0.31)
Overtime ours	7.61	12.80	13.96	14.94	-0.98	1.16
	[15.76]	[17.97]	[18.24]	[10.03]	(0.14)	(0.13)
Total work hours	148.26	161.68	163.82	160.79	3.03	2.15
	[56.46]	[40.80]	[40.22]	[26.69]	(0.48)	(0.36)
Salary paid on a fixed basis	22.91	36.53	36.30	35.97	0.33	-0.23
	[14.84]	[21.08]	[19.86]	[10.60]	(0.11)	(0.19)
Bonus	3.52	10.00	10.44	38.66	-28.21	0.44
	[5.78]	[10.77]	[10.08]	[45.43]	(0.95)	(0.11)
Total salary	26.43	46.53	46.75	74.63	-27.88	0.21
	[18.84]	[29.19]	[27.66]	[50.69]	(0.96)	(0.27)
Observations	172814	7585	4422			

Note: Unit of earnings is 10 thousand JPY/month. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses.

Table F3: Sampling of workers within an establishment: SSPS (Establishments with 500 or more employees)

	All			Non outliers		
Questionnaire	Worker	Establishment	Diff. (1) − (2)	Worker	Establishment	Diff. (4) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	4.011	11.182	-7.171	3.911	3.879	0.032
	[2.602]	[128.831]	(6.543)	[2.308]	[2.411]	(0.037)
p10	1.382	1.414	-0.032	1.379	1.415	-0.036
			(0.048)			(0.047)
p25	2.243	2.136	0.107	2.222	2.138	0.084
			(0.057)			(0.054)
p50	3.586	3.342	0.244	3.549	3.319	0.230
			(0.069)			(0.060)
p75	5.258	5.121	0.137	5.210	5.086	0.124
			(0.067)			(0.050)
p90	7.345	7.403	-0.059	7.120	7.135	-0.015
			(0.281)			(0.196)
Observations	46798	46798	46798	46566	46566	46566

Note: Unit of earnings is 1 million JPY/year. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses. The standard error of each percentile difference was obtained via establishment-level clustering bootstrap with 1000 replication. In the non-outlier sample, we excluded the top and bottom 1 percent of the distribution of a difference between the worker questionnaire and establishment questionnaire.

Table F4: Sampling of workers within an establishment: SSPS (Establishments with less than 500 employees)

	All			Non outliers		
Questionnaire	Worker	Establishment	Diff. (1) − (2)	Worker	Establishment	Diff. (4) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	3.059	3.155	-0.095	3.032	3.024	0.009
	[1.992]	[16.332]	(0.072)	[1.891]	[1.860]	(0.002)
p10	1.180	1.200	-0.020	1.187	1.201	-0.014
			(0.003)			(0.004)
p25	1.772	1.783	-0.011	1.776	1.778	-0.002
			(0.004)			(0.004)
p50	2.700	2.693	0.007	2.689	2.675	0.014
			(0.004)			(0.004)
p75	3.891	3.865	0.026	3.860	3.831	0.029
			(0.006)			(0.005)
p90	5.299	5.237	0.062	5.232	5.153	0.079
			(0.010)			(0.009)
Observations	117195	117195	117195	112600	112600	112600

Note: Unit of earnings is 1 million JPY/year. Standard deviations are reported in the brackets and standard errors clustered by each establishment are reported in parentheses. The standard error of each percentile difference was obtained via establishment-level clustering bootstrap with 1000 replication. In the non-outlier sample, we excluded the top and bottom 1 percent of the distribution of a difference between the worker questionnaire and establishment questionnaire.