Bank of Japan Working Paper Series

"Don't know" Tells: Calculating Non-Response Bias in Firms' Inflation Expectations Using Machine Learning Techniques

Yosuke Uno*
yousuke.uno@boj.or.jp

Ko Adachi**
kou.adachi@boj.or.jp

No.19-E-17
December 2019

Bank of Japan
2-1-1 Nihonbashi-Hongokucho, Chuo-ku, Tokyo 103-0021, Japan

* Research and Statistics Department (currently at the Payment and Settlement Systems Department)

** Research and Statistics Department (currently at the Monetary Affairs Department)

“Don’t Know” Tells: Calculating Non-Response Bias in Firms’ Inflation Expectations Using Machine Learning Techniques*

Yosuke Uno† Ko Adachi‡

December 2019

Abstract

This paper examines the “don’t know” responses for questions concerning inflation expectations in the *Tankan* survey. Specifically, using machine learning techniques, we attempt to extract “don’t know” responses where respondent firms are more likely to “know” in a sense. We then estimate the counterfactual inflation expectations of such respondents and examine the non-response bias based on the estimation results. Our findings can be summarized as follows. First, there is indeed a fraction of firms that respond “don’t know” despite the fact that they seem to “know” something in a sense. Second, the number of such firms, however, is quite small. Third, the estimated counterfactual inflation expectations of such firms are not statistically significantly different from the corresponding official figures in the *Tankan* survey. Fourth and last, based on the above findings, the non-response bias in firms’ inflation expectations likely is statistically negligible.

JEL codes: C55, E31
Keywords: inflation expectations, PU classification, non-response bias

*We thank Kohei Takata, Kosuke Aoki, Shigehiro Kuwabara, Toshitaka Sekine, Junichi Suzuki, Takuto Ninomiya, Hibiki Ichiue, Koki Inamura, Hidetaka Enomoto, Ko Nakayama, and Toshinao Yoshiba for useful comments. All remaining errors are ours. This paper does not necessarily reflect the views of the Bank of Japan.

†Research and Statistics Department (currently at the Payment and Settlement Systems Department), Bank of Japan. E-mail: yousuke.uno@boj.or.jp

‡Research and Statistics Department (currently at the Monetary Affairs Department), Bank of Japan. E-mail: kou.adachi@boj.or.jp

1 Introduction

Surveys often include an option of "don't know" to respond to questions. Imagine the following question:

Question: What country do you think will win the 2022 FIFA World Cup?

Suppose that there are two types of responses: "Brazil" and "don't know." These are observationally different. However, people responding "don't know" might be thinking that Brazil is the most likely to win, but it is too uncertain. On the other hand, people responding "Brazil" may not have much confidence in their answer. This means that although the responses of "don't know" and "Brazil" are observationally different, inherently they might not be that different at all.

This paper examines the response of "don't know" in relation to the question concerning inflation expectations in the Short-term Economic Survey of Enterprises in Japan (*Tankan*). In the *Tankan* survey, there are a significant number of responses of "don't know" in relation to the question which requires respondents to provide their own forecasts for one year, three years, and five years ahead, as pointed out by Uno, Naganuma, and Hara (2018a) and Coibion, Gorodnichenko, Kumar, and Pedemonte (2018). For example, more than 40 percent of firms respond "don't know" with regard to their forecast for five years ahead.

We think that the response of "don't know" in relation to the question concerning inflation expectations in the *Tankan* survey is similar to the example provided above. That is, we think that some of the "don't know" responses in the *Tankan* survey can be treated as inherently similar to a quantitative answer. We therefore assume that the "true" answer in such cases is not "don't know" but a quantitative answer that respondents did not have sufficient confidence in. Based on the example of the 2022 FIFA World Cup question, if a fraction of the "don't know" answers can be regarded as inherently the same answer as "Brazil," we assume that the "true" answer of such responses is "Brazil." Note that we could alternatively assume that the "true" answer is "don't know," not "Brazil." We discuss these assumptions in Section 4.

For the analysis, we divide "true" answers into two classes: "know" and "don't know." Further, we make the following two assumptions: (1) observed "don't know" responses include responses where the unobserved "true class" is "know" (*i.e.*, respondents have, for example, quite a good idea about their inflation expectation for five years hence, but they do not have sufficient confidence in their answer and choose "don't know" instead); and (2) observed "know" responses do *not* include responses where the unobserved "true class" is "don't know" (*i.e.*, we assume that respondents that do not have a good idea about their inflation expectation do *not* simply answer with a random number giving the impression that they do "know" when they do not). Based on these assumptions, we try to extract "don't know" responses that are likely to fall into the "know" category as the "true class" using machine learning techniques. In terms of machine learning, our first assumption means that a fraction of instances labeled "don't know" is mislabeled; our second assumption means that instances labeled "know" are correctly labeled. To deal with this setup, we employ the positive unlabeled (PU) classification developed by Liu, Lee, Yu, and Li (2002), Liu, Dai, Li, Lee, and Yu (2003), and Elkan and Noto (2008). PU classification attempts to learn a binary classifier while only having access to positively labeled and unlabeled data.

In order to solve our PU classification problem, we combine two methods well-known in machine learning: a support vector machine and a logistic regression classifier. Using these methods, the "don't know" responses in the *Tankan* survey are classified into the two "true classes": "know" and "don't know." We then estimate the counterfactual inflation expectations for firms classified as "know" in terms of their "true class." Finally, based on the estimation results, we examine the size of the non-response bias in firms' inflation expectations. Note that, in general, non-response bias is defined as the bias caused by a situation where firms or households chosen for a survey are unwilling or unable to participate. In this paper, we use the term "non-response" not only in this general sense but also for the situation where firms have responded "don't know."

This paper contributes to the literature in two respects. First, we quantitatively examine the non-response bias in firms' inflation expectations. To the best of our knowledge, no studies so far have examined the non-response bias in firms' inflation expectations. Our findings in this regard can be summarized as follows: (1) there is indeed a fraction of firms that respond "don't know" even though they seem to "know" in a certain sense; (2) the number of such firms, however, is quite small; (3) the estimated counterfactual inflation expectations of such firms are not statistically significantly different from the corresponding official figures in the *Tankan* survey; and (4) based on the above findings, the non-response bias in firms' inflation expectations likely is statistically negligible. These figures provide a *quantitative* reply to the *qualitative* criticism by Coibion, Gorodnichenko, Kumar, and Pedemonte (2018) that the *Tankan* survey has a non-response bias.

The second contribution of the paper is that we use machine learning techniques to address the issue of non-response bias. Statistical authorities and some statisticians have discussed this issue for a long time. In particular, for statistical authorities, controlling for non-response bias has played an important role in improving the quality of official statistics. Yet, so far, the innovative techniques developed in the field of machine learning have hardly been applied to practical problems such as non-response bias and missing value imputation, which are two sides of the same coin. Machine learning has been one of the fastest growing fields in statistical analysis in recent years, and applying the innovative methods developed in this field can potentially greatly help in understanding the issue of non-response bias.

The remainder of the paper proceeds as follows. Section 2 provides an overview of the related literature. Section 3 describes the data used. Section 4 formulates the problem addressed in this study as a PU classification problem. Section 5 uses a natural experiment to examine whether our formulation of the problem as a PU classification problem is valid. Section 6 examines our algorithms to discover the "true class," while Section 7 presents the results. Based on the results, Section 8 presents our estimates of the non-response bias. Section 9 concludes.

2 Related Literature

This paper aims to address the problem of non-response bias in a survey using machine learning techniques. The paper therefore primarily relates to studies on non-response bias and missing value imputation, which are two sides of the same coin. Another related field is the PU classification problem, since we apply an algorithm developed to solve the PU classification problem to address non-response bias. This section provides an overview of the related literature on non-response bias and missing value imputation, while the related literature on the PU classification problem will be summarized in Subsection 6.1.

A key study on non-response bias and missing value imputation is that by Little and Rubin (2002), who highlight the importance of missing data mechanisms such as missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). Based on Little and Rubin (2002)'s argument, statistical authorities have exhibited a keen interest in non-response bias and missing value imputation to improve the quality of statistics. For example, economists at the Bank of Japan have proposed a way to impute missing values for annual projections of, for example, sales and fixed investment using the microdata from the *Tankan* survey (Utsunomiya and Sonoda (2004); Hirakawa and Hatogai (2013)). Moreover, the Cabinet Office has examined the non-response bias in the Machinery Orders Statistics (Cabinet Office (2017)). Meanwhile, several academic researchers have used more sophisticated methods to estimate non-response bias using microdata provided by the Institute of Statistical Mathematics, a government-affiliated research institute (Hoshino (2010); Tsuchiya (2010)).

However, whereas these previous studies implicitly assume that there exist true values which should be imputed for the missing values, we do not assume the existence of true values and explicitly allow cases where there are no true values to impute. This can be viewed as a generalization of the previous studies.

At the same time, to the best of our knowledge, this paper is the first attempt to quantitatively evaluate the non-response bias in firms' inflation expectation. Recently, Coibion, Gorodnichenko, Kumar, and Pedemonte (2018) have argued that the mean of firms' inflation expectations in the *Tankan* survey is biased because a considerable number of respondents non-randomly choose the response "don't know." We provide a *quantitative* reply to this *qualitative* criticism of the *Tankan* survey.

3 Data

The data used in this paper consist of the confidential microdata gathered by the Bank of Japan's Research and Statistics Department for the *Tankan* survey, a large-scale firm-level survey. This is the same data set as used by Uno, Naganuma, and Hara (2018a,b).

The *Tankan* gathers firms' inflation expectations with regard to two types of prices: general prices and output prices. The question with respect to general prices is phrased as follows: "What are your institution's expectations of the annual % change in general prices (as measured by the consumer price index) for one year ahead, three years ahead, and five years ahead, respectively? Please select the range nearest to your own expectation from the options below." Respondents are required to choose their responses from the following ten options: (1) around +6% or higher, (2) around +5%, (3) around +4%, and continuing in one-percentage point intervals until (10) around −3% or lower. If respondents have no clear view on the outlook for general prices, they are asked to select one of the following three reasons: (11) uncertainty over the future outlook is high, (12) not really conscious of inflation fluctuations because they should not influence the strategy of the institution, and (13) others. In this paper, "don't know" responses (hereafter DK instances) are defined as responses (11), (12) or (13),[1] while "know" responses (hereafter non-DK instances) are defined as responses (1) to (10).

Regarding output prices, the question is "What are your institution's expectations of the rate of price changes in your mainstay domestic products or services for one year ahead, three years ahead, and five years ahead, respectively? Please select the range nearest to your own expectation from the options below." Respondents are again required to make

[1] This means that we assume that responses (11), (12), or (13) are identical. This assumption does not affect our results.

responses from the following ten options: (1) around +20% or higher, (2) around +15% and more, and continuing in five-percentage point intervals until (9) around −20% or lower, and (10) don't know. DK instances are defined as response (10), while non-DK instances are defined as responses (1) to (9).

Each quarterly *Tankan* survey covers around 10,000 firms selected from a population of approximately 210,000 firms with paid in capital of at least 20 million yen. The question on inflation expectations is included in the survey from March 2014 onwards. Our observation period therefore runs from March 2014 to September 2017. The *Tankan* data are an unbalanced panel in which there is a core of firms that have been surveyed since the start (May 1974), while some have dropped out because they fell below 20 million yen in paid in capital and/or went bankrupt. To maintain the sample size, firms are newly added every two or three years. During our observation period, about 1,000 firms were added in March 2015.

4 Problem Formulation

This paper attempts to uncover the unobservable "true class" ("know" or "don't know") for DK instances in the *Tankan* survey. This task can be formulated as a PU classification problem, which is one of the binary classification problems developed by Liu, Lee, Yu, and Li (2002), Liu, Dai, Li, Lee, and Yu (2003), and Elkan and Noto (2008). In a PU classification problem, binary classifiers are trained on data sets with only positive and unlabeled instances. Training a classifier on positive and unlabeled data, as opposed to on positive and negative data, is an issue of great importance, because in practice the lack of negative instances is quite common. For example, to identify a person's favorite website, one could use websites the person has bookmarked as positive instances and all other websites as unlabeled ones.

Following the notation used by Elkan and Noto (2008), we set three random variables: $\boldsymbol{x}$, y, and s. The k-dimensional vector, $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)$, denotes the characteristics of firms. Let $y = \{1, 0\}$ and $s = \{1, 0\}$ be the binary labels "know" and "don't know." We assume that the variable y is an unobservable variable and call it the "true class." The variable s represents the observed response in the *Tankan* survey. Specifically, instances with $s = 1$ and $s = 0$ represent non-DK instances and DK instances, respectively.

Next, we make the following two assumptions. The first is that the DK instances contain a fraction of mislabeled instances. Using the term *contamination*, as in Mordelet and Vert (2014), the DK instances are assumed to be *contaminated*. This can be written as

$$\Pr(s = 0 \mid \boldsymbol{x}, y = 1) > 0. \tag{1}$$

Note that we do not make any assumptions as to which DK instances are mislabeled. In other words, we do not make any assumptions on the sample selection mechanisms, $\Pr(s = 1 \mid y = 1)$, meaning that Equation 1 allows the sample selection bias first discussed by Heckman (1979).

The second assumption is that the non-DK instances are *uncontaminated*. This can be written as

$$\Pr(s = 1 \mid \boldsymbol{x}, y = 0) = 0. \tag{2}$$

Survey respondents (firms) may have an incentive to answer "don't know" to the questions on inflation expectations even when they have a clear view, since they may be concerned about revealing private information through their answer. In contrast, when firms have no clear view on inflation expectations, they do not have an incentive to provide a quantitative answer. Thus, it is unlikely that our second assumption, Equation 2, is too strong.

Based on the two assumptions, when $s = 1$ it is certain that $y = 1$, but when $s = 0$, then either $y = 1$ or $y = 0$ may be true. Next, let U_P be the set of instances identified as $y = 1$, and let U_N be the set of instances identified as $y = 0$. The first task is to uncover the "true class" ($y = 1$ or $y = 0$) for the DK instances ($s = 0$). The second task is to estimate the counterfactual inflation expectations for firm j classified into the set U_P. Finally, the third task is to evaluate the magnitude of the non-response bias based on the estimation results.

5 Validity of Our Problem Formulation

This section examines the validity of our problem formulation presented in the previous section. Specifically, we propose a natural experiment to identify mislabeled instances, *i.e.*, instances with $y = 1$ and $s = 0$. The results of the experiment show that our first assumption, Equation 1, is empirically supported.

5.1 A Natural Experiment

In order to conduct the *Tankan* survey, the Bank of Japan's Research and Statistics Department keeps various kinds of information, such as the location of firms' head office, their main products/services, contact persons, and so on. In this paper, we treat changes in the contact person as a natural experiment. As is well known among statistical authorities, a change in the contact person can occur for various reasons, such as retirement, transfer, change of the department responsible for cooperating with the statistical authorities, and so on.

The treatment, a change in the contact person, can be regarded as being independent of the "true class" of a firm's response with regard to inflation expectations denoted by the variable y. For example, imagine that the outlook for oil prices becomes extremely uncertain due to an increase in geopolitical risk. In this situation, it may be difficult for some firms to form a clear view on future inflation. That is, a state transition from the state "know" ($y = 1$) to the state "don't know" ($y = 0$) can occur due to the increase in geopolitical risk. However, the state transition is independent of a change in the contact person. The reason is that the change in the contact person obviously does not affect geopolitical risk and, conversely, it is unlikely that a change in geopolitical risk leads to the retirement or transfer of the contact persons. Therefore, the treatment is likely to be independent and possibly random.

Following the notations in Section 4, the state transition from the state "know" at time

$t-1$ to the state "don't know" at time t can be described as follows:

$$\Pr(y_t = 0 \mid \boldsymbol{x}_t, y_{t-1} = 1, z_t = 0) = \Pr(y_t = 0 \mid \boldsymbol{x}_t, y_{t-1} = 1, z_t = 1) \tag{3}$$

where t represents the survey period and $z = \{1, 0\}$ is a binary label which represents whether the contact person changed or not. The assumption that the variable z is independent of the "true class" denoted by y ensures this equality.

Suppose that Equation 1 does not hold, *i.e.*, $\Pr(s = 0 \mid x, y = 1) = 0$. Then the response in the *Tankan* survey, variable s, is equal to the "true class," suggesting that

$$\Pr(s_t = 0 \mid \boldsymbol{x}_t, s_{t-1} = 1, z_t = 0) = \Pr(s_t = 0 \mid \boldsymbol{x}_t, s_{t-1} = 1, z_t = 1). \tag{4}$$

As for the reverse direction, suppose that Equation 4 does not hold; in that case, Equation 1 holds. This implies that we can validate whether Equation 1 holds by comparing the transition probability from the state "know" to the state "don't know" in the case that the contact person changed to the probability in the case that the contact person remained.

5.2 Results of the Experiment

In this subsection, we report the results of the experiment. In our natural experiment, the definition of a change in the contact person is crucial. Here, taking advantage of the *Tankan* survey, which collects detailed information about contact persons, we define a change in the contact person as the case in which the job title of the contact person remains unchanged but the name of the contact person changes. For example, a switch from chief accountant A to chief accountant B falls under this definition, but a switch from chief accountant A to treasurer C does not, so that such a change is excluded from our experiment. This makes the treatment, the variable z, independent of the "true class" denoted by the variable y.

The results of the experiment are presented in Table 1. Two results in the table stand out. First, in the case that the contact person changed, the estimated state transition probabilities from the state "know" to the state "don't know" are statistically significantly higher for all time horizons and for both general prices and output prices than in the case that the contact person remained unchanged. For example, with regard to general prices for one year ahead, the estimated probability rises from 1.7 percent in the case that the contact person remained unchanged to 3.2 percent in the case that the contact person changed.

Table 1: Estimated state transition probabilities from state "know" to state "don't know"

	General prices		
	1-year ahead	3-year ahead	5-year ahead
Treatment ($z = 1$)	0.032	0.059	0.061
[3792]	(0.027,0.037)	(0.052,0.067)	(0.053,0.069)
Control ($z = 0$)	0.017	0.032	0.037
[142709]	(0.016,0.018)	(0.031,0.032)	(0.036,0.038)
Differences	0.015	0.027	0.024
	Output prices		
	1-year ahead	3-year ahead	5-year ahead
Treatment ($z = 1$)	0.021	0.042	0.051
[3792]	(0.017,0.026)	(0.036,0.049)	(0.044,0.058)
Control ($z = 0$)	0.013	0.027	0.037
[142709]	(0.011,0.016)	(0.026,0.028)	(0.036,0.038)
Differences	0.008	0.015	0.014

Notes: The 95% bootstrap confidence intervals are reported in brackets. The resampling sizes for the treatment group and the control group are 2,000 and 200, respectively. The numbers of observations are reported in square brackets.

Second, the differences in the estimated probabilities between the two groups (treatment and control) over longer time horizons and for general prices are larger than those for shorter time horizons and output prices. For example, regarding general prices, the difference between the two groups for five years ahead is 2.4 percentage points, while that for one year ahead is 1.5 percentage points. Moreover, comparing the results for general and for output prices for five years ahead, the difference of 2.4 percentage points for general prices compares with a difference of 1.4 percentage points for output prices.

The results presented in Table 1 suggest that the variable z, which indicates whether the contact person changed, correlates with the variable s, which denotes the response in the *Tankan* survey. Moreover, provided that the setting of our natural experiment assuming that a change in the contact person is independent of a firm's "true" answer is correct, the results imply that a fraction of the DK instances departs from the "true class." That is, using the example presented in Section 1, the "don't know" responses partly include unobservable responses where "Brazil" is the "true" answer. Note that we do not say anything about the case in which the "true" answer is "don't know." In sum, the results of the experiment provide empirical support that Equation 1 holds but are silent with regard to the other assumption represented by Equation 2.

6 Algorithm to Uncover "True Class"

In this section, we examine algorithms to solve our PU classification problem. Let $U = \{U_N, U_P\}$ and $P = \{h \mid s_h = 1, y_h = 1\}$.

6.1 Existing Algorithm

Basically, data set with only positive and unlabeled instances prevents the application of supervised classification algorithms which require negative instances in the data set. According to Yang, Liu, and Yang (2017), algorithms to solve PU classification problems can be roughly categorized into the following four types. The first is algorithms in which $\Pr(y = 1 \mid x)$ is directly estimated with some assumptions on the sample selection mechanism, $\Pr(y = 1 \mid s = 1)$ (Heckman (1979); Lee and Liu (2003); Elkan and Noto (2008)). As mentioned in Section 4, we do not make any assumptions regarding the sample selection mechanism, so this type of algorithms cannot be applied to our problem.

The second type consists of methods based on bootstrap sampling (Mordelet and Vert (2014); Yang, Liu, and Yang (2017)). These methods treat unlabeled instances as negative instances and bootstrap sampling is performed on the set U. A data set consisting of a random subset of U and positive instances is used to train base binary classifiers that form an ensemble. These methods exploit the advantages of Breiman (1996)'s bagging method. A key assumption of these methods is that potential positive instances in the set U and positive instances in the set P are generated from the same distribution. In the context of this paper, this means that the non-response bias is assumed to be zero. Therefore, we cannot employ methods of this type.

The third type are heuristic algorithms consisting of a "two-step strategy" (Liu, Lee, Yu, and Li (2002); Liu, Dai, Li, Lee, and Yu (2003)). The first step of such algorithms extracts a set of reliable negative instances, RN, which are likely to be negative based on criteria from the set U, and uses positive instances in the set P and reliable negative instances in the set RN to solve a PN classification problem. Note that in the first step the PU classification problem is replaced by a PN classification problem. In the second step, the classifier trained in the first step classifies unlabeled instances in the set U. The instances classified as negative are added to the set RN, and the retained positive instances and the newly classified reliable negative instances are used to train the classifier again. The iteration converges when no instances in the set $U \setminus RN$ are classified as negative. Note that the "two-step strategy" does not make any assumptions on the sample selection mechanism. Therefore, it can be applied to our problem if reliable negative instances can be correctly extracted.

The last type of algorithm consists of methods developed to solve "one-class classification," where only positive instances are used for training a classifier (Li, Guo, and Elkan (2011)). Intuitively, the underlying idea behind these methods is that the classifier fitted using only positive instances can identify potential positive instances in the set U. According to Khan and Madden (2014), the algorithm called one-class support vector machine (hereafter "one-class SVM") proposed by Tax and Duin (1999, 2004) has been widely applied to solve "one-class classification" problems. Importantly, one-class SVM allows assuming that potential positive instances in the set U and positive instances in the set P are generated from different distributions. In addition, although one-class SVM performs poorly on data sets with only a small number of positive instances and a large number of unlabeled instances, when a relatively large number of positive instances is available, as in the analysis here, the algorithm can be expected to fit the data set well.

6.2 Our Algorithm

Given the considerations in the previous subsection, we tackle the PU classification problem using an algorithm that combines the "two-step strategy" and one-class SVM as follows.[2]

Step 1 Step 1 extracts a set of reliable negative instances using one-class SVM. In our problem setting, we have no information that would allow us to identify reliable negative instances *a priori*, so the instances mechanically classified by one-class SVM as "outside of the target class" are treated as reliable negatives.

Formally, one-class SVM is trained on instances of two classes: a target class and an outlier class. Suppose that the target class is distributed in a hypersphere characterized by two parameters: a center a and a radius R. One-class SVM solves the following minimization problem:

$$\min \left[R^2 + C \sum_{h}^{M} \xi_h \right], \text{ subject to } \|\boldsymbol{x}_h - a\|^2 \leq R^2 + \xi_h,\ \xi_h \geq 0,\ \forall h,$$

where ξ_h denotes slack variables, C is a misclassification penalty, M indicates the number of instances used for training, and $\|\cdot\|$ represents the l_2 norm.

The hypersphere contains M non-DK instances in the set P. Intuitively, the M non-DK instances within the hypersphere (classified as the target class) are relatively homogenous; the non-DK instances on the boundary or outside the boundary (classified as the outlier class) are different in firms' characteristics from any instances within the hypersphere. We treat the instances outside of the target class as reliable negative instances. Note that, following Tax and Duin (2004), we use the Gaussian kernel for the calculation of l_2 norms to obtain more flexible data descriptions.

Step 2 In Step 2, we first train a classifier on positive instances and the reliable negative instances extracted in Step 1. Here, we employ the following logistic regression

[2] We implement our algorithm in R language. See Appendix A.

classifier:

$$\Pr(y = 1 \mid \boldsymbol{x}) = [1 + \exp(-\{\beta_0 + \beta_1 x_1 +, \ldots, +\beta_k x_k\})]^{-1}.$$

The parameters are estimated employing maximum likelihood estimation. Next, using the trained classifier, we identify negative instances in the set $U \setminus RN$ satisfying $\hat{\Pr}(y = 1 \mid \boldsymbol{x}) < 0.5$ and put them in the set RN. While retaining the size of the set of positive instances denoted by L, we again train the classifier on the set L and the newly defined set RN. The iteration stops when no instances in the set $U \setminus RN$ satisfy $\hat{\Pr}(y = 1 \mid \boldsymbol{x}) < 0.5$.

6.3 Settings of Hyper-Parameters

In machine learning, parameters that cannot be estimated by training classifiers on a data set are called hyper-parameters to distinguish them from model parameters. Our algorithm involves several hyper-parameters in each step. In this subsection, we summarize our settings of the hyper-parameters.

Step 1 To start, we set the misclassification penalty for one-class SVM, C. As shown by Tax and Duin (2004), the parameter is given by $C = 1/\nu M$, where ν indicates the fraction of outliers. We set the parameter ν instead of C to interpret the obtained results intuitively. In our problem setting, the parameter ν should be set at a sufficiently small value to accurately extract reliable negative instances. Here, the parameter is set at $\nu = 0.02$, meaning that we assume that the proportion of outliers in the set P is 2 percent.

Next, we set the width parameter of the Gaussian kernel. Intuitively, a sufficiently small value of the parameter allows us to make a flexible boundary, meaning that reliable negative instances could be correctly classified. Here, the width parameter is set at 0.05.

Finally, the number of non-DK instances used in Step 1 is set at $M = \#P$,[3] which we expect to be large enough to obtain accurate results.

[3] $\#P$ denotes the number of instances in the set P.

Step 2 In Step 2, we set the number of non-DK instances denoted by L to train the logistic regression classifier. In Step 2, we iteratively run the logistic regression classifier using non-DK instances and the reliable negative instances extracted in Step 1 while keeping L constant. If we use all non-DK instances in the set P, *i.e.*, $L = \#P$, when we first run the classifier, the training data with the set L and the set RN could be imbalanced. As discussed by Liu, Lee, Yu, and Li (2002), the optimal size of L is not obvious *a priori*, so we set it at $L = \#RN$. Meanwhile, the non-DK instances in the set L are randomly sampled from the set P. Note that, in the process of iteration in Step 2, the proportion of negative instances increases.

6.4 Independent Variables

Independent variables denoted by the k-dimensional vector $\boldsymbol{x}$ mainly consist of the following firms characteristics: (1) sector (31 categories), (2) firm size (20 categories), (3) head office location (by prefecture: 47 categories), (4) firms' subjective judgment on business conditions, their capacity utilization, and so on (21 questions; each question has four options for answers), and (5) a time dummy (our observation period: 15 quarters). All of these variables are categorical variables. When we count them as binary variables, the number of independent variables is $k = 192$. Note that the set of independent variables is identical in each step of our algorithm.

7 Results

This section reports the "true classes" uncovered by our algorithm described in Subsection 6.2. We then verify the result.

7.1 Uncovered "True Class"

The results are shown in Table 2. Strikingly, a huge number of the DK instances are classified into "don't know" as the "true class," as shown in row $\#U_N$. At the same time, there are indeed a small number of the DK instances classified into "know" as the "true class," as shown in row $\#U_P$. The result that the number of instances classified into the set U_P is not zero is in line with the argument in Subsection 5.2 that a fraction of the DK instances departs from their "true class."

Table 2: Result of uncovered "true class"

	General prices					
	1-year ahead		3-year ahead		5-year ahead	
$\#P$	135600	(0.85)	111884	(0.70)	95291	(0.60)
$\#U$	23958	(0.15)	47674	(0.30)	64267	(0.40)
$\#U_P$	315	(0.01)	407	(0.01)	411	(0.01)
$\#U_N$	23643	(0.99)	47267	(0.99)	63856	(0.99)
	Output prices					
	1-year ahead		3-year ahead		5-year ahead	
$\#P$	146272	(0.92)	122817	(0.77)	102679	(0.64)
$\#U$	13286	(0.08)	36741	(0.23)	56879	(0.36)
$\#U_P$	289	(0.02)	238	(0.01)	188	(0.00)
$\#U_N$	12997	(0.98)	36503	(0.99)	56691	(1.00)

Note: The weights are reported in brackets.

Specifically, the fraction of instances classified into "don't know" as their "true class" is more than 98 percent for all time horizons and for both general prices and output prices. Although not shown in the table, each survey includes an extremely small number of instances that depart from the "true class." For example, regarding five years ahead general prices, the December 2014 survey contains only 16 instances.

7.2 Verification of the Results

In this subsection, we verify the result presented in the previous subsection. Specifically, we check the sectoral and firm-size distributions of the uncovered "true classes" to verify the results obtained in the previous subsection.

Sector Table 3 presents the share of manufacturers and of non-manufacturers in each "true class."

Table 3: Share of manufacturers and non-manufacturers in each "true class"

	General prices					
	1-year ahead		3-year ahead		5-year ahead	
	Mfr.	Non-mfr.	Mfr.	Non-mfr.	Mfr.	Non-mfr.
P	0.40	0.60	0.39	0.61	0.39	0.61
U	0.44	0.56	0.44	0.56	0.43	0.57
U_P	0.25	0.75	0.35	0.65	0.39	0.61
U_N	0.45	0.55	0.44	0.56	0.43	0.57
	Output prices					
	1-year ahead		3-year ahead		5-year ahead	
	Mfr.	Non-mfr.	Mfr.	Non-mfr.	Mfr.	Non-mfr.
P	0.40	0.60	0.39	0.61	0.39	0.61
U	0.45	0.55	0.46	0.54	0.44	0.56
U_P	0.42	0.58	0.43	0.57	0.29	0.71
U_N	0.45	0.55	0.46	0.54	0.44	0.56

As can be seen in Table 3, for all time horizons and for both general prices and output prices, the share of non-manufacturers in the set U_P is higher than in the set U_N. For example, for one year ahead general prices, while the share of non-manufacturers in the set U_N is 55 percent, that in the set U_P is 75 percent. This means that the non-manufacturer shares in the set U_P are closer to those in the set P than in the set U_N. In fact, for five years ahead general prices, the shares of non-manufacturer in the set U_P and in the set P are both 61 percent. The similarity of the shares in the set U_P and the set P suggests that our algorithm works well in that it learns the sectoral distribution of firms in the set P.

Firm Size Next, we compare the size of firms in the "true class." Table 4 presents the mean values of capital stock by "true class."

As can be seen in Table 4, the mean value of the capital stock in the set U_P is statistically significantly lower than that in the set U_N in almost all cases (the only exception is the case of five years ahead general prices). At the same time, the mean values of the capital stock in the set U_P are not significantly different from those in the set P in almost all cases (the exceptions are three years ahead and five years ahead output prices). This

Table 4: Capital stock by "true class"

	General prices		
	1-year ahead	3-year ahead	5-year ahead
P	2610	2088	1932
	(2517,2703)	(2012,2164)	(1854,2010)
U	10805	7953	6670
	(10082,11528)	(7540,8366)	(6356,6983)
U_P	2106	3159	4872
	(957,3253)	(48,6270)	(533,9211)
U_N	10921	7994	6681
	(10188,11654)	(7578,8410)	(6367,6996)
	Output prices		
	1-year ahead	3-year ahead	5-year ahead
P	2663	2174	2151
	(2572,2755)	(2098,2250)	(2068,2234)
U	16798	9412	6890
	(15544,18052)	(8887,9937)	(6544,7236)
U_P	5693	519	483
	(1207,10178)	(201,837)	(182,785)
U_N	17045	9470	6911
	(15767,18323)	(8942,9998)	(6564,7259)

Notes: The table reports the mean with a 95% confidence interval using the normal distribution in brackets. The unit is million yen.

means that the distribution of capital stock in the set U_P is closer to that in the set P than that in the set U_N. This similarity suggests that our algorithm performs well in that the mean of firm size in the set P is learned almost correctly.

8 Non-Response Bias

In this section, we estimate the counterfactual inflation expectations for the DK instances classified into the set U_P. We then examine the magnitude of the non-response bias based on the estimation results.

8.1 Counterfactual Inflation Expectations

We estimate the counterfactual inflation expectations for the DK instances in the set U_P using a matching estimator. Specifically, we reuse the set of independent variables shown in Subsection 6.4 to implement the propensity score matching developed by Rosenbaum and Rubin (1983), which matches the instances in the set U_P to the instances in the set P one by one. This allows us to identify firm h in the set P that is most similar to firm j in the set U_P based on the propensity score values. We define the counterfactual inflation expectations of firm j as the inflation expectations held by firm h.

Figure 1: Estimated counterfactual inflation expectations

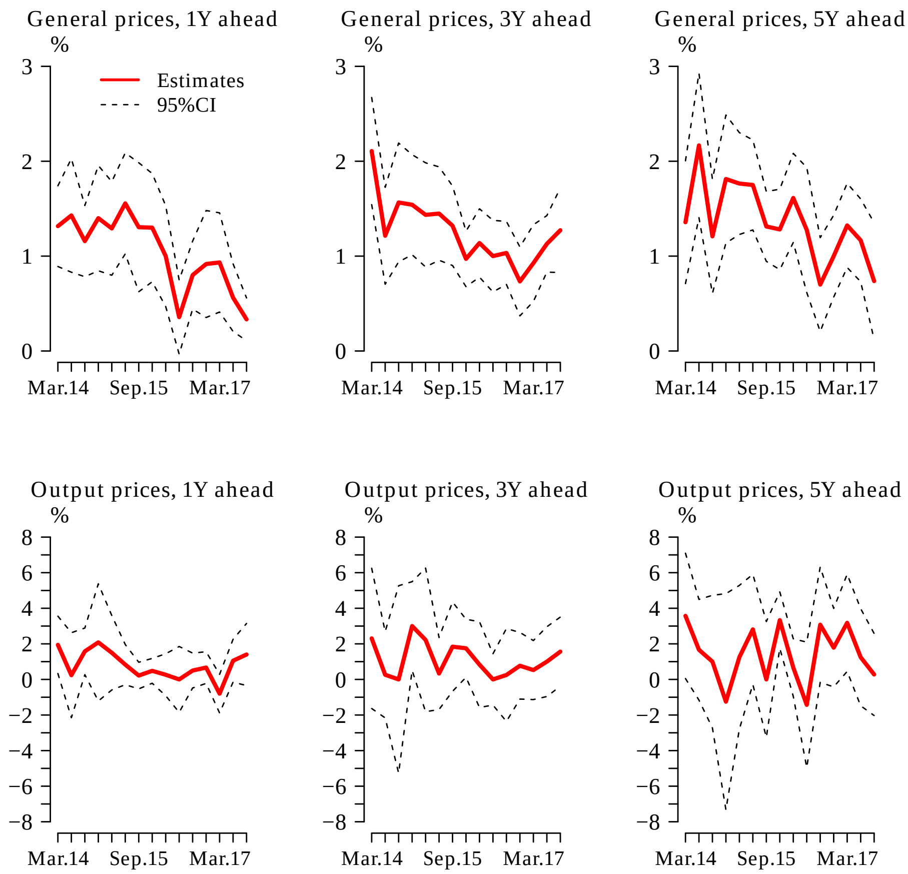

Figure 1 shows the mean value of the estimated counterfactual inflation expectations for the firms in the set U_P (the red line in the figure) and the 95 percent confidence intervals calculated assuming that the parameter is normally distributed (the dotted line in the figure). As can be seen, the estimated counterfactual inflation expectations fluctuate substantially because of the extremely small number of firms classified into the set U_P, as discussed in Subsection 7.1. Moreover, for both general prices and output prices, the longer the time horizon, the larger the extent of the fluctuations. This pattern with regard to the time horizon is similar to the actual inflation expectations. That is, as shown in Uno, Naganuma, and Hara (2018b), with regard to the inflation expectations in the *Tankan* survey, the longer the time horizon, the larger is the standard deviation. Therefore, the pattern with regard to the time horizon in Figure 1 implies that the firms in the set U_P may have been drawn from the same distribution underlying the firms in the set P. We discuss this point further in the next subsection.

8.2 Non-Response Bias

We denote the mean value of the estimated counterfactual inflation expectations of firms in the set U_P and the mean value of inflation expectations of firms in the set P by μ_{U_P} and μ_P, respectively. Note that μ_P corresponds to the official figures for inflation expectations in the *Tankan* survey. When letting q be $q = \#U_P/(\#U_P + \#P)$, the non-response bias, B, is defined as $B = q(\mu_P - \mu_{U_P})$.

Table 5 presents the result of the calculated non-response bias. As shown in the table, the non-response bias in inflation expectations in the *Tankan* survey is zero for all time horizons and for both general prices and output prices. Looking at the results in more detail, the bias is also zero for each firm size and each sector. Two factors contribute to this result. First, the number of firms classified into the set U_P is extremely small, as discussed in Subsection 7.1. That is, the q is quite small for all time horizons and for both general prices and output prices.

The second factor is that the counterfactual inflation expectations, μ_{U_P}, are not statistically significantly different from the corresponding official figures. That is, from a statistical perspective, $\mu_P - \mu_{U_P}$ is equal to zero. This implies that the distribution of firms in the set U_P in terms of firm characteristics such as firm size, sector, and so

Table 5: Non-response bias

	General prices								
	1-year ahead			3-year ahead			5-year ahead		
	μ_P	μ_{U_P}	B	μ_P	μ_{U_P}	B	μ_P	μ_{U_P}	B
Total	1.1	1.1	0.0	1.3	1.2	0.0	1.4	1.4	0.0
	(0.00)	(0.07)		(0.00)	(0.06)		(0.00)	(0.07)	
By firm size									
Large	0.8	0.7	0.0	1.0	1.0	0.0	1.0	1.1	0.0
	(0.01)	(0.13)		(0.01)	(0.11)		(0.01)	(0.13)	
Medium	1.0	1.1	0.0	1.2	1.3	0.0	1.3	1.1	0.0
	(0.01)	(0.14)		(0.01)	(0.11)		(0.01)	(0.15)	
Small	1.2	1.2	0.0	1.5	1.3	0.0	1.5	1.6	0.0
	(0.00)	(0.10)		(0.01)	(0.08)		(0.01)	(0.09)	
By sector									
Mfr. 1	1.1	1.3	0.0	1.3	1.3	0.0	1.3	1.3	0.0
	(0.01)	(0.21)		(0.01)	(0.17)		(0.01)	(0.16)	
Mfr. 2	1.0	1.1	0.0	1.3	1.2	0.0	1.3	1.3	0.0
	(0.01)	(0.16)		(0.01)	(0.11)		(0.01)	(0.15)	
Non-mfr.	1.1	1.1	0.0	1.3	1.2	0.0	1.4	1.4	0.0
	(0.00)	(0.08)		(0.00)	(0.08)		(0.01)	(0.09)	
	Output prices								
	1-year ahead			3-year ahead			5-year ahead		
	μ_P	μ_{U_P}	B	μ_P	μ_{U_P}	B	μ_P	μ_{U_P}	B
Total	0.6	0.7	0.0	1.3	1.1	0.0	1.6	1.5	0.0
	(0.01)	(0.19)		(0.01)	(0.32)		(0.02)	(0.40)	
By firm size									
Large	0.4	0.5	0.0	0.6	0.1	0.0	0.7	1.2	0.0
	(0.02)	(0.60)		(0.03)	(0.84)		(0.04)	(0.66)	
Medium	0.5	0.3	0.0	1.0	0.4	0.0	1.1	1.7	0.0
	(0.01)	(0.26)		(0.02)	(0.50)		(0.03)	(0.79)	
Small	0.8	1.0	0.0	1.7	1.8	0.0	2.2	1.5	0.0
	(0.01)	(0.24)		(0.02)	(0.44)		(0.02)	(0.59)	
By sector									
Mfr. 1	0.8	0.7	0.0	1.5	2.0	0.0	1.8	1.5	0.0
	(0.02)	(0.53)		(0.04)	(0.90)		(0.05)	(0.66)	
Mfr. 2	0.1	−0.3	0.0	0.3	−1.3	0.0	0.3	0.4	0.0
	(0.02)	(0.38)		(0.03)	(0.71)		(0.04)	(1.07)	
Non-mfr.	0.8	1.0	0.0	1.7	1.9	0.0	2.1	1.7	0.0
	(0.01)	(0.24)		(0.02)	(0.34)		(0.02)	(0.56)	

Notes: Large firms are defined as firms with capital of at least 1 billion yen; Medium firms are defined as firms with capital of at least 100 million yen but less than 1 billion yen; Small firms are defined as firms with capital of at least 20 million yen but less than 100 million yen. Mfr. 1, Mfr. 2, and Non-mfr. denote manufacturing (basic materials), manufacturing (processing), and non-manufacturing, respectively. The standard errors are reported in brackets.

on, is similar to that of firms in the set P. In fact, as discussed in Subsection 7.2, the distribution of firms in terms of firm size and sector in the two groups is very similar.

In sum, the very small number of firms classified into the set U_P can be regarded as drawn from the same distribution as the firms in the set P. In terms of the example presented in the introduction, this means that a few randomly chosen firms respond "don't know" even though their "true" answer is "Brazil." In contrast to the qualitative argument by Coibion, Gorodnichenko, Kumar, and Pedemonte (2018) that firms non-randomly choose the option "don't know," we quantitatively show that the choice is random.

9 Conclusion

In this paper, we used machine learning techniques to extract firms that respond "don't know" to questions concerning inflation expectations in the *Tankan* survey even though they seem to have quantitative answers. We then estimated the counterfactual inflation expectations for such firms based on a propensity score matching estimator.

Our findings can be summarized as follows. First, there is indeed a fraction of firms that respond "don't know" even though they seem to "know" something in a sense. Second, the number of such firms is quite small. They are mostly small firms and firms in non-manufacturing. Third, the estimated counterfactual inflation expectations of such firms are not statistically significantly different from the corresponding official figures in the *Tankan* survey. Fourth and finally, based on the above findings, the non-response bias in firms' inflation expectations likely is statistically negligible.

References

Breiman, L. (1996): "Bagging Predictors," *Machine Learning*, 24(2), 123–140.

Cabinet Office (2017): "Research on Missing Value Imputation (in Japanese)," available at: https://www.esri.cao.go.jp/jp/stat/report/report_all_detail.pdf.

Coibion, O., Y. Gorodnichenko, S. Kumar, and M. Pedemonte (2018): "Inflation Expectations as a Policy Tool?," available at: https://sites.google.com/site/ocoibion/CGKP%202018-06-26%20-%20NBER%20version.pdf?attredirects=0&d=1.

Elkan, C., and K. Noto (2008): "Learning Classifiers from Only Positive and Unlabeled Data," *Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 213–220.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.

Hirakawa, T., and J. Hatogai (2013): "A Study of Missing Value Imputation in Business Surveys: The Case of the Short-Term Economic Survey of Enterprises in Japan (Tankan)," Bank of Japan Working Paper Series, No.13-E-9.

Hoshino, T. (2010): "Semiparametric Estimation under Nonresponse in Survey and Sensitivity Analysis: Application to the 12th Survey of the Japanese National Character," *Proceedings of the Institute of Statistical Mathematics*, 58(1), 3–23, (In Japanese).

Khan, S. S., and M. G. Madden (2014): "One-class Classification: Taxonomy of Study and Review of Techniques," *The Knowledge Engineering Review*, 29(3), 345–374.

Lee, W. S., and B. Liu (2003): "Learning with Positive and Unlabeled Examples Using Weighted Logistic Regression," *Proceedings of the 20th international conference on machine learning*, pp. 448–455.

Li, W., Q. Guo, and C. Elkan (2011): "A Positive and Unlabeled Learning Algorithm for One Class Classification of Remote-sensing Data," *IEEE Transactions on Geoscience and Remote Sensing*, 49, 717–725.

Little, R. J. A., and D. B. Rubin (2002): *Statistical Analysis with Missing Data.* Wiley, New York, 2nd edn.

Liu, B., Y. Dai, X. Li, W. S. Lee, and P. S. Yu (2003): "Building Text Classifiers using Positive and Unlabeled Examples," *Proceedings of the third IEEE international conference on data mining*, pp. 179–186.

Liu, B., W. S. Lee, P. S. Yu, and X. Li (2002): "Partially Supervised Classification of Text Documents," *Proceedings of the 19th international conference on machine learning*, pp. 387–394.

Mordelet, F., and J.-P. Vert (2014): "A Bagging SVM to Learn from Positive and Unlabeled Examples," *Pattern Recognition Letters*, 37, 201–209.

Rosenbaum, P. R., and D. B. Rubin (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70(1), 41–55.

Tax, D. M. J., and R. P. W. Duin (1999): "Support Vector Domain Description," *Pattern Recognition Letters*, 20, 1191–1199.

——— (2004): "Support Vector Data Description," *Machine Learning*, 54(1), 45–66.

Tsuchiya, T. (2010): "Two-stage Non-response Bias Adjustment Using Variables on the Survey-orienting Character for the Survey on the Japanese National Character," *Proceedings of the Institute of Statistical Mathematics*, 58(1), 25–38, (In Japanese).

Uno, Y., S. Naganuma, and N. Hara (2018a): "New Facts about Firms' Inflation Expectations: Short- versus Long-term Inflation Expectations," Bank of Japan Working Paper Series, No.18-E-15.

——— (2018b): "New Facts about Firms' Inflation Expectations: Simple Tests for a Sticky Information Model," Bank of Japan Working Paper Series, No.18-E-14.

Utsunomiya, K., and K. Sonoda (2004): "Methodology for Handling Missing Values of "Short-term Economic Survey of All Enterprises in Japan (Tankan)"," *The Economic Review*, 55(3), 217–229.

Yang, P., W. Liu, and J. Yang (2017): "Positive Unlabeled Learning via Wrapper-based Adaptive Sampling," *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp. 3273–3279.

Appendix A R code

```
library(dplyr)
library(kernlab)

data_p <- data %>%
  filter(know == 1)
data_u <- data %>%
  filter(know == 0)

var_names <- colnames(data)
f <- as.formula(paste("know~",
                paste(var_names[var_names != "know"],
                collapse = "+")))

# step 1 #
sigma <- 0.05
outlier <- 0.02
M <- length(data_p$know)
occ_svm <- ksvm(x = f, data = data_p, type = "one-svc",
                kernel = "rbfdot", kpar = list(sigma = sigma),
                C = 1/(outlier*M),
                nu = outlier)
occ_pred <- predict(occ_svm, newdata = data_u)
L <- length(occ_pred) - sum(occ_pred)

# step 2 #
new_positive <- occ_pred
data_u_temp_U <- data_u
dif_U <- 1

repeat{
  if(dif_U > 0){
    U_size_temp <- length(data_u_temp_U$know)
    data_u_temp_N <- data_u_temp_U %>%
      cbind(new_positive) %>%
      filter(new_positive == FALSE) %>%
      select(-new_positive)
    data_u_temp_U <- data_u_temp_U %>%
      cbind(new_positive) %>%
      filter(new_positive == TRUE) %>%
      select(-new_positive)

    Z <- sample(nrow(data_p), size = L)
    glm_fit <- glm(formula = f,
                   data = rbind(data_p[Z,], data_u_temp_N),
                   family = binomial)
    pred <- predict(glm_fit, newdata = data_u_temp_U, type = "response")
    new_positive <- as.data.frame(pred$pred >= 0.5)
    colnames(new_positive) <- c("new_positive")
    dif_U <- U_size_temp - length(data_u_temp_U$know)
  } else {
      break
  }
}
```