Economics Letters 156 (2017) 68–73

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A Bayesian analysis of binary misclassification

Christopher R. Bollinger [a], Martijn van Hasselt [b,*]

[a] Department of Economics, The University of Kentucky, United States

[b] Department of Economics, The University of North Carolina Greensboro, 516 Stirling Street, Greensboro, NC 27412, United States

HIGHLIGHTS

- The mean of a misclassified binary variable is in general only partially identified.
- The exact Bayesian posterior for the mean is derived for several intuitive priors.
- Posterior calculations are feasible without Markov chain Monte Carlo simulation.
- Parts of the identified set for the mean are a posteriori more likely than others.

ARTICLE INFO

Article history:
Received 5 April 2017
Accepted 10 April 2017
Available online 20 April 2017

JEL classification:
C11
C18
C21
C46

Keywords:
Bayesian inference
Partial identification
Misclassification

ABSTRACT

We consider Bayesian inference about the mean of a binary variable that is subject to misclassification error. If the error probabilities are not known, or cannot be estimated, the parameter is only partially identified. For several reasonable and intuitive prior distributions of the misclassification probabilities, we derive new analytical expressions for the posterior distribution. Our results circumvent the need for Markov chain Monte Carlo simulation. The priors we use lead to regions in the identified set that are a posteriori more likely than others.

1. Introduction

We consider the problem of inference for the population mean of a binary variable that suffers from measurement error. That is, there is some nonzero probability that observations are misclassified. This type of model has a long history in both statistics and econometrics (e.g. Neyman, 1950; Bross, 1954; Aigner, 1973). If the misclassification rates are known, the mean is identified and can be estimated without bias. If the rates are unknown but a set of correctly classified observations is available (i.e., validation data), the mean is also identified and estimable (Tenenbein, 1970). In the absence of validation data, however, it is well known that under mild conditions the population mean can be non-trivially bounded. It is then said to be *partially* identified and the collection of feasible parameter values is called the identified set. The bounds of this set can usually be estimated consistently.

In the classical approach to inference (e.g. Bollinger, 1996; Imbens and Manski, 2004; Molinari, 2008), a confidence interval for the parameter takes the form of the estimated bounds, plus a multiple of their standard errors. The resulting region in the parameter space, however, can be quite wide and classical inference provides no additional information about the location of the parameter within the bounds. In particular applications, a researcher's intuition or knowledge of previous studies may lead him or her to believe that the true parameter is, for example, likely to be closer to the estimated upper bound. However, such prior knowledge cannot be easily exploited or incorporated into a classical analysis.

In this paper we take a Bayesian approach to inference. Our analysis relies on key insights of Poirier (1998) and Moon and Schorfheide (2012). Given that some parameters are not identified, extra care must be given to the specification of prior distributions, since even asymptotically these priors will remain an important component of posterior inference. Some previous Bayesian studies of misclassification achieved identification through the prior (Gaba and Winkler, 1992; Joseph et al., 1995; Evans et al., 1996;

* Corresponding author.
E-mail address: mnvanhas@uncg.edu (M. van Hasselt).

http://dx.doi.org/10.1016/j.econlet.2017.04.011

Rahme et al., 2000). In contrast, we consider a variety of priors that explicitly incorporate the parameter bounds inherent in the model. These priors can be considered intermediate between weak information leading only to partial identification, and strong information leading to full identification. A second contribution is that we derive exact, analytical expressions for the posterior and therefore do not have to rely on Markov chain Monte Carlo sampling.

Although sensitivity to the prior distribution is sometimes seen as a weakness of the Bayesian approach, we believe that it facilitates a sensitivity analysis with respect to assumptions about misclassification rates. The analysis examines how additional prior information about these rates affects what the researcher can learn about the population mean. Of course, the identification problem is by no means eliminated through the use of a Bayesian prior. Instead, the prior allows us to easily incorporate varying amounts of information and examine the effect on posterior inferences. Our results show that under a number of reasonable prior assumptions, the posterior is far from uniform and, relative to a classical analysis, provides additional information about the location of the population mean within the identified set.

The remainder of this paper is organized as follows: Section 2 discusses misclassification and partial identification, as well as a number of intuitive prior distributions that range from less to more informative about the probability of a misclassification error. The resulting finite-sample posterior distributions are presented in Section 3. Section 4 provides concluding remarks. Derivations of some of the results are collected in the Appendix.

2. The model

2.1. Misclassification and parameter bounds

Let $Z \in \{0, 1\}$ be a binary random variable with $P(Z = 1) = \pi$. Instead of observing Z, we observe $X \in \{0, 1\}$, which may suffer from misclassification error:

$$P(X = 1|Z) = p(1 - Z) + (1 - q)Z. \tag{1}$$

Here, p is the probability of a false positive, whereas q is the probability of a false negative. We assume, as is typical in the literature, that $p + q < 1$. This ensures that the covariance between Z and X is positive. The mean of X can be written as $\mu = \pi(1 - q) + (1 - \pi)p$, which implies the following bounds on the misclassification rates:

$$0 \le p \le \mu, \quad 0 \le q \le 1 - \mu. \tag{2}$$

The parameter π, however, can take values over the entire unit interval. For example, if $p = \mu$, then $\pi = 0$, regardless of the value of q. Similarly, if $q = 1 - \mu$, then $\pi = 1$. Hence, π is completely unidentified.

Given a random sample $\mathbf{X} = (X_1, \ldots, X_n)$, let $n_1 = \sum_{i=1}^n X_i$ and $n_0 = n - n_1$ be the observed number of ones and zeros, respectively. The likelihood $f(\mathbf{X}|\mu) = \mu^{n_1}(1 - \mu)^{n_0}$ is a function of μ only, so that

$$\begin{aligned} f(\pi, \mu|\mathbf{X}) &\propto f(\mathbf{X}|\mu) \cdot f(\mu) \cdot f(\pi|\mu) \\ &\propto f(\mu|\mathbf{X}) \cdot f(\pi|\mu), \end{aligned} \tag{3}$$

and the posterior is the product of the marginal posterior of the identified parameter and the conditional prior of the unidentified parameter (Poirier, 1998; Moon and Schorfheide, 2012). If the true value of the population mean of X is μ_0, then under standard regularity conditions the posterior distribution of μ will increasingly concentrate around μ_0 as $n \to \infty$ (e.g. Heyde and Johnstone, 1979; Chen, 1985). This has an important implication for the posterior of π. Eq. (3) implies that

$$\begin{aligned} f(\pi|\mathbf{X}) &= \int f(\pi, \mu|\mathbf{X}) d\mu \\ &\propto \int f(\mu|\mathbf{X}) f(\pi|\mu) d\mu, \end{aligned}$$

so that the posterior of π is a mixture of conditional priors. As the sample size increases, the mixing distribution $f(\mu|\mathbf{X})$ – namely the marginal posterior of μ – becomes asymptotically degenerate at $\mu = \mu_0$ and $f(\pi|\mathbf{X})$ converges to $f(\pi|\mu_0)$.[1]

2.2. Prior distributions

In this section we examine a number of prior distributions that are increasingly informative about the misclassification rates. The first prior is a uniform distribution for μ, combined with conditional priors $p|\mu \sim U(0, \mu)$ and $q|\mu \sim U(0, 1 - \mu)$ that are uniform on the identified set:

$$f_1(\mu, p, q) = \frac{1}{\mu(1 - \mu)} \mathbf{1}\{0 \le p \le \mu, 0 \le q \le 1 - \mu\}. \tag{4}$$

It follows that $f_1(\mu, p, \pi) = \frac{\mu - p}{\mu(1-\mu)\pi^2}$. Using the relation between μ, p and q, and letting q range from 0 to $1 - \mu$, it follows that $\max\{0, (\mu - \pi)/(1 - \pi)\} \le p \le \mu$. Since $f_1(\pi|\mu) = f_1(\pi, \mu)$ (because μ has a uniform prior), we find

$$\begin{aligned} f_1(\pi|\mu) &= \mathbf{1}\{\pi > \mu\} \int_0^\mu \frac{(\mu - p)}{\mu(1 - \mu)\pi^2} dp \\ &\quad + \mathbf{1}\{\pi \le \mu\} \int_{\frac{\mu - \pi}{1 - \pi}}^\mu \frac{(\mu - p)}{\mu(1 - \mu)\pi^2} dp \\ &= \mathbf{1}\{\pi > \mu\} \frac{\mu}{2(1 - \mu)\pi^2} + \mathbf{1}\{\pi \le \mu\} \frac{(1 - \mu)}{2\mu(1 - \pi)^2}. \end{aligned} \tag{5}$$

The second prior expresses the belief that, conditional on μ, lower misclassification rates are more likely than higher ones. We combine a uniform prior for μ with 'power-type' conditional priors for p and q (proportional to $p^{-1/2}$ and $q^{-1/2}$ on the identified set). This yields the prior

$$f_2(\mu, p, q) = \frac{1}{4\sqrt{\mu(1 - \mu)pq}} \mathbf{1}\{0 \le p \le \mu, 0 \le q \le 1 - \mu\}. \tag{6}$$

This implies the following joint prior distribution for (μ, p, π):

$$f_2(\mu, p, \pi) = \frac{1}{4\pi\sqrt{\pi\mu(1 - \mu)}} \cdot \frac{\mu - p}{\sqrt{p^2(1 - \pi) + p(\pi - \mu)}},$$

where $\max\{0, (\mu - \pi)/(1 - \pi)\} \le p \le \mu$. It is shown in the Appendix that

$$\begin{aligned} f_2(\pi|\mu) &= \frac{\mu(1 - \pi) + \frac{1}{2}(\pi - \mu)}{4\pi(1 - \pi)\sqrt{\pi(1 - \pi)\mu(1 - \mu)}} \\ &\quad \times \log\left(\frac{\pi - \mu + 2(1 - \pi)\mu + 2\sqrt{\pi(1 - \pi)\mu(1 - \mu)}}{|\pi - \mu|}\right) \\ &\quad - \frac{1}{4\pi(1 - \pi)}. \end{aligned} \tag{7}$$

The third prior expresses the belief that, with probability λ, the misclassification error is *symmetric*. In that case, $p = q$ and false positive and false negatives are equally likely. We maintain the

[1] The argument given here also applies to the unidentified parameter p and q. In large samples $f(p|\mathbf{X})$ and $f(q|\mathbf{X})$ will converge to $f(p|\mu_0)$ and $f(q|\mu_0)$, respectively.

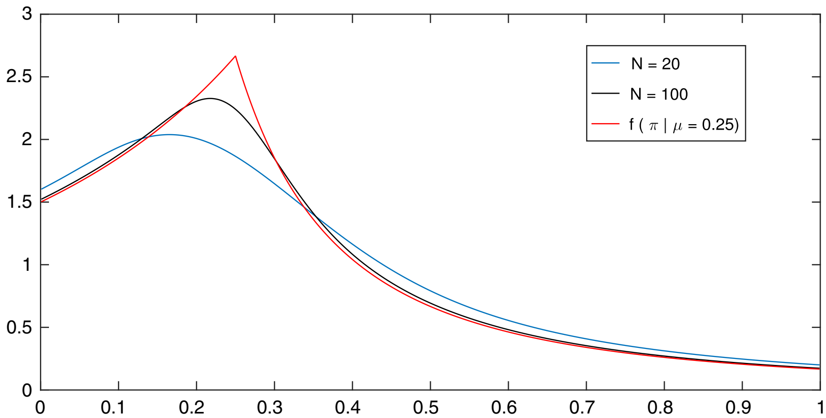

Fig. 1. Finite-sample posteriors $f_1(\pi|\mathbf{X})$ and conditional prior $f_1(\pi|\mu = 0.25)$.

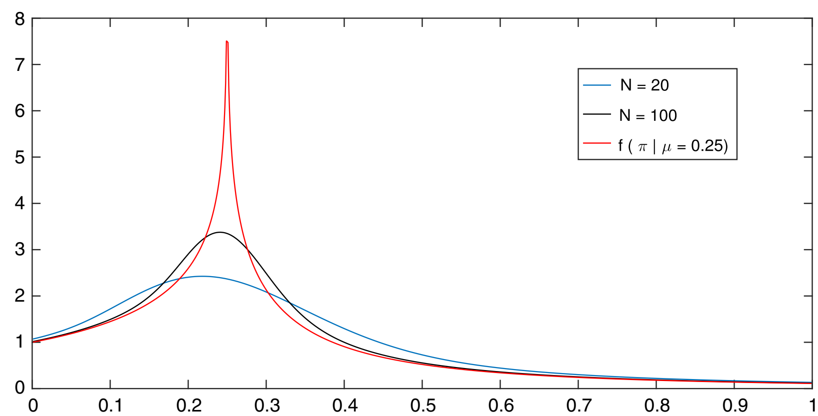

Fig. 2. Finite-sample posteriors $f_2(\pi|\mathbf{X})$ and conditional prior $f_2(\pi|\mu = 0.25)$.

assumption that Z and X are positively correlated, so that $p < \frac{1}{2}$. From $\mu = (1-\pi)p + \pi(1-p)$, it now follows that $\pi \in [0, \mu]$ if $\mu < \frac{1}{2}$ and $\pi \in [\mu, 1]$ if $\mu > \frac{1}{2}$.[2] From (2) it also follows that $p \leq \min\{\mu, 1-\mu\}$. Thus, symmetry of the misclassification error shrinks the identified set. A conditional prior that imposes the restriction $p = q$ and is uniform over the identified set is

$$\tilde{f}(p, q|\mu) = \frac{1}{\mu}\mathbf{1}\left\{p = q, p \leq \mu < \tfrac{1}{2}\right\} + \frac{1}{1-\mu}\mathbf{1}\left\{p = q, \frac{1}{2} < \mu \leq 1-p\right\}. \quad (8)$$

Using a uniform marginal prior for μ, the joint prior is

$$f_3(\mu, p, q) = \frac{\lambda}{\mu}\mathbf{1}\left\{p = q, p \leq \mu < \tfrac{1}{2}\right\} + \frac{\lambda}{1-\mu}\mathbf{1}\left\{p = q, \frac{1}{2} < \mu \leq 1-p\right\} + \frac{(1-\lambda)}{\mu(1-\mu)}\mathbf{1}\{0 \leq p \leq \mu, 0 \leq q \leq 1-\mu\}. \quad (9)$$

Thus, with probability λ the misclassification error is believed to be symmetric (and p has a uniform distribution over the identified set), and with probability $(1-\lambda)$ the error is asymmetric. Using a change of variables to (π, μ), it can be shown that

$$f_3(\pi|\mu) = \frac{\lambda(1-2\mu)}{\mu(1-2\pi)^2}\mathbf{1}\left\{\pi \leq \mu < \tfrac{1}{2}\right\} + \frac{\lambda(2\mu-1)}{(1-\mu)(1-2\pi)^2}\mathbf{1}\left\{\tfrac{1}{2} < \mu \leq \pi\right\} + \frac{(1-\lambda)\mu}{2(1-\mu)\pi^2}\mathbf{1}\{\pi > \mu\} + \frac{(1-\lambda)(1-\mu)}{2\mu(1-\pi)^2}\mathbf{1}\{\pi \leq \mu\}. \quad (10)$$

3. Main results

We now present analytical results for the finite-sample posteriors of π, using the priors discussed in the previous section. Derivations can be found in the Appendix. The posterior corresponding to $f_1(\mu, p, q)$ in Eq. (4) is given by

$$f_1(\pi|\mathbf{X}) = \frac{1}{2\pi^2}\left(\frac{n_1+1}{n_0}\right) I_{n_1+2,n_0}(\pi) + \frac{1}{2(1-\pi)^2}\left(\frac{n_0+1}{n_1}\right)(1 - I_{n_1,n_0+2}(\pi)), \quad (11)$$

where $B_{a,b}$ is the Beta function and $I_{a,b}(t)$ is the cumulative distribution function of the Beta distribution with parameters a and b.[3]

[2] Since $\pi = (\mu - p)/(1-2p)$, it follows that π is identified and equal to $\frac{1}{2}$ if $\mu = \frac{1}{2}$. Identification of π in this case hinges on $p < \frac{1}{2}$ being a strict inequality. If $p = \frac{1}{2}$ is allowed, this result breaks down.

[3] The function $I_{\alpha,\beta}(t)$ is also referred to as the regularized Beta function (Abramowitz and Stegun, 1964).

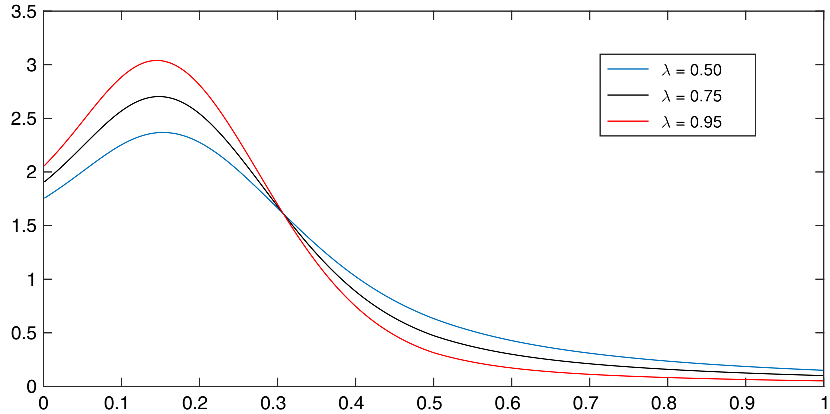

Fig. 3. Finite-sample posteriors $f_3(\pi|\mathbf{X})$ for different probabilities (λ) of error symmetry; $n = 20$.

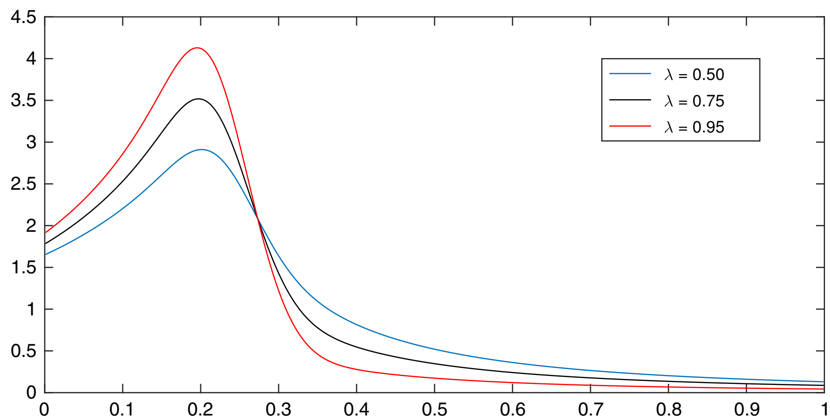

Fig. 4. Finite-sample posteriors $f_3(\pi|\mathbf{X})$ for different probabilities (λ) of error symmetry; $n = 100$.

For prior $f_2(\mu, p, q)$ in Eq. (6), the posterior of π is

$$f_2(\pi|\mathbf{X}) = \frac{1}{B_{n_1+1,n_0+1}} \int_0^1 \mu^{n_1}(1-\mu)^{n_0} f_2(\pi|\mu)d\mu, \tag{12}$$

where $f_2(\pi|\mu)$ is given in Eq. (7). This expression cannot be simplified any further. Finally, using the mixture prior in Eq. (9), the corresponding posterior of π is a mixture distribution

$$f_3(\pi|\mathbf{X}) = \lambda\tilde{f}(\pi|\mathbf{X}) + (1-\lambda)f_1(\pi|\mathbf{X}), \tag{13}$$

where $f_1(\pi|\mathbf{X})$ is the posterior in Eq. (11), $\tilde{f}(\pi|\mathbf{X})$ is given by

$$\tilde{f}(\pi|\mathbf{X}) = \begin{cases} \dfrac{1}{(1-2\pi)^2}\left[\left(\dfrac{n+1}{n_1}\right) I_{n_1,n_0+1}\left(\pi, \dfrac{1}{2}\right) \right. \\ \left. \quad - 2I_{n_1+1,n_0+1}\left(\pi, \dfrac{1}{2}\right)\right] \quad \text{if } \pi < \dfrac{1}{2} \\ \dfrac{1}{(1-2\pi)^2}\left[2\left(\dfrac{n_1+1}{n_0}\right) I_{n_1+2,n_0}\left(\dfrac{1}{2}, \pi\right) \right. \\ \left. \quad - \left(\dfrac{n+1}{n_0}\right) I_{n_1+1,n_0}\left(\dfrac{1}{2}, \pi\right)\right] \quad \text{if } \pi > \dfrac{1}{2}, \end{cases}$$

and $I_{a,b}(s,t) = I_{a,b}(t) - I_{a,b}(s)$.

Graphs of the posteriors $f_1(\pi|\mathbf{X})$ and $f_2(\pi|\mathbf{X})$ in Eqs. (11) and (12) are given in Figs. 1 and 2. We plot the finite-sample posteriors for sample sizes $n = 20$ and $n = 100$, when the observed fraction of ones is 0.25, as well as the conditional priors of π given $\mu = 0.25$. The latter represent the asymptotic posteriors when $\mu_0 = 0.25$. Fig. 1 shows that the posterior $f_1(\pi|\mathbf{X})$ is informative in that it places higher probability on values of π that are close to 0.25 and lower probability on values close to 0 or 1. Fig. 2 shows that under the more informative prior in Eq. (6), the posterior $f_2(\pi|\mathbf{X})$ becomes more concentrated around 0.25.

Figs. 3 and 4 show the mixture posterior $f_3(\pi|\mathbf{X})$ for sample sizes $n = 20$ and $n = 100$, respectively. Within each figure, we consider a range of prior probabilities that the misclassification error is symmetric ($\lambda = 0.5, 0.75, 0.95$). The figures clearly show that as λ increases, the posterior distribution puts more and more mass on values less than 0.25. This occurs because under symmetry, the restriction $\pi \leq \mu$ must hold. In the limit as $n \to \infty$ (see Fig. 5), the posterior becomes discontinuous at $\mu_0 = 0.25$ and values of π less than μ_0 are much more likely than values greater than μ_0.

The classical bounding results do not reveal anything about the location of the parameter within the identified set. Under the posteriors derived here, however, certain parts of the identified set are more likely than others. Also, as expected, the use of stronger information about misclassification rates will lead to a more concentrated posterior distribution.

4. Discussion

In this paper we have derived a number of exact, finite-sample posterior distributions for the mean of a misclassified binary variable. Although this parameter is not identified (unless the probabilities of misclassification errors are known or consistently estimable), the posteriors provide non-trivial information even when weak priors are specified. Classical analyses often consider how the identified set changes when certain model assumptions are either imposed or relaxed. In contrast, a Bayesian analysis allows researchers to impose or relax assumptions in a probabilistic and

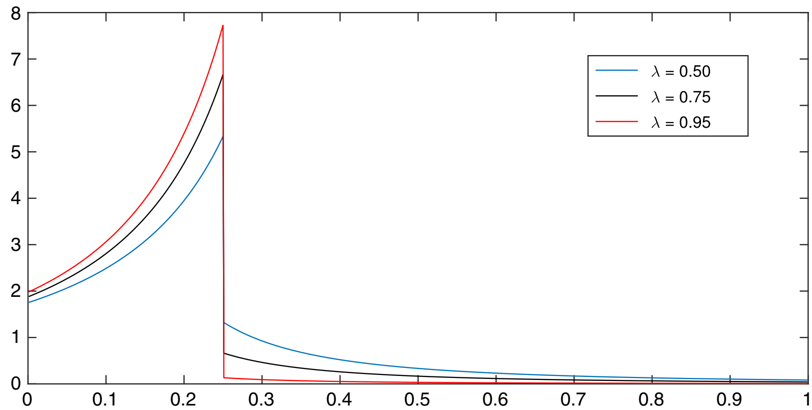

Fig. 5. Asymptotic posterior $f_3(\pi|\mu = 0.25)$ for different probabilities (λ) of error symmetry.

hence, more continuous manner. This facilitates sensitivity analyses and adds to our understanding of the mapping between assumptions and identification.

Acknowledgments

We would like to thank Tony Lancaster and Robin Carter for comments on an earlier version of this paper.

Appendix. Calculating the posteriors

The marginal likelihood $f(\mathbf{X})$ is $\int_0^1 \mu^{n_1}(1-\mu)^{n_0}d\mu = B_{n_1+1,n_0+1}$, so that

$$f_1(\pi|\mathbf{X}) = \frac{1}{B_{n_1+1,n_0+1}} \int_0^1 \mu^{n_1}(1-\mu)^{n_0} f_1(\pi|\mu)d\mu.$$

Substituting Eq. (5) into this expression, we find

$$\begin{aligned} f_1(\pi|\mathbf{X}) &= \frac{1}{2\pi^2 B_{n_1+1,n_0+1}} \int_0^\pi \mu^{n_1+1}(1-\mu)^{n_0-1}d\mu \\ &\quad + \frac{1}{2(1-\pi)^2 B_{n_1+1,n_0+1}} \int_\pi^1 \mu^{n_1-1}(1-\mu)^{n_0+1}d\mu \\ &= \frac{B_{n_1+2,n_0}}{2\pi^2 B_{n_1+1,n_0+1}} I_{n_1+2,n_0}(\pi) \\ &\quad + \frac{B_{n_1,n_0+2}}{2(1-\pi^2)B_{n_1+1,n_0+1}}(1 - I_{n_1,n_0+2}(\pi)). \end{aligned}$$

Substituting $B_{n_1+2,n_0} = B_{n_1+1,n_0+1}(n_1+1)/n_0$ and $B_{n_1,n_0+2} = B_{n_1+1,n_0+1}(n_0+1)/n_1$ into the previous equation yields the posterior in Eq. (11).

Next, we turn to $f_2(\pi|\mu)$. Given the uniform prior for μ, we have $f_2(\pi|\mu) = f_2(\mu, \pi)$. Under the restrictions $\max\{0, (\mu-\pi)/(1-\pi)\} \le p \le \mu$, it follows that

$$\begin{aligned} f_2(\pi|\mu) &= \mathbf{1}\{\pi > \mu\} \int_0^\mu f_2(\mu, p, \pi)dp \\ &\quad + \mathbf{1}\{\pi \le \mu\} \int_{(\mu-\pi)/(1-\pi)}^\mu f_3(\mu, p, \pi)dp \\ &= \frac{\mathbf{1}\{\pi > \mu\}}{4\pi\sqrt{\pi\mu(1-\mu)}} \int_0^\mu \frac{\mu - p}{\sqrt{p^2(1-\pi) + p(\pi-\mu)}}dp \\ &\quad + \frac{\mathbf{1}\{\pi \le \mu\}}{4\pi\sqrt{\pi\mu(1-\mu)}} \\ &\quad \times \int_{(\mu-\pi)/(1-\pi)}^\mu \frac{\mu - p}{\sqrt{p^2(1-\pi) + p(\pi-\mu)}}dp. \end{aligned}$$

The two integrals on the right-hand side can be calculated using the relation

$$\begin{aligned} \int \frac{(a-x)}{\sqrt{bx^2+cx}}dx &= \log\left(\frac{\frac{1}{2}c+bx}{\sqrt{b}} + \sqrt{bx^2+cx}\right) \\ &\quad \times \left(\frac{a}{\sqrt{b}} + \frac{c}{2b^{3/2}}\right) - \frac{\sqrt{bx^2+cx}}{b}. \end{aligned}$$

Substituting $a = \mu$, $b = 1 - \pi$, and $c = \pi - \mu$ and simplifying, it follows that

$$\begin{aligned} f_2(\pi|\mu) &= \frac{\mu(1-\pi) + \frac{1}{2}(\pi-\mu)}{4\pi(1-\pi)\sqrt{\pi(1-\pi)\mu(1-\mu)}} \\ &\quad \times \log\left(\frac{\pi-\mu+2(1-\pi)\mu+2\sqrt{\pi(1-\pi)\mu(1-\mu)}}{|\pi-\mu|}\right) - \frac{1}{4\pi(1-\pi)}, \end{aligned}$$

which is Eq. (7).

Finally, consider $f_3(\pi|\mathbf{X})$. We only need to find the marginal posterior of π under symmetry ($p = q$). From Eq. (8) and a change of variables, it follows that

$$\begin{aligned} \tilde{f}(\pi|\mu) &= \frac{(1-2\mu)}{\mu(1-2\pi)^2}\mathbf{1}\left\{\pi \le \mu < \tfrac{1}{2}\right\} \\ &\quad + \frac{(2\mu-1)}{(1-\mu)(1-2\pi)^2}\mathbf{1}\left\{\tfrac{1}{2} < \mu \le \pi\right\}. \end{aligned}$$

First, consider the case $\pi < \frac{1}{2}$. Then

$$\begin{aligned} \tilde{f}(\pi|\mathbf{X}) &= \frac{1}{f(\mathbf{X})} \int_\pi^{1/2} \frac{(1-2\mu)}{\mu(1-2\pi)^2}\mu^{n_1}(1-\mu)^{n_0}d\mu \\ &= \frac{1}{B_{n_1+1,n_0+1}(1-2\pi)^2} \int_\pi^{1/2} \mu^{n_1-1}(1-\mu)^{n_0}d\mu \\ &\quad - \frac{2}{B_{n_1+1,n_0+1}(1-2\pi)^2} \int_\pi^{1/2} \mu^{n_1}(1-\mu)^{n_0}d\mu \\ &= \frac{1}{(1-2\pi)^2}\left[\frac{B_{n_1,n_0+1}}{B_{n_1+1,n_0+1}} I_{n_1,n_0+1}\left(\pi, \tfrac{1}{2}\right) \right. \\ &\quad \left. - 2I_{n_1+1,n_0+1}\left(\pi, \tfrac{1}{2}\right)\right]. \end{aligned}$$

Using the fact that $B_{n_1+1,n_0+1} = n_1 B_{n_1,n_0+1}/(n+1)$, we find that for $\pi < \frac{1}{2}$:

$$\begin{aligned} \tilde{f}(\pi|\mathbf{X}) &= \frac{1}{(1-2\pi)^2}\left[\left(\frac{n+1}{n_1}\right) I_{n_1,n_0+1}\left(\pi, \tfrac{1}{2}\right) \right. \\ &\quad \left. - 2I_{n_1+1,n_0+1}\left(\pi, \tfrac{1}{2}\right)\right]. \end{aligned}$$

If $\pi > \frac{1}{2}$, then

$$\begin{aligned}\tilde{f}(\pi|\mathbf{X}) &= \frac{1}{(1-2\pi)^2 B_{n_1+1,n_0+1}} \int_{1/2}^{\pi} \mu^{n_1}(1-\mu)^{n_0} \frac{(2\mu-1)}{(1-\mu)} d\mu \\ &= \frac{2}{(1-2\pi)^2 B_{n_1+1,n_0+1}} \int_{1/2}^{\pi} \mu^{n_1+1}(1-\mu)^{n_0-1} d\mu \\ &\quad - \frac{1}{(1-2\pi)^2 B_{n_1+1,n_0+1}} \int_{1/2}^{\pi} \mu^{n_1}(1-\mu)^{n_0-1} d\mu \\ &= \frac{1}{(1-2\pi)^2} \left[\frac{2B_{n_1+2,n_0}}{B_{n_1+1,n_0+1}} I_{n_1+2,n_0}\left(\tfrac{1}{2},\pi\right) \right. \\ &\quad \left. - \frac{B_{n_1+1,n_0}}{B_{n_1+1,n_0+1}} I_{n_1+1,n_0}\left(\tfrac{1}{2},\pi\right) \right].\end{aligned}$$

Since $B_{n_1+2,n_0} = (n_1+1)B_{n_1+1,n_0+1}/n_0$ and $B_{n_1+1,n_+1} = n_0 B_{n_1+1,n_0}/(n+1)$, it follows that for $\pi > \frac{1}{2}$:

$$\begin{aligned}\tilde{f}(\pi|\mathbf{X}) &= \frac{1}{(1-2\pi)^2} \left[2\left(\frac{n_1+1}{n_0}\right) I_{n_1+2,n_0}\left(\tfrac{1}{2},\pi\right) \right. \\ &\quad \left. - \left(\frac{n+1}{n_0}\right) I_{n_1+1,n_0}\left(\tfrac{1}{2},\pi\right) \right],\end{aligned}$$

which completes the derivation of $\tilde{f}(\pi|\mathbf{X})$ in Eq. (13).

References

Abramowitz, M., Stegun, I.A., 1964. Handbook of Mathematical Functions with Formulas. Dover.

Aigner, D.J., 1973. Regression with a binary independent variable subject to errors of observation. J. Econometrics 1, 49–60.

Bollinger, C.R., 1996. Bounding mean regressions when a binary regressor is mismeasured. J. Econometrics 73, 387–399.

Bross, I., 1954. Misclassification in 2x2 tables. Biometrics 10, 478–486.

Chen, C.-F., 1985. On asymptotic normality of limiting density functions with Bayesian implications. J. R. Stat. Soc. Ser. B Stat. Methodol. 47, 540–546.

Evans, M., Guttman, I., Haitovsky, Y., Swartz, T., 1996. Bayesian analysis of binary data subject to misclassification. In: Berry, D.A., Chaloner, K.M., Geweke, J.K. (Eds.), Bayesian Analysis in Statistics and Econometrics: Essays in Honor of Arnold Zellner. Wiley, pp. 67–78.

Gaba, A., Winkler, R.L., 1992. Implications of errors in survey data: A Bayesian model. Manage. Sci. 38 (7), 913–925.

Heyde, C., Johnstone, I., 1979. On asymptotic posterior normality for stochastic processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 41, 184–189.

Imbens, G.W., Manski, C.F., 2004. Confidence intervals for partially identified parameters. Econometrica 72 (6), 1845–1857.

Joseph, L., Gyorkos, T.W., Coupal, L., 1995. Bayesian estimation of disease prevalence and theparametersa of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141 (3), 263–272.

Molinari, F., 2008. Partial identification of probability distributions with misclassified data. J. Econometrics 144 (1), 81–117.

Moon, H.R., Schorfheide, F., 2012. Bayesian and frequentist inference in partially identified models. Econometrica 80 (2), 755–782.

Neyman, J., 1950. First Course in Probability and Statistics. H. Holt.

Poirier, D., 1998. Revising beliefs in nonidentified models. Econometric Theory 14, 483–509.

Rahme, E., Joseph, L., Gyorkos, T.W., 2000. Bayesian sample size determination for estimating binomial parameters from data subject to misclassification. Appl. Stat. 49 (1), 119–128.

Tenenbein, A., 1970. A double sampling scheme for estimating from binomial data with misclassifications. J. Amer. Statist. Assoc. 65, 1350–1361.