Comment on Peek and Rosengren (2005) "Unnatural Selection: Perverse Incentives and the Allocation of Credit in Japan"

Inoue, Hitoshi and Nakashima, Kiyotaka and Takahashi, Koji

Sapporo Gakuin University, Faculty of Economics, Konan University, Faculty of Economics, University of California, San Diego, Department of Economics

25 July 2016

Comment on Peek and Rosengren (2005) "Unnatural Selection: Perverse Incentives and the Allocation of Credit in Japan"

Hitoshi Inoue, Kiyotaka Nakashima, and Koji Takahashi*

* Hitoshi Inoue, Faculty of Economics, Sapporo Gakuin University, 11-banchi Bunkyodai, Ebetsu, Hokkaido, 069-8555, Japan, E-mail: inoue12@sgu.ac.jp. Correspondence to: Kiyotaka Nakashima, Faculty of Economics, Konan University, 8-9-1, Okamoto, Higashinada, Kobe, 658-8501, Japan, E-mail: kiyotaka@center.konan-u.ac.jp. Koji Takahashi, Department of Economics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA, 92093-0508, United States of America, E-mail: ktakahas@ucsd.edu. We are grateful for the helpful comments and discussions by Takashi Hatakeda and Yoshiaki Ogura. This study received financial support in the form of Hirao-research grants from Konan University and Grants-in-Aid from the Japanese Ministry of Education and Science, Japan.

Abstract

Peek and Rosengren (2005) suggested the mechanism of “unnatural selection,” where Japanese banks with impaired capital increase credit to low-quality firms because of their motivation to pursue balance sheet cosmetics. In this study, we reexamine this mechanism in terms of the interaction effect in a nonlinear specification of bank lending, using data from 1994 to 1999. We rigorously demonstrate that their estimation results imply that Japanese banks allocated lending from viable firms to unviable ones regardless of the degree of bank capitalization.

Keywords: interaction effect, nonlinear specification, probit model, forbearance lending

JEL classification: G01, G21, G28

When borrowers become insolvent, a bank may become financially distressed. Such a financially distressed bank has incentives to continue to lend to insolvent borrowers to conceal its predicament, while hoping that the circumstance of insolvent borrowers will improve. This type of bank lending that hopes for a revival in the credit status of borrowers is called forbearance lending, evergreening lending, or zombie lending. If many banks engage in this type of lending, the resulting misallocation of credit to unviable firms that could go bankrupt would damage the macroeconomic situation further still (Hoshi, 2006; Caballero, Hoshi, and Kashyap, 2008). Hence, this practice has been considered to be the source of the prolonged economic stagnation experienced since the 1990s in Japan.

The empirical study by Peek and Rosengren (2005) is the most important and influential piece of research on the misallocation of bank credit in Japan.[1] They

[1] Empirical research on forbearance lending in Japan also includes Sekine, Kobayashi, and Saita (2003) and Watanabe (2010). Sekine, Kobayashi, and Saita (2003) used firm-level panel data from 1986 to 1999 and found that highly indebted firms belonging to nonmanufacturing industries are more likely to increase their bank borrowings for the sample period after 1993 despite their low level of profitability. Watanabe (2010) used bank-level panel data from 1995 to 2000, demonstrating that banks with large capital losses, particularly caused by the regulator's request in 1997 for the rigorous assessment of outstanding bank loans, are more likely to reallocate lending to unhealthy industries with a higher concentration of nonperforming loans. In terms of a theoretical framework, Bruche and Llobet (2014) provided a precondition for avoiding forbearance lending to low-quality

specified forbearance lending in a nonlinear function and used loan-level data from 1994 to 1999. They found that Japanese banks' attempts to avoid the realization of losses on their balance sheets (so-called balance sheet cosmetics herein) induces the mechanism of "unnatural selection," in which Japanese banks with impaired capital are more likely to provide additional credit to unviable firms.

In this study, we reassess this mechanism in terms of the interaction effect in a nonlinear specification of bank lending. Thus, we rigorously demonstrate that Peek and Rosengren's (2005) estimation results imply that Japanese banks allocated lending from viable firms to unviable ones regardless of the degree of bank capitalization.

The rest of the paper is as follows: Section I discusses the potential shortcoming of Peek and Rosengren's (2005) nonlinear specification of bank lending in terms of the interaction effect, Section II reexamines the implication of their estimation results, and the final section provides our conclusions. Appendices A and B explain how to calculate the interaction effect and its standard error for the probit model and for the random probit model, respectively.

---

firms. They suggested that regulators should induce banks to disclose the deterioration of their capital condition.

I. Specifying Lending to Low-quality Firms by using a Nonlinear Function

Peek and Rosengren (2005) specified forbearance lending by using a random effects probit model with an interaction term consisting of a low-capitalized bank indicator and a firm performance variable (i.e., working capital ratio or return on assets). They set random effects terms for each firm as firm unobserved components. This section briefly explains the potential shortcoming of specifying the misallocation of bank credit in such a random effects probit model with an interaction term.[2]

To illustrate the essence of this econometric problem, following Ai and Norton (2003), we use a general functional form $F(\cdot)$ and write the conditional expected values of y as a function of the linear index function $v = \beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2) + X_3 \mathrm{B}_3$:

$$E[y|x_1, x_2, X_3] = F(v) = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2) + X_3 \mathrm{B}_3). \quad (1)$$

Function F could be the logit or probit transformation, the logarithmic or

[2] The same problem of the statistical inference of an interaction term in a nonlinear equation has also been discussed in political science (e.g., Berry, DeMeritt, and Esarey, 2010).

exponential transformation, or any other nonlinear function of the linear index function v. x_1 and x_2 denote the continuous variables to be used to construct the interaction term $x_1 \times x_2$ in nonlinear function F. X_3 denotes a vector variable including other observable control variables.

In nonlinear equation (1), we can express the marginal effect of x_k ($k = 1$ or 2) on the conditional expected value of y as follows:[3]

$$\frac{\partial E[y|x_1, x_2, X_3]}{\partial x_k} = \frac{dF}{dv}\frac{\partial v}{dx_k} = (\beta_k + \beta_{12}x_l)\frac{dF}{dv} \quad \text{for } k \neq l. \tag{2}$$

Then, we can write the cross-partial derivative, or the so-called "interaction effect," in the following equation:

$$\frac{\partial^2 E[y|x_1, x_2, X_3]}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1}\left[\frac{dF}{dv}(\beta_2 + \beta_{12}x_1)\right] = \frac{\partial}{\partial x_2}\left[\frac{dF}{dv}(\beta_1 + \beta_{12}x_2)\right]$$
$$= \left[\frac{dF}{dv}\beta_{12}\right] + \left[\frac{d^2F}{dv^2}(\beta_1 + \beta_{12}x_2)(\beta_2 + \beta_{12}x_1)\right]. \tag{3}$$

[3] To clarify the issue of the nonlinear specifications of bank lending, we assume that x_1 and x_2 are continuous variables. We discuss below the case where one of these two variables is an indicator variable, which is the case where Peek and Rosengren (2005) specified forbearance lending by using a probit specification.

Note that even if the coefficient of the interaction term, β_{12}, is zero, the expression above for the interaction effect, $\partial^2 E[y|x_1, x_2, X_3]/\partial x_1 \partial x_2$, still has a nonzero value. This means that the statistical significance of the interaction effect cannot be tested with that of the estimated coefficient of β_{12}. Further, the sign of β_{12} does not necessarily indicate that of the interaction effect.

In Peek and Rosengren (2005), y corresponds to an indicator variable, $\text{LOAN}_{i,j,t}$, which has a value of one if loans to firm i by bank j increase from year $t-1$ to year t and zero otherwise. Thus, they performed the probit transformation of linear index v in function $E[y|x_1, x_2, X_3] = \Pr(y = 1|x_1, x_2, X_3) = \Phi(v)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Variable x_1 corresponds to a low-capitalized bank indicator $\text{REQ2}_{j,t-1}$, which is a (0,1) dummy variable that has a value of one if the bank's reported risk-based capital ratio is less than two percentage points above the bank's required capital ratio at year $t-1$ and zero otherwise. Variable x_2 corresponds to the lagged variable of firm i's financial capability, measured by using the working capital ratio $\text{FWORKCAP}_{i,t-1}$ or its profitability measured as return on assets $\text{FROA}_{i,t-1}$.

Vector variable X_3 includes other lender-side and borrower-side observables. In their specification of bank lending, if the interaction effects of $\mathrm{REQ2}_{j,t-1} \times \mathrm{FWORKCAP}_{i,t-1}$ and $\mathrm{REQ2}_{j,t-1} \times \mathrm{FROA}_{i,t-1}$ have negative values, it implies that low-capitalized banks provide more credit to unviable firms (i.e., those that have lower financial health and lower profitability) than non-low-capitalized banks do.

A potential shortcoming of the approach of Peek and Rosengren (2005) is that their analysis with the probit estimation of the bank lending equation was based on the estimated coefficients of the bank financial health variable as well as the firm performance variables and their interaction terms, but not on the marginal effects and interaction effects, as expressed in equations (2) and (3). Their estimated coefficients of the bank financial health indicator, $\mathrm{REQ2}_{j,t-1}$, had significantly positive values, while those of the firm variables, $\mathrm{FWORKCAP}_{i,t-1}$ and $\mathrm{FROA}_{i,t-1}$, had significantly negative ones.

More importantly, they showed that the coefficients of the interaction terms, β_{12}, were estimated to be significantly negative, thereby arguing that low-capitalized banks were more likely to lend credit to low-quality firms; in other words, forbearance lending by low-capitalized banks to low-quality firms prevailed during the late 1990s in Japan. However, as discussed above, the negative values of the coefficients of the interaction terms do not necessarily ensure the existence of forbearance lending in terms of the interaction effects, $\partial^2 E[y|x_1, x_2, X_3]/\partial x_1 \partial x_2$. Rather, negative estimates of the interaction effects are necessary for the existence of forbearance lending to low-quality firms by low-capitalized banks, who have window-dressing motives to avoid the realization of losses on their balance sheets. In the next section, we thus critically reassess this mechanism of unnatural selection by reporting the estimated marginal effects and interaction effects obtained with the same probit specification as that in Peek and Rosengren (2005).

II. Critical Evaluation of the Mechanism of Unnatural Selection

Peek and Rosengren (2005) adopted a random effects probit model to specify bank lending. The difficulty in estimating the interaction effect, $\partial^2 E[y|x_1, x_2, X_3]/\partial x_1 \partial x_2$, in a specific nonlinear model with unobserved heterogeneity, including the random effects probit model, involves analytically expressing the cross-partial derivative incorporating the fixed or random effects terms. Peek and Rosengren (2005) did not address this problem. In a nonlinear specification of bank lending, without obtaining firm i's random effects r_i, we cannot compute the conditional probability of $y = 1$ (i.e., $\text{LOAN}_{i,j,t} = 1$) and thus cannot estimate the interaction effect.[4] To estimate the interaction effect, we thus compute three types of conditional probabilities $\Pr(y = 1|x_1, x_2, X_3, r_i)$, each of which has different assumptions about the firm random effects, r_i, as follows:

A1. a prediction conditional on the firm random effects that are set to zeros, $r_i = 0$; that is, $\Pr(y = 1|x_1, x_2, X_3, r_i = 0) = \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2) + X_3 \mathrm{B}_3)$,

A2. a prediction conditional on the empirical Bayes means of the firm random effects $r_i = \hat{r}_i$; that is, $\Pr(y = 1|x_1, x_2, X_3, r_i = \hat{r}_i) = \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2) + X_3 \mathrm{B}_3 + \hat{r}_i)$,

A3. an unconditional prediction with respect to the firm random effects

[4] Another practical issue to overcome when estimating the interaction effect is that we cannot simply use the commands equipped in standard econometric software such as LIMDEP and STATA. Accordingly, we calculated the marginal effect and standard error of the interaction effect by analytically expressing the interaction effect in this nonlinear model. We develop our analytical expression of the interaction effect and its standard error for the probit models with and without the random effects terms in Appendices A and B, respectively.

$r_i \sim N(0, \sigma^2)$, which is computed by integrating a conditional prediction with respect to the firm random effects over their support; that is, $\Pr(y=1|x_1,x_2,X_3) = \int \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2) + X_3 \mathrm{B}_3 + r_i)\ g(r_i|\sigma^2) dr_i$, where $g(r_i|\sigma^2)$ indicates the $N(0,\sigma^2)$ density function.

As discussed above, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. Further, x_1 is the low-capitalized bank indicator, $\mathrm{REQ2}_{j,t-1}$, and x_2 is firm performance measured by using the working capital ratio, $\mathrm{FWORKCAP}_{i,t-1}$, or return on assets, $\mathrm{FROA}_{i,t-1}$. Once we find the conditional probability by using one of the approaches A1–A3 above, we can obtain the interaction effect, expressed in equation (3), in the probit specification proposed by Peek and Rosengren (2005) as follows:

$$\frac{\Delta\partial E[y|x_1,x_2,X_3,r_i]}{\Delta x_1 \partial x_2} = \frac{\Delta\partial \Pr[y=1|x_1,x_2,X_3,r_i]}{\Delta x_1 \partial x_2} = \frac{\partial\Phi(\cdot)}{\partial x_2}\Big|_{x_1=1} - \frac{\partial\Phi(\cdot)}{\partial x_2}\Big|_{x_1=0}$$

$$= \phi(\beta_1 + \beta_2 x_2 + \beta_{12} x_2 + X_3 \mathrm{B}_3 + r_i)(\beta_2 + \beta_{12})$$

$$-\phi(\beta_2 x_2 + X_3 \mathrm{B}_3 + r_i) \cdot \beta_2, \quad (4)$$

where $\phi(\cdot) = \Phi'(\cdot)$ is the probability density function of the standard normal distribution.

Equation (4) clarifies the essence of the empirical analysis based on the interaction terms of the low-capitalized bank indicator and firm performance variables. The first term in this equation represents the marginal effect of bank j's loans to firm i with respect to firm i's profitability (x_2 = $\text{FWORKCAP}_{i,t-1}$ or $\text{FROA}_{i,t-1}$) in the case where bank j is low-capitalized (i.e., x_1 = $\text{REQ2}_{j,t-1} = 1$). The second term indicates the marginal effect in the case where the bank is not low-capitalized (i.e., x_1 = $\text{REQ2}_{j,t-1} = 0$). This decomposition of the interaction effect allows us to rigorously analyze the lending behavior of all banks in Japan regardless of their degree of bank capitalization.[5]

Let us express the former marginal effect evaluated at a hypothetical value of firm performance, $x_2 = \dot{x_2}$, as $\text{ME}_{i,j,t}(\text{REQ2}_{j,t-1} = 1, x_2 = \dot{x_2})$ and the latter marginal effect as $\text{ME}_{i,j,t}(\text{REQ2}_{j,t-1} = 0, x_2 = \dot{x_2})$. Then, we obtain a consistent

[5] Once we introduce the interaction terms, we cannot derive correct inferences about the lending behavior prevailing in the Japanese banking system without comparing the lending behavior of non-low-capitalized and low-capitalized banks. Peek and Rosengren (2005) obtained negative estimates for a coefficient parameter of firm profitability, thus suggesting that Japanese banks lend more credit to low-quality firms through evergreening lending. However, their interpretation of these negative estimates is unsuitable for their empirical analysis based on the interaction terms.

estimator for the average interaction effect evaluated at a hypothetical value of firm performance (hereafter, $\mathrm{AIE}(x_2 = \dot{x_2})$) as the sample mean of the interaction effect (4):

$$\mathrm{AIE}(x_2 = \dot{x_2}) = \frac{1}{n}\sum_{i,j,t}^{n}\left[\frac{\partial\Phi(\cdot)}{\partial x_2}\mid_{x_1=1,x_2=\dot{x_2}}\right] - \frac{1}{n}\sum_{i,j,t}^{n}\left[\frac{\partial\Phi(\cdot)}{\partial x_2}\mid_{x_1=0,x_2=\dot{x_2}}\right]$$
$$= \frac{1}{n}\sum_{i,j,t}^{n}\mathrm{ME}_{i,j,t}\left(\mathrm{REQ2}_{j,t-1} = 1, x_2 = \dot{x_2}\right)$$
$$- \frac{1}{n}\sum_{i,j,t}^{n}\mathrm{ME}_{i,j,t}\left(\mathrm{REQ2}_{j,t-1} = 0, x_2 = \dot{x_2}\right)$$
$$= \mathrm{AME}\left(\mathrm{REQ2}_{j,t-1} = 1, x_2 = \dot{x_2}\right)$$
$$- \mathrm{AME}\left(\mathrm{REQ2}_{j,t-1} = 0, x_2 = \dot{x_2}\right), \qquad (5)$$

where n denotes the number of observations (bank–firm relationships). [6] We calculate the standard errors of the average marginal effects by using the delta method. For the details of the calculation, see Appendices A and B.

To reexamine the implication of Peek and Rosengren's (2005) estimation

[6] More strictly, the first and second terms in equation (5) respectively represent the counterfactual effects in the hypothetical cases where all banks are low-capitalized and non-low-capitalized ones; hence, the interaction effect measures the treatment effect of the bank's low capitalization, expressed in $\mathrm{REQ2}_{j,t-1} = 1$, as long as the confounding factors that can affect bank capitalization are fully controlled for in the bank lending equation by using X_3.

results with the average marginal effects, we start by replicating the estimation results from their dataset of AER Final Data.txt, which is available online at the journal website. The analysis of the average marginal effects presented below is based on the replication of the "Full sample" results reported in Table 5 of Peek and Rosengren (2005). The sample period runs from 1994 to 1999, during which Japanese banks faced increasing pressure to maintain regulatory capital requirements under the Basel I framework.

Table 1 reports the estimated coefficients of and descriptive statistics for the bank's financial health indicator, $x_1 = \mathrm{REQ2}_{j,t-1}$, and firm performance variables measured by using the working capital ratio, $x_2 = \mathrm{FWORKCAP}_{i,t-1}$, and return on assets, $x_2 = \mathrm{FROA}_{i,t-1}$.[7] The descriptive statistics for $\mathrm{REQ2}_{j,t-1}$ indicate that about 70 percent of Japanese banks showed a low degree of capitalization in the late 1990s. The estimation results in this table clearly show that we can replicate Peek and Rosengren's (2005) estimation results: $\mathrm{REQ2}_{j,t-1}$ has a significantly positive coefficient, while $\mathrm{FWORKCAP}_{i,t-1}$ and $\mathrm{FROA}_{i,t-1}$ have significantly

[7] Peek and Rosengren (2005) defined the firm's working capital ratio as $\mathrm{FWORKCAP}_{i,t} = 100 \times (\text{Liquid Asset}_{i,t} - \text{Current Liability}_{i,t}) / \text{ Total Asset}_{i,t}$ and the firm's return on assets as $\mathrm{FROA}_{i,t} = 100 \times \text{Operating Profit}_{i,t}/\text{Total Asset}_{i,t-1}$.

negative ones. The interaction terms, $\text{REQ2}_{j,t-1} \times \text{FWORKCAP}_{i,t-1}$ and $\text{REQ2}_{j,t-1} \times \text{FROA}_{i,t-1}$, appear to have significantly negative coefficients.[8]

Figures 1 and 2 respectively report the estimated interaction effects and marginal effects obtained by using the firm performance variables, $\text{FWORKCAP}_{i,t-1}$ and $\text{FROA}_{i,t-1}$. In each figure, the left-hand side panels show the estimated interaction effects, $\text{AIE}(x_2 = \dot{x_2})$, while the right-hand side panels show the estimated marginal effects for low- and non-low-capitalized banks, $\text{AME}(\text{REQ2}_{j,t-1} = 1, x_2 = \dot{x_2})$ and $\text{AME}(\text{REQ2}_{j,t-1} = 0, x_2 = \dot{x_2})$. Figures A1 to A3 are obtained by using approaches A1 to A3 to compute the conditional probability in the random effects probit specification, respectively. In each figure, the interaction effect and marginal effect are estimated at each hypothetical value of firm performance, $x_2 = \dot{x_2}$, whose range corresponds to the sample range from -120 to 80 for the working capital ratio in Figure 1 and from -25 to 25 for return on assets in Figure 2.

The shaded areas of Figures 1 and 2 also report the histograms of the two firm variables to allow us to analyze the allocation of bank credit in association with

[8] Our random effects probit regression includes all the other control variables and reproduces the same estimation results for these as in Peek and Rosengren's (2005) regression. For the estimation results of the other control variables, see Table 5 in Peek and Rosengren (2005).

the actual performance of borrowing firms in the late 1990s. The histograms illustrate that few Japanese firms borrowing capital in the late 1990s suffered from low financial capability and/or low profitability. Indeed, three-quarters and nine-tenths of the distribution of the working capital ratio and return on assets show positive values, respectively.

The estimates reported in A1 to A3 of Figures 1 and 2 appear to be qualitatively the same, indicating that the firm random effects, r_i, do not play a substantial role in determining the conditional probability. Note also that the estimated interaction effects, $\mathrm{AIE}(x_2 = \dot{x_2})$, in Figures 1 and 2 have significantly negative values for most of the hypothetical values of the firm working capital ratio and return on assets. This finding indicates not only that low-capitalized banks were more likely to increase loans to unviable firms than were non-low-capitalized banks in the study period, as suggested by Peek and Rosengren (2005), but also that low-capitalized banks were more likely to decrease loans to viable firms than non-low-capitalized banks were (or leave them unchanged). Rather, given that most firms borrowing capital performed well in the late 1990s, the negative values of the interaction effects imply that the misallocation of credit from viable firms to unviable ones prevailed because of low-capitalized banks' motivation to pursue balance sheet cosmetics.[9]

More importantly, Figures 1 and 2 show that the marginal effects for low- and non-low-capitalized banks, $\mathrm{AME}(\mathrm{REQ2}_{j,t-1} = 1, x_2 = \dot{x_2})$and $\mathrm{AME}(\mathrm{REQ2}_{j,t-1} = 0, x_2 = \dot{x_2})$, have significantly negative estimates.[10] This finding clearly indicates that the misallocation of bank credit from viable firms to unviable ones prevailed in the Japanese banking sector in the late 1990s; in other words, Japanese banks provided more credit to relatively unviable firms, while decreasing credit to viable ones (or keeping it unchanged) regardless of the degree of bank capitalization. This lending behavior by capitalized banks is not consistent with the balance sheet cosmetics hypothesis.

III. Conclusion

The mechanism of unnatural selection suggested by Peek and Rosengren

---

[9] Sasaki (2011) found that lending to more troubled industries with more nonperforming loans in the 1990s was less sensitive to a bank's capital adequacy ratio than lending to less troubled industries with fewer nonperforming loans. Given this finding, she pointed out that to calculate the risk-weighted assets, Basel I equally weighed all bank loans to firms, regardless of whether they are good borrowers; accordingly, low-capitalized Japanese banks decreased (increased) credit to viable (unviable) firms to maintain adequate capital ratios by avoiding the realization of bankruptcy and nonperforming loans under the Basel I framework.

[10] Although we defined the low-capitalized bank indicator, $\mathrm{REQ2}_{j,t-1}$, by setting the threshold value of bank capital buffers above the minimum requirement to various values less than two percentage points, we confirm that the average marginal effects for both low- and non-low-capitalized banks have significantly negative estimates.

(2005) assumes that forbearance lending by low-capitalized banks to low-quality borrowers prevailed during the late 1990s in Japan, particularly driven by banks' motivation to pursue balance sheet cosmetics. In this study, we reevaluated this mechanism by focusing on the interaction effect instead of the coefficient parameter of the interaction term. More concretely, we discussed a potential shortcoming of specifying bank lending by using nonlinear functions, demonstrating that their estimation results, which are based on the random effects probit model, imply that Japanese banks allocated lending from viable firms to unviable ones in the late 1990s regardless of the degree of bank capitalization, although low-capitalized banks were still more likely to do so than non-low-capitalized banks.

Our finding does not counter the finding of Peek and Rosengren (2005) in that we rigorously show that the bank's balance sheet cosmetics hold for forbearance lending by low-capitalized banks; rather, we complement it in that we also rigorously demonstrate that the misallocation of bank credit from viable firms to unviable ones prevailed in the Japanese banking system in the late 1990s. Other hypotheses to explain why Japanese banks emphasized relationships with relatively low-quality firms were explored by Nakashima and Takahashi (2016) and Ogura, Okui, and Saito (2016).[11]

[11] Nakashima and Takahashi (2016) empirically demonstrated that the long-term contracts in Japanese bank–borrower relationships aim to smoothing loan prices intertemporally by offsetting short-term losses through long-term rents generated by firms with higher uncertainty, as demonstrated by the theoretical literature on relationship banking (Berlin and Mester, 1998; Song and Thakor, 2007). Ogura, Okui, and Saito (2016) theoretically and empirically demonstrated that Japanese banks kept lending to loss-making firms at an interest rate below the prime rate if such firms were located in an influential position in the inter-firm supply network.

References

Ai, Chunrong, and Edward C. Norton. "Interaction terms in logit and probit models." Economics Letters 80, no. 1 (2003): 123-129.

Berry, William D., Jacqueline HR DeMeritt, and Justin Esarey. "Testing for interaction in binary logit and probit models: is a product term essential?" American Journal of Political Science 54, no. 1 (2010): 248-266.

Bruche, Max, and Gerard Llobet. "Preventing zombie lending." Review of Financial Studies 27, no. 3 (2014): 923-956.

Berlin, Mitchell, and Loretta J. Mester. "Deposits and relationship lending." Review of Financial Studies 12, no. 3 (1999): 579-607.

Caballero, Ricardo J., Takeo Hoshi, and Anil K. Kashyap. "Zombie lending and depressed restructuring in Japan." The American Economic Review 98, no. 5 (2008): 1943-1977.

Hoshi, Takeo. "Economics of the living dead." Japanese Economic Review 57, no. 1 (2006): 30-49.

Nakashima, Kiyotaka, and Koji Takahashi. "Termination of Bank-Firm Relationships." mimeo (2016).

Ogura, Yoshiaki, Ryo Okui, and Yukiko Saito. “Network-motivated Lending Decisions”, mimeo (2016).

Peek, Joe, and Eric S. Rosengren. "Unnatural selection: Perverse incentives and the misallocation of credit in Japan." The American Economic Review 95, no. 4 (2005): 1144-1166.

Sasaki, Yuri. “The Effect of Basel Accord on Japanese Banks (in Japanese: Jikoshihonhiritsu no Hougin heno Eikyo).” In the Annual Report of Research Institute of Industry and Economy, vol. 28. Tokyo: Research Institute of Industry and Economy, Meiji Gakuin University, (2011) 53-63.

Sekine, Toshitaka, Keiichiro Kobayashi, and Yumi Saita. "Forbearance lending: the case of Japanese firms." Monetary & Economic Studies 21 (2003): 2.

Song, Fenghua, and Anjan V. Thakor. "Relationship banking, fragility, and the asset-liability matching problem." Review of Financial Studies 20, no. 6 (2007): 2129-2177.

Watanabe, Wako. "Does a large loss of bank capital cause evergreening? Evidence from Japan." Journal of the Japanese and International Economies 24, no. 1 (2010): 116-136.

Appendix A. Probit Model with an Interaction Term

In Appendix A, we explain how to calculate the average marginal effects and their standard errors for the probit model with an interaction term consisting of an indicator variable and a continuous variable. In Appendix B, we consider the random effects probit model with an interaction term.

First, we consider a probit model with an interaction term as follows:

$$\begin{aligned} \Pr(y = 1 | x_1, x_2, \mathrm{X}_3) &= \Phi(v), \\ v &= \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathrm{X}_3 \mathrm{B}_3, \end{aligned} \tag{A-1}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and $x_1 x_2$ is an interaction term of an indicator variable x_1 and a continuous variable x_2. In our lending equation specified in the main text, x_1 is a dummy variable that indicates whether an individual bank's capital buffers are less than two percentage points, $\mathrm{REQ2}_{j,t-1}$. x_2 is a firm performance variable measured by using the working capital ratio, $\mathrm{FWORKCAP}_{i,t-1}$, or return on assets, $\mathrm{FROA}_{i,t-1}$. X_3 denotes a $(1 \times k)$ vector including the other control variables. y corresponds to $\mathrm{LOAN}_{i,j,t}$. β_1, β_2, β_{12}, and B_3 are the coefficient parameters.

A-I. Average Marginal Effects with respect to an Indicator Variable

We start by calculating the marginal effect of $\Pr(y = 1 | x_1, x_2, \mathrm{X}_3)$ with respect to the indicator variable x_1 as expressed in the following:

$$\mathrm{ME}_1 = \frac{\Delta \Pr(y = 1 | x_1, x_2, \mathrm{X}_3)}{\Delta x_1} = \Pr(y = 1 | x_1 = 1) - \Pr(y = 1 | x_1 = 0). \tag{A-2}$$

Note that ME_1 is a random variable because the coefficient estimators $\hat{\beta}$'s and $\widehat{\mathrm{B}}_3$ are random variables. A consistent estimator for the marginal effect is calculated as

$$\widehat{\mathrm{ME}}_1 = \Phi(\hat{v}_1) - \Phi(\hat{v}_0), \tag{A-3}$$

where

$$\hat{v}_1 = \hat{\beta}_1 + \hat{\beta}_2 x_2 + \hat{\beta}_{12} x_2 + \mathrm{X}_3 \widehat{\mathrm{B}}_3, \qquad \text{(A-4)}$$
$$\hat{v}_0 = \hat{\beta}_2 x_2 + \mathrm{X}_3 \widehat{\mathrm{B}}_3.$$

The average marginal effect is calculated as a sample mean of the marginal effect $\widehat{\mathrm{ME}}_1$, namely,

$$\begin{aligned} \widehat{\mathrm{AME}}_1 &= \frac{1}{n} \sum^{n} \left[\Phi(\hat{v}_1) - \Phi(\hat{v}_0) \right] \\ &= \frac{1}{n} \sum^{n} \left[\Phi(\hat{v}_1) \right] - \frac{1}{n} \sum^{n} \left[\Phi(\hat{v}_0) \right], \end{aligned} \qquad \text{(A-5)}$$

where n denotes the number of observations.

We can compute the variance of the average marginal effect by using the delta method as follows:

$$\mathrm{Var}(\widehat{\mathrm{AME}}_1) = \mathbf{J_1} \widehat{\mathbf{V}} \mathbf{J_1'}, \qquad \text{(A-6)}$$

where $((3+k) \times (3+k))$ matrix $\widehat{\mathbf{V}}$ is the estimated variance-covariance matrix of the coefficient column vector $\widehat{\mathbf{b}} = [\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_{12}, \widehat{\mathrm{B}}_3']'$ and a $(1 \times (3+k))$ vector $\mathbf{J_1}$ is the Jacobian vector as expressed in the following derivatives:

$$\mathbf{J_1} = \left[\left. \frac{\partial \mathrm{AME}_1}{\partial \mathbf{b}'} \right|_{\mathbf{b}=\hat{\mathbf{b}}} \right]. \qquad \text{(A-7)}$$

A $(1,1)$ element of the Jacobian vector $\mathbf{J_1}$ is

$$\begin{aligned} \left. \frac{\partial \mathrm{AME}_1}{\partial \beta_1} \right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n} \sum^{n} \left\{ \frac{\partial \Phi(v_1)}{\partial v_1} \left. \frac{\partial v_1}{\partial \beta_1} \right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} - \frac{1}{n} \sum^{n} \left\{ \frac{\partial \Phi(v_0)}{\partial v_0} \left. \frac{\partial v_0}{\partial \beta_1} \right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} \\ &= \frac{1}{n} \sum^{n} \left\{ \phi(\hat{v}_1) \cdot 1 \right\} - \frac{1}{n} \sum^{n} \left\{ \phi(\hat{v}_0) \cdot 0 \right\}; \end{aligned} \qquad \text{(A-8)}$$

a $(1,2)$ element is

$$\begin{aligned}\left.\frac{\partial \mathrm{AME}_1}{\partial \beta_2}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \beta_2}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} - \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \beta_2}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} \\ &= \frac{1}{n}\sum^n \{\phi(\hat{v}_1)\cdot x_2\} - \frac{1}{n}\sum^n \{\phi(\hat{v}_0)\cdot x_2\};\end{aligned} \tag{A-9}$$

a $(1,3)$ element is

$$\begin{aligned}\left.\frac{\partial \mathrm{AME}_1}{\partial \beta_{12}}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \beta_{12}}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} - \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \beta_{12}}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} \\ &= \frac{1}{n}\sum^n \{\phi(\hat{v}_1)\cdot x_2\} - \frac{1}{n}\sum^n \{\phi(\hat{v}_0)\cdot 0\};\end{aligned} \tag{A-10}$$

and the $(1,4)$ to $(1,3+k)$ elements are

$$\begin{aligned}\left.\frac{\partial \mathrm{AME}_1}{\partial \mathrm{B}_3'}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \mathrm{B}_3'}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} - \frac{1}{n}\sum^n \left\{ \left.\frac{\partial \Phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \mathrm{B}_3'}\right|_{\mathbf{b}=\hat{\mathbf{b}}} \right\} \\ &= \frac{1}{n}\sum^n \{\phi(\hat{v}_1)\cdot \mathrm{X}_3\} - \frac{1}{n}\sum^n \{\phi(\hat{v}_0)\cdot \mathrm{X}_3\},\end{aligned} \tag{A-11}$$

where $\phi(\cdot) = \Phi'(\cdot)$ is the probability density function of the standard normal distribution.

A-II. Average Marginal Effects with respect to a Continuous Variable

Here, we calculate the marginal effect with respect to the continuous variable x_2 in the following:

$$\mathrm{ME}_2 = \frac{\partial \Pr(y=1|x_1, x_2, \mathrm{X}_3)}{\partial x_2} = \frac{\partial \Phi(v)}{\partial v}\cdot\frac{\partial v}{\partial x_2} = \phi(v)(\beta_2 + \beta_{12}x_1). \tag{A-12}$$

A consistent estimator for the average marginal effect is calculated as the sample mean of the marginal effect, $\widehat{\mathrm{ME}}_2$, namely

$$\widehat{\mathrm{AME}}_2 = \frac{1}{n}\sum^n \left[\phi(\hat{v})(\hat{\beta}_2 + \hat{\beta}_{12}x_1)\right]. \tag{A-13}$$

The variance of the average marginal effect is constructed by using the delta method as follows:

$$\mathrm{Var}(\widehat{\mathrm{AME}_2}) = \mathbf{J_2}\widehat{\mathbf{V}}\mathbf{J_2'}. \tag{A-14}$$

In equation (A-14), the Jacobian vector $\mathbf{J_2}$ has the following elements: a $(1,1)$ element is

$$\begin{aligned} \left.\frac{\partial \mathrm{AME}_2}{\partial \beta_1}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \beta_1}\left[\phi(v)(\beta_2+\beta_{12}x_1)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial \phi(v)}{\partial v}\frac{\partial v}{\partial \beta_1}(\beta_2+\beta_{12}x_1)\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{-\hat{v}\cdot\phi(\hat{v})(\hat{\beta}_2+\hat{\beta}_{12}x_1)\cdot x_1\right\}; \end{aligned} \tag{A-15}$$

a $(1,2)$ element is

$$\begin{aligned} \left.\frac{\partial \mathrm{AME}_2}{\partial \beta_2}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \beta_2}\left[\phi(v)(\beta_2+\beta_{12}x_1)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{\left.\left[\frac{\partial \phi(v)}{\partial v}\frac{\partial v}{\partial \beta_2}(\beta_2+\beta_{12}x_1)+\phi(v)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{-\hat{v}\cdot\phi(\hat{v})(\hat{\beta}_2+\hat{\beta}_{12}x_1)\cdot x_2\right\}+\frac{1}{n}\sum^{n}\left\{\phi(\hat{v})\right\}; \end{aligned} \tag{A-16}$$

a $(1,3)$ element is

$$\begin{aligned} \left.\frac{\partial \mathrm{AME}_2}{\partial \beta_{12}}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \beta_{12}}\left[\phi(v)(\beta_2+\beta_{12}x_1)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{\left.\left[\frac{\partial \phi(v)}{\partial v}\frac{\partial v}{\partial \beta_{12}}(\beta_2+\beta_{12}x_1)+\phi(v)x_1\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\ &= \frac{1}{n}\sum^{n}\left\{-\hat{v}\cdot\phi(\hat{v})(\hat{\beta}_2+\hat{\beta}_{12}x_1)\cdot x_1x_2\right\}+\frac{1}{n}\sum^{n}\left\{\phi(\hat{v})x_1\right\}; \end{aligned} \tag{A-17}$$

and the $(1,4)$ to $(1,3+k)$ elements are

$$\begin{aligned}
\left.\frac{\partial \mathrm{AME}_2}{\partial \mathrm{B}_3'}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \mathrm{B}_3'}\left[\phi(v)(\beta_2+\beta_{12}x_1)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial\phi(v)}{\partial v}\frac{\partial v}{\partial \mathrm{B}_3'}(\beta_2+\beta_{12}x_1)\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{-\hat{v}\cdot\phi(\hat{v})(\hat{\beta}_2+\hat{\beta}_{12}x_1)\cdot \mathrm{X}_3\right\}.
\end{aligned} \tag{A-18}$$

A-III. Average Marginal Effects with respect to the Interaction Term

Finally, we calculate the average interaction effects, or the average marginal effect with respect to the interaction term x_1x_2. To this end, we define the difference between the partial derivatives with respect to continuous variable x_2 evaluated at $x_1=1$ and $x_1=0$ as follows:

$$\begin{aligned}
\mathrm{ME}_{12} = \frac{\Delta\partial\Pr(y=1|x_1,x_2,\mathrm{X}_3)}{\Delta x_1\partial x_2} &= \left.\frac{\partial\Phi(v)}{\partial x_2}\right|_{x_1=1} - \left.\frac{\partial\Phi(v)}{\partial x_2}\right|_{x_1=0} \\
&= \phi(v)(\beta_2+\beta_{12}x_1)\big|_{x_1=1} - \phi(v)(\beta_2+\beta_{12}x_1)\big|_{x_1=0} \\
&= \phi(v_1)(\beta_2+\beta_{12}) - \phi(v_0)\cdot\beta_2,
\end{aligned} \tag{A-19}$$

where

$$\begin{aligned}
v_1 &= \beta_1+\beta_2x_2+\beta_{12}x_2+\mathrm{X}_3\mathrm{B}_3, \\
v_0 &= \beta_2x_2+\mathrm{X}_3\mathrm{B}_3.
\end{aligned} \tag{A-20}$$

Then, the average marginal effect of the interaction term can be constructed as the sample mean of the marginal effect ME_{12}. We call this average marginal effect the average interaction effect AIE. The consistent estimator of the average interaction effect is

$$\begin{aligned}
\widehat{\mathrm{AIE}} &= \frac{1}{n}\sum^{n}\left[\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12}) - \phi(\hat{v}_0)\cdot\hat{\beta}_2\right] \\
&= \frac{1}{n}\sum^{n}\left[\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12})\right] - \frac{1}{n}\sum^{n}\left[\phi(\hat{v}_0)\cdot\hat{\beta}_2\right].
\end{aligned} \tag{A-21}$$

The variance of the average interaction effect is computed by using the delta method as follows:

$$\mathrm{Var}(\widehat{\mathrm{AIE}}) = \mathbf{J_{12}}\widehat{\mathbf{V}}\mathbf{J}'_{\mathbf{12}}. \tag{A-22}$$

The Jacobian vector $\mathbf{J_{12}}$ has the following elements: a $(1,1)$ element is

$$\begin{aligned}
\left.\frac{\partial \mathrm{AIE}}{\partial \beta_1}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\frac{\partial}{\partial \beta_1}\left[\phi(v_1)(\beta_2+\beta_{12})\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\frac{\partial}{\partial \beta_1}\left[\phi(v_0)\cdot\beta_2\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{\frac{\partial\phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \beta_1}(\beta_2+\beta_{12})\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\frac{\partial\phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \beta_1}\cdot\beta_2\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{-\hat{v}_1\cdot\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12})\cdot 1\right\} - \frac{1}{n}\sum^{n}\left\{-\hat{v}_0\cdot\phi(\hat{v}_0)\cdot\hat{\beta}_2\cdot 0\right\};
\end{aligned} \tag{A-23}$$

a $(1,2)$ element is

$$\begin{aligned}
\left.\frac{\partial \mathrm{AIE}}{\partial \beta_2}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\frac{\partial}{\partial \beta_2}\left[\phi(v_1)(\beta_2+\beta_{12})\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\frac{\partial}{\partial \beta_2}\left[\phi(v_0)\cdot\beta_2\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{\left[\frac{\partial\phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \beta_2}(\beta_2+\beta_{12})+\phi(v_1)\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&\qquad - \frac{1}{n}\sum^{n}\left\{\left[\frac{\partial\phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \beta_2}\cdot\beta_2+\phi(v_0)\right]\Big|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{-\hat{v}_1\cdot\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12})\cdot x_2\right\} + \frac{1}{n}\sum^{n}\left\{\phi(\hat{v}_1)\right\} \\
&\qquad - \frac{1}{n}\sum^{n}\left\{-\hat{v}_0\cdot\phi(\hat{v}_0)\cdot\hat{\beta}_2\cdot x_2\right\} - \frac{1}{n}\sum^{n}\left\{\phi(\hat{v}_0)\right\};
\end{aligned} \tag{A-24}$$

a $(1,3)$ element is

$$
\begin{aligned}
\left.\frac{\partial \mathrm{AIE}}{\partial \beta_{12}}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \beta_{12}}\left[\phi(v_1)(\beta_2+\beta_{12})\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \beta_{12}}\left[\phi(v_0)\cdot\beta_2\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{\left.\left[\frac{\partial \phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \beta_{12}}(\beta_2+\beta_{12})+\phi(v_1)\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&\qquad\qquad - \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial \phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \beta_{12}}\cdot\beta_2\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{-\hat{v}_1\cdot\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12})\cdot x_2\right\} + \frac{1}{n}\sum^{n}\left\{\phi(\hat{v}_1)\right\} \\
&\qquad\qquad - \frac{1}{n}\sum^{n}\left\{-\hat{v}_0\cdot\phi(\hat{v}_0)\cdot\hat{\beta}_2\cdot 0\right\};
\end{aligned}
\tag{A-25}
$$

and the $(1,4)$ to $(1,3+k)$ elements are

$$
\begin{aligned}
\left.\frac{\partial \mathrm{AIE}}{\partial \mathrm{B}_3'}\right|_{\mathbf{b}=\hat{\mathbf{b}}} &= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \mathrm{B}_3'}\left[\phi(v_1)(\beta_2+\beta_{12})\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial}{\partial \mathrm{B}_3'}\left[\phi(v_0)\cdot\beta_2\right]\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial \phi(v_1)}{\partial v_1}\frac{\partial v_1}{\partial \mathrm{B}_3'}(\beta_2+\beta_{12})\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} - \frac{1}{n}\sum^{n}\left\{\left.\frac{\partial \phi(v_0)}{\partial v_0}\frac{\partial v_0}{\partial \mathrm{B}_3'}\cdot\beta_2\right|_{\mathbf{b}=\hat{\mathbf{b}}}\right\} \\
&= \frac{1}{n}\sum^{n}\left\{-\hat{v}_1\cdot\phi(\hat{v}_1)(\hat{\beta}_2+\hat{\beta}_{12})\cdot \mathrm{X}_3\right\} - \frac{1}{n}\sum^{n}\left\{-\hat{v}_0\cdot\phi(\hat{v}_0)\cdot\hat{\beta}_2\cdot \mathrm{X}_3\right\}.
\end{aligned}
\tag{A-26}
$$

A-IV. Average Interaction Effects at a Hypothetical Value

The average interaction effect can take various values because of the nonlinearity of the standard normal distribution function $\Phi(\cdot)$. Hence, we compute the average interaction effects evaluated at a hypothetical value of the continuous variable $x_2 = \dot{x}_2$ as follows:

$$
\begin{aligned}
\mathrm{AIE}(x_2=\dot{x}_2) &= \frac{1}{n}\sum^{n}\left[\frac{\Delta\partial \Pr(y=1|x_1,x_2=\dot{x}_2,\mathrm{X}_3)}{\Delta x_1 \partial x_2}\right] \\
&= \frac{1}{n}\sum^{n}\left[\left.\frac{\partial\Phi(v)}{\partial x_2}\right|_{x_1=1,x_2=\dot{x}_2}\right] - \frac{1}{n}\sum^{n}\left[\left.\frac{\partial\Phi(v)}{\partial x_2}\right|_{x_1=0,x_2=\dot{x}_2}\right] \\
&= \frac{1}{n}\sum^{n}\left[\phi(\dot{v}_1)(\beta_2+\beta_{12})\right] - \frac{1}{n}\sum^{n}\left[\phi(\dot{v}_0)\cdot\beta_2\right],
\end{aligned}
\tag{A-27}
$$

where

$$\dot{v}_1 = \beta_1 + \beta_2 \dot{x}_2 + \beta_{12} \dot{x}_2 + \mathrm{X}_3 \mathrm{B}_3, \\ \dot{v}_0 = \beta_2 \dot{x}_2 + \mathrm{X}_3 \mathrm{B}_3. \tag{A-28}$$

Appendix B. Random Effects Probit Model with an Interaction Term

Next, we consider a random effects panel probit model with an interaction term as follows:

$$\Pr(y = 1 | x_1, x_2, \mathrm{X}_3, r_i) = \Phi(v + r_i), \\ v = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathrm{X}_3 \mathrm{B}_3, \tag{A-29}$$

where r_i denotes the firm random effects.

There are at least three ways of computing predictions when the model includes the firm random effects $r_i \sim N(0, \sigma^2)$ as follows:

1. $\Pr(y = 1 | x_1, x_2, \mathrm{X}_3, r_i = 0) = \Phi(v)$, a prediction conditional on the random effects that are set to zeros.

2. $\Pr(y = 1 | x_1, x_2, \mathrm{X}_3, r_i = \hat{r}_i) = \Phi(v + \hat{r}_i)$, a prediction conditional on the random effects that are estimated by the empirical Bayes means. The empirical Bayes means of the random effects, $\hat{r}_i$, are calculated as

$$\hat{r}_i = \frac{\int r_i \Phi(y | r_i, x_1, x_2, \mathrm{X}_3; \hat{\beta}_1, \hat{\beta}_2, \hat{\mathrm{B}}_3) \phi(r_i; \hat{\sigma}) dr_i}{\int \Phi(y | r_i) \phi(r_i) dr_i}, \tag{A-30}$$

   where $\hat{\beta}_1, \hat{\beta}_2, \hat{\mathrm{B}}_3$, and $\hat{\sigma}$ denote the estimated model parameters in equation (A-29). The integration is conducted by using the adaptive Gaussian quadrature method.

3. $\Pr(y = 1 | x_1, x_2, \mathrm{X}_3) = \int \Phi(v + r_i) g(r_i | \sigma^2) dr_i$, where $g(r_i | \sigma^2)$ is the $N(0, \sigma^2)$ density function. This is the unconditional prediction with respect to the random effects, which is calculated by integrating the conditional prediction with respect to the random effects over their support. The integration is conducted by using the adaptive

Gaussian quadrature method.

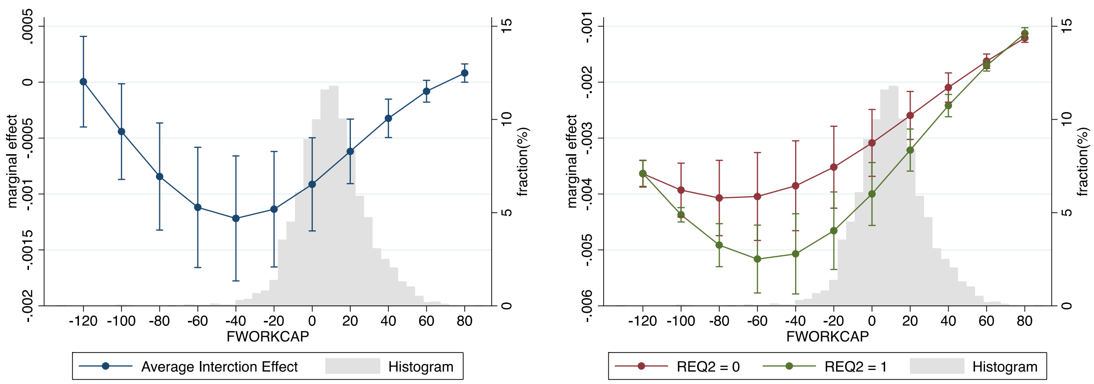
Table 1: Estimated Coefficients of the Bank Financial Health Variable, Firm Performance Variable, and their Interaction Term (1994–1999)

	Estimated coefficients (standard error)		**Descriptive statistics**		
	Our estimation	Peek and Rosengren	Mean (standard deviation)	Min	Max
REQ2	0.0556** (0.016)	0.0582** (0.0157)	0.714 (0.452)	0	1
FWORKCAP	-0.0107** (0.0010)	-0.0097** (0.0006)	10.283 (17.431)	-133.592	89.354
FROA	-0.0086* (0.0040)	-0.0075** (0.0028)	2.970 (3.292)	-27.291	26.193
REQ2×FWORKCAP	-0.0029** (0.0008)	-0.0030** (0.0006)	-	-	-
REQ2×FROA	-0.0088* (0.0039)	-0.0095* (0.0034)	-	-	-
Number of Observations			95,566		
Number of Firms			1,215		

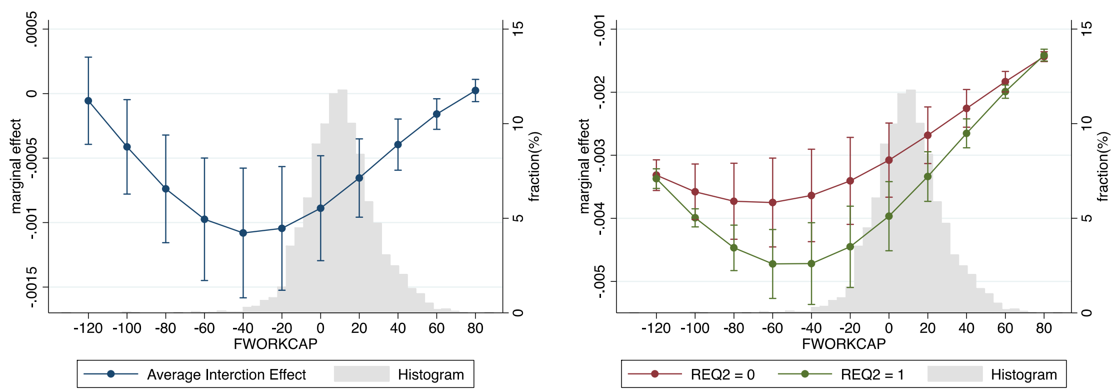
Notes: The estimation results are obtained from the random effects probit regression. The regression also includes all the other control variables of the full sample model reported in Table 5 of Peek and Rosengren (2005). Peek and Rosengren's (2005) results are obtained from their Table 5. For the estimation results, standard errors are in parentheses. *Significant at the 5 percent level; ** significant at the 1 percent level.

Figure 1: Average Interaction Effects and Average Marginal Effects of FWORKCAP

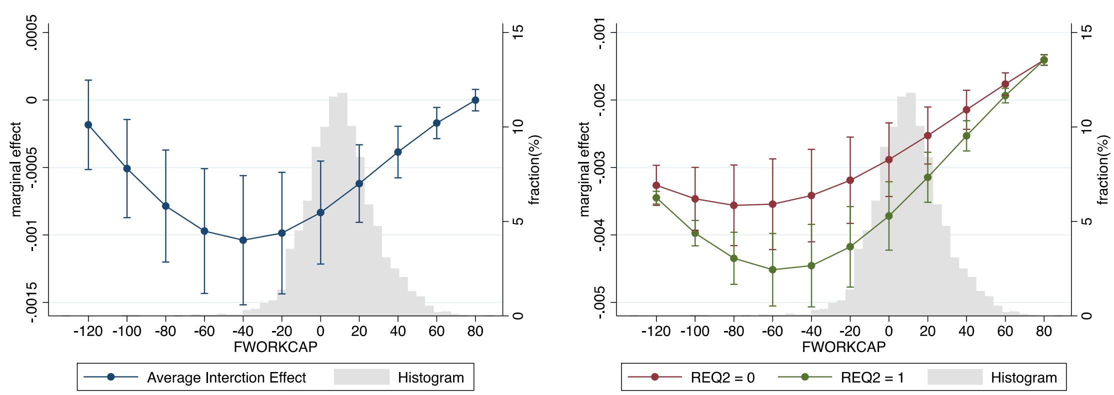
A1. The Firm Random Effects are Zero



A2. The Firm Random Effects are the Empirical Bayes Means



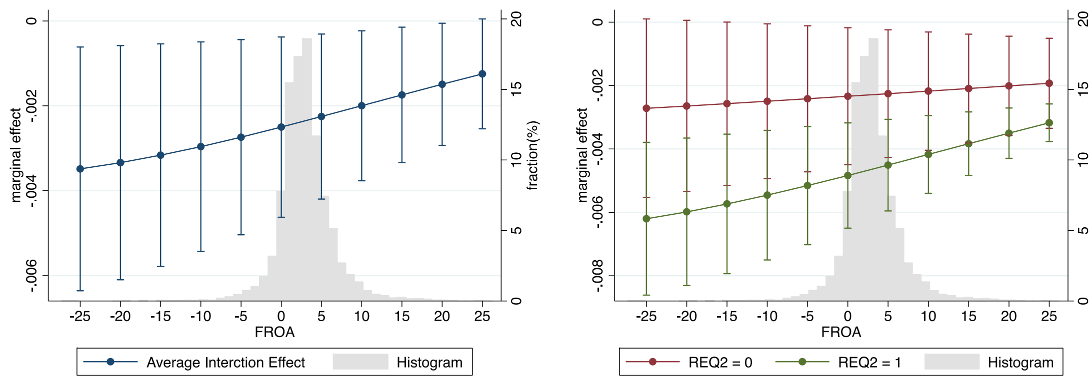
A3. The Case of Unconditional Probability



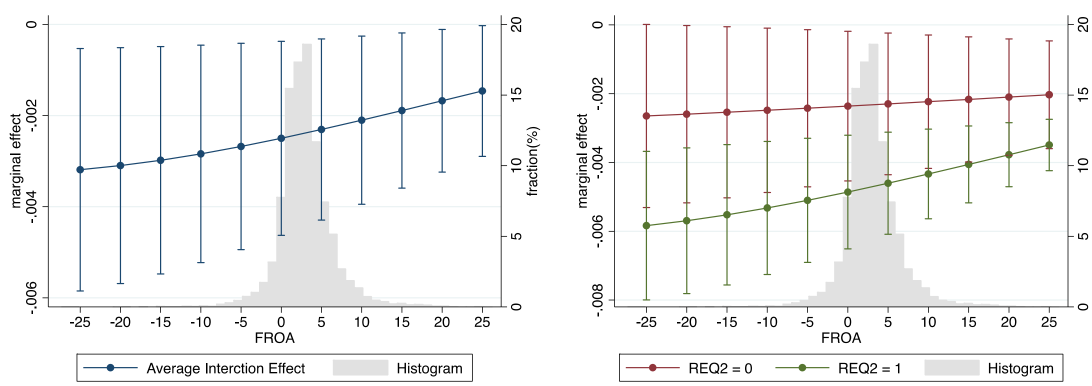
Notes: The dots (left axis) indicate the estimated average effects and the capped spikes indicate their 95% confidence intervals. The shaded areas (right axis) report a histogram of the firm performance variable.

Figure 2: Average Marginal Effects and Average Interaction Effects of FROA

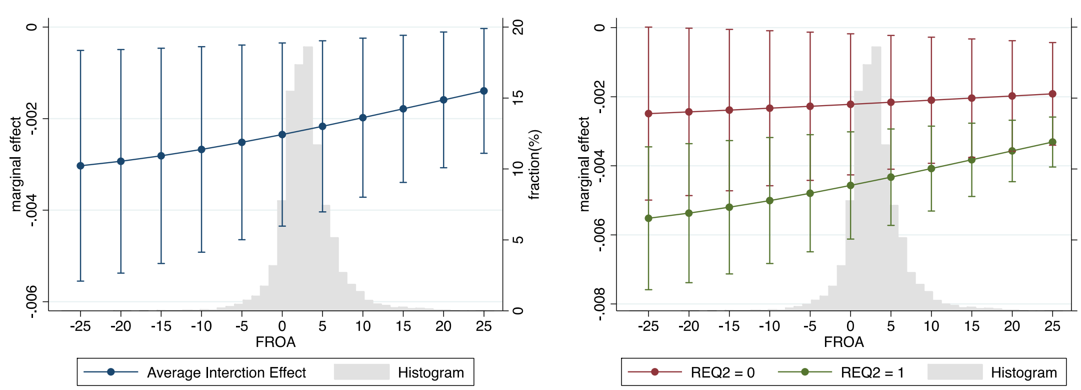
A1. The Firm Random Effects are Zero



A2. The Firm Random Effects are the Empirical Bayes Means



A3. The Case of Unconditional Probability



Notes: The dots (left axis) indicate the estimated average effects and the capped spikes indicate their 95% confidence intervals. The shaded areas (right axis) report a histogram of the firm performance variable.