Is the Treatment Worse than the Disease? Average Marginal Effects, Linear Probability Models and the Incomparability of Coefficients in Logistic Regression

Martin Elff

Zeppelin Universität Friedrichshafen (martin.elff@zu.de)

2026 Annual Meeting of the DVPW Section "Methods of Political Science"

Revision date: 2026-03-06 23:47:00+01:00

Abstract

While being a standard instrument in the toolbox of statisticians in the life sciences for a long time, the adoption of logistic regression for social science research was long delayed by the difficulties in the interpretation of its parameters, which led some scholars to not feel "comfortable" with using it (e.g. Franklin, Mackie, and Valen 1992: 438). More recently, the use of logistic regression was questioned as a means of describing effects of independent variables on binary predictors (Mood 2010) on grounds that coefficients of nested models are incomparable even when estimated with the same sample. Such comparisons are necessary to assess omitted variable biases and to distinguish between direct and indirect effects of independent variables. This has led many scholars to turn to linear probability models instead of logistic regression. Using analytical techniques and Monte Carlo simulations, the paper examines whether the comparability problem can be solved or avoided by the use of linear probability models. It further discusses the KHB-technique (Karlson, Holm, and Breen 1992), which directly addresses the comparability problem, but has not yet gained the same attention in political science that it has in sociology.

1 Introduction

In one the most cited article in social sciences,[1] Carina Mood (Mood, 2010) points a couple of problems in the interpretation of results from logistic regression. These problems have led many authors to move away from interpreting logistic regression coefficients, but base their conclusions on average marginal effects (AMEs) instead. Some authors even abandon logistic regression as a method to analyse binary dependent variables altogether and use linear probability models (LPMs) instead (e.g. Maneuvrier-Hervieu, Azzollini and Jeannet, 2025). In this and the following blog posts I'll take a look at Mood's claims and ask how serious the problems uncovered by Mood are and how helpful her suggested solutions are.

The problems Mood claims to have uncovered are:

[1] *Semantic Scholar* reports, at the time of writing, no less than 3,100 citations. for this article

1. Coefficients of nested logistic regression models cannot be compared, because coefficients will change when predictor variables are added to a model, even if these predictor variables are uncorrelated. This is problem does not occur with linear regression.

2. Coefficients of logistic regression models fitted to samples from different populations cannot be compared, because populations may have different error variances in the dependent variable.

For these reasons, Mood explicitly recommends the use of average marginal effects (AMEs) for drawing conclusions about the influence of predictor variables on binary. She also remarks that if the size of true logistic regression coefficients is small enough so that the non-linearity in the link between the probability of a positive outcome and the predictor variables is (almost) inconsequential, then the coefficients of a linear probability model are very close to the coefficients of a linear probability model fitted by OLS. So if one is interested in AMEs and the influence of the predictor variables is not "too strong", one can just avoid the computational and interpretative complications of logistic regression and use linear regression instead. However, she does not explicitly recommend the use of linear regression for binary dependent variables.

For scholars who would rather avoid the use of logistic regression and related models, e.g. because they are not "confortable" with the complications of their interpretation (e.g. Franklin, Mackie and Valen, 1992, 438), such recommendations may be reason for relief. However, the linear probability model has already for some time considered as a "defective" model of binary data (e.g. Agresti, 2002, 120). The question therefore arises whether the reasons given by Mood are strong enough to abandon logistic regression in favour of linear probability models, or whether the proposed solutions to the problems that logistic regression is allegedly afflicted by may have any misleading implications.

This question is addressed in the present paper, which is composed as follows. The next sections outlines logistic regression (or logit) modelling, the problems uncovered by Mood (2010), and her proposed solutions. Most of the rest of the paper consists on reports over several simulation studies. The third section reports the results of a simulation study that addresses Mood's first claim, i.e., that coefficients are not comparable between nested models and that average marginal effects (AMEs) and coefficients of linear probability models are not afflicted by this problem. The fourth section reports simulation results pertaining the claimed incomparability of logistic regression coefficients due to error variance heterogeneity and whether AMEs and linear regression coefficients are unafflicted by this. The fifth section demonstrates that AMEs and LPM coefficients are incomparable between populations that differ in the distribution of the independent variables of a model, a problem that does not affect logistic regression coefficients. The sixth session shows that OLS estimates are misleading even if a linear probability model describes the true population. The seventh section shows how confounding and mediation can be treated within the framework of logistic regression, without having to rely on average marginal effects or linear probability models. As usual, a concluding section provides a summary.

2 Logistic regression, average marginal effects, and linear probability models

Logistic regression is a statistical technique to describe the influence of one or more predictor variables on a binary response variable. Mathematically, the binary response variable can be represented by a sequence of random variables Y_i that take a positive value of unity ($Y_i = 1$) with some probability $\Pr(Y_i = 1) = \pi_i$, where $0 < \pi_i < 1$, and a value of zero with probability $\Pr(Y_i = 0) = 1 - \pi_i$. In a logistic regression model, the relation between the binary response and a set of m predictor variables is expressed in the following form:

$$\ln \frac{\Pr(Y_i = 1)}{\Pr(Y_i = 0)} = \ln \frac{\pi_i}{1 - \pi_i} = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi}$$

or alternatively

$$\Pr(Y_i = 1) = \pi_i = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi})}.$$

Here, x_{1i} is the value of the first predictor variable for observation (or individual or case) i, x_{2i} is the value of the second predictor variable, etc.

The logarithmic ratio $\ln \frac{\pi_i}{1-\pi_i}$ is also known as log-odds or logit, while the function $\frac{\exp(x)}{1+\exp(x)}$ is also known as the logistic function. It can also be written as $\frac{1}{1+\exp(-x)}$

Logistic regression also has a latent variable interpretation: This interpretation starts with a latent variable with values Y_i^* with

$$Y_i^* = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi} + \epsilon_i.$$

The binary response variable Y_i is created by the following dichotomization:

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{if } Y_i^* \leq 0 \end{cases}$$

If ϵ_i has a logistic distribution, i.e. a distribution with cumulative distribution function

$$\Pr(\epsilon < t) = \frac{\exp(t)}{1 + \exp(t)}$$

then the logistic regression model applies to $\Pr(Y_i) = \Pr(Y_i^* > 0)$.

In a (linear or non-linear) regression model, the *marginal effect* of a predictor variable is defined as the partial derivative of the expected value of the response regarding the values of the predictor variable.

In a linear regression model without interaction terms, a marginal effect of a variable equals its coefficient. For numeric dependent variable Y_i a linear regression model can written as

$$E(Y_i) = \mu_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi}$$

(i.e. instead of the "raw" response variable we have its expected value on the left-hand side and drop the error term on the right-hand side). The marginal effect of the h-th predictor variable with values $x_{h1}, \ldots, x_{hn}$ is then defined as:

$$\frac{\partial \mu_i}{\partial x_{hi}} = \beta_h,$$

that is, the marginal effect is identical for each observation and each potential value of the j independent variable.

In a logistic regression model, the marginal effect of an independent variable is also defined as the partial derivative of the expected value of the response variable:

$$\frac{\partial \pi_i}{\partial x_{hi}} = \frac{\partial \pi_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial x_{hi}} = \pi_i(1 - \pi_i)\beta_h.$$

In contrast to linear regression, the marginal effect is not constant, but varies with the values of the independent variables via π_i, which is a function the independent variables. It should be noted the marginal effect of the h-th independent variable not only varies with its values x_{hi} but also with the value of the other independent variables, a fact that has lead to quite some confusion.[2] It is noteworthy that $\pi_i(1-\pi_i)$ has a maximum at $\pi_i = \frac{1}{2}$, so that the maximum attainable value of the marginal effect is $\beta_h/4$.

The *average marginal effect* (AME) of the h-th independent variable in a logistic regression model can be defined as the (sample) average of the marginal effects:

$$\text{AME}_h = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial \pi_i}{\partial x_{hi}} = \frac{1}{n}\sum_{i=1}^{n}\pi_i(1-\pi_i)\beta_h$$

If the values of the predictor variables can be interpreted as independent identically distributed random variables with joint density function $f(x_1,\dots,x_m)$ then one could interpret an AME as the estimate of a population value

$$\tau_h = \int \frac{\partial \pi}{\partial x_h} f(x_1,\dots,x_m)\partial x_1 \cdots \partial x_m.$$

In the special case of a logistic regression model in which the single independent variable has as a uniform distribution

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{if } x < a \text{ or } x > b \end{cases}$$

this "population" AME equals a difference ratio:

$$\tau_h = \int_{-\infty}^{\infty}\frac{\partial \pi}{\partial x} f(x)\partial x = \frac{1}{b-a}\int_a^b \frac{\partial \pi}{\partial x}\partial x = \frac{\pi(b)-\pi(a)}{b-a}$$

While this is a very narrow special case, it may give some intuition of the meaning of average marginal effects, as well as some of their properties that emerge in the discussion below.

Linear regression involves the estimation of the coefficients of the linear model of the relation between a response variable Y_i and one or several predictor variables $x_{1i},\dots,x_{mi}$:

$$\text{E}(Y_i) = \mu_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi}$$

If Y_i is dichotomous with values 0 and 1 then $\text{E}(Y_i)$ equals $\Pr(Y_i = 1)$, at least as long the value of the expected value is between 0 and 1. This leads to the *linear probability model*:

$$\Pr(Y_i = 1) = \pi_i = g(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_m x_{mi})$$

with

$$g(x) = \begin{cases} x & \text{if } 0 < x < 1 \\ 0 & \text{if } x \le 0 \\ 1 & \text{if } x \ge 1. \end{cases}$$

As long as μ_i is between 0 and 1, the partial derivatives with respect to any value of the independent variables will equal their respective coefficients, but whenever μ_i is either less than 0 or greater than 1 the partial derivative zero.

[2]This confusion culminates in the claim that product terms are neither necessary nor sufficient for interaction effects in logistic regression (see e.g. Berry, DeMeritt and Esarey (2010)) *if* interaction effects are defined as systematic changes in marginal effects.

3 Nested models

It has been common practice in empirical research using linear regression models to compare coefficients of nested models, i.e. models that share the same dependent variable and at least one independent variable. One common use of such comparisons is to clarify the role of control variables or to demonstrate spurious relationships. Another common use is to distinguish between direct and indirect effects of certain independent variables. These uses rest on the assumption that differences in estimates of the coefficients of the same variable across different models can be attributed to the correlation of this variable with another variable that is excluded in one of the nested models and included in the other model. For this assumption to work, the coefficient of a variable should not change if an uncorrelated variable is added to the regression. While this holds for linear regression estimated by OLS, it does not for logistic regression, as Mood (2010) argues and the following simulation demonstrates.

The simulation study simulation study repeatedly fits two logistic regression models (using a maximum likelihood estimator) and two linear regression models (using an OLS estimator) to artificial data. These data are created such that two pairwise independent predictor variables of i.i.e. random numbers affect a binary response variable. Of the two logistic regression models, named `glm1` and `glm2`, and the two linear regression models, named `lm1` and `lm2`, the first includes only a single predictor variable and the other includes variables. From each logistic regression model fit to an artificial data set also the average marginal effect is computed for the predictor that is present in both models. The variable that is present in both variables is named `x1` and its true coefficient `b1`, the variable that is present only in the full model is named `x2` and its true coefficient `b2`. The values of both variables are random numbers drawn from a standard normal distribution. In each replication, the artificial data has 5000 observations in order to minimize the consequences of sampling error. For each of the nine setting of the parameters `b1` and `b2`, 100 replications are conducted, which makes $9 \cdot 9 \cdot 100 = 8100$ replications.

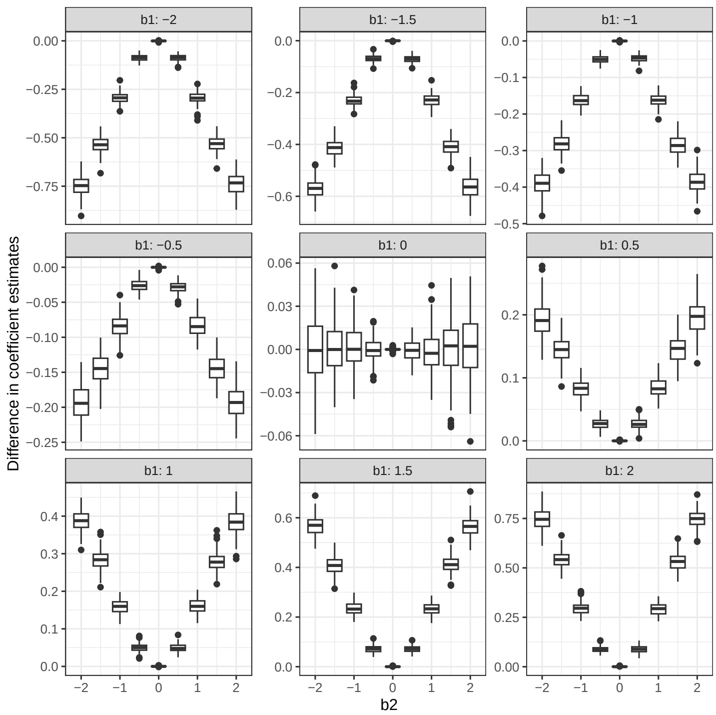
Figure 1 illustrates the differences between the estimates for `b1` based on the logistic regression models `glm1` and `glm2`. It is clear that the difference in estimates between the two models is related both to the coefficient of the variable present in both models and to the coefficient of the variable added in the full model (the model with both independent variables). If either of the coefficients is zero in the population, no systematic change in the parameter estimators occurs. Otherwise, the greater absolute value the true coefficients, the greater the change in the coefficient estimates. Further, the direction of the change is such that the absolute value of the coefficient increases after the inclusion of the additional variable, irrespective of whether the added variable has a positive or negative sign.

Whereas Figure 1 shows differences in coefficient estimates for `x1`, Figure 2 shows the differences in the average marginal effects. It strongly suggests that there are no systematic variations of differences in the AME values related to the true coefficient values, but more importantly the average change in AME values between models is always close to zero. It is only its sample variation that seems to be related to the size of the coefficient of the added variable.

Figure 3 shows the differences between OLS estimates for the coefficients of a linear probability model and the average marginal effects from a logistic regression. It makes obvious that the OLS coefficients and AMEs tend to be almost identical, except for very minor sampling variations.

The simulation study in this confirms the first set of claims made by Mood (2010): When an independent variable is added to a logistic regression model, it changes the size of the coef-

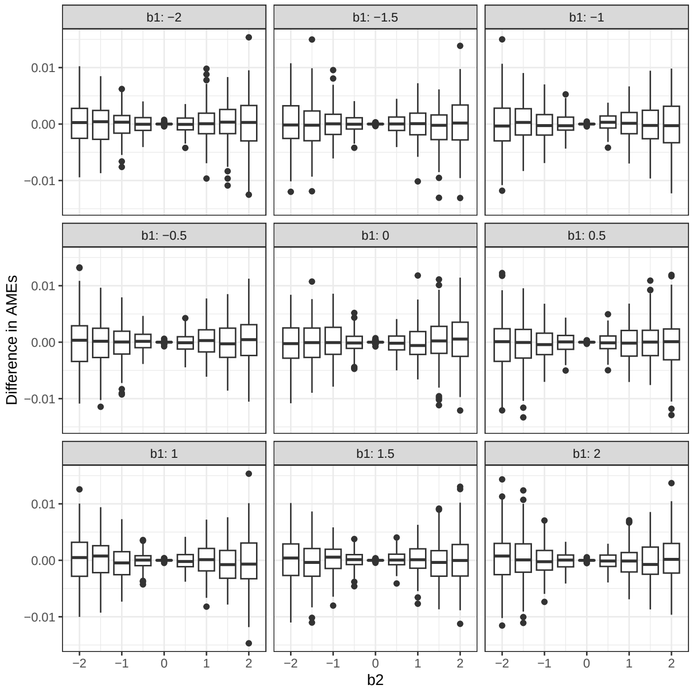
Figure 1: Change in coefficient estimates by including an additional predictor variable



Note: `b1` is the true coefficient of the predictor variable present in both models, `b2` is the true coefficient of the added variable.

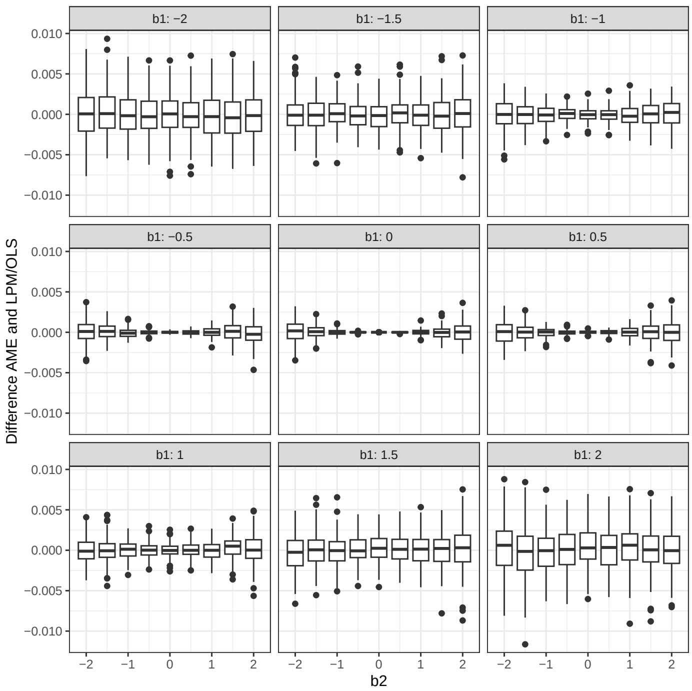
ficients already present in the model, even if the added variable is uncorrelated with them – unless the does not have a relevant influence on the dependent variable. Further, average marginal effects do not change systematically, when a variable is added to a model. Finally, OLS estimates of a linear probability model and average marginal effects from a logistic regression model tend to be almost equal.

Figure 2: Change in AMEs by including an additional predictor variable



Note: b1 is the true coefficient of the predictor variable present in both models, b2 is the true coefficient of the added variable.

Figure 3: Differences between AMEs and OLS linear probability model coefficients



Note: b1 is the true coefficient of the predictor variable present in both models, b2 is the true coefficient of the added variable.

4 Variance heterogeneity

As indicated earlier, logistic regression can be interpreted as a latent variable model: If one assumes that the observed binary variable is the result of a latent variable being dichotomized, where the conditional distribution of the latent variable is a logistic distribution with scale parameter equal to 1 and location given by the linear part of the model, then this implies that a logistic regression model applies to the binary response variable. With scale parameter equal to unity, the variance of the latent dependent variable is equal to $\pi^2/3$.

It could be argued that it is a quite strong assumption that the error term has an identical variance across all (sub-)populations. But if, for example, we have

$$Y^* = \alpha + \beta x + \epsilon, \quad \Pr(Y = 1) = \Pr(Y^* > 0)$$

where ϵ has a logistic distribution with scale parameter σ then

$$\Pr(Y = 1) = \Pr(Y^* > 0) = \Pr(\alpha + \beta x + \epsilon > 0) = \Pr(\alpha + \beta x > \epsilon) = \frac{\exp\left(\frac{\alpha+\beta x}{\sigma}\right)}{1 + \exp\left(\frac{\alpha+\beta x}{\sigma}\right)}$$

which is equivalent to a logistic regression with intercept α/σ and coefficient β/σ. For Mood (2010) this is a reason to claim that logistic regression coefficients are incomparable between (sub-)populations, because it can never be verified that the scale parameter is equal in all (sub-)populations. A counterargument against this reasoning is that setting the scale parameter $\sigma = 1$ is an identifying assumption, without which logistic regression analysis would be impossible.

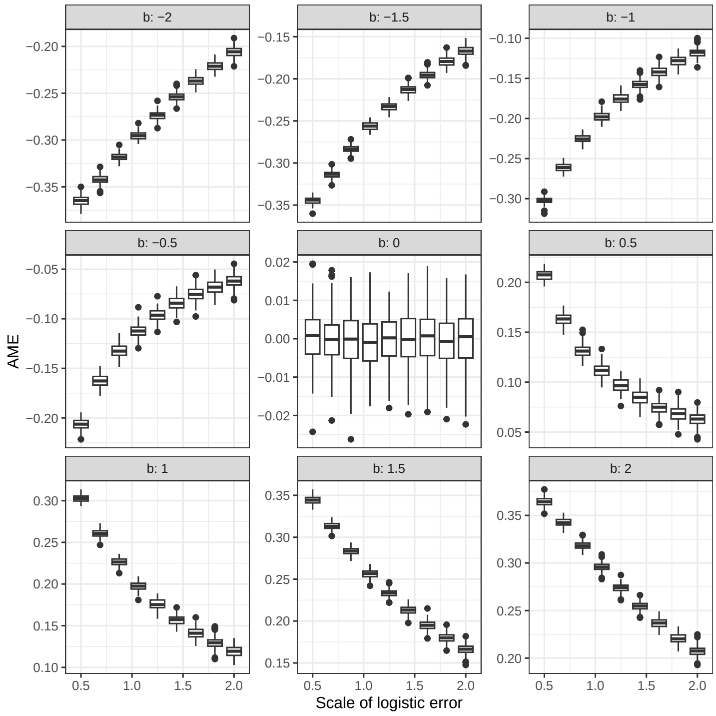
Because of the simple mathematical fact just discussed, it would be pointless to use a simulation study the influence of the scale parameter on coefficient estimates in logistic regression. However, given that average marginal effects and linear probability models may seem to have been advertised by Mood (2010) as a remedy to the alleged shortcomings of logistic regression, it may be interesting to examine the influence of the scale parameter on AMEs and OLS coefficients of linear probability models.

In the following the results of a simulation study are discussed, where in each replication a binary dependent variable is created based on the above discussed latent variable model, where the true coefficient `b` is varied, as is the scale parameter `sigma`. The single independent variable `x` has, as before, a standard normal distribution with expectation 0 and variance 1. In each replication, a logistic regression model is fitted to the binary dependent variable (with a maximum likelihood estimator) and a linear regression model is fitted (using OLS). From the logistic regression model the AME of `x` is obtained and from the linear regression model the coefficient of `x`. For each setting of `b` and `sigma` 100 replications are computed, so that the total number of replication is again $9 \cdot 9 \cdot 100 = 8100$.

Figure 4 shows the simulated distribution of the average marginal effect of the single predictor variable of a simple bivariate logistic regression. It makes clear that not only logistic regression coefficients are affected by the settings of the logistic scale parameter, but also the average marginal effects. The greater the scale value, the smaller are the AME values (unless the true logistic regression coefficient is zero) in absolute value. Figure 5 shows how OLS estimates of the coefficient in a linear probability model are affected by the scale parameter of the logistic distribution from which the data are generated. The relation between the OLS estimates, the values of scale parameter and the true values of the logistic regression coefficient is almost the same as the relation between the AMEs and the two parameters.

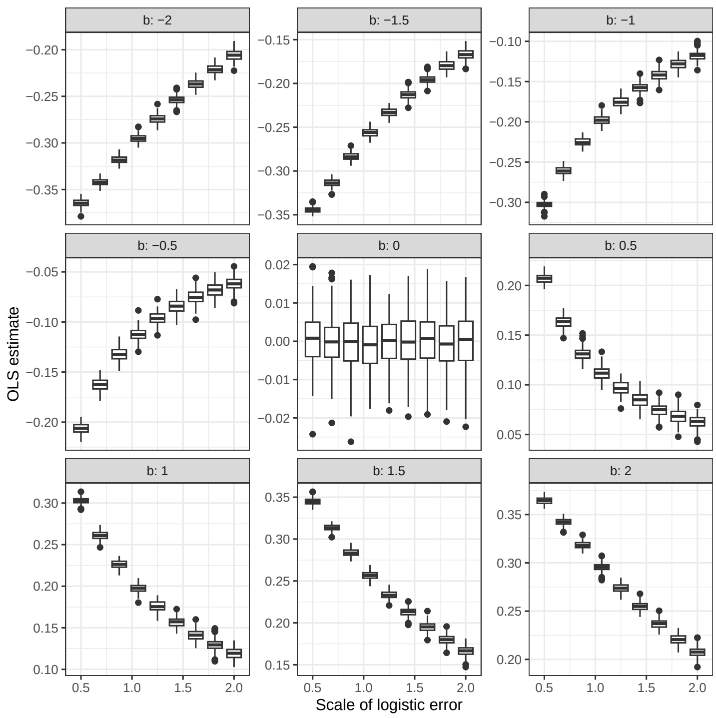
That the findings regarding the influence of the scale parameter on AME values and OLS estimates is so similar, is hardly surprising, given that we found earlier that AME values and

Figure 4: Scale of the error variance and average marginal effects



Note: b is the true coefficient of the predictor variable.

Figure 5: Scale of the error variance and OLS estimates of linear probability models



Note: b is the true coefficient of the predictor variable.

OLS estimates of a linear probability model tend to be very close. It is nevertheless non-trivial that AME values and OLS estimates vary with the logistic scale parameter, even though AME values are derived from estimated logistic regression coefficient estimates. Earlier when we examined the case of nested models, we saw that the bias in logistic regression coefficient estimates created by the absence of a relevant independent variable is just great enough so that the AME is identical in to the one of the full model. So one could have hoped that AME value are similarly immune to the variation in the logistic scale parameter. However, this hope is in vain, as the simulation study of this section indicates.

5 Means of predictor variables

Regression models, be they linear regression models, logistic regression models, or any other generalized linear models, are supposed to describe the *conditional* distribution of a response variable or dependent variable for given values of one or more independent or predictor variables. This description should be independent from the distribution of the independent variables, otherwise it would not describe a conditional distribution.[3] That is, coefficients of a linear regression or a logistic regression model will not be systematically affected by the means or variance of the independent variables, provided that the relevant model adequately describes the data. However, this is not so clear if a linear model is fitted to a binary dependent variable for which a logistic regression model is the correct one. Since average marginal effects and coefficients of linear probability models tend to be related, the same can be asked about AMEs. This question also arises if one considers the construction of an AME: The formula for the AME of the h-th independent variable in a logistic regression is

$$\text{AME}_h = \frac{1}{n}\sum_{i=1}^{n} \pi_i(1-\pi_i)\beta_h$$

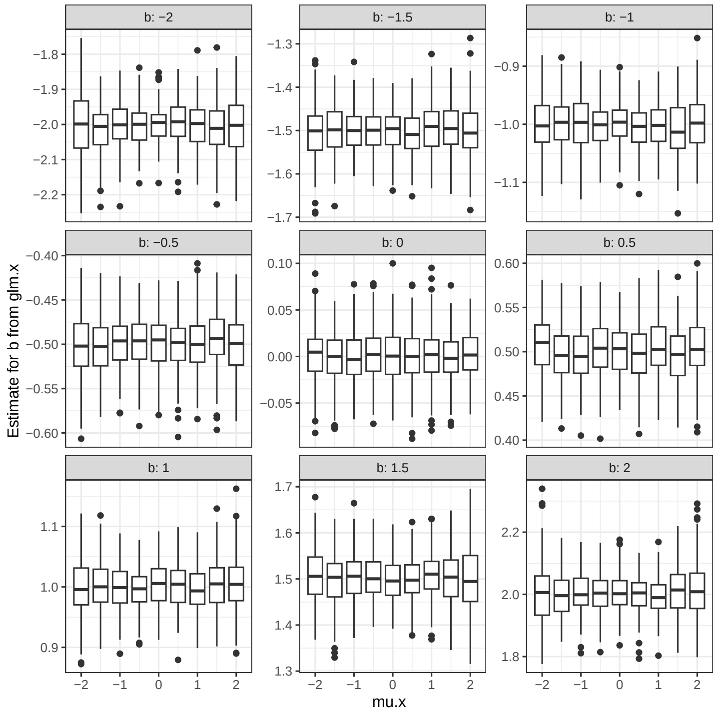
Since the predicted values π_i not only depend on the estimate on the coefficient(s) but also the values of the independent variable(s), its distribution affects the AME.

The simulation study discussed in this section proceeds as follows: In each replication values for an independent variable `x` are generated that follow a normal distribution with expected value set to `mu.x` and standard deviation (or variance) set to 1. Then a binary dependent variable is created that satisfies a logistic regression with intercept zero and coefficient `b`. To this binary dependent variable a logistic regression model named `glm.x` is fitted (using a maximum likelihood estimator), as well as a linear regression model (with OLS estimator), named `lm.x`. The sample size is 5000, like in the previous simulation studies. Based on the estimated `glm.x`, a value for the average marginal effect is computed, while the regression coefficient of `x` is obtained from the linear probability model estimated with OLS. The simulation study involves 100 replications for each setting of the parameters `mu.x` and `b`. Again, the simulation study involves $9 \cdot 9 \cdot 100 = 8100$ replications in total.

Figure 6 shows the simulated distribution of the logistic regression coefficients for the different settings of the true value of the coefficient and the expected value of the independent variable. It becomes quite clear that the estimates show only systematic variation with the true coefficient values, but are unrelated to the expected value of the independent variable. That is, the coefficient estimates behave as one should be able to expect from a description of a conditional distribution.

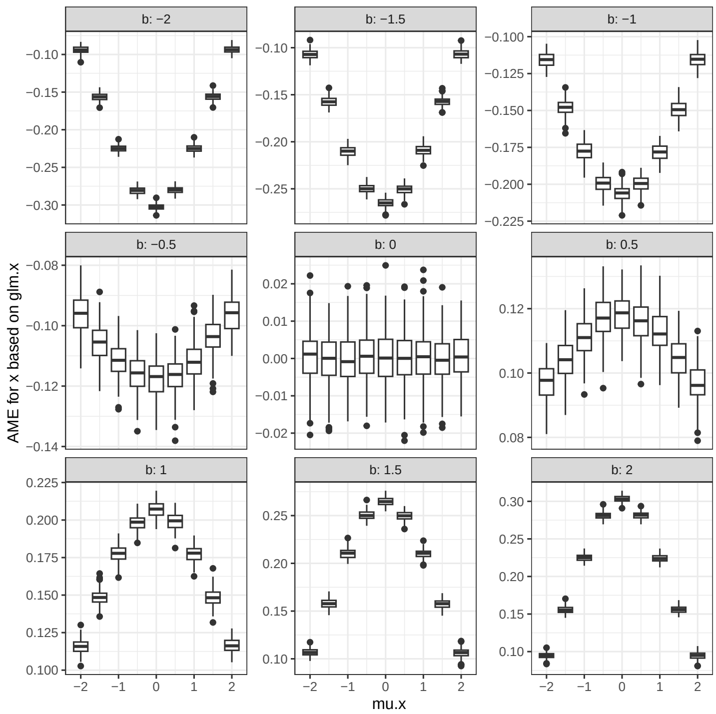
[3] The distribution of the independent variables nevertheless often has an influence on the precision in which the parameters of the conditional distribution can be estimated. For example, the variances and covariance among the independent variables in a linear regression affects the stand error of the OLS estimates.

Figure 6: Means of predictor a variable and logit coefficient estimates



Note: `b` is the true coefficient of the predictor variable, `mu.x` expected value of the predictor variable.

Figure 7: Means of predictor a variable and average marginal effects



Note: `b` is the true coefficient of the predictor variable, `mu.x` expected value of the predictor variable.

Figure 7 shows the simulated distribution of the average marginal effect for the different settings of the true value of the coefficient (b) and the expected value of the independent variable (mu.x). In contrast to the logistic regression coefficient estimates, the average marginal effect is not only influenced by the true coefficient value b, but also by the expected value of the independent variable mu.x. It appears that the absolute value of the AME reaches a maximum when mu.x is zero and gets smaller the greater the absolute value of mu.x is.[4]

The conclusions that can be drawn from this simulation study are pretty straightforward: Logistic regression coefficients can be compared between samples or populations that differ in the distribution of the independent variables, while average marginal effects (and for the same reason, OLS coefficient estimates of linear probability models) cannot. The use of average marginal effects and of linear probability coefficients can by highly misleading.

6 A linear probability model as data generator

It can be argued that in the previously discussed simulation studies the cards are stacked against the linear probability model, because the data are generated as following a (scaled) logistic regression model. One may thus ask whether OLS linear regression "works" if the data are generated according to a linear probability model. This is explored in the simulation study of the present section.

In this simulation study, in each replication values of a standard normal distributed independent variable x are generated. With an intercept a and a slope b the linear function value of the independent variable are created as `p.star = a + b*x`. The linear function values are then censored to remain between zero and unity: A linear probability value p is computed that is equal to p.star if it is greater than zero and less than unity. If p.star is greater than or equal to unity, p is set to unity. If p.star is less than or equal to zero, p is set to zero. Finally, a binary dependent variable is generated that equals unity with probability p. To this binary dependent variable a linear regression model is fitted using OLS. Again the simulation study is run with 100 replications for each of the $9 \cdot 9$ parameter settings.

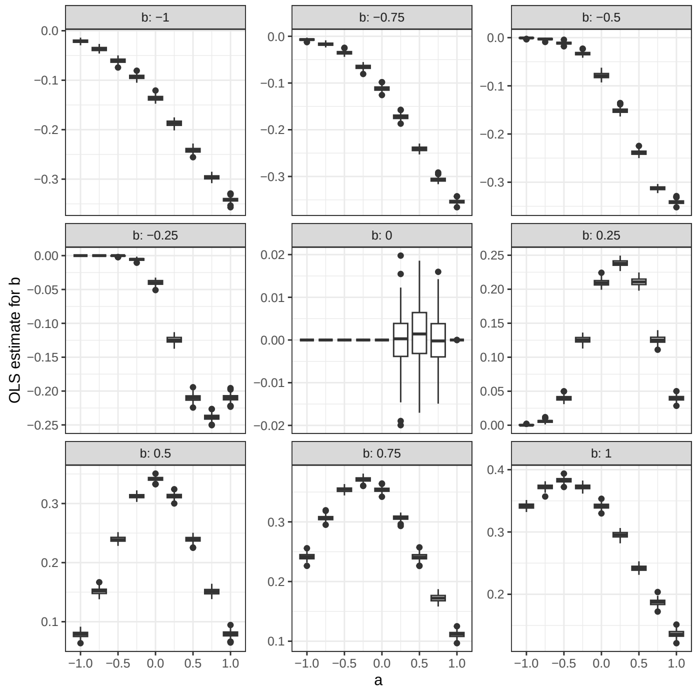
Figure 8 illustrates the distribution of the OLS estimates of the coefficient of a linear probability model for various settings of the true values of the intercept and slope coefficient. It suggests that coefficients are biased towards zero and that their average not only depends on the true coefficient values but also on the trues value of the intercept.

Figure 9 illustrates the distribution of the intercept estimates obtained by OLS. The estimates of the intercept do not show the clear bias towards zero as the estimates of the coefficients of the predictor variable.

The results of this simulation study provide quite dismal news for enthusiasts of linear probability models: The often used OLS estimator leads to strongly biased estimates of its coefficients. Even if the linear probability model is structurally correct, the preferred estimator is likely to be strongly misleading.

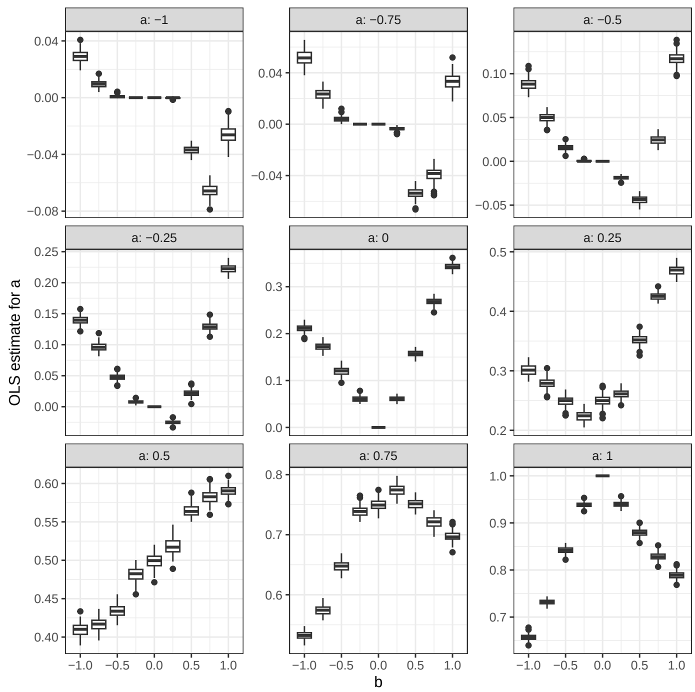
[4] In this simulation study, the true intercept in the data generating model is zero. Further simulations not reported here indicate that the maximum is attained when mu.x equals the negative of the true intercept.

Figure 8: True values and OLS estimates of the coefficient of a single predictor variable in a linear probability model



Note: b is the true coefficient of the predictor variable, a is the true intercept.

Figure 9: True values and OLS estimates of the intercept of a linear probability model



Note: b is the true coefficient of the predictor variable, a is the true intercept.

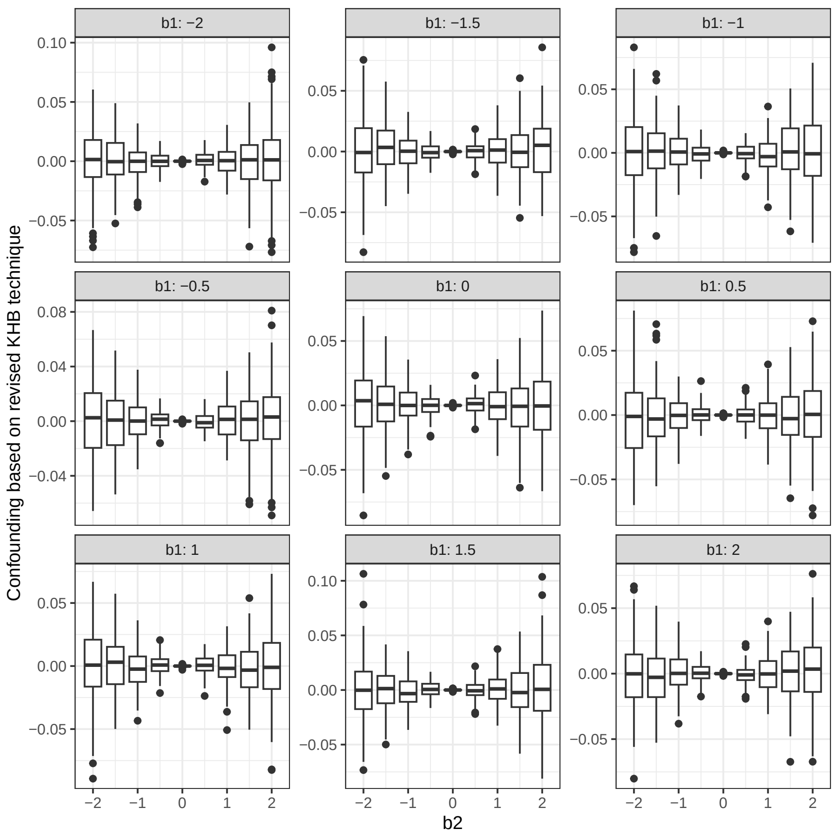
7 The KHB technique

The problem of comparing coefficients of nested logistic regression models for the purpose of uncovering confounding or describing mediation was not also addressed by other authors than Mood (2010). The most promising and influential approach (other than using AMEs or OLS regression) stems from Karlson, Holm and Breen (2012). They propose to compare logistic regression coefficients after some suitable post-hoc re-scaling that takes into account the adding variables to a logistic regression model should reduce the error variance in the dichotomized latent dependent variable. While in its original presentation the approach appears a bit complex, the revised presentation of the technique (Breen, Karlson and Holm, 2021) appears quite straightforward and easy to understand.

As an example for the application of this technique consider the comparison of two models with the same dependent variable Y (with values Y_i for the individuals in the sample) and predictor variables x_1 and x_2 (with individual values x_{1i} and x_{2i}, respectively), where the full model includes x_1 and x_2 while the simpler model, nested in the full model includes only x_1 as independent variable. Thus one may be interested to compare the coefficients of x_1 in both models, e.g. to check how much of the influence of x_1 is confounded by x_2 or how much of the influence of x_2 is mediated by x_1. Call the coefficient of x_1 in the nested smaller model $\beta_1^{(1)}$ and the coefficients of x_1 and x_2 in the full model $\beta_1^{(2)}$ and $\beta_1^{(2)}$ and their MLEs $\hat{\beta}_1^{(1)}$, $\hat{\beta}_1^{(2)}$, and $\hat{\beta}_2^{(2)}$, respectively. Call the intercepts in the two models $\alpha^{(1)}$ and $\alpha^{(2)}$ and their estimates $\hat{\alpha}^{(1)}$ and $\hat{\alpha}^{(2)}$. This leads to the linear components $\hat{\eta}^{(1)} = \hat{\alpha}^{(1)} + \hat{\beta}_1^{(1)} x_1$ and $\hat{\eta}^{(2)} = \hat{\alpha}^{(2)} + \hat{\beta}_1^{(2)} x_1 + \hat{\beta}_2^{(2)} x_2$. In this situation, the KHB approach will be based on running a linear regression with $\hat{\eta}^{(2)}$ as dependent variable and x_1 as independent variable, with coefficient estimate $\hat{\beta}_1^{(2;1)}$. Based on this "second-stage" linear regression, the KHB technique is completed by comparing $\hat{\beta}_1^{(2;1)}$ with $\hat{\beta}_1^{(1)}$ instead of comparing $\hat{\beta}_1^{(2)}$ with $\hat{\beta}_1^{(1)}$

With the data obtained from the simulation study discussed in the third section, Figure 10 allows to examine whether a modified comparison of coefficients according to the KHB approach works as expected. If the two independent variables are uncorrelated, then KHB approach should lead to equality of the (re-scaled) coefficients of the variable present in both models. That is, the difference between (re-scaled) coefficients should be zero, at least barring sampling fluctuations. The figure indicates that this is indeed the case.

Figure 10: Illustration of the (revised) KHB technique



Note: b1 is the true coefficient of the predictor variable present in both models, b2 is the true coefficient of the added variable.

8 Conclusion

The results of the present paper can be summarized as follows:

- The first simulation study confirms Mood's 2010 claim that coefficients of nested logistic regression models are indeed incomparable. It is also confirmed that average marginal effects are instead comparable between nested models. However, the KHB approach provides another technique to obtain quantities that can be compared between nested models, alternative to AMEs.

- The second simulation study confirms Mood's second claim that coefficients of logistic regression models are incomparable between populations if one relaxes the assumption of a fixed scale parameter of the logistic distribution in the latent variable interpretation of logistic regression. However, neither average marginal effects nor linear probability models can help solving this problems, since AMEs are as much affected by the scale parameter of the logistic distribution as logistic regression coefficients are.

- The third simulation study shows that AMEs (and OLS coefficients of linear probability models) are incomparable, between samples and (sub-)populations that vary in the distribution of independent variables.

- The fourth simulation study shows that OLS estimators are severely biased even if the data come from a linear probability model.

In sum, average marginal effects and linear probability models may be severely misleading in applications other than the comparison of nested models. But in they are dispensable for such comparisons. Not only is the treatment of AMEs and LPMs worse than the disease, there is no need for the treatment at all.

References

Agresti, Alan. 2002. *Categorical Data Analysis*. 2 ed. New York: Wiley.

Berry, William D., Jacqueline H. R. DeMeritt and Justin Esarey. 2010. "Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential?" *American Journal of Political Science* 54(1):248–266.

Breen, Richard, Kristian Bernt Karlson and Anders Holm. 2021. "A Note on a Reformulation of the KHB Method." *Sociological Methods & Research* 50(2):901–912.

Franklin, Mark N., Thomas T. Mackie and Henry Valen. 1992. *Electoral Change: Responses to Evolving Social and Attitudinal Structures in Western Countries*. Cambridge: Cambridge University Press.

Karlson, Kristian Bernt, Anders Holm and Richard Breen. 2012. "Comparing Regression Coefficients Between Same-Sample Nested Models Using Logit and Probit A New Method." *Sociological Methodology* 42(1):286–313.

Maneuvrier-Hervieu, Paul, Leo Azzollini and Anne-Marie Jeannet. 2025. "Cultivating Trust? The Role of European Union Investments in Bridging Rural-Urban Divides." *Electoral Studies* 98:102987.

Mood, Carina. 2010. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It." *European Sociological Review* 26(1):67–82.