

# 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](mailto:martin.elff@zu.de))

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

# Motivation I: The Problem

## Voting for Pinochet in the 1988 referendum – Logistic regression

	(1)	(2)	(3)
(Intercept)	-0.814*** (0.133)	-0.529*** (0.141)	0.233 (0.296)
Age	0.021*** (0.003)	0.022*** (0.003)	0.009 (0.007)
Male/female		-0.610*** (0.098)	-0.696*** (0.198)
Status-quo preference			3.195*** (0.144)
Log-likelihood	-1197.036	-1177.407	-368.591
N	1757	1757	1754

Significance: \*\*\*  $\equiv p < 0.001$ ; \*\*  $\equiv p < 0.01$ ; \*  $\equiv p < 0.05$

## Motivation II: Should we abandon logit coefficients?

- In a widely cited paper (> 3000 citations) Mood (2010) (published in *Europ. Soc. Rev.*) claims that
  - coefficients in logistic regression cannot be compared across nested models
  - average marginal effects (AMEs) *are* comparable across nested models
  - coefficients in logistic regression cannot be compared across populations if variance heterogeneity is possible
  - coefficients from linear regression are almost identical to AMEs if influence indep. var. is not strong – linear probability models are useful in many instances
- As a consequence many scholars (mostly sociologists) have
  - abandoned logit coefficients in favor of AMEs
  - abandoned logistic regression altogether in favour of a linear probability model
- Students in my Essex Summer School course tell me that their supervisors advise against using logistic regression and recommend linear probability models

# Concepts

## ■ Logistic regression

- Generalized linear model for binary dependent variables
- Linear part: intercept and coefficients multiplied with independent variables
- Non-linear part: Logit-link between probability of positive outcome and linear part
- Latent variable interpretation: Observed dependent variable is dichotomized latent variable – linear regression where the error term has a standard logistic distribution

## ■ Average marginal effect

- Marginal effect: Partial derivative of positive outcome probability for an indep. variable
- Average marginal effect: Sample average of marginal effects

## ■ Linear probability model

- Positive outcome probability is a *linear* function of independent variables if result is between zero and unity, otherwise censored towards zero or unity
- Usually fit with OLS like regular linear regression

## AME example: Voting for Pinochet in the 1988 referendum

	(1)	(2)	(3)
Age	0.005	0.005	0.001
Male/female		-0.146	-0.041
Status-quo preference			0.190

# LPM/OLS example: Voting for Pinochet in the 1988 referendum

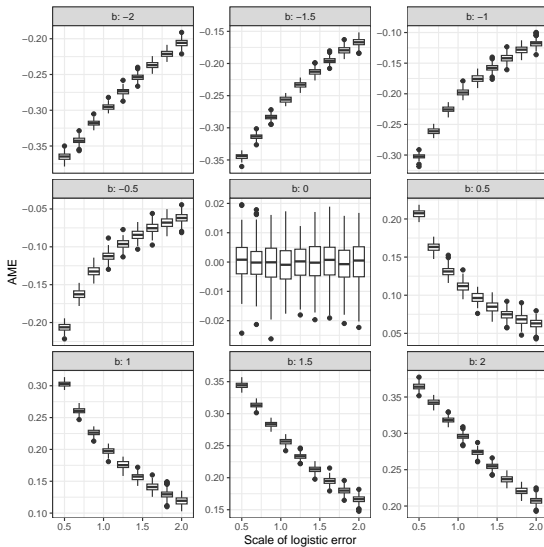
	(1)	(2)	(3)
(Intercept)	0.299*** (0.032)	0.373*** (0.034)	0.495*** (0.018)
Age	0.005*** (0.001)	0.005*** (0.001)	0.001 (0.000)
Male/female		-0.148*** (0.023)	-0.053*** (0.012)
Status-quo preference			0.390*** (0.006)
R-squared	0.023	0.045	0.733
N	1757	1757	1754

Significance: \*\*\*  $\equiv p < 0.001$ ; \*\*  $\equiv p < 0.01$ ; \*  $\equiv p < 0.05$

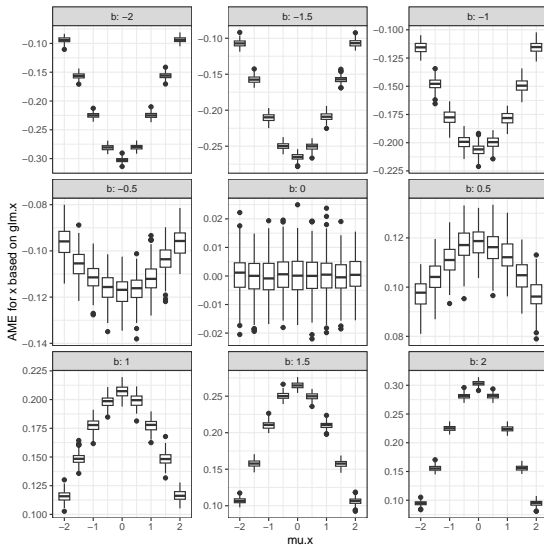
## Problems with Moods' solution – and the ensuing (mal-)practice

- AMEs and LPM/OLS *seem* to enable to assess confounding and mediation, but do not solve the problem of fixed error variance
- AMEs and LPM/OLS values are affected by distribution of independent variables and are therefore *not* comparable between (sub-)populations
- Even if LPM is the data generating process
  - OLS estimates are way off
  - Logistic regression predicts dependent variable as least as well if not better

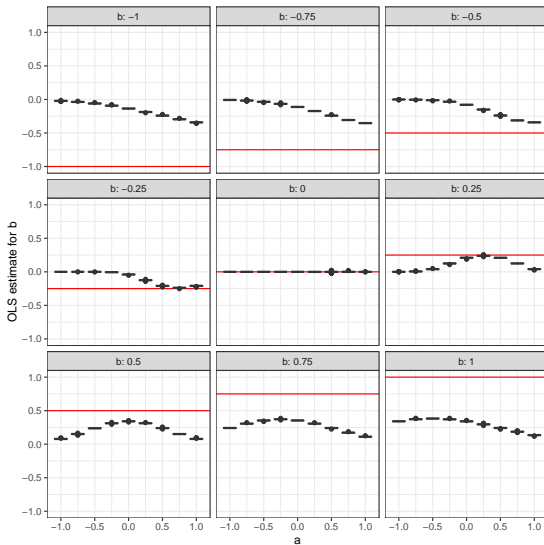
# AME and non-standard scale of error variance



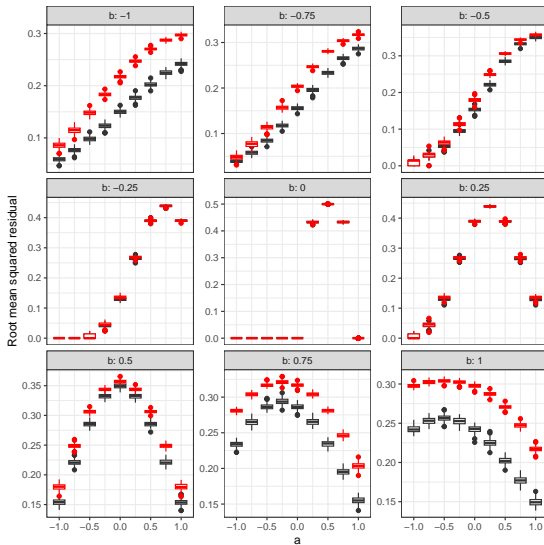
# AME and mean of independent variable



# OLS coefficients estimates and true LPM coefficients



# Root mean square residual of OLS (red) and logistic regression (black)



# The KHB approach

- Karlson, Holm and Breen (2012) develop an alternative approach to uncover confounding and/or mediation.
- Reformulation by Breen, Karlson and Holm (2021) is clear and straightforward
- Steps:
  - 1 Fit/estimate "full" model with all relevant independent variables
  - 2 Created predictions for the linear part – multiply independent variables with their coefficients and add intercept
  - 3 Use predicted linear part as dependent variable in linear regressions with reduced sets of independent variables
- Advantages
  - Does not break the "logic" of logistic regression
  - No additional uncertainty
  - Results will be correct for correctly specified "full" model

# Voting for Pinochet in the 1988 referendum – Example for the KHB approach

## Logistic regression coefficients

	(1)	(2)	(3)
(Intercept)	-0.814*** (0.133)	-0.529*** (0.141)	0.233 (0.296)
Age	0.021*** (0.003)	0.022*** (0.003)	0.009 (0.007)
Male/female		-0.610*** (0.098)	-0.696*** (0.198)
Status-quo preference			3.195*** (0.144)
Log-likelihood	-1197.036	-1177.407	-368.591
N	1757	1757	1754

Significance: \*\*\*  $\equiv p < 0.001$ ; \*\*  $\equiv p < 0.01$ ; \*  $\equiv p < 0.05$

# Voting for Pinochet in the 1988 referendum – Example for the KHB approach

Effect comparison using the KHB approach

	(1)	(2)	(3)
(Intercept)	-1.486*** (0.271)	-0.762** (0.293)	0.233 (0.296)
Age	0.045*** (0.007)	0.046*** (0.007)	0.009 (0.007)
Male/female		-1.461*** (0.205)	-0.696*** (0.198)
Status-quo preference			3.195*** (0.144)
Log-likelihood			-368.591
N			1754

Significance: \*\*\*  $\equiv p < 0.001$ ; \*\*  $\equiv p < 0.01$ ; \*  $\equiv p < 0.05$

## References

- 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.
- 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.
- 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.