# VERGLEICHENDE SOZIALFORSCHUNG MIT MEHREBENENMODELLEN IN R

Forschungspraktikum I und II
Dr. Christian Czymara
Logistic multi-level models

# AGENDA

- So far: hierarchical linear models for continuous outcomes
- And logistic regression
- Today: the combination of both
- And hierarchical linear probability models
- Exercise: Predicting political protest

# LOGISTIC MULTI-LEVEL MODELS

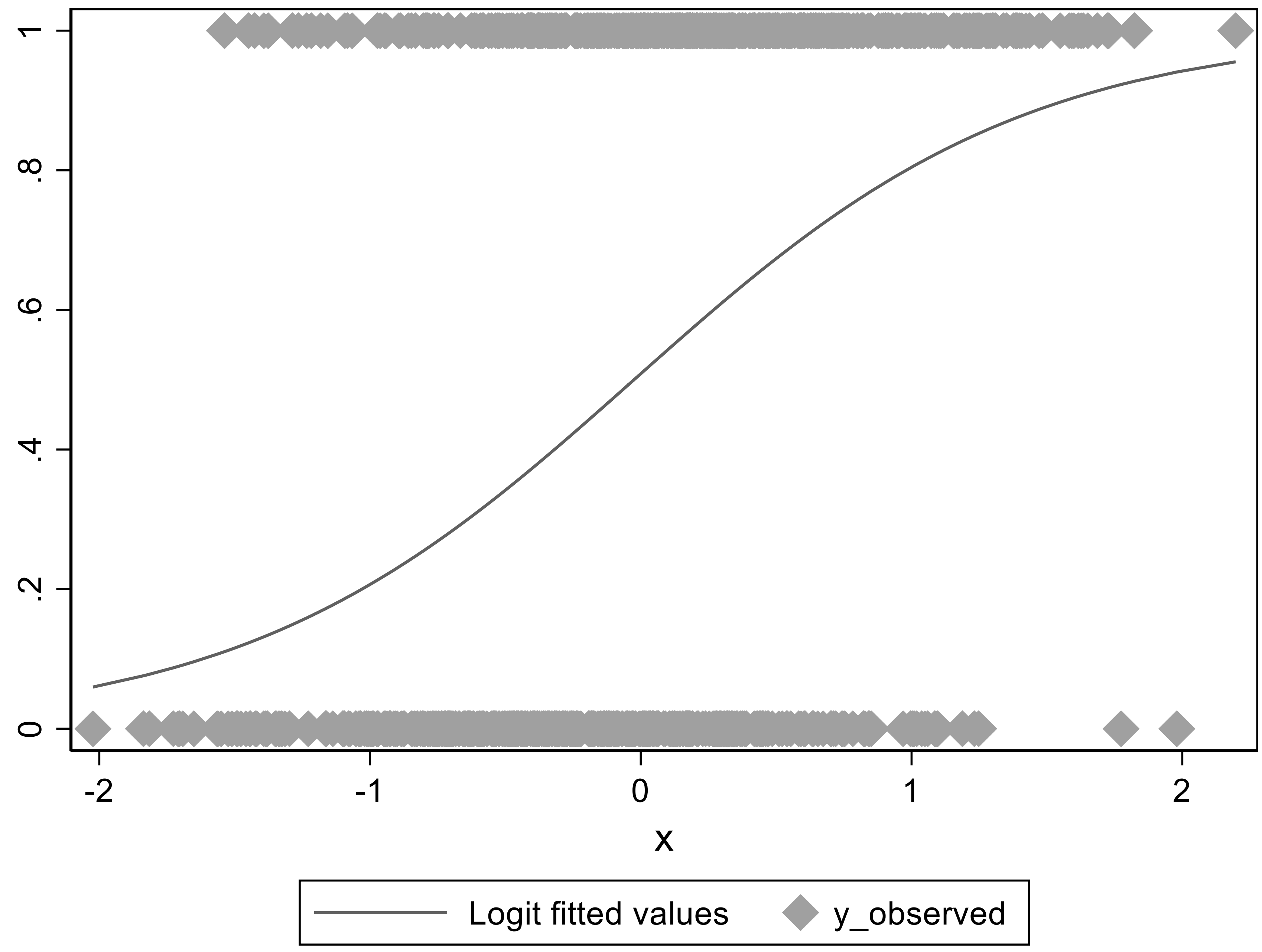
# RECAP: LOGISTIC REGRESSION

- Binary outcome
- Model probability of $y = 1$ (given $x$)

→Logistic distribution function

→$pr(y = 1) = \frac{e^{\beta_0+\beta_1 x_1}}{1+e^{\beta_0+\beta_1 x_1}} \Leftrightarrow ln\left(\frac{pr(y=1)}{1-pr(y=1)}\right) = \beta_0 + \beta_1 x_1$

- Does not minimize residuals, but maximize the likelihood



1
.8
.6
.4
.2
0
-2
-1
0
1
2
x
Logit fitted values
y_observed

# LOGISTIC LINK FUNCTION FOR MULTI-LEVEL MODELS

- Logit link function ensures predicted probabilities of [0, 1]

$$ln\left(\frac{pr(y=1)}{1-pr(y=1)}\right) = \gamma_0 + \beta x_{ij} + \ldots + \gamma z_j + u_j$$

Grand mean

Individual level variable

Country level variable

Random intercept

No individual-level error $(e_{ij})$

# LOGISTIC MULTI-LEVEL MODELS

- Interpretation of coefficients follow the same scheme of general logistic regression

1. Logit coefficients

$$ln\left(\frac{pr(y=1)}{1-pr(y=1)}\right) = \gamma_0 + \beta x_{ij} + \cdots + \gamma z_j + u_j$$

2. Odds ratios

$$\frac{pr(y=1)}{1-pr(y=1)} = e^{\gamma_0+\beta x_{ij}+\ldots+\gamma z_j+u_j} = e^{\gamma_0} * e^{\beta x_{ij}} * \cdots * e^{\gamma z_j} * e^{u_j}$$

3. Average marginal effects

4. Predicted probabilities

$$pr(y=1) = \frac{e^{\gamma_0+\beta x_{ij}+\ldots+\gamma z_j+u_j}}{1+e^{\gamma_0+\beta x_{ij}+\ldots+\gamma z_j+u_j}}$$

# INTERPRETING LOGISTIC REGRESSION COEFFICIENTS

| Coefficient | Interpretation |
|---|---|
| $\gamma_0$ | logged odds of $y = 1$ (instead of $y = 0$) when $x = 0$ and $z = 0$ (and $u_j = 0$) |
| $e^{\gamma_0}$ | odds of $y = 1$ (instead of $y = 0$) when $x = 0$ and $z = 0$ (and $u_j = 0$) |
| $\beta_1$ (or $\gamma_1$) | change in logged odds of $y = 1$ (instead of $y = 0$) for an increase in $x$ (or $z$) by one unit |
| $e^{\beta_1}$ (or $e^{\gamma_1}$) | change in odds  of $y = 1$ (instead of $y = 0$) for increase in $x$ (or $z$) by one unit → odds ratio |
| $\frac{e^{\gamma_0+\beta x_{ij}+\ldots+\gamma z_j+u_j}}{1+e^{\gamma_0+\beta x_{ij}+\ldots+\gamma z_j+u_j}}$ | Probability of  $y = 1$ given $x$, $z$ and $u_j$ |

# VARIANCE COMPONENTS FOR LOGISTIC REGRESSION

# LOGISTIC REGRESSION AND VARIANCE

- Logistic regression fixes ${\sigma_e}^2$ (variance of residual part / of level 1 error) to $\frac{\pi^2}{3}$ (probit models fixes ${\sigma_e}^2$ to 1)
- Variance components difficult to interpret
- Moreover, effects of all variables are rescaled for each model (cf. Hox 2010, p. 133 ff.)

→*Do not compare variance components between models*

→No useful interpretation of within country variance ${\sigma_e}^2$

# EXAMPLE: NULL MODEL

- Data: ESS 2002/03
- Outcome: (not) being in paid work (0: yes, 1: no)
- Cluster variable: `cntry`
- Question: How large is share of between country variance?

→estimate null model

→R command (for binary outcomes):
```
glmer(nipw ~ 1 + (1 | cntry), data = ESS, family = binomial)
```

# ICC FOR LOGISTIC MODELS

| Being unemployed | **Model 0** |
| --- | --- |
| Intercept | -0.147 * |
| *Random effects* | |
| Intercept | 0.0902 |

* p<0.05, ** p<0.01, *** p<0.001

- $\sigma_u{}^2 = 0.09$
- $\sigma_e{}^2 = \frac{\pi^2}{3}$
- $\mathrm{ICC} = \frac{0.09}{0.09 + \frac{\pi^2}{3}} = 0.027$

→ About three percent of the variance of being unemployed is between countries

# EXAMPLE: ADDING EXPLANATORY VARIABLES

- Individual level: education (`eduyrs`)
- Country level: GDP/c (`rgdpc`)
- Question: Is unemployment related to education and national economic wealth?

→
```
glmer(nipw ~ eduyrs + rgdpc + (1 | cntry), data =
ESS, family = binomial)
```

# EXAMPLE: LOGITS

| Being unemployed | Model 0 | **Model 1** |
|---|---|---|
| Education | | **-0.180 *** ** |
| GDP | | **-0.00** |
| Intercept | -0.147 * | **2.078 *** ** |
| *Random effects* | | |
| Intercept | 0.0902 | **0.0918** |

* p<0.05, ** p<0.01, *** p<0.001

- One year more of education decreases logged odds of being unemployed by -0.18
- GDP does not have a statistically significant influence
- *Differences in the random intercept ($\sigma_u{}^2$) can not be interpreted because the latent $y$ is rescaled between Model 0 and Model 1*

# HIERARCHICAL LINEAR PROBABILITY MODEL

# HIERARCHICAL LINEAR PROBABILITY MODELS

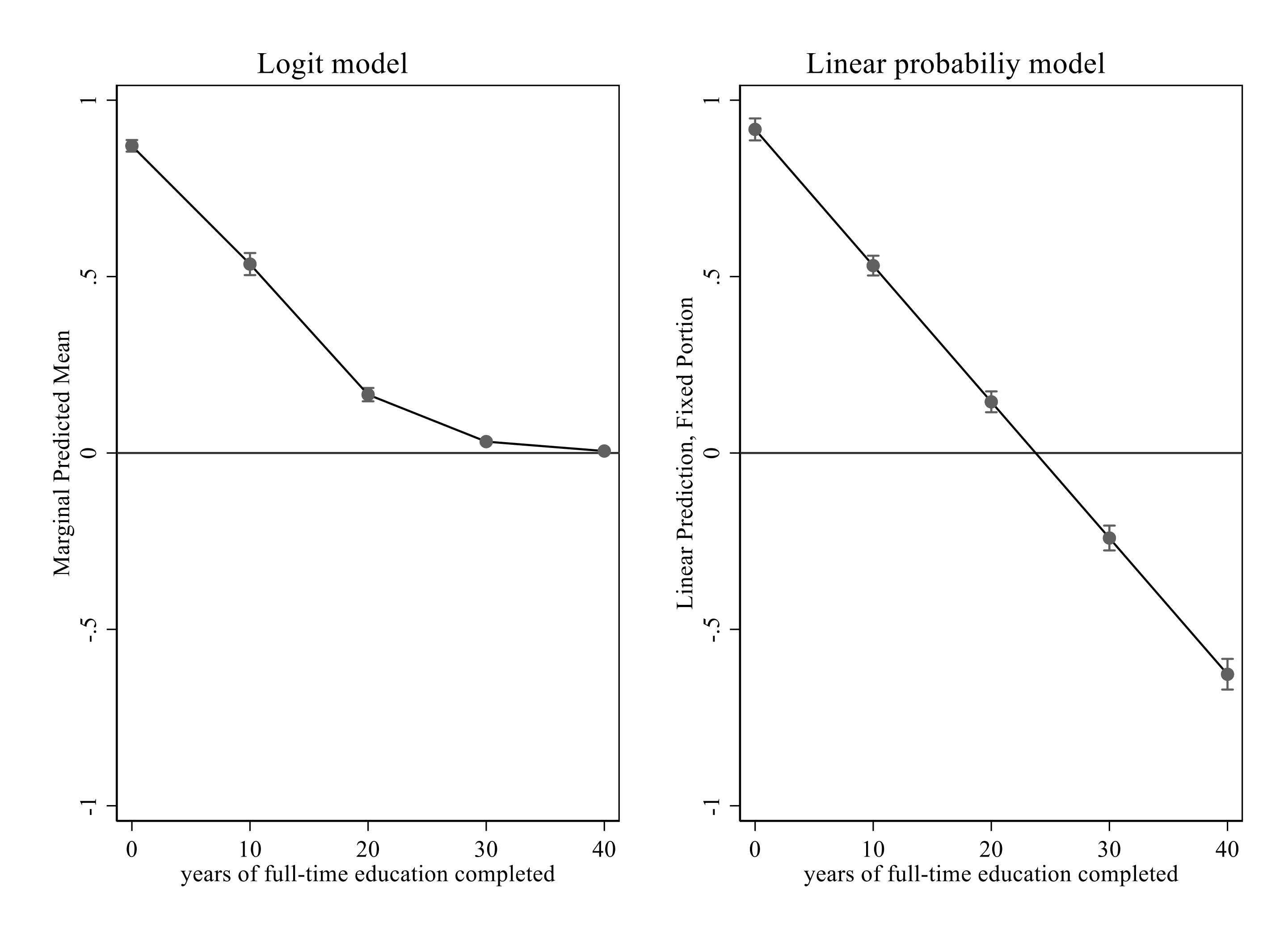
- Everything we learned about hierarchical linear models applies to hierarchical linear probability models
- Only difference: Interpretation of coefficients as "*percentage point change in probability of $y = 1$*" (see session on linear probability model)
- Main benefits
  - Easy interpretation of effects
  - Meaningful (changes of) variance components
  - Comparability of coefficients between models
  - Fast estimation
- Main drawbacks
  - Predicted probabilities can be $< 0$ or $> 1$
  - Effects do not depend on the level of $x$ (relevant when effects for "extreme" observations are calculated)

# EXAMPLE: LPM

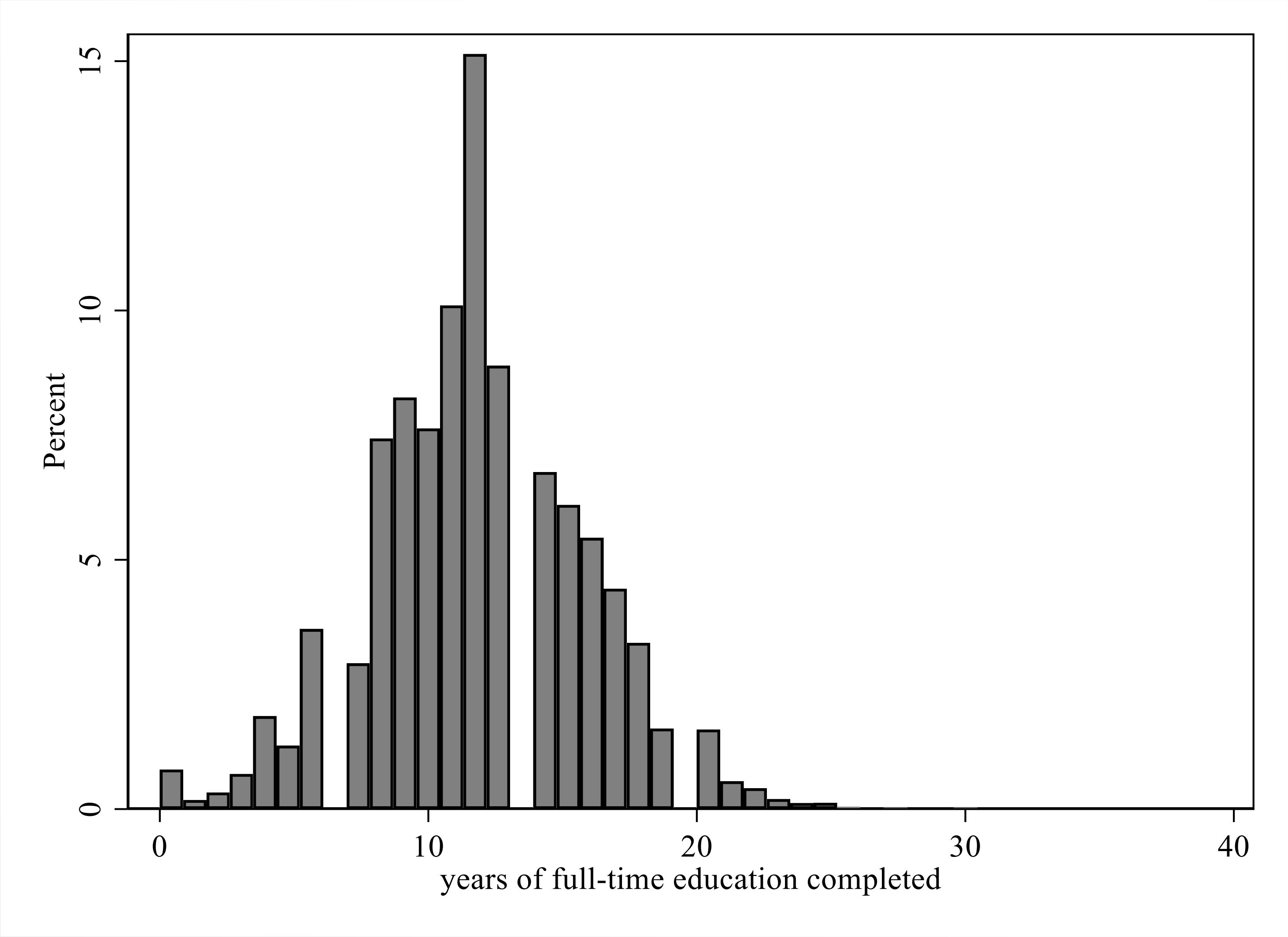
| Being unemployed | Model 0 | Logit | **LPM** |
|---|---|---|---|
| Education | | -0.180 *** | **-0.0386*** ** |
| GDP | | -0.00 | **-0.00** |
| Intercept | -0.147 * | 2.078 *** | **0.948*** ** |
| *Random effects* | | | |
| Intercept | 0.090 | 0.092 | **0.004** |
| Residual | | | **0.221** |

- One year more of education decreases probability of unemployment by 4 percentage points
- Still no significant effect of GDP

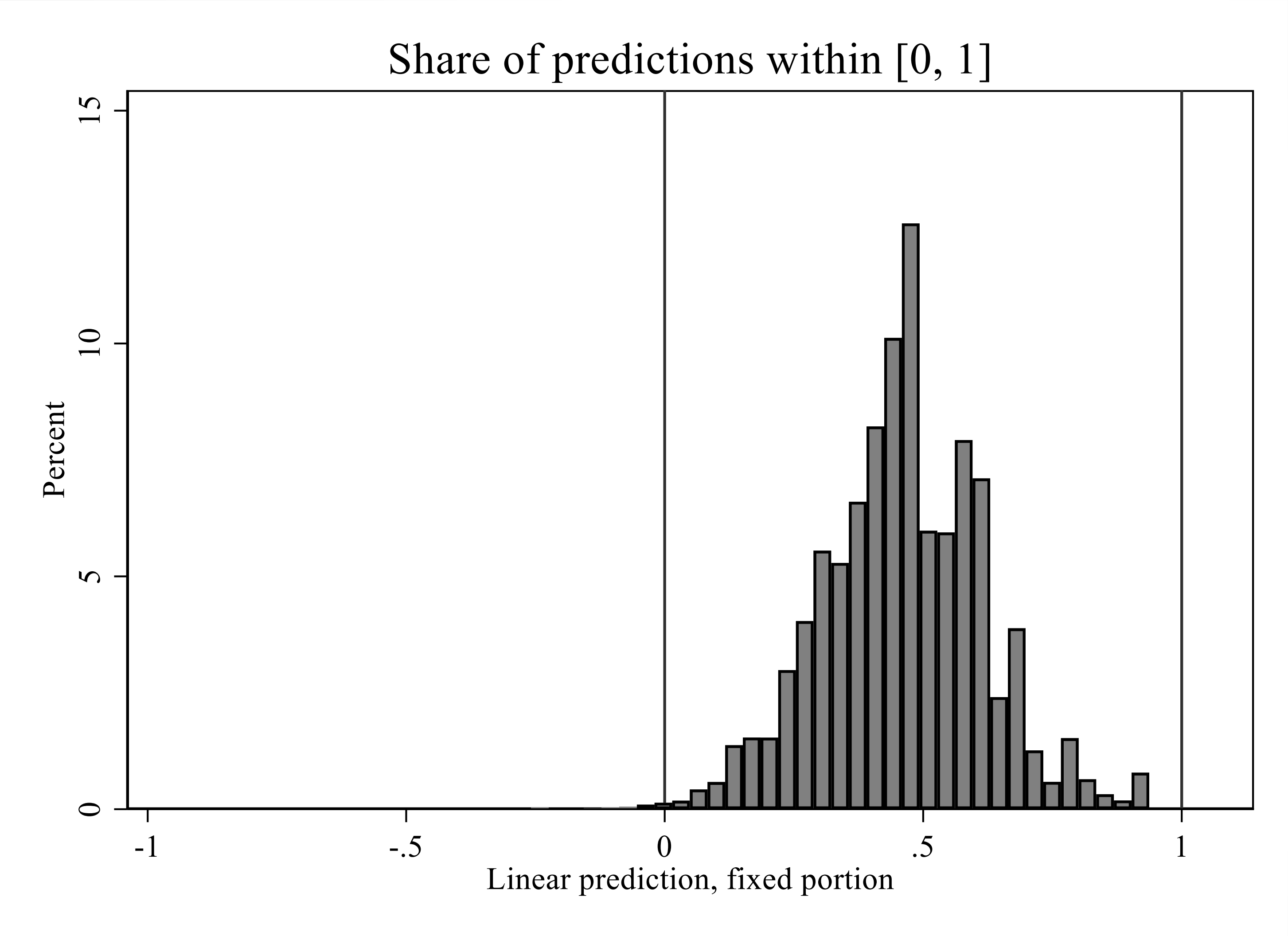
# EXAMPLE: COMPARING LOGISTIC REGRESSION AND LPM



# DISTRIBUTION OF YEARS OF EDUCATION



# PROBLEMATIC?



# EXTENSIONS OF LOGISTIC MULTI-LEVEL MODELS

# POSSIBLE EXTENSIONS

- Ordinally scaled outcome: Ordered multi-level models
- Nominally scaled outcome: Multinominal multi-level models
- Logistic random slope models

→The logic is similar to what we learned in our sessions on logistic and random slopes models, respectively

- Computational demand usually increases significantly for more complex nonlinear multilevel models

# LITERATURE

- Hox (2002): Chapter 6 in: Multilevel Analysis. Techniques and Applications. Routledge