Marginal odds ratios

What they are, how to compute them, and why we might want to use them

Ben Jann and Kristian Bernt Karlson

Methods Seminar at the Institute of Epidemiology & Health Care
University College London, January 18, 2024

Outline

Acknowledgements: This presentation is based on work which is part of a project by Kristian Bernt Karlson that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 851293).

Background

- Odds ratios form the backbone of much quantitative research in social sciences, epidemiology, and other disciplines.
- But: Falling out of favor!
  - Magnitude of odds ratios depends on unmeasured covariates orthogonal to the predictor of interest.
  - Noncollapsibility (rescaling bias).
  - Invalid cross-model and subgroup coefficient comparisons.
  - See, e.g., Allison (1999), Mood (2010), Pang et al. (2016), Breen et al. (2018), Norton, Edward and Dowd (2018), Schuster et al. (2021), Bloome and Ang (2022).

Background

- Solutions?
- KHB for cross-model comparisons (Karlson et al. 2012)
- Compare sign not magnitude
- Average marginal effects based on nonlinear probability model (AME)
- Linear probability models (LPM)

Background

- At least in social sciences, AMEs are now the standard. Some even suggest simply applying LPM, as the difference to AMEs is typically small.
  - For illustration: Papers published in the *American Sociological Review*

	2010–2015	2016–2021
"odds ratio"	32	9
"marginal effect"	2	11
"linear probability model"	3	16

- But this might be throwing out the baby with the bathwater, because ...

  ... magnitudes of AMEs likely depend on the margin,

  ... AMEs focus on absolute probability differences, not relative differences, which are key to much theory and applied research.

Background

- What we suggest (Karlson and Jann 2023):

  Use (covariate-adjusted) **marginal (log) odds ratios**, which …

  … behave like AMEs but retain the (relative) odds ratio interpretation!

  - ✔ unaffected by noncollapsibility
  - ✔ an average effect (population-averaged)
  - ✔ comparable across populations/studies

Marginal odds ratio

- Following Zhang (2008) and Daniel et al. (2021) we use potential outcomes notation to define the marginal odds ratio.
- Y_t: Potential outcome that would realize if treatment T was set to level t by manipulation (i.e., without changing anything else).
- We focus on *binary* outcomes only, that is, $Y_t \in \{0, 1\}$ (failure or success).
- Thus:

  $\Pr(Y_t = 1) = E[Y_t]$ is the (marginal) probability that Y_t will be equal to 1 (probability of success).

Marginal odds ratio

- Consider a binary treatment $T \in \{0, 1\}$.
- The marginal odds ratio (MOR) of the alternative treatment ($T = 1$) versus the standard treatment ($T = 0$) is defined as

$$\text{MOR} = \frac{\text{odds}(\Pr[Y_1 = 1])}{\text{odds}(\Pr[Y_0 = 1])}$$

  where odds(p) stands for $p/(1 - p)$.
- Interpretation of MOR: The ratio of the odds of success if everyone would receive the alternative treatment versus the odds of success if everyone would receive the standard treatment (assuming that there are no general equilibrium effects, i.e., SUTVA holds).

"Marginal" refers to how a predictor affects the "marginal distribution" of an outcome (i.e., not to a marginal change in a predictor). "Unconditional" would be another term but we use "marginal" because the term is established in the literature (Stampf et al. 2010; Karlson, Popham, and Holm 2021).

Adjusting for covariates

- The probability of success may not only depend on T, but also on other factors $\mathbf{X}$.
- Assume that $\mathbf{X}$ has a specific distribution in the population and let $\Pr(Y_t = 1|\mathbf{X} = \mathbf{x}) = E[Y_t|\mathbf{X} = \mathbf{x}]$ be the conditional success probability given $\mathbf{X} = \mathbf{x}$.
- By the law of iterated expectations,

$$\Pr(Y_t = 1) = E_{\mathbf{X}}[\Pr(Y_t = 1|\mathbf{X} = \mathbf{x})]$$

where $E_{\mathbf{X}}$ is the expectation over the distribution of $\mathbf{X}$.

Adjusting for covariates

- The marginal odds ratio, adjusting for **X**, can then be written as

$$\text{MOR} = \frac{\text{odds}(\Pr[Y_1 = 1])}{\text{odds}(\Pr[Y_0 = 1])} = \frac{\text{odds}\{E_{\mathbf{X}}[\Pr(Y_1 = 1|\mathbf{X} = \mathbf{x})]\}}{\text{odds}\{E_{\mathbf{X}}[\Pr(Y_0 = 1|\mathbf{X} = \mathbf{x})]\}}$$

- We term this the *adjusted* MOR.
- Note:
  - The adjusted MOR is the same as the unadjusted MOR by definition (i.e., same estimand)!
  - However, estimation based on the adjusted MOR formulation can be used to address confounding bias in observational data. It can also be used to increase efficiency in analyses of randomized experiments.
  - Close relationship to AME, which is defined as

$$\text{AME} = E_{\mathbf{X}}[\Pr(Y_1 = 1|\mathbf{X} = \mathbf{x})] - E_{\mathbf{X}}[\Pr(Y_0 = 1|\mathbf{X} = \mathbf{x})]$$

Continuous treatments

- In case of a continuous treatment, the MOR may depend on the level of the treatment (i.e., the MOR may not be constant).

- We define the level-specific marginal log odds ratio as the derivative of the marginal log odds by the treatment:

$$
\begin{aligned}
\ln \mathrm{MOR}(t) &= \lim_{\epsilon \to 0} \frac{\ln \mathrm{odds}[\Pr(Y_{t+\epsilon} = 1)] - \ln \mathrm{odds}[\Pr(Y_t = 1)]}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{\begin{array}{c}\ln \mathrm{odds}\{E_{\mathbf{X}}[\Pr(Y_{t+\epsilon} = 1 | \mathbf{X} = \mathbf{x})]\} \\ \qquad - \ln \mathrm{odds}\{E_{\mathbf{X}}[\Pr(Y_t = 1 | \mathbf{X} = \mathbf{x})]\}\end{array}}{\epsilon}
\end{aligned}
$$

- We can then obtain the average MOR by integrating over the distribution of T:

$$
\overline{\mathrm{MOR}} = \exp\{E_T[\ln \mathrm{MOR}(t)]\}
$$

Continuous treatments

- Another possibility is to integrate over T when obtaining the population-averaged probabilities, that is,

$$
\begin{aligned}
\ln \mathrm{MOR}' &= \lim_{\epsilon \to 0} \frac{\begin{array}{r}\ln \mathrm{odds}\{E_T[\Pr(Y_{t+\epsilon}=1)]\} \\ -\ln \mathrm{odds}\{E_T[\Pr(Y_t=1)]\}\end{array}}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{\begin{array}{r}\ln \mathrm{odds}\{E_{T,\mathbf{X}}[\Pr(Y_{t+\epsilon}|\mathbf{X}=\mathbf{x})]\} \\ -\ln \mathrm{odds}\{E_{T,\mathbf{X}}[\Pr(Y_t|\mathbf{X}=\mathbf{x})]\}\end{array}}{\epsilon}
\end{aligned}
$$

- This corresponds to the marginal odds ratio that is obtained if treatment is slightly increased for each population member, given the member's existing values for T and $\mathbf{X}$.

Relationship to the logistic model

- Consider a simple logistic model

$$\Pr(Y_t = 1) = \text{logistic}(\alpha + \delta t) \quad \text{where} \quad \text{logistic}(z) = \frac{\exp(z)}{1 + \exp(z)}$$

  which implies

$$\ln \text{odds}(\Pr[Y_t = 1]) = \alpha + \delta t$$

- Assume T is binary. We then recover the MOR as

$$\begin{aligned} \text{MOR} &= \exp\{\ln \text{odds}(\Pr[Y_t = 1]) - \ln \text{odds}(\Pr[Y_0 = 1])\} \\ &= \exp\{(\alpha + \delta) - (\alpha)\} = \exp(\delta) \end{aligned}$$

- Meaning: the (exponent of the) slope coefficient in a simple logistic regression estimates the MOR

  (The same also holds in case of a continuous treatment, which is easy to show.)

Relationship to the logistic model

- If we condition on $\mathbf{X}$, then

$$\Pr(Y_t = 1|\mathbf{X} = \mathbf{x}) = \text{logistic}(\alpha + \delta t + \mathbf{x}\beta)$$

- Here $\exp(\delta)$ is the *conditional* odds ratio (i.e., the odds ratio within a subgroup defined by a specific value of $\mathbf{X}$). For a binary treatment:

$$\text{COR} = \frac{\text{odds}\{\text{logistic}(\alpha + \delta + \mathbf{x}\beta)\}}{\text{odds}\{\text{logistic}(\alpha + \mathbf{x}\beta)\}} = \exp(\delta)$$

- This conditional odds ratio (COR) is different from the (covariate-adjusted) MOR, which has a more involved form. For example, in case of a binary treatment:

$$\text{MOR} = \frac{\text{odds}\{E_{\mathbf{X}}[\text{logistic}(\alpha + \delta + \mathbf{x}\beta)]\}}{\text{odds}\{E_{\mathbf{X}}[\text{logistic}(\alpha + \mathbf{x}\beta)]\}}$$

  This will be different from COR when $\beta \neq \mathbf{0}$.

Marginal vs. conditional odds ratios

- The difference between the COR and the (covariate-adjusted) MOR is referred to as *noncollapsibility* or *rescaling bias*.

- "Noncollapsibility of the OR derives from the fact that when the expected probability of outcome is modeled as a nonlinear function of the exposure, the marginal effect cannot be expressed as a weighted average of the conditional effects" (Pang et al. 2016).

- The (covariate-adjusted) MOR will be attenuated compared to the COR, what is commonly referred to as rescaling effect.
  - For example, if there is just a single covariate, the relationship between MOR and COR can be approximated by

$$\ln \text{MOR} = \frac{\ln \text{COR}}{\sqrt{1 + 0.35\beta^2 \text{var}(X)}}$$

Key message

MOR and COR correspond to different estimands!
They are conceptually different.

- Both are valid estimands. Why should we prefer one over the other?
  1. While there exists only one MOR, there are many CORs, as the latter depend on the conditioning set **X**. That is, the interpretation of the COR depends on the covariates included in the regression equation.
  2. The MOR has an interpretation similar to an AME on the probability scale: it quantifies the "population response" to a treatment.
  3. Because the MOR is unaffected by noncollapsibility, it can be used to compare results from same-sample models including different covariates (e.g. effect decompositions in mediation analysis).
  4. The MOR is straightforward to compare across different studies or populations as it does not depend in arbitrary ways on the conditioning set.

Two illustrations

1. Academic ability and intergenerational college mobility
   - This example illustrates the difference between the conditional OR and the (covariate-adjusted) marginal OR.
2. College gap in attitudes toward racial segregation
   - This example illustrates the difference between average marginal effects (AMEs) and marginal odds ratios.

Academic ability and intergenerational college mobility

- Comparison of “secondary effects” of family background on educational attainment between the United States (National Longitudinal Survey of Youth, 1979 cohort) and Denmark (Danish Longitudinal Survey of Youth).
- How much does educational attainment (here: going to college) depend one whether parents have college education, once we control for academic ability (the “primary effect”)?

Academic ability and intergenerational college mobility

Odds ratios of parental college attainment gap in college attainment unadjusted and adjusted for academic ability. The United States and Denmark. Standard errors in parentheses.

	USA (N = 10,068)	**DNK** (N = 2,185)	**USA/DNK**
Unadjusted odds ratio	7.7 (0.46)	3.8 (0.55)	2.03*
Conditional odds ratio adjusted for academic ability	3.4 (0.23)	2.9 (0.45)	1.17
Marginal odds ratio adjusted for academic ability	2.5 (0.13)	2.5 (0.33)	1.00

Notes: US data are from the NLSY79; the Danish data are from the Danish Longitudinal Survey of Youth. * indicates that the country difference in log odds ratios is statistically significant at a five percent level.

(Karlson and Jann 2023)

Academic ability and intergenerational college mobility

- The unadjusted or gross marginal odds ratio is about twice as large in the US as in Denmark, meaning that Denmark is significantly more educationally mobile.
- However, the “secondary effects” of family background are of similar magnitude in the two countries (adjusted marginal odds ratio). This means that academic ability “mediates” a significantly larger portion of the overall effect in the US than in Denmark.
- Had we been using the conditional odds ratio, we would have concluded that, net of academic ability, Denmark is a (albeit only slightly) more educationally mobile country than the US.
- This is because the attenuating impact of noncollapsibility is more pronounced in the US than in Denmark (since academic ability is a much stronger predictor of college attainment in the US).

College gap in attitudes toward racial segregation

- Analysis of how the gap in attitudes toward racial segregation between respondents with and without a four-year college degree changed over time.
- General Social Survey cumulative file, 1976–1996
- Outcome: Agreement with claim that white people have a right to keep black people out of their neighborhoods.
- Predictor of interest: college attainment
- Controls: survey year, age, gender, race, marital status, liberal–conservative scale
- Joint model across time points with survey year fully interacted with college attainment and all other covariates; we obtain AMEs and MORs from the model at different years.

College gap in attitudes toward racial segregation

Average marginal effects and marginal odds ratios of the college gap in the attitude toward racial segregation in 1976, 1981, 1986, 1991, and 1996. Standard errors in parentheses.

	AME	**ln(MOR)**
1976	-0.213	-1.006
1981	-0.181	-0.983
1986	-0.147	-0.959
1991	-0.115	-0.934
1996	-0.088	-0.909
1976–1996 difference	0.125 (0.028)	0.097 (0.211)
1976–1996 proportional reduction	58.5%	9.7%

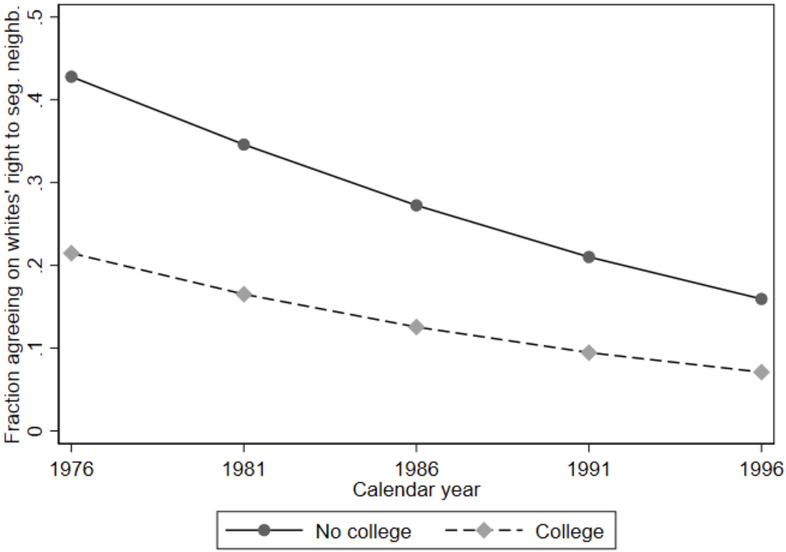
Note: MOR is marginal odds ratio; AME is average marginal effect. Estimates are adjusted for gender, race, age, marital status, and overall political view. Data are from General Social Surveys Cumulative File, N = 12,239.

(Karlson and Jann 2023)

College gap in attitudes toward racial segregation

- Main finding is that the absolute college gap (as measured by the AME) in the attitude toward racial segregation has declined significantly over the 20-year period, whereas the relative gap (as measured by the MOR) remained practically unchanged.
- The absolute gap reduced over time because there is a general decline in support of racial segregation, and this decline is steeper among the non-college educated in absolute terms because they start at a higher level than the college educated. However, the relative difference does not change much (see figure).
- The example illustrates how the MOR can be an informative complement to the AME.

College gap in attitudes toward racial segregation



(Karlson and Jann 2023)

Estimation

- Estimand $\Rightarrow$ Estimation
- There are several approaches how we can estimate the adjusted MOR.
  - G-computation (using predictions from a model)
  - Inverse probability weighting
  - Unconditional logistic regression (RIF regression)
- All are discussed in detail in Jann and Karlson (2023) (for binary/categorical as well as continuous treatments; including formulas for analytic standard errors based on influence functions).
- Our preferred method is G-computation as it closely resembles the formulation of the adjusted MOR above. That is, G-computation obtains the MOR that is *implied* by the chosen logit model. The other methods follow a somewhat different logic.

G-computation

- G-computation estimates the MOR using counterfactual predictions from a logit model (or any other model in principle).

- For example, for a binary treatment, the procedure is as follows.
  1. Regress Y on T and $\mathbf{X}$ using logistic regression (or, in principle, any other model).
  2. Use the model estimates to generate two predictions for each observation, one with T set to 0 and one with T set to 1, that is, $\widehat{p}_i^0 = \widehat{Pr}(Y = 1 | T = 0, \mathbf{X} = \mathbf{x}_i)$ and $\widehat{p}_i^1 = \widehat{Pr}(Y = 1 | T = 1, \mathbf{X} = \mathbf{x}_i)$.
  3. Average the predictions across the sample to obtain estimates of population-averaged success probability by treatment level, that is, $\overline{p}^0 = \frac{1}{W}\sum_{i=1}^n w_i \widehat{p}_i^0$ and $\overline{p}^1 = \frac{1}{W}\sum_{i=1}^n w_i \widehat{p}_i^1$ where w_i are sampling weights and W is the sum of weights.
  4. Finally, plug these averaged predictions into the formula for the MOR:

$$\ln \widehat{\mathrm{MOR}} = \ln \mathrm{odds}(\overline{p}^1) - \ln \mathrm{odds}(\overline{p}^0)$$

- Note that, in case of a binary treatment, `margins` followed by `nlcom` can be used to do the above computations.

G-computation

- For continuous treatments we evaluate level-specific MORs (using analytic derivatives) at each level of the treatment (possibly using an approximation grid) and then average over the treatment distribution (not directly possible with `margins`).
- An alternative approach is based on applying fractional logit to averaged counterfactual predictions at each value of T. For binary/categorical treatments this leads to the same results as the procedure above. For continuous treatments results slightly differ (due to the different implicit averaging). Nonetheless we prefer this approach due to its generality and flexibility.

Software

- We provide three Stata packages for the estimation of marginal odds ratios, each implementing one of the three approaches (Jann and Karlson 2023). All packages provide consistent standard errors and support complex survey estimation.
  - `lnmor`: G-computation (https://github.com/benjann/lnmor)
  - `ipwlogit`: Inverse probability weighting (https://github.com/benjann/ipwlogit)
  - `riflogit`: Unconditional logistic regression (https://github.com/benjann/riflogit)
- Installation:
  - Type

    ```
    ssc install name, replace
    ```

    where *name* is the package name.

Example application: Gender gap in STEM

- Data:
  - cohort of school leavers in Switzerland in 2016 (9th grade, age 15)
  - dependent variable (two years later): in educational track leading to a STEM (Science, Technology, Engineering, and Math) profession
  - Swiss panel study on "Transitions from Education to Employment" (www.tree.unibe.ch, Hupka-Brunner et al. 2021, TREE 2021)

```
. use stem, clear
(Excerpt from TREE cohort 2)

. describe

Contains data from stem.dta
 Observations:         6,809                  Excerpt from TREE cohort 2
    Variables:             7                  22 Oct 2023 21:07
--------------------------------------------------------------------------------

Variable      Storage   Display    Value
    name         type    format    label      Variable label
--------------------------------------------------------------------------------
stem            byte    %8.0g                 Is in STEM training in 2018
male            byte    %8.0g                 Is male
mathscore       double  %10.0g                Math score in 2016
repeat          byte    %8.0g                 Ever repeated a grade
books           byte    %19.0g     books      Number of books at home
wt              double  %10.0g                Sampling weight
psu             int     %8.0g                 Sampling unit
--------------------------------------------------------------------------------

Sorted by:
```

Example application: Gender gap in STEM

- Probability difference

```
. mean stem [pw=wt], over(male) cluster(psu)
Mean estimation                               Number of obs = 6,809

                  (Std. err. adjusted for 800 clusters in psu)
------------------------------------------------------------
             |             Robust
             |       Mean   std. err.     [95% conf. interval]
-------------+----------------------------------------------
 c.stem@male |
          0  |    .163234   .0093646      .1448519    .1816161
          1  |   .2748687   .0145161      .2463745    .3033629
------------------------------------------------------------

. regress stem i.male [pw=wt], cluster(psu) noheader
(sum of wgt is 78,600.1929332293)
                              (Std. err. adjusted for 800 clusters in psu)
----------------------------------------------------------------------------
             |               Robust
        stem | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+--------------------------------------------------------------
      1.male |   .1116347   .0142969     7.81   0.000     .0835708    .1396987
       _cons |    .163234   .0093653    17.43   0.000     .1448506    .1816174
----------------------------------------------------------------------------
```

Example application: Gender gap in STEM

- Unadjusted (marginal) odds ratio

```
. logit stem i.male [pw=wt], or cluster(psu) nolog
Logistic regression                                     Number of obs =  6,809
                                                        Wald chi2(1)  =  67.37
                                                        Prob > chi2   = 0.0000
Log pseudolikelihood = -40949.278                       Pseudo R2     = 0.0172
                                  (Std. err. adjusted for 800 clusters in psu)
------------------------------------------------------------------------------
             |               Robust
        stem | Odds ratio   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
      1.male |   1.943131   .1572663     8.21   0.000     1.658099     2.27716
       _cons |   .1950773   .0133746   -23.84   0.000     .1705485    .2231338
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

Example application: Gender gap in STEM

- How does the gender gap change once covariates such as math skills are taken into account?
- Conventional approach: conditional odds ratio

```
. logit stem i.male mathscore i.repeat books [pw=wt], or cluster(psu) nolog
Logistic regression                                     Number of obs =  6,809
                                                        Wald chi2(4)  = 596.03
                                                        Prob > chi2   = 0.0000
Log pseudolikelihood = -31905.554                       Pseudo R2     = 0.2343
                                  (Std. err. adjusted for 800 clusters in psu)
------------------------------------------------------------------------------
             |               Robust
        stem | Odds ratio   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
      1.male |   1.959295   .1675426     7.87   0.000      1.65696    2.316794
   mathscore |   2.606164   .1252437    19.93   0.000     2.371897     2.86357
    1.repeat |   .6563627   .0965248    -2.86   0.004     .4920011    .8756321
       books |   1.087051   .0341241     2.66   0.008     1.022185    1.156034
       _cons |   .1058314   .0166897   -14.24   0.000     .0776926    .1441616
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

Example application: Gender gap in STEM

- Computation of covariate-adjusted marginal odds ratio using `lnmor`
  - `lnmor` is a post-estimation command, i.e. first estimate the model, then apply `lnmor`

```
. lnmor i.male, or
Enumerating predictions: male..done
Marginal odds ratio                              Number of obs     =      6,809
                                                 Command           =      logit
                                  (Std. err. adjusted for 800 clusters in psu)
------------------------------------------------------------------------------
             |               Robust
        stem | Odds Ratio   std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      1.male |   1.677032   .1103015     7.86   0.000     1.473911    1.908145
------------------------------------------------------------------------------
```

Example application: Gender gap in STEM

- Compare results (SEs in parentheses)

ln(OR)	Unadjusted	COR	MOR
1.male	0.664	0.673	0.517
	(0.0809)	(0.0855)	(0.0658)

OR	Unadjusted	COR	MOR
1.male	1.943	1.959	1.677
	(0.157)	(0.168)	(0.110)

Example application: Gender gap in STEM

- `lnmor` allows you to store the influence functions of the estimates.
- Influence functions (IFs) are awesome!
- For example, here is how you could construct a test for confounding using the IFs.

```
. quietly logit stem i.male [pw=wt], cluster(psu)
. quietly lnmor i.male, nodots rif(RIF*)
. quietly logit stem i.male mathscore i.repeat books [pw=wt], cluster(psu)
. quietly lnmor i.male, nodots rif(RIFadj*)
. quietly total RIF2 RIFadj2 [pw=wt], cluster(psu)
. lincom RIFadj2 - RIF2
 ( 1)  - RIF2 + RIFadj2 = 0
```

Total	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
(1)	-.1472752	.0420574	-3.50	0.000	-.2298313	-.0647192

```
. drop RIF*
```

Some further functionality

- Obtain results for several predictors in one call

```
. logit stem i.male mathscore i.repeat books [pw=wt], cluster(psu)
```

(*output omitted*)

```
. lnmor i.male mathscore i.repeat books, or
(mathscore has 491 levels; using 82 binned levels)
Enumerating predictions: male..mathscore.....................................
............................................repeat..books.......done
Marginal odds ratio                                 Number of obs     =      6,809
                                                    Command           =      logit
                                  (Std. err. adjusted for 800 clusters in psu)
```

stem	Odds Ratio	Robust std. err.	t	P>|t|	[95% conf.	interval]
1.male	1.677032	.1103015	7.86	0.000	1.473911	1.908145
mathscore	2.544257	.1227974	19.35	0.000	2.314279	2.797088
1.repeat	.7244256	.0839026	-2.78	0.006	.5771099	.9093457
books	1.065542	.025807	2.62	0.009	1.01607	1.117423

Some further functionality

- Using `at()` to evaluate interactions

```
. logit stem i.male##c.mathscore##c.mathscore##i.repeat##c.books [pw=wt], ///
>     cluster(psu)
```

(*output omitted*)

```
. lnmor i.male, nodots or at(repeat)

Marginal odds ratio                                     Number of obs     =      6,809
                                                        Command           =      logit

Evaluated at:
    1: repeat = 0
    2: repeat = 1

                                  (Std. err. adjusted for 800 clusters in psu)
```

stem	Odds Ratio	Robust std. err.	t	P>|t|	[95% conf.	interval]
1						
1.male	1.700763	.1254126	7.20	0.000	1.471573	1.965648
2						
1.male	1.514474	.3348553	1.88	0.061	.9812346	2.337496

Some further functionality

- Using at() to evaluate interactions

```
. lnmor i.male, nodots or at(mathscore = -2(2)2)
Marginal odds ratio                                Number of obs     =      6,809
                                                   Command           =      logit

Evaluated at:
    1: mathscore = -2
    2: mathscore = 0
    3: mathscore = 2
                                  (Std. err. adjusted for 800 clusters in psu)
```

stem	Odds Ratio	Robust std. err.	t	P>|t|	[95% conf.	interval]
1						
1.male	1.697829	.6740845	1.33	0.183	.7788088	3.701323
2						
1.male	1.890954	.2009018	6.00	0.000	1.535003	2.329448
3						
1.male	1.991302	.3565062	3.85	0.000	1.401245	2.829831

Some further functionality

- Nonlinear effects: polynomials

```
. logit stem i.male c.mathscore##c.mathscore i.repeat c.books [pw=wt], ///
>     cluster(psu)
```

(*output omitted*)

```
. lnmor c.mathscore##c.mathscore
(mathscore has 491 levels; using 82 binned levels)

Enumerating predictions: mathscore.........................................
.....................................done

Marginal log odds ratio                         Number of obs     =      6,809
                                                Command           =      logit

                                 (Std. err. adjusted for 800 clusters in psu)
------------------------------------------------------------------------------
             |               Robust
        stem | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   mathscore |   1.022883   .0698188    14.65   0.000     .8858334    1.159933
             |
 c.mathscore#|
 c.mathscore |  -.0757118   .0276166    -2.74   0.006    -.1299215   -.0215022
------------------------------------------------------------------------------
```

Some further functionality

- Nonlinear effects: level-specific MORs using option `dx()`

```
. lnmor mathscore, or dx(-3(1)3)
Enumerating predictions: mathscore.......done
Marginal odds ratio                               Number of obs     =      6,809
                                                  Command           =      logit
                                                  Type of dx()      =     levels
                                    (Std. err. adjusted for 800 clusters in psu)
```

stem	Odds Ratio	Robust std. err.	t	P>|t|	[95% conf.	interval]
mathscore@l1	4.423561	1.002176	6.56	0.000	2.835543	6.900933
mathscore@l2	3.805173	.6532834	7.78	0.000	2.716539	5.33007
mathscore@l3	3.259279	.3848859	10.01	0.000	2.584945	4.109527
mathscore@l4	2.779502	.1942306	14.63	0.000	2.423232	3.18815
mathscore@l5	2.378501	.1148555	17.94	0.000	2.163402	2.614986
mathscore@l6	2.056937	.1577385	9.40	0.000	1.769484	2.391087
mathscore@l7	1.789337	.2248394	4.63	0.000	1.398208	2.289877

```
Terms affected by dx(): mathscore
Levels of dx(): -3 -2 -1 0 1 2 3
```

Discussion

- We provide a clear definition of the marginal OR (clarification of estimand).

- We illustrate the advantages of the marginal odds ratio over the conditional odds ratio; we illustrate the value of the marginal odds ratio as a complement to AMEs.

- We provide flexible software that can estimate the marginal OR for categorical as well as continuous predictors (including support for complex surveys).

- But ...

  ... is it worth the hassle? How much do applied researchers love odds ratios?

  ... will it change practice?

References I

Allison, P.D. 1999. Comparing Logit and Probit Coefficients across Groups. *Sociological Methods & Research* 28:186–208.

Bloome, D., S. Ang. 2022. Is the Effect Larger in Group A or B? It Depends: Understanding Results from Nonlinear Probability Models. *Demography* 59:1459–88.

Breen, R., K.B. Karlson, A. Holm. 2018. Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models. *Annual Review of Sociology* 44:39–54.

Daniel, R., J. Zhang, and D. Farewell. 2021. Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets. *Biometrical Journal* 63(3): 528–557.

Hupka-Brunner, S., B. Jann, M. Koomen, D. Krebs-Oesch, T. Meyer, B. Müller, C. von Rotz, S. Sacchi, B. Wilhelmi. 2021. TREE2 study design. Bern: TREE. https://doi.org/10.7892/boris.152018

Jann, B. 2022. VIOLINPLOT: Stata module to draw violin plots. Available from https://ideas.repec.org/c/boc/bocode/s459132.html.

Jann, B., K.B. Karlson. 2023. Estimation of marginal odds ratios. University of Bern Social Sciences Working Papers 44. https://ideas.repec.org/p/bss/wpaper/44.html

Karlson, K.B., B. Jann. 2023. Marginal Odds Ratios: What They Are, How to Compute Them, and Why Sociologists Might Want to Use Them. *Sociological Science* 10: 332–347. https://doi.org/10.15195/v10.a10

References II

Karlson, K. B., F. Popham, and A. Holm. 2021. Marginal and Conditional Confounding Using Logits. *Sociological Methods & Research*.

Mood, C. 2010. Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do about It. *European Sociological Review* 26:67–82.

Norton, E.C., B.E. Dowd. 2018. Log Odds and the Interpretation of Logit Models. *Health Services Research* 52(2):859–78.

Pang, M., J.S. Kaufman, R.W. Platt. 2016. Studying Noncollapsibility of the Odds Ratio with Marginal Structural and Logistic Regression Models. *Statistical Methods in Medical Research* 25:1925–1937.

Schuster, N.A., J.W.R. Twisk, G. ter Riet, M.W. Heymans, J.J.M. Rijnhart. 2021. Noncollapsibility and Its Role in Quantifying Confounding Bias in Logistic Regression. *BMC Medical Research Methodology* 21:136.

Stampf, S., E. Graf, C. Schmoor, and M. Schumacher. 2010. Estimators and confidence intervals for the marginal odds ratio using logistic regression and propensity score stratification. *Statistics in Medicine* 29: 760–769.

TREE. 2021. Transitions from Education to Employment, Cohort 2 (TREE2), Panel waves 0-2 (2016-2018) [Dataset]. University of Bern. Distributed by FORS, Lausanne. https://doi.org/10.23662/FORS-DS-1255-1

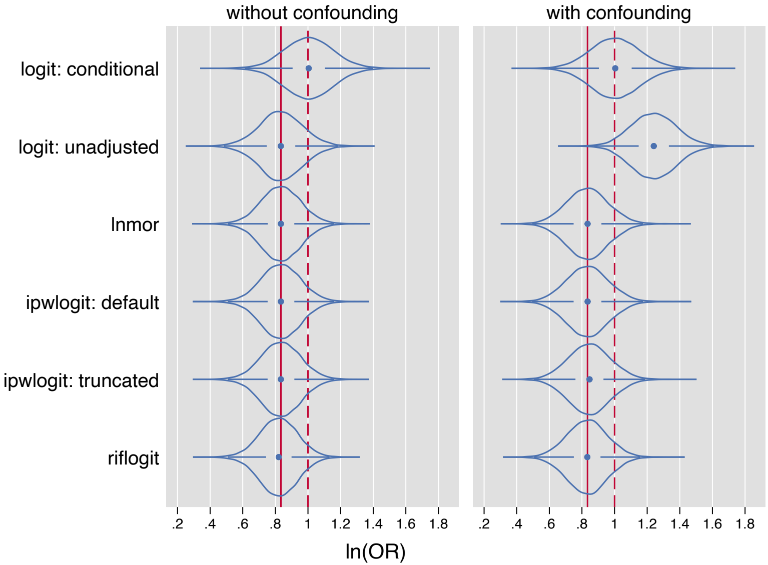
Zhang, Z. 2008. Estimating a Marginal Causal Odds Ratio Subject to Confounding. *Communications in Statistics – Theory and Methods* 38(3): 309–321.

Appendix: Some simulation results

- Setup
  - Binary outcome Y depends on treatment T and control variable X through a logistic model.
  - The effects of T and X on Y (the conditional log odds ratios) are set to 1 in all simulations (intercept is 0).
  - X has a standard normal distribution.
  - T is either binary or continuous.
  - Two scenarios:
    1. unconfounded: T is independent from X (X has an even distribution if binary and a standard normal distribution if continuous)
    2. confounded: T depends on X (binary: logistic model with slope 0.5; continuous: linear model with slope 0.5 and standard normal errors)
  - 10'000 replications.
  - Using `violinplot` (Jann 2022) to display results.

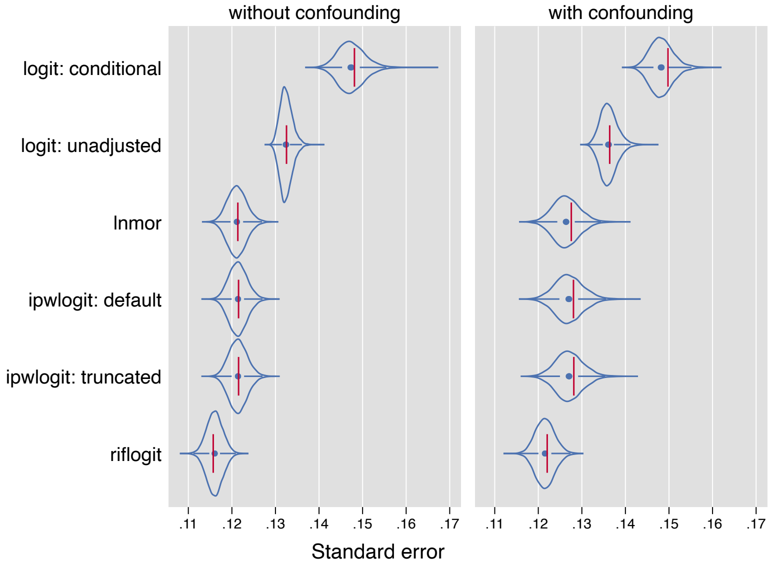
Appendix: Some simulation results

- Distribution of effect estimates for binary treatment



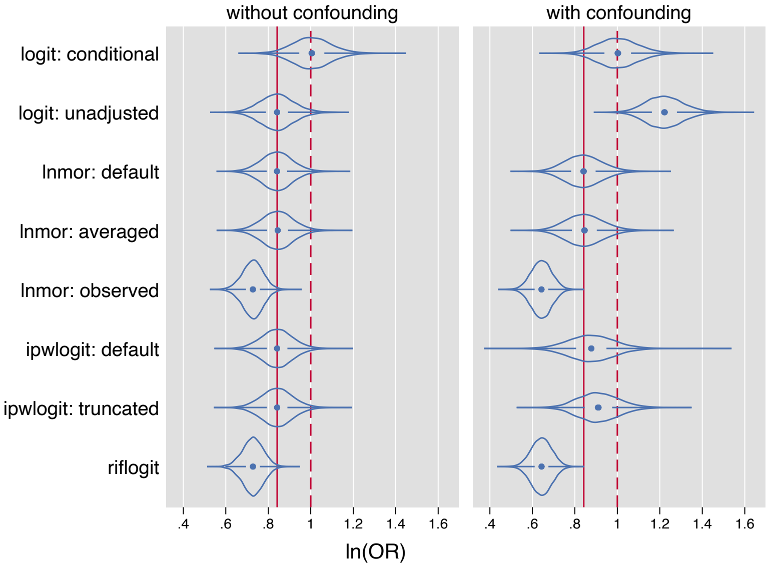
Appendix: Some simulation results

- Distribution of standard errors for binary treatment



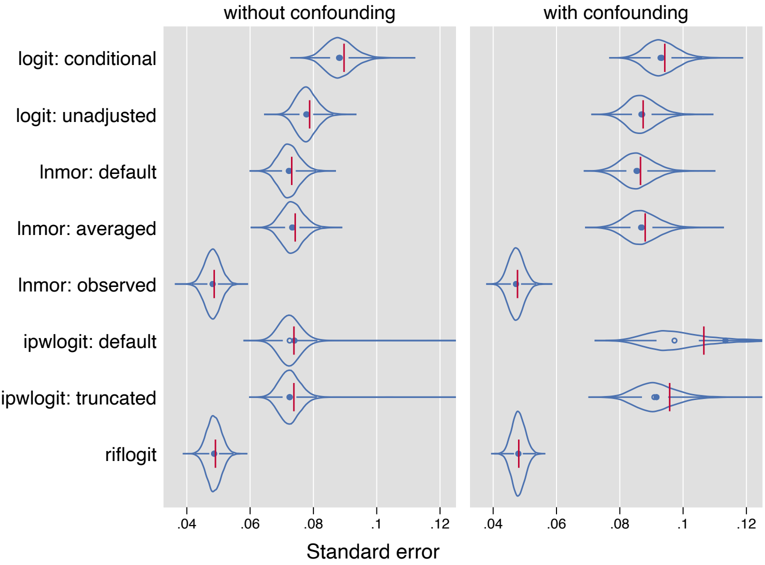
Appendix: Some simulation results

- Distribution of effect estimates for continuous treatment



Appendix: Some simulation results

- Distribution of standard errors for continuous treatment



Appendix: Some simulation results

- Binary treatment:
  - All estimators appear to work well.
  - However, note that the treatment has an even distribution in these simulations; may need to do more simulations with uneven distributions.
- Continuous treatment:
  - IPW does not fully remove confounding. Furthermore, stability of IPW becomes problematic. Truncation helps somewhat but also increases bias.
  - MOR' ("observed") is a different estimand than $\overline{\mathrm{MOR}}$ ("averaged"). RIF logit appears to approximate MOR', not $\overline{\mathrm{MOR}}$.