

Profesor: Héctor Bahamonde, PhD.

e: `hector.bahamonde@uoh.cl`

w: `www.hectorbahamonde.com`

Curso: MLE.

TA: Gonzalo Barria.

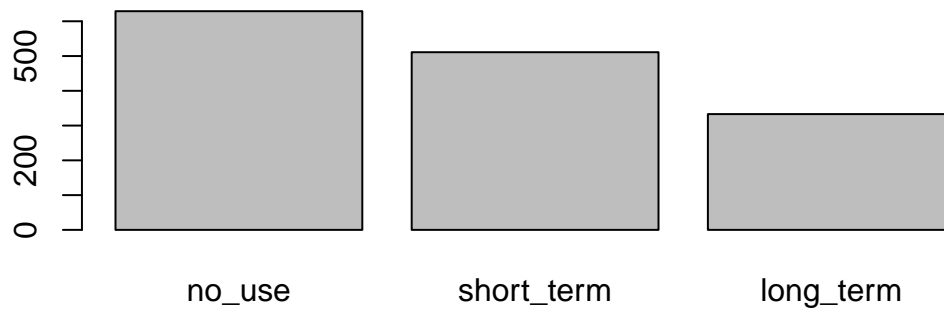
I. 1

Carguemos los datos

```
load(url("https://github.com/hbahamonde/MLE/raw/master/Datasets/Contraceptive_Method_Choice.RData"))
```

Ahora, veamos de qué se trata la variable dependiente.

```
plot(dat$cmc)
```



- 1.1 Es un multinomial ordenado. Corresponde a un ordered probit o un ordered logit.
- 1.2 La diferencia es que en la especificación probit **el error** está distribuido normalmente (con promedio 0 y varianza 1), mientras que la distribución logit **el error** tiene un promedio 0 y una varianza $\frac{\pi^2}{3}$. **El estudiante debe mencionar todo esto.**

II. 2

En este caso, estimaremos el modelo según las variables que se especifican. El estudiante puede hacer su propia selección de variables. **Siempre deben estar justificadas.** Ambos modelos deben ser iguales en el número y tipo de variables independientes.

- 2.1 Variables se muestran abajo. No incluyo justificación aquí porque esto dependerá de la selección del estudiante. Se incluye tabla.

```
p_load(MASS)
o.logit = polr(cmc ~ wife_edu + wife_age + num_child + islam + wife_working,
data = dat, method = "logistic")
o.probit = polr(cmc ~ wife_edu + wife_age + num_child + islam + wife_working,
data = dat, method = "probit")
```

Hagamos la tabla.

```
p_load(texreg)
texreg(list(o.logit, o.probit)) # usa "screenreg" no "texreg".
```

- 2.2 Es esperable que el estudiante use el **BIC** o **AIC**. Números más pequeños indican mejor *fit*. En este caso, el modelo probit (“Model 2”) hace un mejor trabajo. Pero esto dependerá de la selección de variables del estudiante. El estudiante también puede usar el log-lik (numeros más cercanos a cero son mejores).

III. 3

- 3 El estudiante no necesariamente debe interpretar la tabla. Si lo hace, debe ser capaz de:

- (a) Decir “cuando x sube en una unidad, y sube en β “log-odds”, manteniendo el resto de las variables constantes en su media.”
- (b) El estudiante **debe** identificar correctamente la unidad de cambio de los coeficientes: “log-odds”.

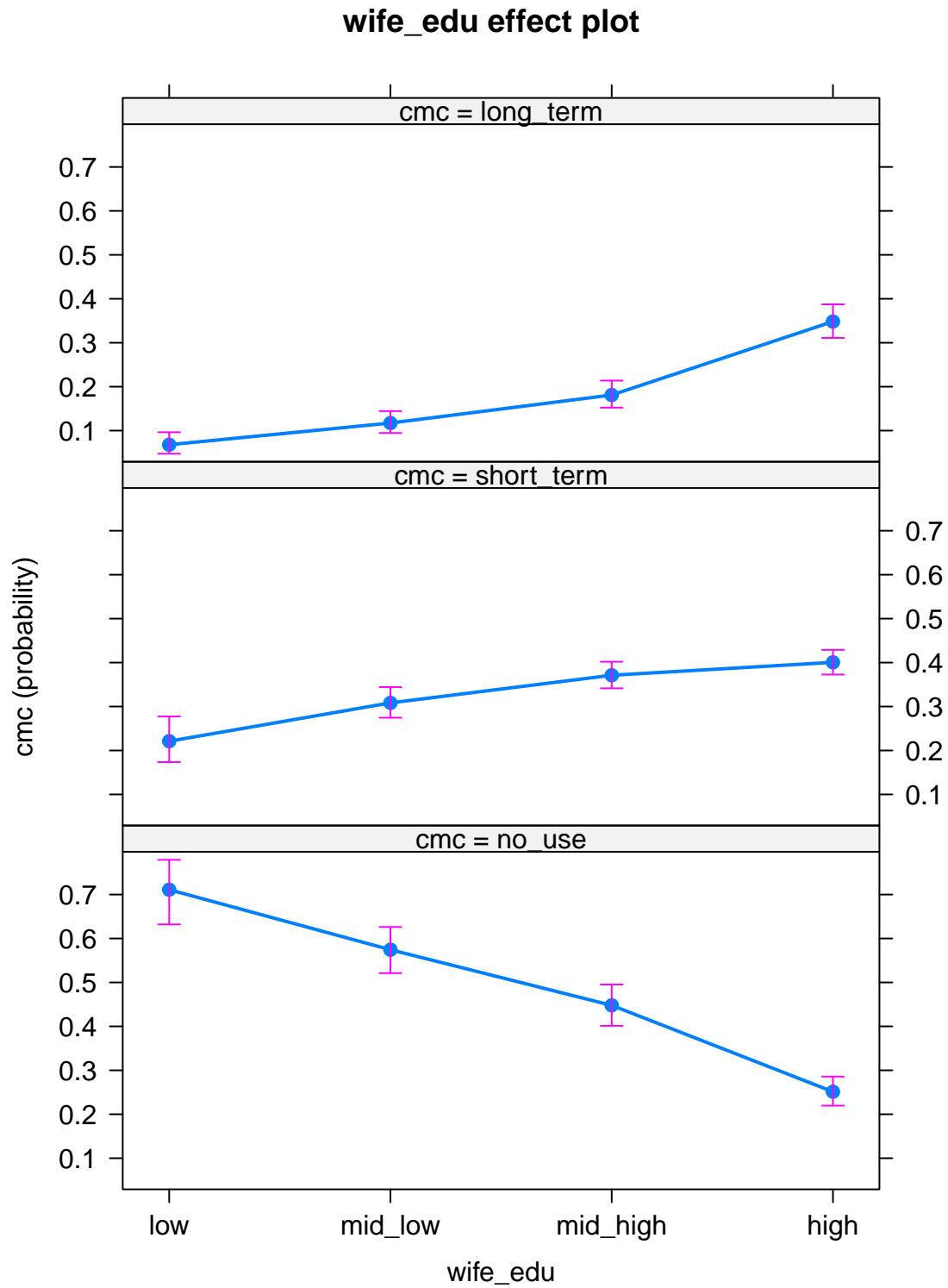
	Model 1	Model 2
wife_edumid_low	0.60** (0.21)	0.36** (0.12)
wife_edumid_high	1.11*** (0.21)	0.68*** (0.12)
wife_eduhigh	1.99*** (0.21)	1.21*** (0.12)
wife_age	-0.05*** (0.01)	-0.03*** (0.00)
num_child	0.28*** (0.03)	0.16*** (0.02)
islam	-0.46** (0.15)	-0.27** (0.09)
wife_working	0.05 (0.12)	0.03 (0.07)
no_use short_term	-0.13 (0.37)	-0.03 (0.22)
short_term long_term	1.58*** (0.37)	1.00*** (0.22)
AIC	2925.06	2924.57
BIC	2972.71	2972.23
Log Likelihood	-1453.53	-1453.29
Deviance	2907.06	2906.57
Num. obs.	1473	1473

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

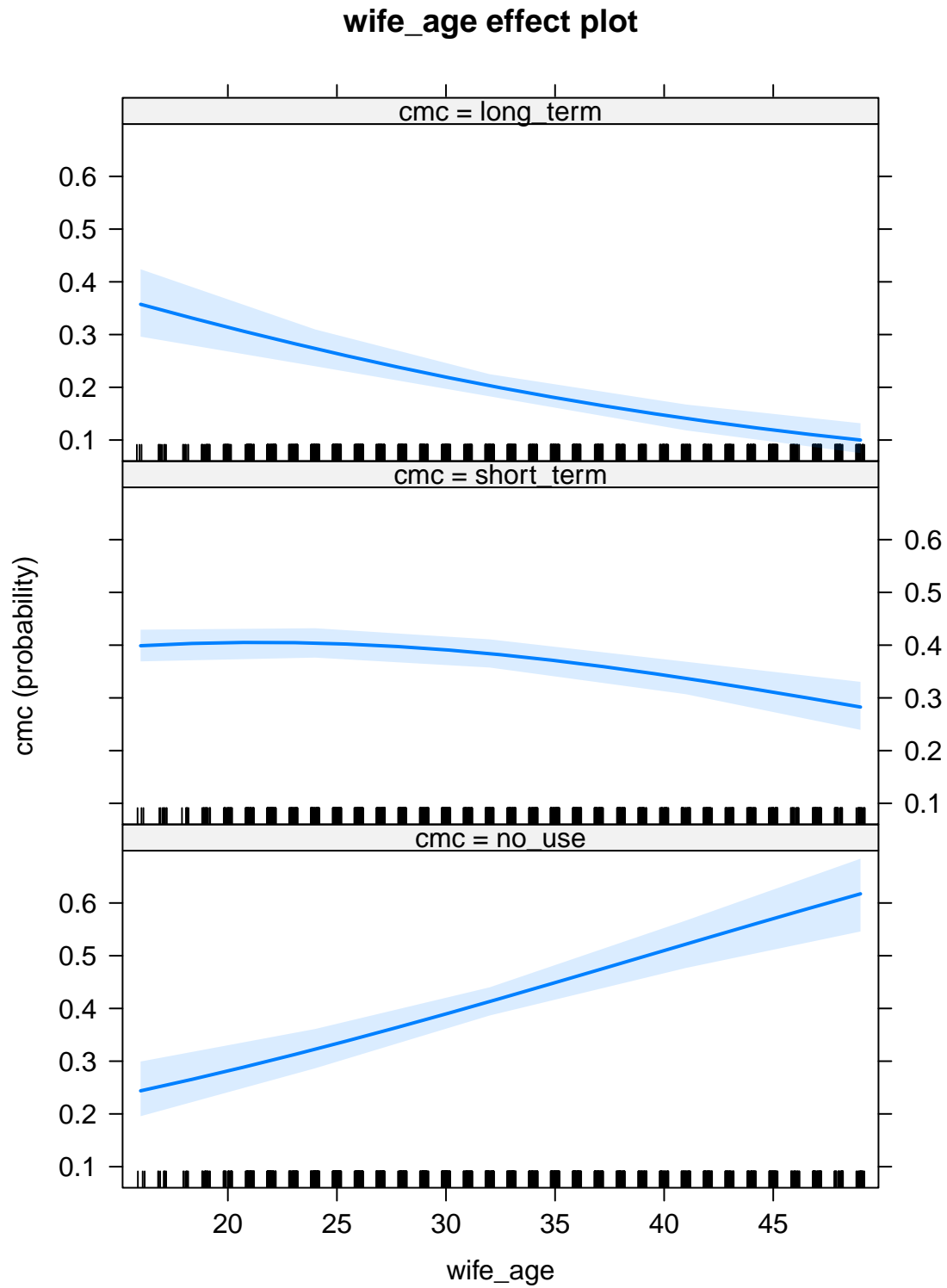
Table 1: *Statistical models*

Calculemos ahora los *predicted probabilities*.

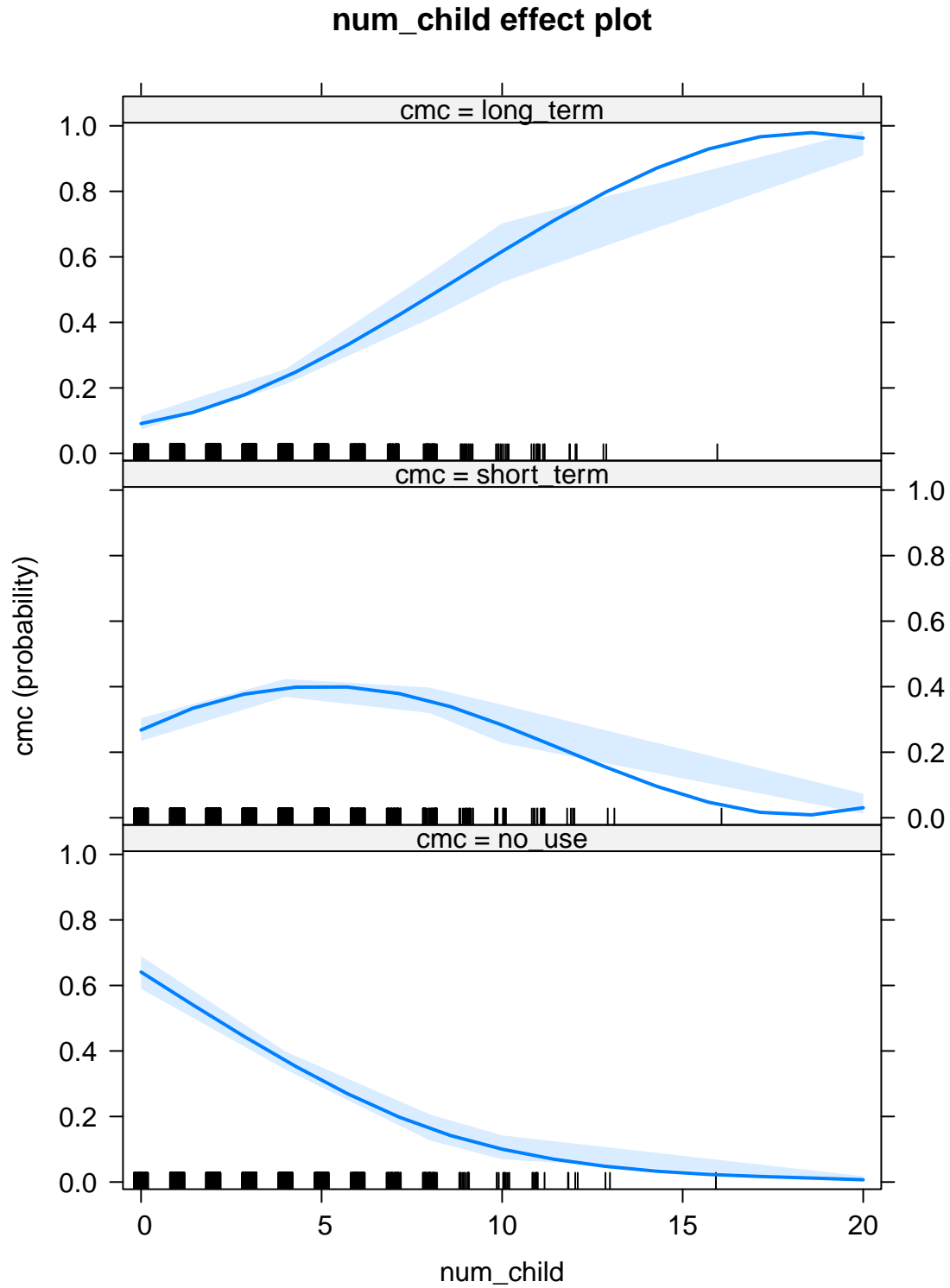
```
p_load(effects)
plot(effect("wife_edu", o.logit))
```



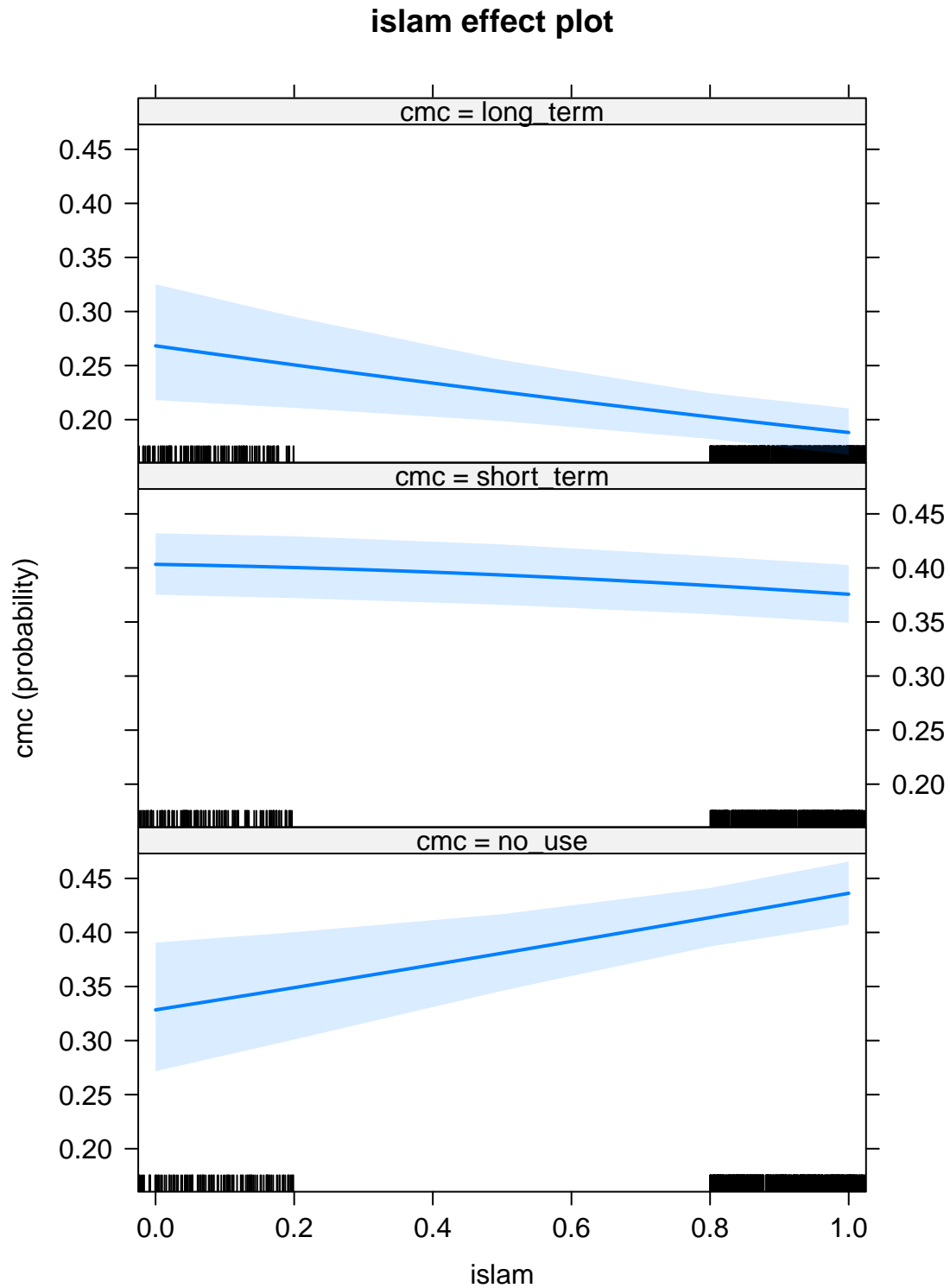
```
plot(effect("wife_age", o.logit))
```



```
plot(effect("num_child", o.logit))
```

```
plot(effect("islam", o.logit))
```



```
plot(effect("wife_workin", o.logit))

## Error in Analyze.model(focal.predictors, mod, xlevels, default.levels, : the following
predictor is not in the model: wife_workin
```

No interpretaré las *predicted probabilities* aquí (ya que es contingente en el modelo del estudiante).
El estudiante debe:

1. Interpretar todos los parámetros de manera substantiva.
2. Seleccionar el set de interpretaciones en base al mejor modelo (en este caso, el modelo probit).
3. Identificar las unidades de los ejes de los gráficos: el eje Y es probabilidad, el eje X son los niveles de las variables dependientes.