

# Tech 2. Pre-Analysis 5% Sample.

This is the analysis carried out in the ‘Data Finder’ phase of the SCORE project used for the pre-registration.

```
df_orig_complete <- read_csv(file = here::here("data/df_orig_complete.csv"))
```

Here I select out 5% of the sample at random and run their models as part of the pre-registration.

```
set.seed(90210)

df_orig_complete <- df_orig_complete %>%
  mutate(rand = runif(NROW(df_orig_complete$y1)),
         p5 = ifelse(rand > 0.05, NA, 1))

df_orig_complete_5 <- df_orig_complete[!is.na(df_orig_complete["p5"]),]
```

```
Tbl5_m1 <- polr(factor(stdliv_next5) ~ factor(cat_age) + factor(wave) + factor(cat_age)*factor
(wave), data = df_orig_complete_5, Hess = T)

Tbl5_m2 <- polr(factor(stdliv_next5) ~ factor(cat_age) + factor(wave) + factor(cat_age)*factor
(wave) + female + factor(education) + factor(EGP6) + factor(income), data = df_orig_complete_5
, Hess = T)

Tbl5_m3 <- polr(factor(stdliv_next5) ~ factor(cat_age) + factor(wave) + factor(cat_age)*factor
(wave) + female + factor(education) + factor(EGP6) + factor(income) + pensions + unemployed +
car_owner + factor(cntry), data = df_orig_complete_5, Hess = T)

tab_model(Tbl5_m1, Tbl5_m2, Tbl5_m3, p.style = "stars", show.ci = F, show.loglik = T, pred.la
bels = c("k1", "k2", "k3", "k4", "Age 30-44", "Age 44-59", "Age >60", "Year 2007", "Age 30-44*year'07"
, "Age 44-59*year'07", "Age >60*year'07", "Female", "Educ mid", "Educ high", "EGP: routine non
-man", "EGP: Self", "EGP: Skilled", "EGP: Unskilled", "EGP: Farmers", "EGP: Never had a job",
"Income mid", "Income high", "Income missing", "Pensions and benefits", "Unemployed", "Car", "bul
garia", "czech", "estonia", "hungary", "latvia", "lithuania", "moldova", "poland", "romania", "r
ussia", "slovakia", "ukraine"), file = "results/Tbl5_rep.htm")
```

	factor(stdliv next 5)	factor(stdliv next 5)	factor(stdliv next 5)
Predictors	Odds Ratios	Odds Ratios	Odds Ratios
k1	0.08 ***	0.09 ***	0.10 ***
k2	0.20 ***	0.22 ***	0.26 ***
k3	1.72 ***	1.93 **	2.44 **
k4	27.71 ***	32.20 ***	43.24 ***
Age 30-44	0.76	0.74	0.76
Age 44-59	0.46 ***	0.46 ***	0.46 ***
	***	***	***

Age >60	0.34	0.40	0.45
Year 2007	2.45 ***	2.72 ***	2.91 ***
Age 30-44*year'07	0.63	0.57	0.54 *
Age 44-59*year'07	0.65	0.58	0.57
Age >60*year'07	0.61	0.53 *	0.51 *
Female		0.86	0.89
Educ mid		1.12	1.07
Educ high		1.19	1.23
EGP: routine non-man		0.88	0.94
EGP: Self		1.39	1.41
EGP: Skilled		0.95	0.96
EGP: Unskilled		0.79	0.84
EGP: Farmers		0.98	1.08
EGP: Never had a job		0.54 *	0.67
Income mid		1.10	1.10
Income high		1.52 ***	1.50 ***
Income missing		0.76	0.74
Pensions and benefits			0.93
Unemployed			1.02
Car			1.36 **
bulgaria			0.64
czech			1.30
estonia			1.46
hungary			0.57 *
latvia			2.29 **
lithuania			1.35
moldova			0.56 *
poland			0.94
romania			1.10

russia			1.01
slovakia			0.66
ukraine			0.88
Observations	1774	1774	1774
R <sup>2</sup> Nagelkerke	0.071	0.095	0.141
log-Likelihood	-2266.751	-2245.657	-2202.994
•		<i>p</i> <0.05	** <i>p</i> <0.01   *** <i>p</i> <0.001

Output Table

```
#not used in the end

#webshot("results/Tbl5_rep.htm", file = "results/Tbl5_rep.png")

#knitr::include_graphics(here::here("results/Tbl5_rep.png"))
```

Colophon

```
sessionInfo()

## R version 4.1.3 (2022-03-10)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19043)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] knitr_1.38      webshot_0.5.2   sjPlot_2.8.10   MASS_7.3-55
## [5] kableExtra_1.3.4 forcats_0.5.1   stringr_1.4.0   dplyr_1.0.8
## [9] purrr_0.3.4     readr_2.1.2     tidyr_1.2.0     tibble_3.1.6
## [13] ggplot2_3.3.5   tidyverse_1.3.1
##
## loaded via a namespace (and not attached):
## [1] nlme_3.1-155     fs_1.5.2        bit64_4.0.5     lubridate_1.8.0
## [5] insight_0.17.0   httr_1.4.2      rprojroot_2.0.2  tools_4.1.3
## [9] backports_1.4.1  bslib_0.3.1     utf8_1.2.2      R6_2.5.1
## [13] sjlabelled_1.1.8 DBI_1.1.2       colorspace_2.0-3 withr_2.5.0
```

## [17]	tidyselect_1.1.2	emmeans_1.7.3	bit_4.0.4	compiler_4.1.3
## [21]	performance_0.9.0	cli_3.2.0	rvest_1.0.2	pacman_0.5.1
## [25]	xml2_1.3.3	bayestestR_0.11.5	sass_0.4.1	scales_1.1.1
## [29]	mvtnorm_1.1-3	systemfonts_1.0.4	digest_0.6.29	minqa_1.2.4
## [33]	rmarkdown_2.13	svglite_2.1.0	pkgconfig_2.0.3	htmltools_0.5.2
## [37]	lme4_1.1-28	dbplyr_2.1.1	fastmap_1.1.0	rlang_1.0.2
## [41]	readxl_1.4.0	rstudioapi_0.13	jquerylib_0.1.4	generics_0.1.2
## [45]	jsonlite_1.8.0	vroom_1.5.7	magrittr_2.0.2	parameters_0.17.0
## [49]	Matrix_1.4-0	Rcpp_1.0.8.3	munSELL_0.5.0	fansi_1.0.3
## [53]	lifecycle_1.0.1	stringi_1.7.6	yaml_2.3.5	snakecase_0.11.0
## [57]	grid_4.1.3	parallel_4.1.3	sjmisc_2.8.9	crayon_1.5.1
## [61]	lattice_0.20-45	ggeffects_1.1.1	haven_2.4.3	splines_4.1.3
## [65]	sjstats_0.18.1	hms_1.1.1	pillar_1.7.0	boot_1.3-28
## [69]	estimability_1.3	effectsize_0.6.0.1	reprex_2.0.1	glue_1.6.2
## [73]	evaluate_0.15	modelr_0.1.8	vctrs_0.3.8	nloptr_2.0.0
## [77]	tzdb_0.3.0	cellranger_1.1.0	gtable_0.3.0	assertthat_0.2.1
## [81]	datawizard_0.4.0	xfun_0.30	xtable_1.8-4	broom_0.7.12
## [85]	viridisLite_0.4.0	ellipsis_0.3.2	here_1.0.1	