Summary of Synthetic Variables Estimation

EES 2019 Voter Study (Danish, Estonian, German, Luxembourgian, Maltese, Dutch, Spanish and British (UK) samples)

Wilhelmine Haeussling

15.11.2021

1 Denmark

Synthetic variables have been estimated for seven of ten of Danish parties available in the original 2019 EES Danish voter study selected according to the criteria stated in the EES 2019 SDM codebook (for the criteria see Sect. XXX; for the relevant parties see Table 1.1).

Table 1.1: Danish relevant parties

Dep. Var.	Party	Party name (eng)
stack_701	701	Social Democratic Party
stack_702	702	Liberals
stack_703	703	Danish People's Party
stack_704	704	Radical Party
stack_705	705	Socialist People's Party
stack_706	706	Red-Green Unity List
stack_707	707	Conservative People's Party

Full OLS models converge and coefficients do not show any particular issue (see Table 1.4). In terms of model fit, the adjusted coefficient of determination (R^2) values vary between a minimum value of 0.006 for party 703 (Danish People's Party) and a maximum of 0.084 for party 707 (Conservative People's Party). Moreover, the difference between Akaike Information Criterion (AIC) values for full OLS models and null models shows that in 1 case out of 7 null models perform better than full ones (see Table 1.2).

Table 1.2: Akaike Information Criterion values for OLS full and null models

Dep. Var.	Party	Full Mod.	Null Mod.	Diff. (Full-Null)
stack_701	701	602.803	650.130	-47.327
stack_702	702	664.540	698.147	-33.607
stack_703	703	707.938	702.351	5.587
stack_704	704	320.451	385.068	-64.616
stack_705	705	606.771	637.521	-30.750
stack_706	706	652.782	694.182	-41.400
stack_707	707	366.698	431.607	-64.908

Furthermore, there were no unusual standard errors for any coefficients in the logistic regression models. (see Table 1.5)

In terms of model fit, adjusted McFadden's pseudo R^2 values for the logistic full models vary between a minimum value of -0.044 for party 704 (Radical Party) and a maximum of 0.016 for party 703 (Danish People's Party). Moreover, the difference between Akaike Information Criterion (AIC) values for logistic full models and null models shows that in 3 cases out of 7 null models perform better than full ones (see Table 1.3).

Table 1.3: Akaike Information Criterion values for logistic full and null models

Dep. Var.	Party	Full Mod.	Null Mod.	Diff. (Full-Null)
stack_701	701	759.240	771.073	-11.833
stack_702	702	746.506	757.681	-11.175
stack_703	703	591.607	602.927	-11.320
stack_704	704	394.156	379.564	14.592
stack_705	705	524.297	523.116	1.181
stack_706	706	408.469	412.696	-4.227
stack_707	707	316.194	314.621	1.573

Table 1.4: Propensity to vote for a relevant party according to respondents' socio-demographic characteristics (OLS regression models)

| | 701 | 702 | 703 | 704 | 705 | 706 | 707 |
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
D3_rec2	0.083***	−0.051*	−0.068**	0.034	0.088***	0.044	−0.086***
	(0.023)	(0.024)	(0.024)	(0.020)	(0.023)	(0.024)	(0.020)
D8_rec1	0.068*	0.002	0.036	0.001	0.064*	0.052	0.020
	(0.028)	(0.029)	(0.030)	(0.024)	(0.028)	(0.029)	(0.025)
D5_rec1	−0.007	0.037	0.026	−0.025	−0.021	−0.038	−0.004
	(0.024)	(0.025)	(0.026)	(0.021)	(0.025)	(0.026)	(0.022)
EDU_rec2	−0.061	0.113*	0.037	−0.034	−0.070	−0.056	0.071
	(0.051)	(0.053)	(0.055)	(0.044)	(0.053)	(0.054)	(0.046)
EDU_rec3	−0.023	0.038	−0.005	0.004	−0.004	0.004	0.048
	(0.048)	(0.049)	(0.051)	(0.041)	(0.049)	(0.050)	(0.042)
D1_rec1	0.139***	−0.056*	−0.003	0.005	0.068*	0.041	−0.007
	(0.026)	(0.027)	(0.028)	(0.022)	(0.026)	(0.027)	(0.023)
D7_rec1	−0.001	0.100***	−0.047	0.045*	−0.019	−0.021	0.063**
	(0.026)	(0.027)	(0.028)	(0.022)	(0.027)	(0.027)	(0.023)
D7_rec2	−0.116**	0.172***	−0.053	0.048	−0.074*	−0.112**	0.192***
	(0.036)	(0.037)	(0.038)	(0.031)	(0.036)	(0.037)	(0.032)
D6_une1	0.053	−0.056	−0.003	−0.022	−0.043	−0.033	−0.089*
	(0.048)	(0.049)	(0.051)	(0.041)	(0.048)	(0.050)	(0.042)
D4_age	0.0001	−0.002*	0.001	−0.005***	−0.002**	−0.004***	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
D10_rec	0.011	0.023**	0.006	0.012*	0.007	0.005	0.027***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)
Constant	0.358***	0.389***	0.311***	0.512***	0.385***	0.494***	0.275***
	(0.062)	(0.064)	(0.066)	(0.053)	(0.063)	(0.065)	(0.055)
N	879	878	877	873	863	861	863
R-squared	0.076	0.061	0.019	0.094	0.059	0.071	0.096
Adj. R-squared	0.064	0.049	0.006	0.083	0.047	0.059	0.084

***p < .001; **p < .01; *p < .05

2   Estonia

Synthetic variables have been estimated for seven of twelve Estonian parties available in the original 2019 EES Estonian voter study selected according to the criteria stated in the EES 2019 SDM codebook (for the

Table 1.5: Vote choice for a relevant party according to respondents' socio-demographic characteristics (Logistic regression models)

	701	702	703	704	705	706	707
	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14
D3_rec2	0.515**	−0.193	−0.731**	−0.077	0.870***	0.187	−0.358
	(0.192)	(0.193)	(0.231)	(0.296)	(0.261)	(0.284)	(0.347)
D8_rec1	0.117	0.010	0.420	−0.132	−0.416	−0.015	−0.478
	(0.233)	(0.231)	(0.296)	(0.361)	(0.276)	(0.348)	(0.374)
D5_rec1	0.076	0.409	−0.311	−0.336	−0.073	−0.037	0.197
	(0.201)	(0.212)	(0.234)	(0.310)	(0.259)	(0.296)	(0.374)
EDU_rec2	−0.600	−0.456	0.782	0.759	0.058	−0.279	0.423
	(0.417)	(0.453)	(0.652)	(0.815)	(0.628)	(0.635)	(0.830)
EDU_rec3	−0.378	−0.331	0.420	0.924	0.550	0.012	−0.252
	(0.372)	(0.409)	(0.630)	(0.777)	(0.575)	(0.585)	(0.795)
D1_rec1	0.726**	−0.201	−0.242	0.532	−0.219	0.237	0.011
	(0.232)	(0.216)	(0.241)	(0.365)	(0.272)	(0.320)	(0.392)
D7_rec1	0.128	0.504*	0.127	0.014	0.101	−0.010	0.530
	(0.212)	(0.244)	(0.257)	(0.335)	(0.290)	(0.299)	(0.468)
D7_rec2	−0.556	1.002***	−0.184	0.083	0.558	−2.402*	1.372**
	(0.329)	(0.290)	(0.372)	(0.458)	(0.352)	(1.035)	(0.510)
D6_une1	0.170	0.190	−0.182	−0.597	−0.247	−0.887	−0.474
	(0.363)	(0.407)	(0.496)	(0.744)	(0.545)	(0.741)	(1.039)
D4_age	0.003	0.012	0.024**	−0.012	−0.006	0.004	0.011
	(0.006)	(0.006)	(0.008)	(0.010)	(0.008)	(0.010)	(0.011)
D10_rec	0.140**	0.072	−0.010	−0.016	−0.001	−0.327**	0.174
	(0.052)	(0.053)	(0.067)	(0.085)	(0.070)	(0.114)	(0.089)
Constant	−2.591***	−2.585***	−3.560***	−3.101***	−2.631***	−2.471**	−4.042***
	(0.513)	(0.537)	(0.757)	(0.925)	(0.700)	(0.758)	(0.999)
N	874	874	874	874	874	874	874
Log Likelihood	−367.620	−361.253	−283.804	−185.078	−250.148	−192.234	−146.097
AIC	759.240	746.506	591.607	394.156	524.297	408.469	316.194

*** $p < .001$; ** $p < .01$; * $p < .05$

criteria see Sect. XXX; for the relevant parties see Table 2.1).

Table 2.1: Estonian relevant parties

Dep. Var.	Party	Party name (eng)
stack_901	901	Estonian Reform Party
stack_902	902	Estonian Center Party
stack_903	903	Conservative People's Party of Estonia
stack_904	904	Union for the Republic – Res Publica
stack_905	905	Social Democratic Party
stack_906	906	Estonia 200
stack_907	907	Estonian Greens

Full OLS models converge and coefficients do not show any particular issue (see Table **??**). In terms of model fit, the adjusted coefficient of determination (R^2) values vary between a minimum value of 0.028 for party 906 (Estonia 200) and a maximum of 0.063 for party 905 (Social Democratic Party). Moreover, the difference between Akaike Information Criterion (AIC) values for full OLS models and null models shows that in 0 cases out of 7 null models perform better than full ones (see Table 2.2).

Table 2.2: Akaike Information Criterion values for OLS full and null models

Dep. Var.	Party	Full Mod.	Null Mod.	Diff. (Full-Null)
stack_901	901	645.297	681.225	-35.928
stack_902	902	468.318	482.379	-14.061
stack_903	903	523.081	555.418	-32.337
stack_904	904	316.225	331.505	-15.280
stack_905	905	364.945	406.867	-41.922
stack_906	906	321.048	332.960	-11.912
stack_907	907	171.422	186.800	-15.378

On the contrary, three out of seven logistic regression models (see Table **??**) show inflated standard errors for some of the coefficients of interest, in particular:

- Model 9: D6_une;
- Model 13a: EDU_rec;
- Model 14a: D5_rec, EDU_rec, D1_rec, D6_une.

Nevertheless, model's 9 constant terms and other regression coefficients are not affected by said inflated standard errors, whereas model 13a and 14a present a more problematic profile.

Model 13a's and 14a's inflated standard errors are due to separation issues. In short, no respondents from respondents with low education did vote for party 906 (see Table **??**) and no respondents with low education, with high subjective social status, no members of trade unions, and unemployed and only very few respondents married or in partnership as well as married or in a partnership (2 and 9) did vote for party 907 (see Tables **??**, **??**, **??**, **??**).

As a consequence, a constrained version of model 13 (namely, Model 13b) without said variable was estimated and contrasted with the original (Model 13a), full model. Likelihood-ratio test results show that H_0 (namely,

that the constrained model fits better than the full model) cannot be rejected (see Table 2.3). Consequently, synthetic variables for respondents' vote choice for party 906 have been predicted relying on the constrained model (Model 13b).

Table 2.3: Likelihood-ratio Test between Model 13a (Unconstrained) and Model 13b (Constrained)

Model	Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
Constrained	800	202.6392			
Unconstrained	798	197.9122	2	4.726983	0.0940911

In **??** there is no 0 disrupting our logit regression. As a consequence, a constrained version of model 14 (namely, Model 14b_1) without said variables and without D5_rec was estimated and contrasted with the original (Model 14a), full model. Likelihood-ratio test results show that H_0 (namely, that the constrained model fits better than the full model) cannot be rejected (see Table **??**).

Furthermore, another constrained version of model 14 (namely, Model 14b_2) without said variables, but with D5_rec, was estimated and contrasted with the original (Model 14a), full model. Model 14b_2 is therefore less constrained than 14b_1 (by D5_rec). Likelihood-ratio test results show that H_0 (namely, that the constrained model fits better than the full model) cannot be rejected (see Table **??**).

Then we compared the fit of 14b_1 and 14b_2. Likelihood-ratio test results show that H_0 (namely, that the 'fuller' constrained model with D5_rec fits better than the constrained model without D5_rec) cannot be rejected (see Table **??**). Consequently, synthetic variables for respondents' vote choice for party 907 have been predicted relying on the less constrained model with D5_rec (Model 14b_2).