Questions Concerning Model Choice when Facing Collinearity

Nicolai Berk

4 3 2022

Dear Will,

apologies for annoying you with this, but I am increasingly confused by the estimates of models I tried to present in last week's course.

The basic problem I am interested in is whether news exposure affects issue attitudes. More specifically, I am interested in whether increased emphasis of crime in immigration news changes migration attitudes, and whether this effect is conditioned by past attitude on the issue (I described this as values in class, but I believe the prior issue stance is a more appropriate measure of the concept I am trying to capture). I believe the model for the direct effect would look something like this:

$$y_{it} = \alpha + \beta_1 * CrimeExposure_{it} + \rho_i + \phi_t + \epsilon_{it}$$

With y_{it} being the immigration attitude of individual i at time t, $CrimeExposure_{it}$ symbolising the estimated exposure of individual i to crime news in immigration coverage in the week preceding the interview at time t, and ρ_i & ϕ_t describing individual fixed effects.

The above mentioned model returns a weak positive correlation of crime exposure and migration attitude that might well arise by chance:

```
summary(feglm(dv ~ crime_share | wave + lfdn,
       cluster = c("lfdn"), data = merged_data))
```

```
## NOTE: 3,736 observations removed because of NA values (LHS: 3,736).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 26,585
## Fixed-effects: wave: 6,  lfdn: 10,357
## Standard-errors: Clustered (lfdn)
##              Estimate Std. Error t value Pr(>|t|))
## crime_share 0.504611   0.299751  1.6834  0.092321 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -26,189.8   Adj. Pseudo R2: 0.315051
##           BIC: 157,958.9      Squared Cor.: 0.870535
```

Now what about the conditioning effect of past migration opinions? First, I tried the model discussed in class:

$$y_{it} = \alpha + \beta_1 * CrimeExposure_{it} + \beta_2 * PastOpinion + \beta_3 * CrimeExposure * PastOpinion + \rho_i + \phi_t + \epsilon_{it}$$

Using the initial migration attitude of a respondent as conditioning variable creates collinearity issues:

```
summary(feglm(dv ~ crime_share*init_mig | wave + lfdn,
       cluster = c("lfdn"), data = merged_data))
```

```
## NOTE: 3,736 observations removed because of NA values (LHS: 3,736, RHS: 23).

## The variable 'init_mig' has been removed because of collinearity (see $collin.var).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 26,585
## Fixed-effects: wave: 6,  lfdn: 10,357
## Standard-errors: Clustered (lfdn)
##                       Estimate Std. Error t value Pr(>|t|))
## crime_share           1.082200   0.361769  2.9915  0.002783 **
## crime_share:init_mig -0.352347   0.098975 -3.5600  0.000373 ***
## ... 1 variable was removed because of collinearity (init_mig)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -26,178.8   Adj. Pseudo R2: 0.315238
##           BIC: 157,947.1     Squared Cor.: 0.870642
```

We find a significant decrease with more restrictive (higher) migration attitudes, but the direct effect of initial migration attitudes isn't taken into account. I wouldn't consider this a huge problem, as the individual fixed effects should take care of this, but I am cautious so I try different models.

```
summary(feglm(dv ~ crime_share*init_mig | wave,
      cluster = c("lfdn"), data = merged_data))
```

```
## NOTE: 3,736 observations removed because of NA values (LHS: 3,736, RHS: 23).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 26,585
## Fixed-effects: wave: 6
## Standard-errors: Clustered (lfdn)
##                       Estimate Std. Error  t value Pr(>|t|))
## crime_share           3.473100   0.436810   7.9511 2.04e-15 ***
## init_mig              0.855017   0.006517 131.1900 < 2.2e-16 ***
## crime_share:init_mig -0.830479   0.116954  -7.1009 1.32e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -38,105.5   Adj. Pseudo R2: 0.285787
##           BIC:  76,302.6     Squared Cor.: 0.682707
```

Dropping the fixed effects and controlling for initial migration opinion (init_mig) returns an even larger estimate. So far, so good - but this entire thing could also be modelled with a lagged dependent variable instead of the initial migration opinion. This also solves the multicollinearity issue.

```
summary(feglm(dv ~ crime_share*dv_lag | wave + lfdn,
      cluster = c("lfdn"), data = merged_data))
```

```
## NOTE: 16,370 observations removed because of NA values (LHS: 3,736, RHS: 13,787).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 13,951
## Fixed-effects: wave: 5,  lfdn: 5,689
## Standard-errors: Clustered (lfdn)
##                      Estimate Std. Error    t value Pr(>|t|))
## crime_share          0.149868   0.452492   0.331206  0.740501
## dv_lag              -0.172739   0.016735 -10.322000 < 2.2e-16 ***
## crime_share:dv_lag   0.077975   0.114736   0.679608   0.49678
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Log-Likelihood: -12,592.1    Adj. Pseudo R2: 0.350242
##             BIC:  79,542.8      Squared Cor.: 0.892425
```

In this model, the interaction effect disappears. The lagged dependent has a negative effect, which i interpret as regression to the mean around some underlying migration opinion captured by the fixed effects. However, given this I am also worried that controlling for the lagged dependent is overdoing it (since the individual fixed effects should take care of this). I go against the holy wisdom of Brambor et al and drop a constitutive term:

```
summary(feglm(dv ~ crime_share:dv_lag + crime_share | wave + lfdn,
      cluster = c("lfdn"), data = merged_data))
```

```
## NOTE: 16,370 observations removed because of NA values (LHS: 3,736, RHS: 13,787).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 13,951
## Fixed-effects: wave: 5,  lfdn: 5,689
## Standard-errors: Clustered (lfdn)
##                     Estimate Std. Error t value Pr(>|t|))
## crime_share         1.567000   0.431706  3.6299  2.86e-04 ***
## crime_share:dv_lag -0.719561   0.104965 -6.8553  7.87e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -12,770.6    Adj. Pseudo R2: 0.343933
##             BIC:  79,890.4      Squared Cor.: 0.889636
```

Which returns a significant direct and interaction effect again. I am however still unsure about the preceding model, which seems to control for too much but might also be showing some cross-temporal variation I should take into account - the news content of last week should not be predictable from some opinion four months ago. I also added estimated the same model with the initial opinion as a moderator, which again returned significant effects:

```
feglm(dv ~ crime_share:init_mig + crime_share | wave + lfdn,
      cluster = c("lfdn"), data = merged_data)
```

```
## NOTE: 3,736 observations removed because of NA values (LHS: 3,736, RHS: 23).

## GLM estimation, family = gaussian, Dep. Var.: dv
## Observations: 26,585
## Fixed-effects: wave: 6,  lfdn: 10,357
## Standard-errors: Clustered (lfdn)
##                       Estimate Std. Error t value Pr(>|t|))
## crime_share           1.082200   0.361769  2.9915  0.002783 **
## crime_share:init_mig -0.352347   0.098975 -3.5600  0.000373 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Log-Likelihood: -26,178.8    Adj. Pseudo R2: 0.315238
##             BIC: 157,947.1      Squared Cor.: 0.870642
```

I hoped for some guidance here. My main idea for now is to present the four significant models and add an analysis where I regress the news content of the week *following* the interview on the migration opinion given in the interview to show that this is not a result of respondents actively seeking out news content on crime. However, if the model including the lagged dependent variable shows some serious flaw, I might dig deeper to see what is going on - I am just not sure where yet. Additionally, I see the reviewer comment "but Brambor et al tell us to include constitutive terms" coming (could be me) and would love to have a convincing answer.

Thank you so much for taking a look, all the best, Nico