Appendix: Replication Materials

The American National Election Studies (ANES) 2008 Time Series Study data and codebook were downloaded from The American National Election Study webpage at https://electionstudies.org/data-center/2008-time-series-study/ (ANES 2015). From the ANES file anes_timeseries_2008.dta we created an extract called Obama2008.dta that contains only the variables needed to replicate our analysis. A slightly different version of the data is available from the Inter-university Consortium for Political and Social Research (ICPSR) webpage https://www.icpsr.umich.edu/web/ICPSR/studies/25383.

All analyses were done with Stata 16 or Stata 17 (but our programs may work with earlier versions of Stata as well). The three programs we used (reproduced below) are called **hypothetical.do** and **Obama2008.do**, which prepared the data, and **nested.ado**, which is called by the other two programs to do the actual analysis. These are not designed to be general-purpose programs – e.g. they do not check for syntax errors – so anyone who wants to do similar analysis of their own may have to make some modifications.

In order to run the programs as is, you will also need to install the Stata user-written routines **estadd** (Jann 2005), **esttab (**Jann 2007), **khb** (Kohler et al, 2011) and **spost13_ado** (Long & Freese, 2014). Spost13 contains the **listcoef** program and several other useful utilities. All can be found and installed for free using the **findit** or **search** commands in Stata.

Note that the programs issue clear commands and write and overwrite some files, so be careful not to lose any work you want saved.

7.1 Listing of hypothetical.do

```
* Hypothetical Data
* Create uncorrelated X Vars
version 16
clear
input X1 X2 freq
0 0 125
0 1 125
1 0 125
```

```
1 1 125
end
expand freq
drop freq
set seed 123456789
gen Y = (4.5 * X1) + (4.5 * X2) - 4.0 + rlogistic() > 0
order Y X1 X2
sort X1 X2 Y
corr Y X1 X2

* nested.ado does the actual analyes
nested Y i.X1 i.X2, stub(Hypothetical) description(Hypothetical Data)
```

7.2 Listing of Obama2008.do

```
version 16
use https://www3.nd.edu/~rwilliam/statafiles/Obama2008, clear

* American National Election Study (ANES) 2008 Time Series Study
* Only necessary variables for this analysis are included in the extract.
* Complete dataset and documentation can be obtained from
* https://electionstudies.org/data-center/2008-time-series-study/
* ANES gave permission for us to circulate an extract using its data.

* ICPSR has a slightly different version of the data at
* https://www.icpsr.umich.edu/web/ICPSR/studies/25383
* The following minor changes are made for consistency between the
*    ICPSR & ANES versions of the data
rename *, upper
capture gen CASEID = _n

* Clearer names for some variables
clonevar idnum = CASEID
clonevar wgtvar = V080102A

* All numeric MD codes are recoded to the Stata system missing value
quietly mvdecode _all, mv(-9/-1)

* Create race dichotomy
recode V081102 (1 = 1 "White")(. = .)(else = 0 "NonWhite"), gen(white) label (white)
recode white (1 = 0) if V081103 == 1
label variable white "Race recoded to 0 = NonWhite, 1 = White"
* More specifically, 1 = White Non-Hispanic, 0 = NonWhite and/or Hispanic

* Presidential vote -- only a few people voted for other candidates
```

* and they are recoded to missing
recode V085044A (1 = 1 "Obama") (3 = 0 "McCain") (else = .), gen(Obama) label(Obama)
label variable Obama "Who did R vote for in 2008? (0 = McCain, 1 = Obama)"

* Feminist Thermometer
clonevar feminist= V085064D
label variable feminist "Feeling thermometer: FEMINISTS"

* Bush thermometer
clonevar Bush = V085063A
label variable Bush "Feeling thermometer: President GW Bush"

* Age
clonevar age = V081104
label variable age "Age"

* Income. Ordinal categories are recoded to their midpoints, with the values
* capped at $150,000
 gen income = V083309A
 recode income (1=1.5)(2=4)(3=6.25)(4=8.75)(5=10.5)(6=11.75) ///
   (7=13.75)(8=16)(9=18.5)(10=21)(11=23.5)(12=27.5)(13=32.5)(14=37.5) ///
   (15=42.5)(16=47.5)(17=55)(18=67.5)(19=82.5)(20=95)(21=105)(22=115) ///
   (23=127.5)(24=142.5)(25=150) (98 = .)

* Rescale variables to make output easier to read
replace age = age/ 10.0
replace income = income/ 10.0
replace Bush = Bush/ 10.0
replace feminist = feminist/ 10.0
label variable income "Estimated Family Income in $10,000s"

order Obama white age income Bush feminist idnum wgtvar

* We do not use weights in our analysis, but if they are used the changes in results
* are very minor. Weights perhaps should be used for other analyses,
* e.g. descriptive statistics
svyset idnum [pw = wgtvar]

* nested.ado does the actual analyes
nested Obama i.white age income Bush feminist, stub(Obama) description(Obama 2008
Presidential Vote) zshow

7.3 Listing of nested.ado

program nested, rclass

```
syntax [varlist(default=none fv)], [stub(string) DESCription(string) ZSHOW]
* syntax:
* yvar xvar othervars otheroptionsifdesired
* yvar should be a 0/1 dependent variable
* xvar should be a dichotomous categerorical variable or a continuous variable
* othervars can be categerorical or continuous
* stub is user to partially specify the outpot file name, i.e.
*    the output file is named ssr`stub'.rtf
* description is used in the table titles
* zshow is used in the KHB analysis. By default, the
*   Z-residual variables are temporary variables and are not
*   shown in the output. If zshow is specified then the
*   Z-residual variables are included in the output
* NOTE: Any variables and estimation results created by the program
*   are not saved. You can modify the program if you want them.
* NOTE: rtf Output files will likely need some additional editing.

* Only cases with nonmissing data are kept
marksample touse
preserve
keep if `touse'

gettoken y z: varlist
gettoken x z: z
if "`description'" == "" local description TBA
if "`stub'" == "" local stub TBA

* This is a check to see if x is a factor variable and if so,
* use the 2nd column for its coefficient value
quietly reg `y' `x'
if e(b)[1,1] != 0 {
        local startcol = 1
}
else local startcol = 2

*****************************************
* 1. Run Logistic regressions
display " "
display as result "1. Logistic regressions"
display " "

logit `y' `x', nolog
AddInfo `startcol'
est store Xonly
local xbefore = e(xcoef)
```

```
logit `y' `z', nolog
AddInfo `startcol'
est store Zonly

logit `y' `x' `z', nolog
AddInfo `startcol'
est store Both
local xafter = e(xcoef)
AddNote `x' `xbefore' `xafter'

esttab Xonly Zonly Both using ssr`stub'.rtf, ///
        pr2 b(%8.2f) scalar(VarYstar SDYstar) nodepvars nonumbers replace ///
        noconstant obslast nobase z  ///
        title(Table LPM: Logistic Regressions - `description') ///
        addnote(`=r(note)')

*****************************************
* 2. OLS/LPM regressions
display " "
display as result "2. Linear Probability Models"
display " "

sum `y'
local vary = r(Var)
local sdy = r(Var)^.5

reg `y' `x'
estadd scalar VarY = `vary'
estadd scalar SDY = `sdy'
est store LPMXonly
local xbefore = e(b)[1,`startcol']

reg `y' `z'
estadd scalar VarY = `vary'
estadd scalar SDY = `sdy'
est store LPMZonly

reg `y' `x' `z'
estadd scalar VarY = `vary'
estadd scalar SDY = `sdy'
est store LPMBoth
local xafter = e(b)[1,`startcol']

AddNote `x' `xbefore' `xafter'

esttab LPMXonly LPMZonly LPMBoth using ssr`stub'.rtf, ///
```

```
            r2 b(%8.2f) scalar(VarY SDY) nodepvars nonumbers append ///
            noconstant obslast nobase  ///
            mtitles (LPM-Xonly LPM-Zonly LPM-Both) ///
            title(Table LPM: Linear Probability Models - `description') ///
            addnote(`=r(note)')


*****************************************
* 3. Y-standardization
display " "
display as result "3. Y-standardization"
display " "

est restore Xonly
local xbefore = e(xcoef)/ e(SDYstar)
est restore Both
local xafter = e(xcoef)/ e(SDYstar)
AddNote `x' `xbefore' `xafter'

esttab Xonly Zonly Both using ssr`stub'.rtf, ///
        main(bStdY %8.2f) nodepvars nonumbers append ///
        noconstant obslast nobase z  ///
        title(Table Ystd: Y Standardized Coefficients - `description') ///
        addnote(`=r(note)')

* Confirm with listcoef
est restore Xonly
listcoef, std
est restore Zonly
listcoef, std
est restore Both
listcoef, std


*****************************************
* 4. Do it yourself KHB
display " "
display as result  "4. KHB"
display " "

*** Do NOT show z-residual coefficients ***
*** Z-residuals will be created as temporary variables
*** and not shown in the table.
if "`zshow'" == "" {
        fvrevar `z'
        foreach zvar of varlist `r(varlist)' {
                quietly sum `zvar'
                if r(min) != r(max) {
```

```stata
                        quietly reg `zvar' `x'
                        tempvar zresid
                        predict `zresid', residual
                        local zresids `zresids' `zresid'
                }
        }
}
*** DO show z-residual coefficientS ***
*** Z-residuals will be created as permanent variables
*** and shown in the table.
else if "`zshow'" != "" {
        fvrevar `z', list
        foreach zvar of varlist `r(varlist)' {
                quietly reg `zvar' `x'
                predict `zvar'_resid, residual
                local zresids `zresids' `zvar'_resid
        }
}

* KHB Reduced model
quietly logit `y' `x' `zresids', nolog
AddInfo `startcol'
est store Reduced
local xbefore = e(xcoef)

* KHB Full model
logit `y' `x' `z', nolog
AddInfo `startcol'
est store Full
local xafter = e(xcoef)

AddNote `x' `xbefore' `xafter'

* Modified Note is added to the table when the user has opted to not show the Z-residuals
if "`zshow'" == "" {
        local znote ""The z-residual variables are included in the Reduced Model but
coefficients are not shown" "`=r(note)'""
}
else if "`zshow'" != "" {
        local znote `=r(note)'
        local zresids
}

esttab Reduced Full using ssr`stub'.rtf, ///
        pr2 b(%8.2f) scalar(VarYstar SDYstar) nodepvars nonumbers append ///
        noconstant obslast nobase z  ///
```

```
        drop(`zresids') ///
        title(Table KHB: KHB Models - `description') ///
        addnote (`znote')

* Confirm with KHB Program
khb logit `y' `x' || `z', nolog


******************************************
* 5. Marginal Effects
display " "
display as result "5A. Conventional Marginal Effects"
display " "

foreach estimates in Xonly Zonly Both Reduced Full {
        est restore `estimates'
        margins, dydx(_all) post
        estadd scalar xcoef = e(b)[1,`startcol']
        est store `estimates'marg
}

est restore Xonlymarg
local xbefore = e(xcoef)
est restore Bothmarg
local xafter = e(xcoef)
AddNote `x' `xbefore' `xafter'

esttab Xonlymarg Zonlymarg Bothmarg using ssr`stub'.rtf, ///
        b(%8.2f) nodepvars nonumbers append ///
        noconstant obslast nobase z   ///
        mtitles (Xonly Zonly Both) ///
        title(Table Marg: Marginal Effects - `description') ///
        addnote(`=r(note)')

display " "
display as result "5B. KHB Marginal Effects"
display " "

est restore Reducedmarg
local xbefore = e(xcoef)
est restore Fullmarg
local xafter = e(xcoef)
AddNote `x' `xbefore' `xafter'

esttab Reducedmarg Fullmarg using ssr`stub'.rtf, ///
        b(%8.2f) nodepvars nonumbers append ///
        noconstant obslast nobase z   ///
```

```
                drop(`zresids') ///
                mtitles (Reduced Full) ///
                title(Table KHBMarg: KHB Marginal Effects - `description') ///
                addnote ("The z-residual variables are included in the Reduced Model but
coefficients are not shown" "`=r(note)'")

        * Confirm with KHB Program APE option
        khb logit `y' `x' || `z', nolog ape


        ******************************************
end

program AddInfo
        * Adds information involving ystar, y-standardized coefficients
        * to the saved results for a model
        args startcol
        quietly {
                tempname ystar bStdY
                predict `ystar', xb
                quietly sum `ystar'
                local vary = r(Var) + _pi^2/3
                local sdy = `vary' ^ .5
                matrix `bStdY' = e(b) / `sdy'
                local xcoef = e(b)[1, `startcol']

                estadd scalar VarYstar = `vary'
                estadd scalar SDYstar = `sdy'
                estadd matrix bStdY = `bStdY'
                estadd scalar xcoef = `xcoef'
        }
end

        ******************************************
program AddNote, rclass
        * Adds notes to tables indicating the numeric and percentage changes
        * in coefficients across nested models
        args x xbefore xafter

        local xdiff = `xafter' - `xbefore'
        local xpctchg = round(`xdiff'/ `xbefore', .001) * 100
        local xdiff = round(`xdiff', .01)

        if `xpctchg' < 0 {
                local xpctchg = abs(`xpctchg')
```

```
            local note "With controls added the coefficient for `x' decreases (in magnitude) by
`xdiff', i.e. by `xpctchg'%"
        }
        else if `xpctchg' > 0 {
            local note "With controls added the coefficient for `x' increases by `xdiff', i.e. by
`xpctchg'%"
        }
        else {
            local note "When controls are added the coefficient for `x' does not change"
        }
        return local note "`note'"
end
```

7.4 References

The American National Election Studies (ANES). ANES 2008 Time Series Study. Inter-

university Consortium for Political and Social Research [distributor], 2015-11-10.

https://doi.org/10.3886/ICPSR25383.v3.

Jann, Ben. (2005). Making Regression Tables from Stored Estimates. The Stata Journal:

Promoting Communications on Statistics and Stata, 5(3), 288–308.

https://doi.org/10.1177/1536867x0500500302.

Jann, Ben. (2007). Making Regression Tables Simplified. The Stata Journal: Promoting

Communications on Statistics and Stata, 7(2), 227–244.

https://doi.org/10.1177/1536867x0700700207.

Kohler, U., Karlson, K. B., & Holm, A. (2011). Comparing Coefficients of Nested Nonlinear

Probability Models. The Stata Journal: Promoting Communications on Statistics and Stata, 11(3),

420–438.http://www.stata-journal.com/article.html?article=st0236.

Long, J. S., and J. Freese. 2014. Regression Models for Categorical Dependent Variables Using Stata. 3rd ed. College Station, TX: Stata Press.