Online Appendix

A. Why hybrid models do not effectively separate within and between effects in the presence of interaction terms

Schunck (2013: p. 72) highlights that the following specification is the appropriate way to estimate hybrid models (i.e., models that present separate estimates of within and between effects) with interaction terms.

$$y_{it} = \beta_1(x1_{it} - \overline{x1_i}) + \beta_2(x2_{it} - \overline{x2_i}) + \beta_3\left[(x1_{it} * x2_{it}) - \overline{(x1 * x2)_i}\right] + \beta_4\overline{x1_i} + \beta_5\overline{x2_i} + \beta_6\overline{(x1 * x2)_i} + u_i + \varepsilon_{it} \quad \text{[A1]}$$

In equation A1, the terms are as previously defined in the body of the paper and u_i is a random effect for firm *i*. The first three terms on the right-hand side of the equation are said to capture the within-firm effects. However, because these terms are identical to equation 5 in the body of the paper; this approach confounds between and within variation when identifying the coefficient estimates for these terms. Therefore, the within-firm coefficient estimates from this approach will not accurately isolate within-firm from between-firm effects.

To demonstrate this, I present simulations from the above specification in Table A1. The data used to generate this table is the same as data used to generate Table 1 in the body of the paper. The simulation results in Table A1 parallel the simulation results in Table 1. Holding constant the level of within-firm variation, the coefficient estimates of $\widehat{\beta_3}$ approach their true value and more coefficient estimates take p-values of less than 0.05 as the between-firm variation increases. One can see this by moving from left to right in the first and second rows of results as in Table 1. In the bottom row, between-firm variation aids identifying the relationship. Although it is not noticeable in the table due to low p-values of estimates in all cells, the average t-statistic is larger and the estimates more concentrated around zero as presented in Figure 1 in the body of the paper.

B. Properties of the estimator that interacts within-firm effects.

In order to demonstrate that the estimator I present in equation 6 in the body of the paper (i.e., interacting the within-firm differences of *x*1 and *x*2) does not generate meaningful coefficient estimates for a fixed-effect model, I turn to simulated data. The simulated data are the same data that I use in Table 1 in the body of the paper.

Table A2 summarizes the average coefficient estimates and the percentage of observations where the p-value of the estimates is less than 0.05 with a two-tailed t-test. Figure A1 plots the coefficient estimates of the interaction term and their t-values for the four corner cells of Table A2.

In the top left cell of Table A2 – where there is minimal within-firm and between-firm variation – the coefficient estimate of $\widehat{\beta_3}$ is far away from the true value (7.92 compared to the true value of 5) and only 2.3 percent of the simulated samples have a coefficient that tests different from zero ($p<0.05$). As in Table 1 in the body of the paper, this reflects that there is little variation in the data. Figure A1 demonstrates that the distribution of coefficient estimates is diffuse and the t-values concentrate around zero.

However, unlike Table 1 in the body of the paper, this estimate does not approach the true value as within-firm or between-firm variation increases. The average coefficient estimate in the bottom right cell – the condition where within-firm and between-firm variation are large – is also very different from the true value (3.75 compared to the true value of 5). In addition, while the coefficient tests different from zero in all simulations, in 79 percent of these simulations the coefficient estimate tests different from the true value of 5. Figure A1 demonstrates a concentration of coefficient estimates around that mean value, with a small proportion in the right tail around the true value.

Therefore, in contrast to Table 1 in the body of the paper, the coefficient estimates do not approach the true value as within-firm variation or between-firm variation increase. It is clear that the estimates from this approach are biased. In addition, as the within and between-firm variation increase, the estimates of $\widehat{\beta_1}$ and $\widehat{\beta_2}$ move further away from their true values. In the bottom right cell, the estimates of $\widehat{\beta_1}$ and $\widehat{\beta_2}$ are 1000 and 500, compared to their true values of 500 and 0. Not only is the interaction term biased with this estimator, so are the other coefficient estimates.

These simulations demonstrate that interacting within-firm differences of independent variables does not provide a meaningful way to address the concern that fixed-effect models with interaction terms can capture between-firm variation.

C. Stata code for simulations by Table (one parameter setting presented)

Table 1:

```
program xtr1_1
	drop _all
	set obs 500
	gen e=(rnormal())
	gen u1= uniform()
	gen u2= uniform()
	gen u3= uniform()
	gen u4= uniform()
	gen f=mod(_n,5)
	replace f=5 if f==0
	gen firm=ceil(_n/5)
	xtset firm f
	replace u1=L1.u1 if f==2
	replace u1=L2.u1 if f==3
	replace u1=L3.u1 if f==4
	replace u1=L4.u1 if f==5
```

```
	replace u3=L1.u3 if f==2
	replace u3=L2.u3 if f==3
	replace u3=L3.u3 if f==4
	replace u3=L4.u3 if f==5
	gen x1=1*u1+1*u2
	gen x2=1*u3+1*u4
	gen y=100+500*x1+5*x2*x1+1000*e
	gen x1_x2=x1*x2
	xtreg y x1 x2 x1_x2, fe
	end

set seed 12345
simulate _b _se, reps(10000): xtr1_1

gen t_x1=_b_x1/_se_x1
gen t_x2=_b_x2/_se_x2
gen t_x1_x2=_b_x1_x2/_se_x1_x2


sum _b_x1 _b_x2 _b_x1_x2
sum t_x1 t_x2 t_x1_x2

gen p_x1=0
gen p_x2=0
gen p_x1_x2=0

replace p_x1=1 if t_x1>=1.96
replace p_x1=1 if t_x1<=-1.96
replace p_x2=1 if t_x2>=1.96
replace p_x2=1 if t_x2<=-1.96
replace p_x1_x2=1 if t_x1_x2>=1.96
replace p_x1_x2=1 if t_x1_x2<=-1.96

sum p_x1 p_x2 p_x1_x2
```

Table 2

```
program unobs1_1_10
	drop _all
	set obs 500
	gen e=(rnormal())
	gen u1= uniform()
	gen u2= uniform()
	gen u3= uniform()
	gen u4= uniform()
	gen u5= uniform()
	gen f=mod(_n,5)
	replace f=5 if f==0
	gen firm=ceil(_n/5)
	xtset firm f
	replace u1=L1.u1 if f==2
	replace u1=L2.u1 if f==3
	replace u1=L3.u1 if f==4
```

```
		replace u1=L4.u1 if f==5
		replace u3=L1.u3 if f==2
		replace u3=L2.u3 if f==3
		replace u3=L3.u3 if f==4
		replace u3=L4.u3 if f==5
		gen x1=1*u1+1*u2
		gen x2=1*u3+1*u4
		gen x3=1*u3+10*u5
		gen y=100+500*x1+5*x3*x1+1000*e
		gen x1_x2=x1*x2
		xtreg y x1 x2 x1_x2, fe
		end

set seed 12345
simulate _b _se, reps(10000): unobs1_1_10

gen t_x1=_b_x1/_se_x1
gen t_x2=_b_x2/_se_x2
gen t_x1_x2=_b_x1_x2/_se_x1_x2


sum _b_x1 _b_x2 _b_x1_x2
sum t_x1 t_x2 t_x1_x2

gen p_x1=0
gen p_x2=0
gen p_x1_x2=0

replace p_x1=1 if t_x1>=1.96
replace p_x1=1 if t_x1<=-1.96
replace p_x2=1 if t_x2>=1.96
replace p_x2=1 if t_x2<=-1.96
replace p_x1_x2=1 if t_x1_x2>=1.96
replace p_x1_x2=1 if t_x1_x2<=-1.96

sum p_x1 p_x2 p_x1_x2
```

Table 3 – quartile 1 of x1

```
program unobs100_100_10_q1
		drop _all
		set obs 500
		gen e=(rnormal())
		gen u1= uniform()
		gen u2= uniform()
		gen u3= uniform()
		gen u4= uniform()
		gen u5= uniform()
		gen f=mod(_n,5)
		replace f=5 if f==0
		gen firm=ceil(_n/5)
		xtset firm f
		replace u1=L1.u1 if f==2
```

```
            replace u1=L2.u1 if f==3
            replace u1=L3.u1 if f==4
            replace u1=L4.u1 if f==5
            replace u3=L1.u3 if f==2
            replace u3=L2.u3 if f==3
            replace u3=L3.u3 if f==4
            replace u3=L4.u3 if f==5
            gen x1=100*u1+100*u2
            gen x2=100*u3+100*u4
            gen x3=100*u3+10*u5
            gen y=100+500*x1+5*x2*x1+1000*e
            sort x1
            egen x1_quartile=cut(x1), group(4)
            drop if x1_quartile~=0
            xtreg y x1 x2, fe
            end

set seed 12345
simulate _b _se, reps(10000): unobs100_100_10_q1

gen t_x1=_b_x1/_se_x1
gen t_x2=_b_x2/_se_x2



sum _b_x1 _b_x2
sum t_x1 t_x2

gen p_x1=0
gen p_x2=0


replace p_x1=1 if t_x1>=1.96
replace p_x2=1 if t_x2>=1.96


sum p_x1 p_x2
```

Table 4 – quartile 1 of x1

```
program unobs100_100_10_q1
            drop _all
            set obs 500
            gen e=(rnormal())
            gen u1= uniform()
            gen u2= uniform()
            gen u3= uniform()
            gen u4= uniform()
            gen u5= uniform()
            gen f=mod(_n,5)
            replace f=5 if f==0
            gen firm=ceil(_n/5)
            xtset firm f
```

```
    replace u1=L1.u1 if f==2
    replace u1=L2.u1 if f==3
    replace u1=L3.u1 if f==4
    replace u1=L4.u1 if f==5
    replace u3=L1.u3 if f==2
    replace u3=L2.u3 if f==3
    replace u3=L3.u3 if f==4
    replace u3=L4.u3 if f==5
    gen x1=100*u1+100*u2
    gen x2=100*u3+100*u4
    gen x3=100*u3+10*u5
    gen y=100+500*x1+5*x3*x1+1000*e
    sort x1
    egen x1_quartile=cut(x1), group(4)
    drop if x1_quartile~=0
    xtreg y x1 x2, fe
    end

set seed 12345
simulate _b _se, reps(10000): unobs100_100_10_q1

gen t_x1=_b_x1/_se_x1
gen t_x2=_b_x2/_se_x2



sum _b_x1 _b_x2
sum t_x1 t_x2

gen p_x1=0
gen p_x2=0


replace p_x1=1 if t_x1>=1.96
replace p_x2=1 if t_x2>=1.96


sum p_x1 p_x2
```

Table A2:

```
program wf1_1
    drop _all
    set obs 500
    gen e=(rnormal())
    gen u1= uniform()
    gen u2= uniform()
    gen u3= uniform()
    gen u4= uniform()
    gen f=mod(_n,5)
    replace f=5 if f==0
    gen firm=ceil(_n/5)
```

```
		xtset firm f
		replace u1=L1.u1 if f==2
		replace u1=L2.u1 if f==3
		replace u1=L3.u1 if f==4
		replace u1=L4.u1 if f==5
		replace u3=L1.u3 if f==2
		replace u3=L2.u3 if f==3
		replace u3=L3.u3 if f==4
		replace u3=L4.u3 if f==5
		gen x1=1*u1+1*u2
		gen x2=1*u3+1*u4
		gen y=100+500*x1+5*x2*x1+500*e
		gen x1_x2=x1*x2
		sort firm
		by firm: egen firm_avg_y=mean(y)
		by firm: egen firm_avg_x2 =mean(x2)
		by firm: egen firm_avg_x1=mean(x1)
		by firm: egen firm_avg_x1_x2 =mean(x1_x2)
		gen y_dev=y-firm_avg_y
		gen x2_dev=x2-firm_avg_x2
		gen x1_dev=x1-firm_avg_x1
		gen x2_dev_x1=x2_dev*firm_avg_x1
		gen x1_dev_x2=x1_dev*firm_avg_x2
		gen x1_dev_x2_dev=x1_dev*x2_dev
		reg y_dev x1_dev x2_dev x1_dev_x2_dev, nocon
		end

set seed 12345
simulate _b _se, reps(10000): wf1_1

replace _se_x1_dev=_se_x1_dev*sqrt((497/397))
replace _se_x2_dev=_se_x2_dev*sqrt((497/397))
replace _se_x1_dev_x2_dev=_se_x1_dev_x2_dev*sqrt((497/397))

gen t_x1_dev=_b_x1_dev/_se_x1_dev
gen t_x2_dev=_b_x2_dev/_se_x2_dev
gen t_x1_dev_x2_dev=_b_x1_dev_x2_dev/_se_x1_dev_x2_dev


sum _b_x1_dev _b_x2_dev _b_x1_dev_x2_dev
sum t_x1_dev t_x2_dev t_x1_dev_x2_dev

gen p_x1_dev=0
gen p_x2_dev=0
gen p_x1_dev_x2_dev=0

replace p_x1_dev=1 if t_x1_dev>=1.96
replace p_x1_dev=1 if t_x1_dev<=-1.96
replace p_x2_dev=1 if t_x2_dev>=1.96
replace p_x2_dev=1 if t_x2_dev<=-1.96
replace p_x1_dev_x2_dev=1 if t_x1_dev_x2_dev>=1.96
replace p_x1_dev_x2_dev=1 if t_x1_dev_x2_dev<=-1.96
```

sum p_x1_dev p_x2_dev p_x1_dev_x2_dev

Table A1: Simulation Results for standard fixed-effects model with an interaction

Average coefficient estimates [percent with p<0.05 – two tailed in square brackets]

	Degree of between-firm variation in *x*1 and *x*2		
Degree of within-firm variation in *x*1 and *x*2	Low	Medium	High
Low	Average $\widehat{\beta_3}$ = 8.30 [5.3%] Average $\widehat{\beta_1}$ = 497 [25.9%] Average $\widehat{\beta_2}$ = -2.58 [5.7%] Average $\widehat{\beta_4}$ = 501 [18.3%] Average $\widehat{\beta_5}$ = 5.02 [4.8%] Average $\widehat{\beta_6}$ = 3.92 [4.8%]	Average $\widehat{\beta_3}$ = 5.09 [5.6%] Average $\widehat{\beta_1}$ = 500 [41.3%] Average $\widehat{\beta_2}$ = 0.10 [5.5%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = 0.58 [4.6%] Average $\widehat{\beta_6}$ = 4.97 [13.9%]	Average $\widehat{\beta_3}$ = 5.01 [22.3%] Average $\widehat{\beta_1}$ = 500 [45.1%] Average $\widehat{\beta_2}$ = 0.33 [5.4%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = 0.06 [4.5%] Average $\widehat{\beta_6}$ = 5.00 [100%]
Medium	Average $\widehat{\beta_3}$ = 5.08 [14.0%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = -0.35 [5.1%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = -0.33 [5.0%] Average $\widehat{\beta_6}$ = 5.14 [7.0%]	Average $\widehat{\beta_3}$ = 5.03 [31.1%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = -0.26 [5.7%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = 0.50 [4.8%] Average $\widehat{\beta_6}$ = 4.99 [19.5%]	Average $\widehat{\beta_3}$ = 5.00 [100%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = 0.01 [5.5%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = 0.06 [4.6%] Average $\widehat{\beta_6}$ = 5.00 [100%]
High	Average $\widehat{\beta_3}$ = 5.00 [100%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = -0.03 [5.3%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = -0.04 [5.0%] Average $\widehat{\beta_6}$ = 5.00 [100%]	Average $\widehat{\beta_3}$ = 5.00 [100%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = -0.04 [5.1%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = -0.03 [4.9%] Average $\widehat{\beta_6}$ = 5.00 [100%]	Average $\widehat{\beta_3}$ = 5.00 [100%] Average $\widehat{\beta_1}$ = 500 [100%] Average $\widehat{\beta_2}$ = -0.03 [5.7%] Average $\widehat{\beta_4}$ = 500 [100%] Average $\widehat{\beta_5}$ = 0.05 [4.8%] Average $\widehat{\beta_6}$ = 5.00 [100%]

$\widehat{\beta_3}$, $\widehat{\beta_6}$ are the estimates of the interaction term (true value=5)

$\widehat{\beta_1}$, $\widehat{\beta_4}$ are the estimates of *x*1 (true value=500)

$\widehat{\beta_2}$, $\widehat{\beta_5}$ are the estimates of *x*2 (true value=0)

Estimates are from 10,000 simulations of the underlying data structure

Data are the same data used to estimate Table 1 in the text of the paper

Table A2 – Simulation Results for interacting with-firm effects of *x*1 and *x*2

Average coefficient estimates [percent with p<0.05 – two tailed in square brackets]

	Degree of between-firm variation in *x*1 and *x*2		
Degree of within-firm variation in *x*1 and *x*2	Low	Medium	High
Low	Average $\widehat{\beta_3}$ = 7.92 [2.3%] Average $\widehat{\beta_1}$ = 505 [100%] Average $\widehat{\beta_2}$ = 5.42 [5.1%]	Average $\widehat{\beta_3}$ = 7.86 [2.4%] Average $\widehat{\beta_1}$ = 527 [100%] Average $\widehat{\beta_2}$ = 27.0 [6.2%]	Average $\widehat{\beta_3}$ = 7.21 [2.46%] Average $\widehat{\beta_1}$ = 752 [100%] Average $\widehat{\beta_2}$ = 253 [82%]
Medium	Average $\widehat{\beta_3}$ = 3.80 [16.7%] Average $\widehat{\beta_1}$ = 527 [100%] Average $\widehat{\beta_2}$ = 27.5 [89%]	Average $\widehat{\beta_3}$ = 3.79 [16.5%] Average $\widehat{\beta_1}$ = 550 [100%] Average $\widehat{\beta_2}$ = 50.0 [100%]	Average $\widehat{\beta_3}$ = 3.73 [9.3%] Average $\widehat{\beta_1}$ = 775 [100%] Average $\widehat{\beta_2}$ = 275 [100%]
High	Average $\widehat{\beta_3}$ = 3.76 [100%] Average $\widehat{\beta_1}$ = 752 [100%] Average $\widehat{\beta_2}$ = 252 [100%]	Average $\widehat{\beta_3}$ = 3.76 [100%] Average $\widehat{\beta_1}$ = 775 [100%] Average $\widehat{\beta_2}$ = 275 [100%]	Average $\widehat{\beta_3}$ = 3.75 [100%] Average $\widehat{\beta_1}$ = 1000 [100%] Average $\widehat{\beta_2}$ = 500 [100%]

$\widehat{\beta_3}$ is the estimate of the interaction term (true value=5)

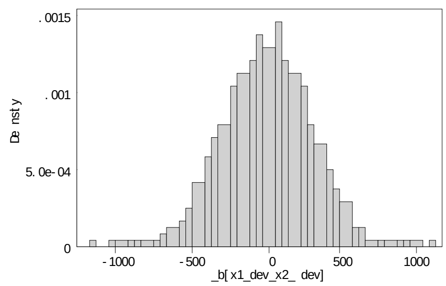
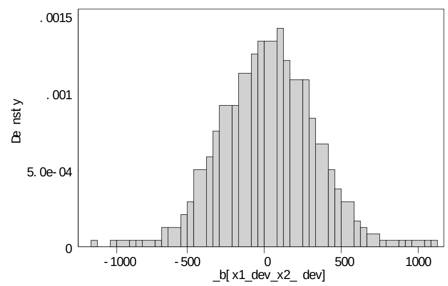
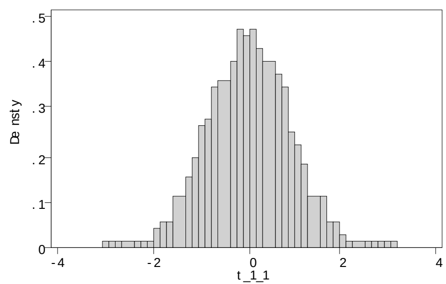
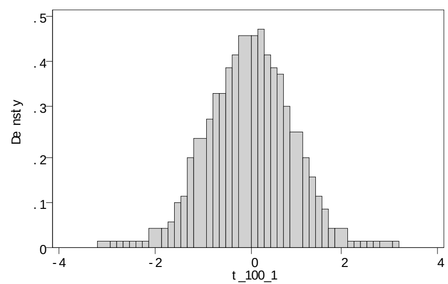
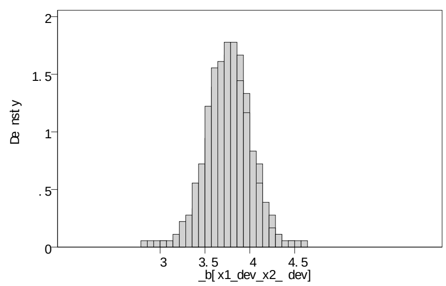
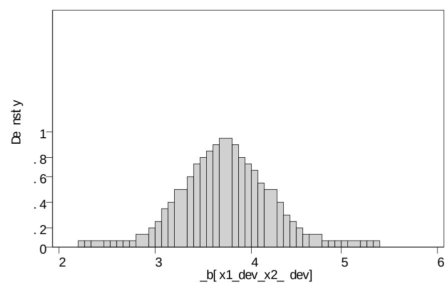
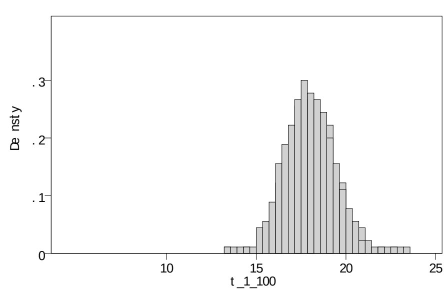
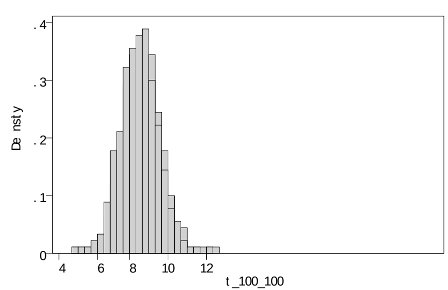
$\widehat{\beta_1}$ is the estimate of *x*1 (true value=500)

$\widehat{\beta_2}$ is the estimate of *x*2 (true value=0)

Estimates are from 10,000 simulations of the underlying data structure

Precise data definitions appear in Appendix 2.

Figure A1 – Distribution of interaction coefficient estimates from Table 2

<table>
<tr><td></td><td></td><td colspan="2">Between-firm variation</td></tr>
<tr><td>Within-firm variation</td><td></td><td>Low</td><td>High</td></tr>
<tr><td>Low</td><td>$\widehat{\beta_3}$</td><td>
</td><td>
</td></tr>
<tr><td></td><td>t</td><td>
</td><td>
</td></tr>
<tr><td>High</td><td>$\widehat{\beta_3}$</td><td>
</td><td>
</td></tr>
<tr><td></td><td>t</td><td>
</td><td>
</td></tr>
</table>