BUSN 41903 50: Applied Econometrics

Andrew McKinley, Lauren Mostrom, Pietro Ramella,
Francisco Ruela, Bohan Yang

Problem Set 3

1.

(a) Adding aggregate time effects captures macro trends that affects all countries in certain periods, for example a recession that interests all countries in a given year.

(b) The country fixed effects α_i capture unobserved heterogeneity that is constant within country over the examined period, for example levels of natural resources, or institutional factors.

(c) Economic reasoning suggests a negative δ_1: because higher tax rates makes investment less profitable, we would expect that a decrease in investment would follow higher marginal tax rates.

(d) Rewrite the model in matrix form:

$$Y = \theta R + \alpha D + \gamma Z + \delta_1 tax + \delta_2 disaster + u = \alpha D + \beta X + u$$

where $Y_{it} = log(investment_{it})$, $\beta = [\delta_1, \delta_2, \gamma', \theta']$, $X = [tax', disaster', Z', R']'$.
Because fixed effects, $\delta = [\delta_1, \delta_2]$, and an intercept cannot be jointly identified, I need to choose a normalization: set to zero the intercept and the first time fixed effect. If we are willing to assume strict exogeneity, that is tax and $disaster$ are uncorrelated with ϵ at every time period, we can then estimate using OLS the coefficients

$$(\hat{\beta}_{FE}, \hat{\alpha}) = (W'W)^{-1}(W'Y)$$

where $W = [X \ D]$.
To compute the standard errors, assuming observations are independent across countries, $\hat{V} = \hat{\sigma}^2 \hat{Q}^{-1}$, where

$$\hat{\sigma}^2 = \frac{1}{NT} \sum_{i,t} e_{it}^2, \quad e_{it} = y_{it} - \bar{y}_i - (x_{it} - \bar{x}_i)' \hat{\beta}_{FE}, \quad k = dim(x_{it})$$

$$\hat{Q} = \frac{1}{NT} \sum_{i,t} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)'$$

.

(e) It seems unlikely that we can rule out dynamics between new capital investments and changes in marginal taxes or natural disasters, yet, tax_{it} may still violate strict exogeneity because the taxing authority may choose tax_{it} in response of low/excessive levels of observed investment, therefore $E[x_{it}, uit - j] \neq 0$ for some $j = 1, ...$ There are not real issues with natural disasters if we rule out persistent effects, and therefore dynamics.

2.

(a)

In the table below we report the IID standard errors, heteroskedasticity-robust standard errors, and clustered standard errors. Robust SEs were calculated as follows:

$$\hat{\Omega} = \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{it} \right)^{-1} \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{it} \hat{\epsilon}_{it}^2 \right) \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{it} \right)^{-1}$$

Where $\hat{\epsilon}$ is the POLS residual. We also include clustered standard errors; however, if we suspect the observations are serially correlated (as would be implied by the use of clustered SEs) we should not be using POLS anyway. Nevertheless, clustered SEs were calculated as follows:

$$\hat{\Omega} = \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{it} \right)^{-1} \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{i,90} \hat{\epsilon}_{it} \hat{\epsilon}_{i,90} + x_{it} x'_{i,93} \hat{\epsilon}_{it} \hat{\epsilon}_{i,93} \right) \left(\frac{1}{NT} \sum_{i,t} x_{it} x'_{it} \right)^{-1}$$

Table 1: Pooled OLS

	IID	Robust	Clustered
(Intercept)	-5.2780	-5.2780	-5.2780
	(4.4278)	(5.3868)	(6.6760)
d93	-2.0674	-2.0674	-2.0674
	(2.1446)	(1.9981)	(1.3066)
exec	0.1277	0.1277	0.1277
	(0.2632)	(0.1342)	(0.1678)
unem	2.5289**	2.5289**	2.5289**
	(0.7817)	(1.1076)	(1.5047)
R^2	0.1016	0.1016	0.1016
Adj. R^2	0.0741	0.0741	0.0741
Num. obs.	102	102	102

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

(b)

In this setting FE and FD are numerically identical because there are only two time periods, 1990 and 1993. To see this, recall the fixed effect model:

$$y_{i1} - \bar{y}_i = (x_{i1} - \bar{x}_i)'\beta + \epsilon_{i1} - \bar{\epsilon}_i$$

$$y_{i1} - \left(\frac{y_{i1} + y_{i0}}{2}\right) = \left(x_{i1} - \frac{x_{i1} + x_{i0}}{2}\right)'\beta + \epsilon_{i1} - \left(\frac{\epsilon_{i1} + \epsilon_{i0}}{2}\right)$$

$$\frac{y_{i1} - y_{i0}}{2} = \left(\frac{x_{i1} - x_{i0}}{2}\right)'\beta + \frac{\epsilon_{i1} - \epsilon_{i0}}{2}$$

$$y_{i1} - y_{i0} = (x_{i1} - x_{i0})'\beta + \epsilon_{i1} - \epsilon_{i0}$$

Which is exactly the first differences model.

In the table below we report only the results for FD. We chose FD because it reduces the procedure down to a cross-sectional dataset and only requires heteroskedasticity-robust standard errors, rather than maintaining the panel data structure and requiring clustered standard errors. Clustered standard errors are reported in the table below, but only to illustrate that they are identical to robust in the first differences method when we only have two periods.

There does not appear to be any deterrant effect of capital punishment; based on these results we fail to reject the null hypothesis that the coefficient on *exec* is equal to zero. In this setting FE and FD are numerically identical because there are only two time periods, 1990 and 1993. To see this, recall the fixed effect model:

$$y_{i1} - \bar{y}_i = (x_{i1} - \bar{x}_i)'\beta + \epsilon_{i1} - \bar{\epsilon}_i$$

$$y_{i1} - \left(\frac{y_{i1} + y_{i0}}{2}\right) = \left(x_{i1} - \frac{x_{i1} + x_{i0}}{2}\right)'\beta + \epsilon_{i1} - \left(\frac{\epsilon_{i1} + \epsilon_{i0}}{2}\right)$$

$$\frac{y_{i1} - y_{i0}}{2} = \left(\frac{x_{i1} - x_{i0}}{2}\right)'\beta + \frac{\epsilon_{i1} - \epsilon_{i0}}{2}$$

$$y_{i1} - y_{i0} = (x_{i1} - x_{i0})'\beta + \epsilon_{i1} - \epsilon_{i0}$$

Which is exactly the first differences model.

In the table below we report only the results for FD. We chose FD because it reduces the procedure down to a cross-sectional dataset and only requires heterskedasticity-robust standard errors, rather than maintaining the panel data structure and requiring clustered standard errors. Clustered standard errors are reported in the table below, but only to illustrate that they are identical to robust in the first differences method when we only have two periods.

There does not appear to be any deterrant effect of capital punishment; based on these results we fail to reject the null hypothesis that the coefficient on *exec* is equal to zero.

(c)

Table 2: First Differences	IID	Robust	Clustered
(Intercept)	0.4133	0.4133	0.4133
	(0.2094)	(0.2000)	(0.2000)
delta_exec	−0.1038*	−0.1038*	−0.1038*
	(0.0434)	(0.0170)	(0.0170)
delta_unem	−0.0666	−0.0666	−0.0666
	(0.1587)	(0.1469)	(0.1469)
R^2	0.1097	0.1097	0.1097
Adj. R^2	0.0727	0.0727	0.0727
Num. obs.	51	51	51

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

As explained above, we used heteroskedasticity-robust SEs because the assumptions required for IID standard errors are too strong, and in a first differences model with only two periods, clustered and robust SEs are identical (as demonstrated in the table above).

From these results the coefficient on *exec* is negative and statistically significant, so it does appear that there is a deterrant effect of capital punishment.

(d)

FD is preferred in this context because POLS requires the assumption that there is no serial correlation between observations. This is an unrealistic assumption in this setting because there could be many persistent features of states that are correlated both with executions and with the murder rate.

3.

Compute the bias and root-mean-squared-error of each estimator. Form a histogram for each estimator. Do the histograms appear to be approximately normal centered over the true parameter value? Comment on what you think you learn that is generalizable from this exercise.

First, we include the tables of the bias and RMSE for the instance where T = 6.

Table 1: T=6

	Fixed-Effects	First-Difference	Hahn-Kuersteiner	Jackknife
Bias_rho= 0	-0.164	0.006	-0.025	0.006
Bias_rho = 0.5	-0.260	0.007	-0.054	0.148
Bias_rho = 0.95	-0.262	0.023	0.020	0.026
RMSE_rho = 0	0.168	0.078	0.051	0.089
RMSE_rho = 0.5	0.264	0.132	0.073	0.162
RMSE_rho = 0.95	0.264	0.239	0.049	0.043
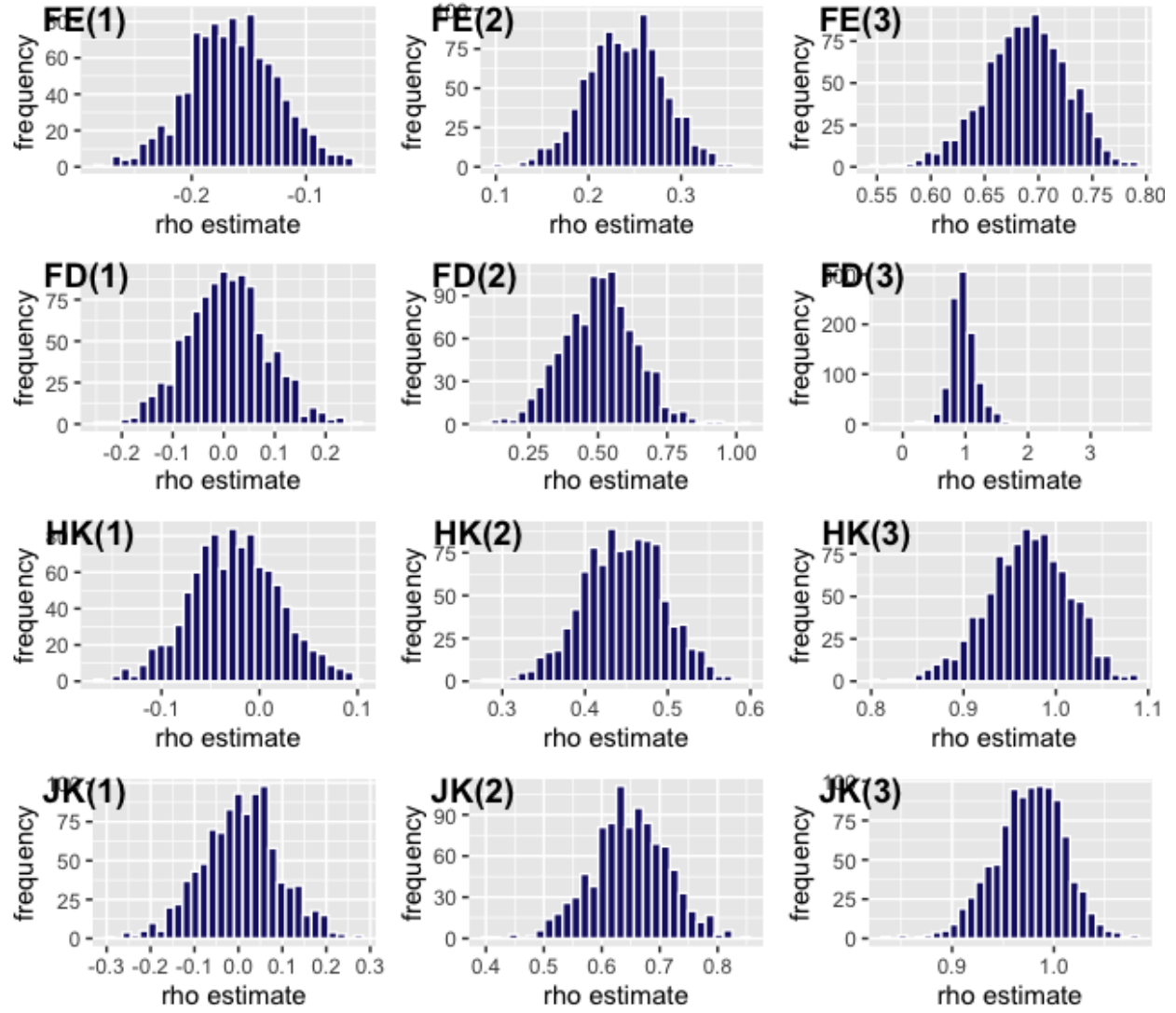
Next, we include the same for the case where T = 25.

Table 2: T=25

	Fixed-Effects	First-Difference	Hahn-Kuersteiner	Jackknife
Bias_rho = 0	-0.040	-0.0001	-0.002	-0.0004
Bias_rho = 0.5	-0.062	0.003	-0.004	0.230
Bias_rho = 0.95	-0.045	0.003	0.031	0.049
RMSE_rho = 0	0.045	0.031	0.021	0.046
RMSE_rho = 0.5	0.064	0.040	0.020	0.231
RMSE_rho = 0.95	0.046	0.068	0.032	0.049
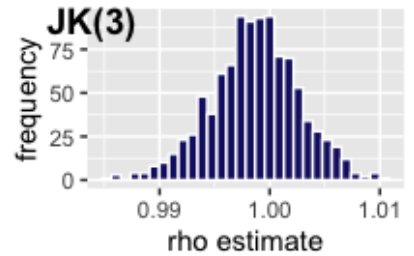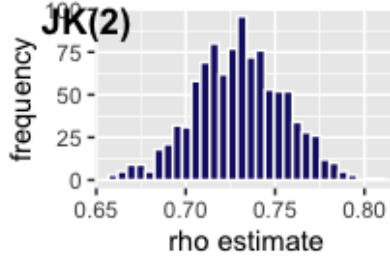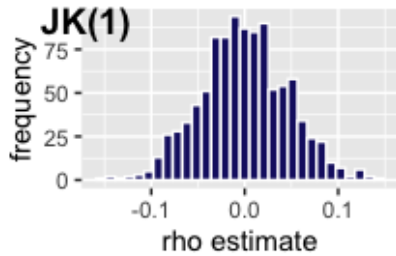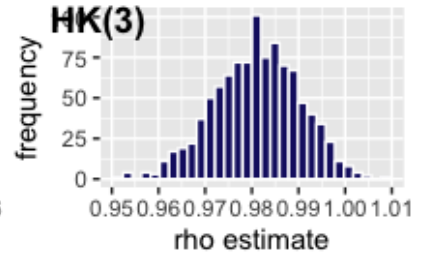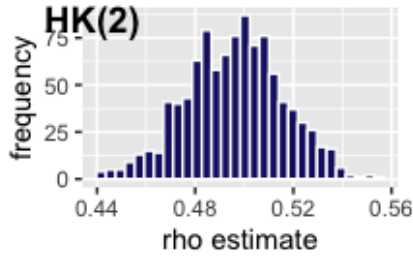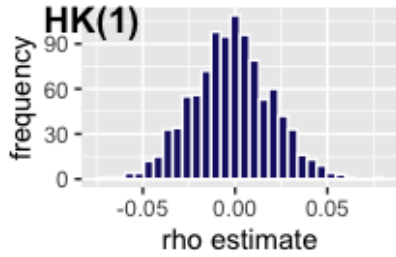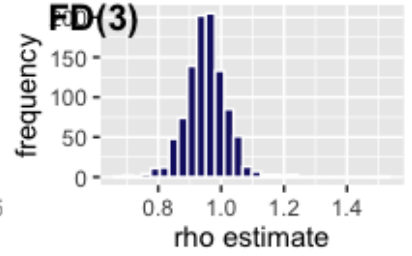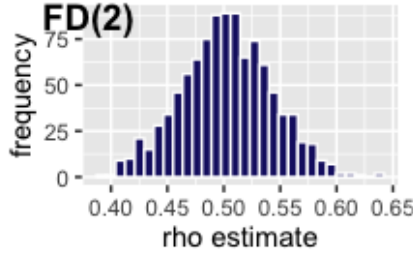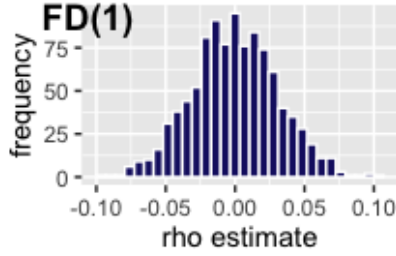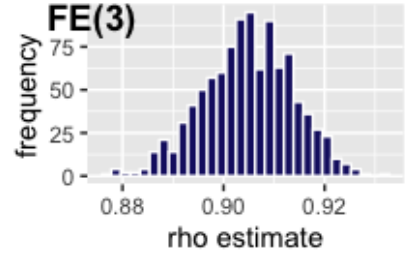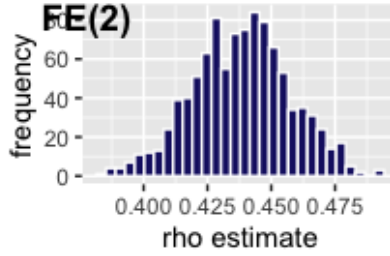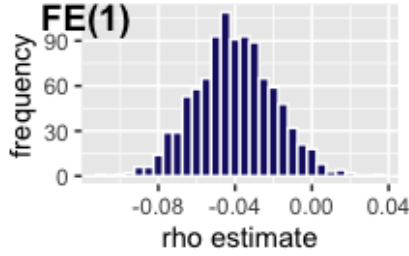
We note that the distributions do appear to be approximately centered over the true value in a few cases, notably those that are not the fixed effects estimator which appears to have a wider spread and weaker identification. However, some are still a bit off to either direction. We can also see, as we learned in class, that as $\frac{N}{T}$ gets smaller that our inferences are better. In this case when T = 25.

It also appears that the First Difference IV estimator works better for lower values of ρ. This makes sense as the first instrument becomes weaker as ρ increases. Turning to the bias corrected estimates, it is clear that both out perform the fixed effects estimator. We note, however, that the distribution for the Hahn-Kuersteiner estimator does appear to have wider tails than the Jackknife estimator. This could be caused by the fact that it corrects only for the first order bias. We also note that for T=25 and $\rho = 0.5$ our Jackknife estimator appears to be off, while it is fine else where. This however does not appear to be generalizable.

4.

(a)

Assumption: after controlling for *Log of state nonfarm employment*, *state fixed effect* (all the non-observable state-level factors influencing both outcome and the implication of policy, and don't change over time) and *year fixed effect* (all the non-observable year-level macroeconomics factors influencing both outcome and the implication of policy), implication of policy is independent on the error term.

Implication: the policy increases THS employment by 12.8%, while this is not statistically significant - essentially, the policy has no causal effect.

Dependent Variable:	lnths
Model:	(1)
Variables	
mico	0.1280
	(0.0888)
lnemp	2.014***
	(0.4236)
Fixed-effects	
s	Yes
t	Yes
Fit statistics	
Observations	850
R^2	0.97270
Within R^2	0.13857

Clustered (s) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

(b)

Assumption: after controlling for *Log of state nonfarm employment*, *state fixed effect* (all the non-observable state-level factors influencing both outcome and the implication of policy, and don't change over time), *year fixed effect* (all the non-observable year-level macroeconomics factors influencing both outcome and the implication of policy) and *state-level time trends* (states can have their own linear time trends over the years), implication of policy is independent on the error term.

Implication: the policy increases THS employement by 14.5%, which is statistically significant.

Dependent Variable:	lnths
Model:	(1)
Variables	
mico	0.1451**
	(0.0566)
lnemp	1.500***
	(0.4139)
Fixed-effects	
s	Yes
t	Yes
Varying Slopes	
t (s)	Yes
Fit statistics	
Observations	850
R^2	0.98848
Within R^2	0.08042

Clustered (s) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

(c)

Assumption: 1) After controlling for *state fixed effect* (all the non-observable state-level factors influencing both outcome and the implication of policy, and don't change over time), *year fixed effect* (all the non-observable year-level macroeconomics factors influencing both outcome and the implication of policy), *lnemp* is not correlated with other omitted variables and only influences the outcome through implication of the policy. 2) *lnemp* is correlated with implication of the policy. 3) The labor market law cannot have a direct effect until it is passed.

Implication: the policy decreases THS employment by 10.18%, while this is not statistically significant - essentially, the policy has no causal effect.

Dependent Variable:	lnths
Model:	(1)
Variables	
mico	-10.18
	(23.04)
Fixed-effects	
s	Yes
t	Yes
Fit statistics	
Observations	850
R²	-1.6737
Within R²	-83.358

Clustered (s) standard-errors in parentheses
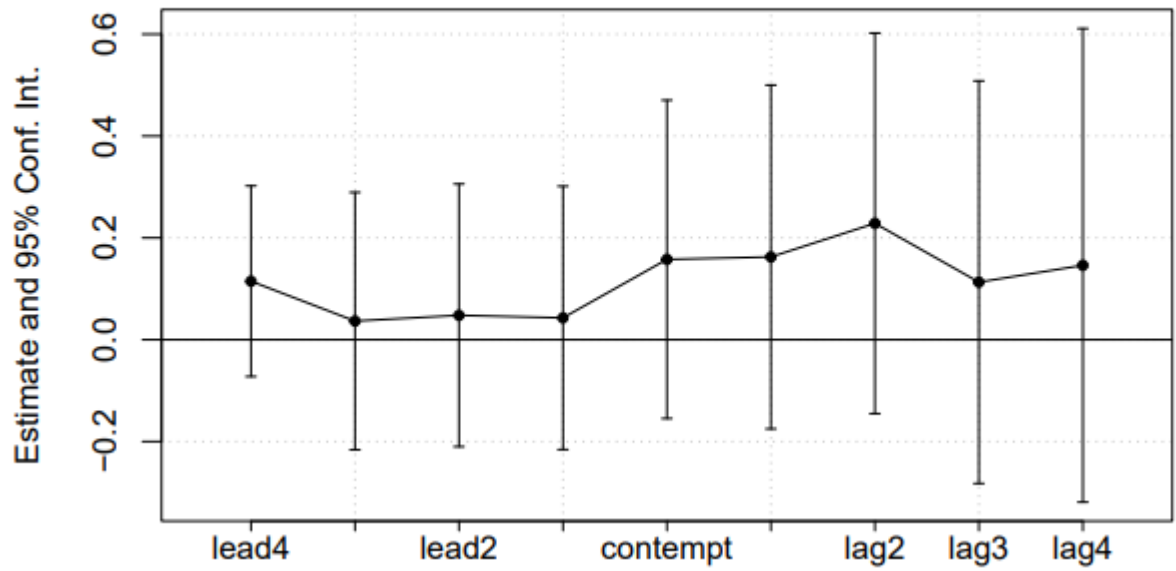*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

(d)

The results for model *a* and *b* are quite similar: lead terms are close to zero and not significant, which validates the assumptions behind them, since there's no pre-trend, or anticipatory response.

According to the contemporaneous and lag terms, the treatment effect is stronger in first two years (especially in the second year), and relatively weaker after the third year. Need to mention, even though the magnitude of coefficients make sence, they are not statistically significant.

Dependent Variable:	lnths	
Model:	(1)	(2)
Variables		
contempt	0.1578	0.1493
	(0.1595)	(0.1379)
lnemp	2.028***	1.526***
	(0.4303)	(0.4067)
lead1	0.0429	0.0093
	(0.1319)	(0.1149)
lead2	0.0479	-0.0064
	(0.1316)	(0.1170)
lead3	0.0366	-0.0348
	(0.1290)	(0.1134)
lead4	0.1149	0.0494
	(0.0955)	(0.0802)
lag1	0.1625	0.1616
	(0.1722)	(0.1422)
lag2	0.2285	0.2344
	(0.1906)	(0.1587)
lag3	0.1130	0.1270
	(0.2017)	(0.1688)
lag4	0.1461	0.2084
	(0.2373)	(0.1857)
Fixed-effects		
s	Yes	Yes
t	Yes	Yes
Varying Slopes		
t (s)		Yes
Fit statistics		
Observations	850	850
R^2	0.97282	0.98868
Within R^2	0.14253	0.09646

Clustered (s) standard-errors in parentheses
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Model A

Model B