

**UW CSSS/POLS 512:
Time Series and Panel Data for the Social Sciences**

Heteroskedasticity in Panel Data

Christopher Adolph

Department of Political Science

and

Center for Statistics and the Social Sciences

University of Washington, Seattle

Review of heteroskedasticity

Recall that in cross-sectional LS, heteroskedasticity

- is assumed away
- if present, biases our standard errors

We noted two approaches

- Model the heteroskedasticity directly with an appropriate ML model, *or*
- Less optimally, continue to use the wrong method (LS), but try to correct the se's; these are known as Huber-White, sandwich, or robust standard errors

How do these approaches transfer to the time series context? to panel data?

Dynamic heteroskedasticity

As with cross-sectional models, we can model heteroskedasticity directly

One possibility is to let heteroskedasticity evolve dynamically

We can let heteroskedasticity be (sort-of) “ARMA”, under the name “GARCH”
Generalized Autoregressive Conditional Heteroskedasticity:

$$y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \sim f_{\mathcal{N}}(0, \sigma_t^2)$$

where

$$\mu_t = \alpha + x_t\beta + \sum_{p=1}^P y_{t-p}\phi_p + \sum_{q=1}^Q \varepsilon_{t-q}\theta_q$$

$$\sigma_t^2 = \exp(\eta + z_t\gamma) + \sum_{c=1}^C \sigma_{t-c}^2\lambda_c + \sum_{d=1}^D \varepsilon_{t-d}^2\xi_d$$

In words, y_t is an ARMA(P, Q)-GARCH(C, D) distributed time-series

(Of course, we could leave out x and/or z if we wanted)

Dynamic heteroskedasticity

$$y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \sim f_{\mathcal{N}}(0, \sigma_t^2)$$

where

$$\mu_t = \alpha + x_t\beta + \sum_{p=1}^P y_{t-p}\phi_p + \sum_{q=1}^Q \varepsilon_{t-q}\theta_q$$

$$\sigma_t^2 = \exp(\eta + z_t\gamma) + \sum_{c=1}^C \sigma_{t-c}^2\lambda_c + \sum_{d=1}^D \varepsilon_{t-d}^2\xi_d$$

Models like the above are workhorses of financial forecasting

Can estimated by ML as usual

In R, `garch()` in the `tseries` package does GARCH

May have to look around a bit for ARMA-GARCH

Dynamic and panel heteroskedasticity

Panel data allows for more complex forms of heteroskedasticity and serial correlation than cross-sectional data. For example. . .

- Serial correlation: $E(\varepsilon_{is}\varepsilon_{it}) = \sigma_{st} \neq 0$

(Reduced/eliminated by appropriate ARMA specification)

- Contemporaneous correlation: $E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij} \neq 0$

(Globally reduced by year fixed effects, but what about pairs of correlated units?)

- Panel heteroskedasticity: $E(\varepsilon_{is}^2) = E(\varepsilon_{it}^2) = \sigma_i^2$, but $\sigma_i^2 \neq \sigma_j^2$

(How to fix?)

- Dynamic heteroskedasticity: $\sigma_{it}^2 = f(\sigma_{i,t-k}^2)$

(How to fix?)

Dynamic and panel heteroskedasticity

Consider a Panel ARMA-GARCH model:

$$y_{it} = \mu_{it} + \varepsilon_{it} \quad \varepsilon_{it} \sim f_{\mathcal{N}}(0, \sigma_{it}^2)$$

$$\mu_{it} = \alpha_i + \tau_t + x_{it}\beta + \sum_{p=1}^P y_{i,t-p}\phi_p + \sum_{q=1}^Q \varepsilon_{i,t-q}\theta_q$$

$$\sigma_{it}^2 = \exp(\eta_i + \zeta_t + z_{it}\gamma) + \sum_{c=1}^C \sigma_{i,t-c}^2 \lambda_c + \sum_{d=1}^D \varepsilon_{i,t-d}^2 \xi_d$$

Suppose we estimated y_{it} as a function of μ_{it} ,
but ignored the structure of the error term

That is, we estimate a panel ARMA or GMM,
but assume ε_{it} is homoskedastic and serially uncorrelated,
conditional on the covariates and lags in μ_{it}

Dynamic and panel heteroskedasticity

If we ignore this:

$$\sigma_{it}^2 = \exp(\eta_i + \zeta_t + z_{it}\gamma) + \sum_{c=1}^C \sigma_{i,t-c}^2 \lambda_c + \sum_{d=1}^D \varepsilon_{i,t-d}^2 \xi_d$$

We would miss three non-standard features of the error variance-covariance:

- Panel heteroskedasticity, from η_i , the unit random effect in the variance function
- Contemporaneous correlation, from ζ_t , the time random effect in the variance function
- Conditional heteroskedasticity: λ_i and ξ_i make the variance time dependent

Thankfully, few reasonable models are this complex. . .

Dynamic and panel heteroskedasticity

Suppose we think this AR(1) with panel heteroskedasticity is appropriate:

$$y_{it} = \mu_{it} + \varepsilon_{it} \quad \varepsilon_{it} \sim f_{\mathcal{N}}(0, \sigma_i^2)$$

$$\mu_{it} = \alpha_i + x_{it}\beta + y_{i,t-p}\phi$$

$$\sigma_i^2 = \exp(\eta_i)$$

Only source of heteroskedasticity is now η_i :
panel heteroskedasticity, not dynamic heteroskedasticity

We could switch this to contemporaneous correlation, by swapping ζ_t for η_i

Roughly the model Beck & Katz advocate as a baseline for comparative politics

Suggest estimating by LS then correcting se's for omission of η_i & contemp. corr.

This procedure yields “panel-corrected standard errors”, PCSEs

What are they, and how do we compute them?

What would we do if we had a plain-vanilla cross-sectional regression and suspected or detected heteroskedasticity?

Recall the standard errors from LS are the square roots of the diagonal elements of

$$\hat{V}(\hat{\beta}) = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$$

So if σ^2 varies by i , these will be badly estimated

Instead, of the usual σ^2 estimator, we could use the residual from each observation as a robust estimate of its variance:

$$\hat{\sigma}_i^2 = \hat{\varepsilon}_i^2$$

A “heteroskedasticity robust” formula for the Var-Cov matrix follows:

$$\hat{V}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1} \left(\sum_i \hat{\varepsilon}_i x_i' x_i \right) (\mathbf{X}'\mathbf{X})^{-1}$$

The standard errors of our parameters (β 's) are the square roots of the diagonal of this matrix

Review: Adjusting standard errors for heteroskedasticity

$$\hat{V}(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1} \left(\sum_i \hat{\varepsilon}_i x_i' x_i \right) (\mathbf{X}'\mathbf{X})^{-1}$$

SE's calculated from this equation are known by many names:

- Huber-White standard errors
- robust standard errors
- sandwich standard errors
- heteroskedasticity consistent standard errors

If you have a single time series, Newey-West standard errors generalize this concept to include robustness to serial correlation

For panel data there are many further options, leading to a vast literature exploring refinements to this basic concept

Panel-corrected standard errors

To calculate panel-corrected standard errors, we need to estimate the correct variance-covariance matrix

Why not just use Huber-White? That would ignore panel structure, which is inefficient if we know how that structure affects heteroskedasticity:

$$\hat{\sigma}_{it}^2 = \hat{\varepsilon}_{it}^2$$

Beck and Katz's panel correction produces sharper estimates of $\hat{\sigma}_{it}^2$ by borrowing strength across the observations from a single unit:

$$\hat{\sigma}_{it}^2 = \hat{\sigma}_i^2 = \frac{1}{T}(\hat{\varepsilon}_{i,1}^2 + \hat{\varepsilon}_{i,2}^2 + \cdots + \hat{\varepsilon}_{i,T}^2)$$

Beck and Katz's panel correction also accounts for contemporaneous correlations across units:

$$\hat{\sigma}_{i,j} = \frac{1}{T}(\hat{\varepsilon}_{i,1}\hat{\varepsilon}_{j,1} + \hat{\varepsilon}_{i,2}\hat{\varepsilon}_{j,2} + \cdots + \hat{\varepsilon}_{i,T}\hat{\varepsilon}_{j,T})$$

Note the above will work better if T is large relative to N

Panel-corrected standard errors

Building this intuition out into a variance-covariance matrix involves a bit of algebra

To make PCSEs,

suppose the variance-covariance matrix Ω is $NT \times NT$ block-diagonal with an $N \times N$ matrix Σ of contemporaneous covariances on diagonal

In other words, allow for unit or contemporaneous heteroskedasticity that stays the same over time

Visualizing this large matrix is tricky

Note that “ $NT \times NT$ block-diagonal” means we are ordering the observations first by time, then by unit (reverse of our usual practice)

Panel-corrected standard errors

Instead, suppose Ω is $NT \times NT$ block-diagonal with an $N \times N$ matrix Σ of contemporaneous covariances on diagonal

In other words, allow for unit or contemporaneous heteroskedasticity that stays the same over time

Beck and Katz (1995) estimate Σ using LS residuals $e_{i,t}$:

$$\hat{\Sigma}_{i,j} = \sum_{t=1}^T \frac{e_{i,t}e_{j,t}}{T}$$

And then use $\hat{\Sigma}$ to construct the covariance matrix

Panel-corrected standard errors

Monte Carlo experiments show panel-corrected standard errors are “correct” unless contemporaneous correlation is very high or T is small relative to N

(Note: alternative is to estimate random effects in variance by ML.)

Beck and Katz suggest using LS with PCSEs and lagged DVs as a baseline model

Most practitioners think fixed effects should also be used

Most important: getting the right lag structure & including FEs where appropriate

PCSEs (or other var-cov correction) is a second-order concern

In R, package `pcse` will calculate PCSEs for a linear regression

Even easier: In the `plm` package,
`vcovBK()` will produce a panel corrected var-cov matrix from a `plm` object

If N is large relative to T , consider the Driscoll and Kraay alternative, `vcovSCC()`

Panel-corrected standard errors: Application

Let's apply Beck-Katz PCSEs to our panel ARIMA/plm example: we'll replace the usual variance-covariance matrix with the panel corrected variance covariance matrix

We must make this substitution *manually* after estimation to get corrected standard errors, confidence intervals, and var-cov matrices:

1. to print the `summary()` of a `plm` model

Example: `summary(plm.res, .vcov=vcovBK(plm.res))`

2. to use the `coeftest()` function on a `plm` model

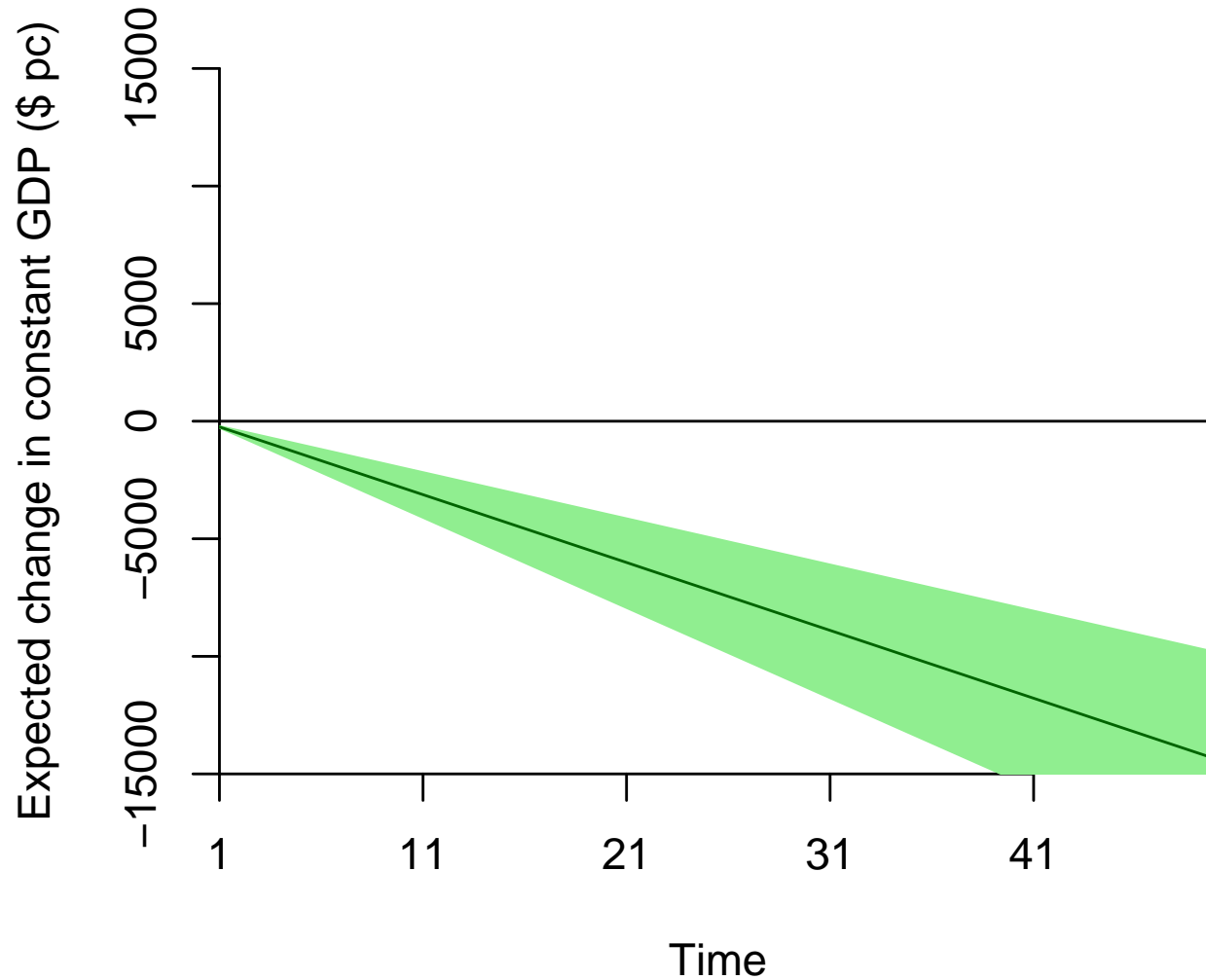
Example: `coeftest(plm.res, .vcov=vcovBK(plm.res))`

3. to simulate parameters with `mvrnorm()` for computing counterfactuals

Example: `mvrnorm(10000, coef(plm.res), vcovBK(plm.res))`

	Model				
	RE	FE	FE-pcse	FE	ME
Education _{<i>it</i>}	23.96	-75.56	-75.56	-86.68	-84.75
	5.59	12.16	13.48	12.45	14.56
Democracy _{<i>it</i>}	110.94	-12.90	-12.90	-26.15	-3.63
	34.63	47.69	50.58	47.97	55.22
Oil-Producer _{<i>it</i>}	-26.89	—	—	—	—
	44.84	—	—	—	—
GDP _{<i>i,t-1</i>}	0.23	0.15	0.15	0.17	0.20
		0.02	0.02	0.02	
GDP _{<i>i,t-2</i>}				-0.12	
				0.02	
σ_α	0.14	—	—	—	309.10
Fixed effects		x	x	x	x
Random effects	x				x
<i>N</i>	113	113	113	113	113
<i>T</i>	328	328	328	228	328
observed <i>N</i> × <i>T</i>	2794	2794	2794	2741	2794
AIC	43376				42112
LM test <i>p</i> -value		<0.001	<0.001	0.131	

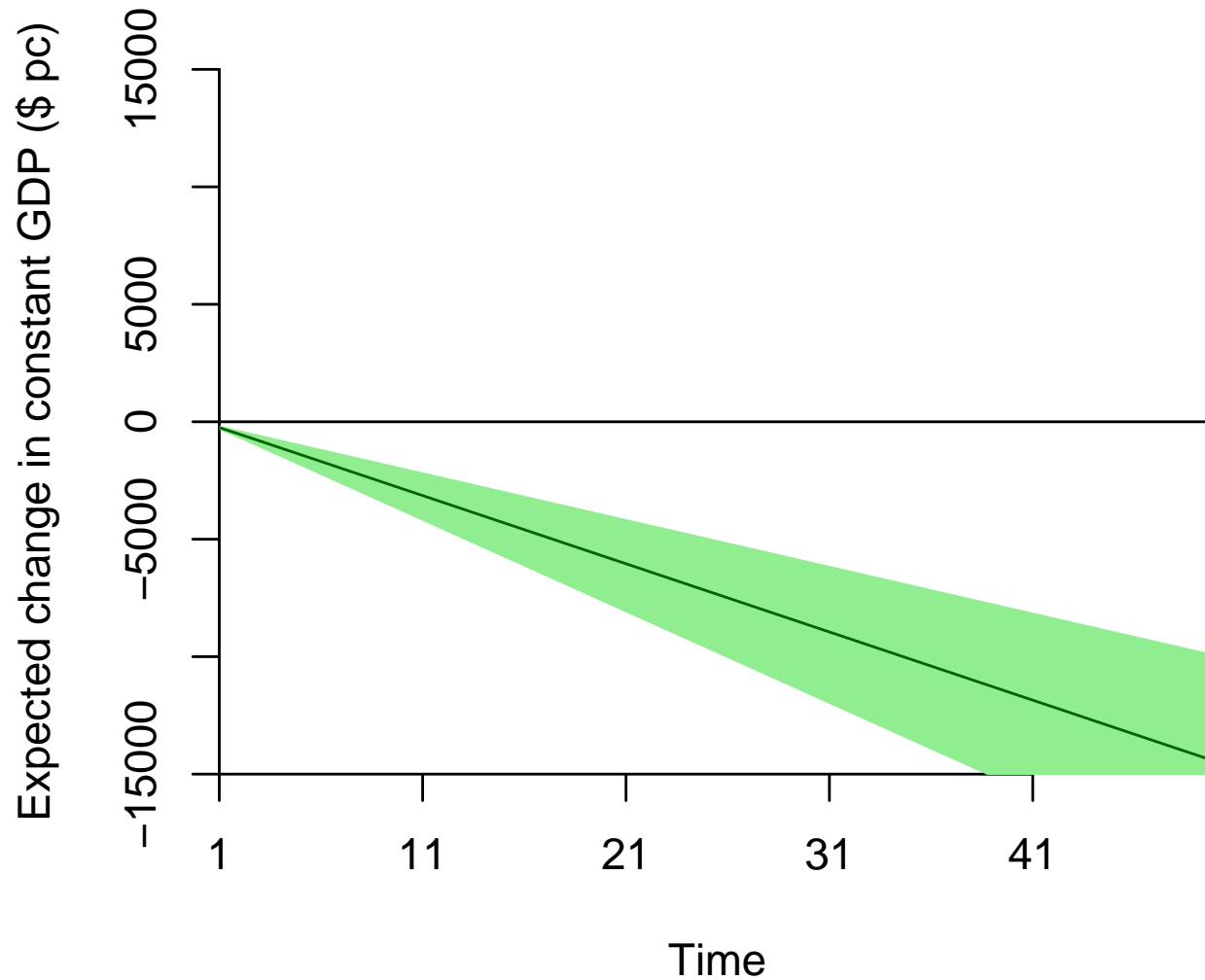
Fixed effects ARIMA(1,1,0)



Recall the fixed effects results. . .

These are uncorrected for panel heteroskedasticity or contemporaneous correlation

Fixed effects ARIMA(1,1,0), pcse



Panel correction usually makes little difference in long T small N contexts

But in short T , robust standard errors can be quite important. . .

Heteroskedastic and serial correlation consistent Var-Cov

In the PCSEs approach, the focus is on panel heteroskedasticity

It is assumed that serial correlation has been adequately modeled and purged

A reasonable check when we have a few dozen periods of data, though similar in most cases to either ordinary SEs or White SEs

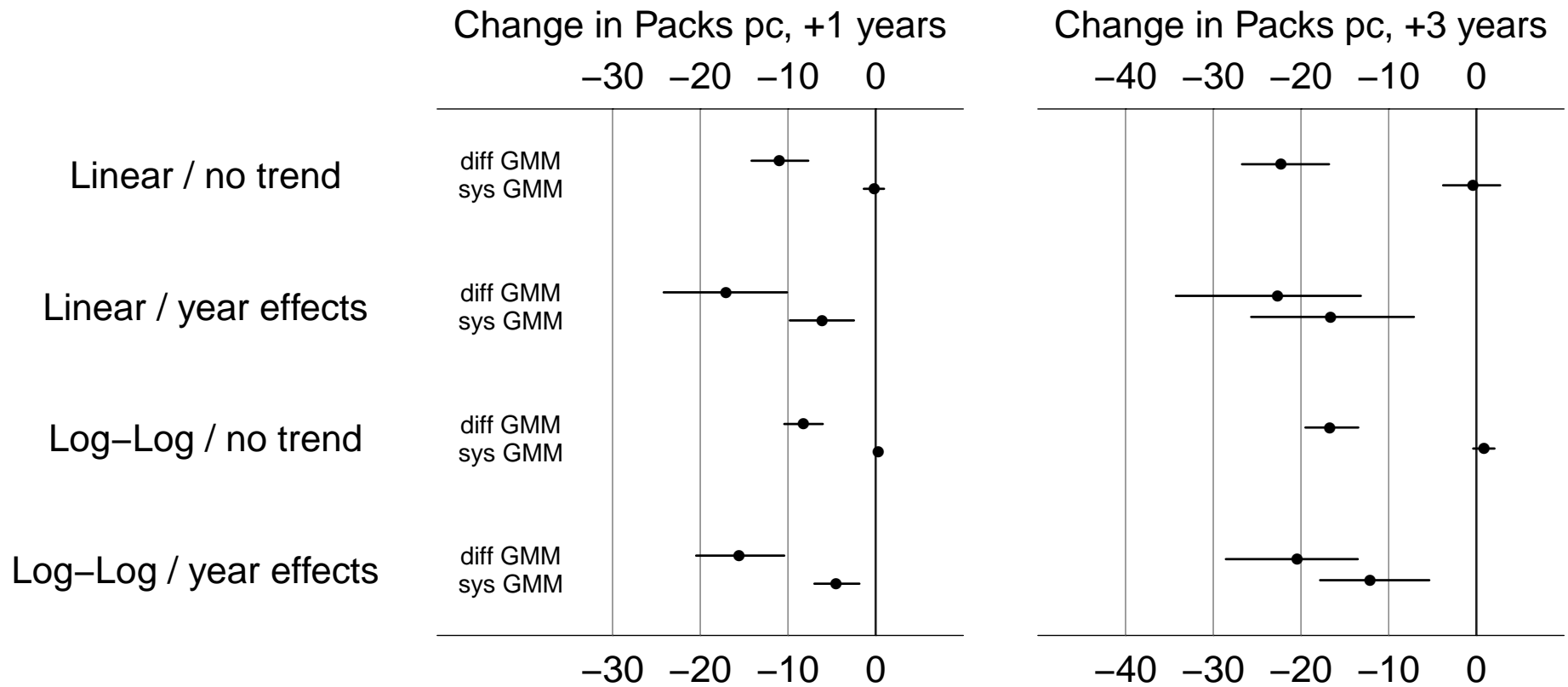
But what if we have a low T ? We might be more worried about residual serial correlation (and don't have practical access to ARMA diagnostics or fitting)

Now there is more need for a correction to the variance covariance that corrects for observed error correlation across units and across periods

Arellano (1987) provides a heteroskedastic and autocorrelation consistent variance-covariance matrix: in `plm`, `vcovHC()`

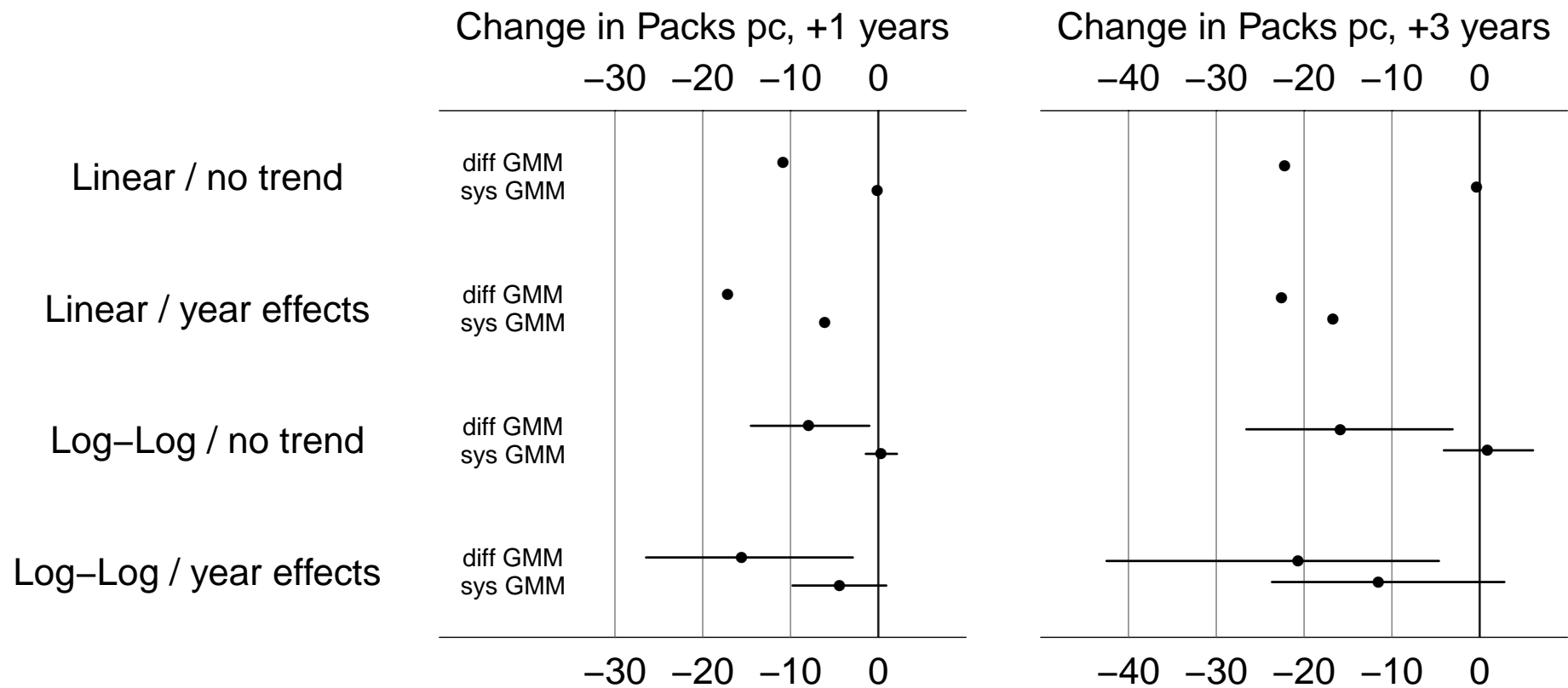
Use the same commands as above, but with `vcovHC()` instead of `vcovBK()`

Particularly important to correct with panel GMM estimators



Our prior results for cigarette taxes used the Arellano heteroskedastic and serial correlation consistent var-cov matrix

What would happen if we had used the ordinary, homoskedastic var-cov matrix to compute CIs?

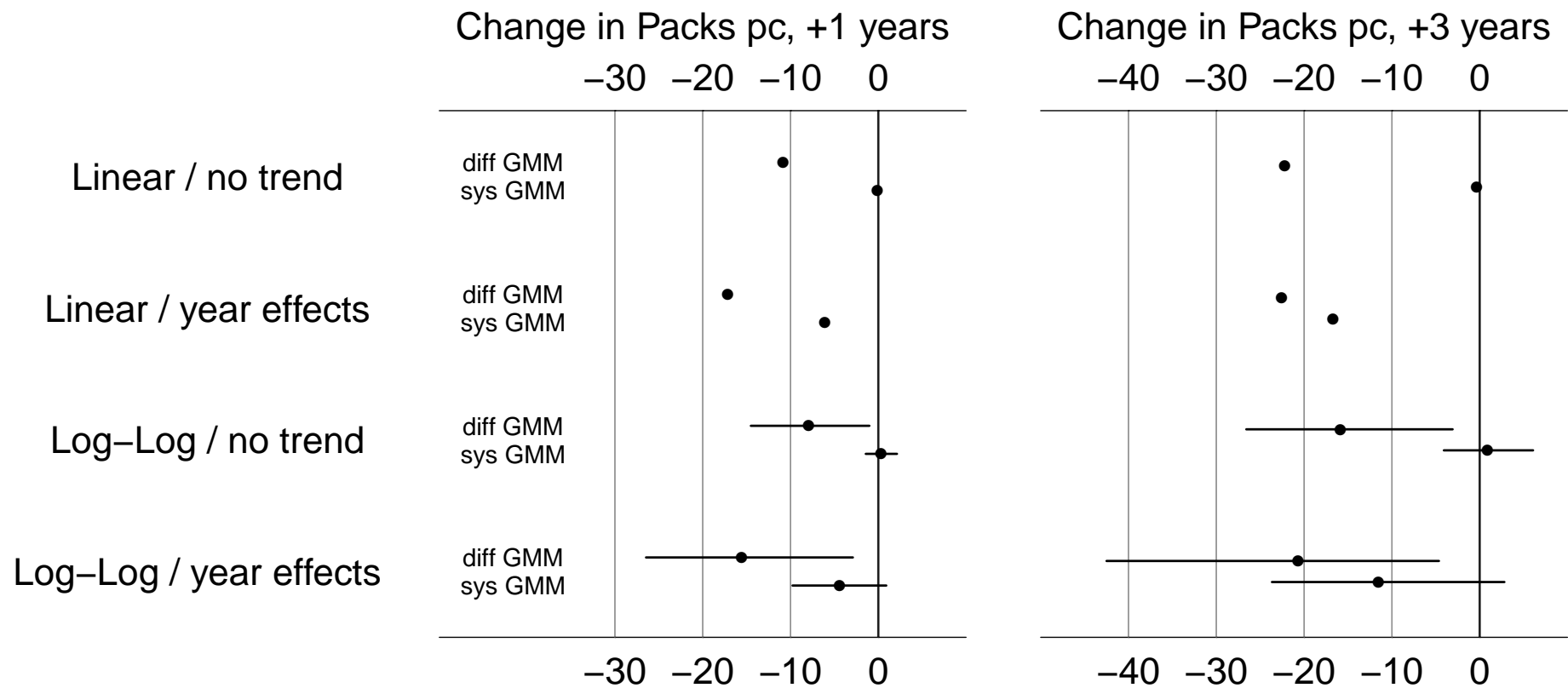


The effects sizes are mostly unchanged:
 adjustments to standard errors affect CIs, not point estimates

But the CIs are radically different under the traditional var-cov estimator

Far too small (invisible even!) for the misspecified linear models

And too large for the more correctly specified log-log models!



Just as panel GMM point estimate are sensitive to assumptions, so are the standard errors

Use caution, and prefer `vcovHC()` to `vcov()` in PGMM models

Be sure to check which var-cov matrix your functions are using: the default may be wrong!