"Climate change and U.S. agriculture: Accounting for multidimensional slope heterogeneity in panel data": Corrigendum

Michael Keane
Department of Economics, University of New South Wales

Timothy Neal
Department of Economics, University of New South Wales

Theorem 3 in our paper in *Quantitative Economics* Volume 11 Issue 4 contains errors in the expressions for both the asymptotic variance of $\hat{\boldsymbol{\beta}}_{MO}$ and for the variance estimator in equation (22). This note corrects those errors.

Importantly, the computer code used in the article, and posted online in the supplement accompanying the article, implements the correct standard error calculation. Thus, the standard errors reported in the paper, as well as any subsequent research that relied on the code, are unaffected by this correction.

Definitions. Given the assumed additive structure of $\boldsymbol{\beta}_{it} = \boldsymbol{\beta} + \boldsymbol{\lambda}_i + \boldsymbol{\theta}_t$, we can write without loss of generality that $\hat{\boldsymbol{\beta}}_{it} = \boldsymbol{\beta} + \boldsymbol{\lambda}_i + \boldsymbol{\theta}_t + \boldsymbol{\Omega}_i + \boldsymbol{\Phi}_t$, where $\boldsymbol{\Omega}_i$ and $\boldsymbol{\Phi}_t$ are error components that vary over i and t, respectively. These error components are only identified up to a mean shift. But their variances $\text{Var}(\boldsymbol{\Omega}_i)$ and $\text{Var}(\boldsymbol{\Phi}_t)$ are invariant to such shifts. We also define $\bar{\boldsymbol{\beta}} = \boldsymbol{\beta} + E(\boldsymbol{\lambda}_i) + E(\boldsymbol{\theta}_t)$. The the mean observation estimate of $\bar{\beta}$ is defined by $\hat{\boldsymbol{\beta}}_{MO} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\boldsymbol{\beta}}_{it}$.

Remark. It follows from equations (29) and (30) that we may write $\boldsymbol{\Omega}_i = \boldsymbol{Q}_{xu,T} - \boldsymbol{Q}_{xu,NT} + \sum_{\ell=0}^{L}(-1)^{\ell+1}(\boldsymbol{Q}_{xx,T}^{-1}\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\Lambda_{2,\ell} - \boldsymbol{Q}_{xx,NT}^{-1}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\Lambda_{2,\ell})$ and $\boldsymbol{\Phi}_t = \boldsymbol{Q}_{xu,N} + \sum_{\ell=0}^{L}(-1)^{\ell+1}(\boldsymbol{Q}_{xx,N}^{-1}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\Lambda_{1,\ell} - \boldsymbol{Q}_{xx,NT}^{-1}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\Lambda_{1,\ell})$, where $\boldsymbol{\Lambda}_{1,\ell} = \boldsymbol{Q}_{xx,T}^{-1}(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\boldsymbol{\Lambda}_{2,\ell-1})$ and $\boldsymbol{\Lambda}_{2,\ell} = \boldsymbol{Q}_{xx,N}^{-1}(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\boldsymbol{\Lambda}_{1,\ell-1})$ for $\ell > 0$, $\boldsymbol{\Lambda}_{1,0} = \boldsymbol{Q}_{xu,N}$ and $\boldsymbol{\Lambda}_{2,0} = \boldsymbol{Q}_{xu,T}$. This is one way to decompose the error into i and t components. But this decomposition is not unique as it is arbitrary how the term $\boldsymbol{Q}_{xu,NT}$ is allocated between $\boldsymbol{\Omega}_i$ and $\boldsymbol{\Phi}_t$. Nevertheless, Theorem 1 (consistency of $\hat{\beta}_{it}$) implies that $\boldsymbol{\Omega}_i \xrightarrow{p} 0$ and $\boldsymbol{\Phi}_i \xrightarrow{p} 0$ as $N, T \to \infty$ regardless of how $\boldsymbol{Q}_{xu,NT}$ is allocated, as $\boldsymbol{Q}_{xu,NT} \xrightarrow{p} 0$.

Theorem 3 (Asymptotic normality of $\hat{\boldsymbol{\beta}}_{MO}$). *For the model in* (12) *of the paper, with* **A.1–A.4** *and assuming* $\boldsymbol{\theta}_t$ *is stationary, if* $L \to \infty$ *and subsequently* $(N, T) \xrightarrow{j} \infty$ *such that*

Michael Keane: m.keane@unsw.edu.au
Timothy Neal: timothy.neal@unsw.edu.au

$N/T \to \chi$ and $0 < \chi < \infty$, then

$$\sqrt{NT}(\hat{\boldsymbol{\beta}}_{MO} - \bar{\boldsymbol{\beta}}) \xrightarrow{d} N(0, NT \cdot \boldsymbol{\Sigma}_{MO}),$$

where $\boldsymbol{\Sigma}_{MO} = \frac{\text{Var}(\boldsymbol{\lambda}_i)}{N} + \frac{\text{Var}(\boldsymbol{\theta}_t)}{T} + \frac{\bar{\text{Var}}(\boldsymbol{\Omega}_i)}{N} + \frac{\bar{\text{Var}}(\boldsymbol{\Phi}_t)}{T}$.

The asymptotic variance can be consistently estimated nonparametrically by

$$\hat{\boldsymbol{\Sigma}}_{MO} = \frac{1}{(NT-1)N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})' + \frac{1}{(NT-1)T} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})', \tag{1}$$

where $\hat{\boldsymbol{\beta}}_{\bar{i}} = \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{\beta}}_{it}$ and $\hat{\boldsymbol{\beta}}_{\bar{t}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\beta}}_{it}$.

Remark 1. The asymptotic variance of $\hat{\boldsymbol{\beta}}_{MO}$ arises from two components: (i) the variance of true slope heterogeneity in the unit ($\boldsymbol{\lambda}_i$) and time ($\boldsymbol{\theta}_t$) dimensions, and (ii) the sampling variance that arises from imprecision in estimating the i, t-level parameters ($\boldsymbol{\Omega}_i$ and $\boldsymbol{\Phi}_t$).

Remark 2. The MO-OLS algorithm gives estimates of $\boldsymbol{\beta}_{it}$ identical to a "brute force" OLS estimator (that interacts all regressors with a complete set of i and t dummies). Thus, $\hat{\boldsymbol{\beta}}_{MO}$ can (in principle) also be obtained by averaging the OLS estimates. The asymptotic distribution is identical, as are conditions required for consistency and asymptotic normality. Of course, the motivation for MO-OLS is that construction of OLS estimates and standard errors may be infeasible in large panels.

Remark 3. We require stationarity of $\boldsymbol{\theta}_t$ so that $E(\theta_t)$ and $\text{Var}(\theta_t)$ exist and can be consistently estimated as as $N, T \to \infty$.

Proof. Given that $\hat{\boldsymbol{\beta}}_{it} = \boldsymbol{\beta} + \boldsymbol{\lambda}_i + \boldsymbol{\theta}_t + \boldsymbol{\Omega}_i + \boldsymbol{\Phi}_t$ and $\bar{\boldsymbol{\beta}} = \boldsymbol{\beta} + E(\boldsymbol{\lambda}_i) + E(\boldsymbol{\theta}_t)$, we can write

$$(\hat{\boldsymbol{\beta}}_{MO} - \bar{\boldsymbol{\beta}}) = N^{-1} \sum_{i=1}^{N} \big(\boldsymbol{\lambda}_i - E(\boldsymbol{\lambda}_i)\big) + T^{-1} \sum_{t=1}^{T} \big(\boldsymbol{\theta}_t - E(\boldsymbol{\theta}_t)\big) + N^{-1} \sum_{i=1}^{N} \boldsymbol{\Omega}_i + T^{-1} \sum_{t=1}^{T} \boldsymbol{\Phi}_t$$

and thus we have

$$\sqrt{NT}(\hat{\boldsymbol{\beta}}_{MO} - \bar{\boldsymbol{\beta}}) = \sqrt{T}\left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \big(\boldsymbol{\lambda}_i - E(\boldsymbol{\lambda}_i)\big)\right] + \sqrt{T}\left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\Omega}_i\right] + \sqrt{N}\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \big(\boldsymbol{\theta}_t - E(\boldsymbol{\theta}_t)\big)\right] + \sqrt{N}\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{\Phi}_t\right].$$

Under our assumptions, we can apply the classical CLT to the $\boldsymbol{\lambda}_i$ and $\boldsymbol{\theta}_t$ terms and the Lindeberg–Feller CLT to the $\boldsymbol{\Omega}_i$ and $\boldsymbol{\Phi}_t$ terms to obtain

$$\sqrt{NT}(\hat{\boldsymbol{\beta}}_{MO} - \bar{\boldsymbol{\beta}}) \sim N\big(0, T \cdot \text{Var}(\boldsymbol{\lambda}_i) + N \cdot \text{Var}(\boldsymbol{\theta}_t) + T \cdot \bar{\text{Var}}(\boldsymbol{\Omega}_i) + N \cdot \bar{\text{Var}}(\boldsymbol{\Phi}_t)\big). \tag{2}$$

Thus $\sqrt{NT}(\hat{\boldsymbol{\beta}}_{MO} - \bar{\boldsymbol{\beta}}) \xrightarrow{d} N(0, NT \cdot \boldsymbol{\Sigma}_{MO})$ as stated, and we obtain the associated small sample approximation $\hat{\boldsymbol{\beta}}_{MO} \sim N(\bar{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{MO})$.

Now consider the nonparameteric estimate of Σ_{MO} that was proposed in (1):

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{MO} = {} & \frac{1}{(NT-1)N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})' \\ & + \frac{1}{(NT-1)T} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})'.\end{aligned}$$

Combining the fact that $\hat{\boldsymbol{\beta}}_{it} = \boldsymbol{\beta} + \boldsymbol{\lambda}_i + \boldsymbol{\theta}_t + \boldsymbol{\Omega}_i + \boldsymbol{\Phi}_t$ with the definitions $\hat{\boldsymbol{\beta}}_{\bar{i}} = \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{\beta}}_{it}$ and $\hat{\boldsymbol{\beta}}_{\bar{t}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\beta}}_{it}$ we have that

$$(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}}) = \left(\boldsymbol{\lambda}_i - \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\lambda}_i\right) + \left(\boldsymbol{\Omega}_i - \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\Omega}_i\right) \xrightarrow{p} \big(\boldsymbol{\lambda}_i - E(\boldsymbol{\lambda}_i)\big) + \boldsymbol{\Omega}_i,$$

$$(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}}) = \left(\boldsymbol{\theta}_t - \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\theta}_t\right) + \left(\boldsymbol{\Phi}_t - \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\Phi}_t\right) \xrightarrow{p} \big(\boldsymbol{\theta}_t - E(\boldsymbol{\theta}_t)\big) + \boldsymbol{\Phi}_t,$$

Therefore, $\frac{1}{(NT-1)N}\sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{t}})' \xrightarrow{p} \frac{\text{Var}(\boldsymbol{\lambda}_i)}{N} + \frac{\bar{\text{Var}}(\boldsymbol{\Omega}_i)}{N}$ and $\frac{1}{(NT-1)T} \times \sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})(\hat{\boldsymbol{\beta}}_{it} - \hat{\boldsymbol{\beta}}_{\bar{i}})' \xrightarrow{p} \frac{\text{Var}(\boldsymbol{\theta}_t)}{T} + \frac{\bar{\text{Var}}(\boldsymbol{\Phi}_t)}{T}$ so $\hat{\boldsymbol{\Sigma}}_{MO} \xrightarrow{p} \boldsymbol{\Sigma}_{MO}$ as required. □

Monte Carlo simulations

Here, we examine the finite sample performance of the estimate of the MO-OLS standard error proposed in (1). The performance is assessed by comparing the mean of the estimate against the empirical standard deviation of the estimator across Monte Carlo replications in a range of scenario environments. We consider environments with both slope/intercept homogeneity and heterogeneity. In each of the five scenarios, MO-OLS is also compared against Mean Group OLS and its proposed standard error.

We consider a data generating process where the dependent variable is generated by

$$y_{it} = c_{it} + \beta_{it} x_{it} + \epsilon_{it}, \tag{3}$$

where $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The distribution of the error term is $\epsilon_{it} \sim N(0, 16)$. The fixed effects in the intercept are generated by $c_{it} = 1 + k_i + f_t$, while the slope parameter is generated by $\beta_{it} = 1 + \lambda_i + \theta_t$. The distribution of the heterogeneity terms in both the slope and the intercept will differ by scenario (see below).

Table 1. Monte Carlo scenarios.

	k_i	f_t	λ_i	θ_t	α
Scenario 1	0	0	0	0	0
Scenario 2	N(0,0.25)	0	N(0,0.25)	0	0
Scenario 3	N(0,1)	N(0,1)	N(0,0.25)	N(0,0.25)	0
Scenario 4	N(0,1)	N(0,1)	N(0,1)	N(0,1)	0
Scenario 5	N(0,1)	N(0,1)	N(0,1)	N(0,1)	1

The regressor x is generated by

$$x_{it} = c_{it} + \alpha\beta_{it} + v_i + u_t + e_{it}. \tag{4}$$

The inclusion of c_{it} generates correlation between the regressor and intercept (i.e., a fixed effects scenario), whenever the intercept varies across i and/or t. The parameter α governs the correlation between the slope coefficients β_{it} and the regressor (i.e., fixed effects in slopes). We set $v_i \sim u_t \sim N(0, 0.25)$ and $e_{it} \sim N(0, 1)$.

We consider five scenarios that vary the distribution of the heterogeneity and the degree of correlation between x_{it} and the slope coefficient. These are described in Table 1.

Table 2 contains the results of this Monte Carlo analysis. The columns on the left report results for a sample size of $N, T = 50$, while the columns on the right increase the sample size to $N, T = 150$. The mean of the coefficient, the standard deviation of the coefficient, and the mean estimated standard error are reported across replications for each scenario. The true average value of β_{it} is 1 in all scenarios.

Scenario 1 considers a case with no true intercept or slope heterogeneity. Bias is negligible for both estimators in this case. MO-OLS is less efficient than MG-OLS in this case, as it adds extraneous parameters. The estimated standard error for MO-OLS is also slightly conservative, with the bias decreasing as the sample size increases.[1]

Scenario 2 introduces intercept and slope heterogeneity across i alone, which is an environment suited to the MG-OLS model. In this case, MO-OLS is only slightly less efficient than MG-OLS, and the estimated standard error is slightly conservative when $N, T = 50$.

Scenario 3 adds heterogeneity across t as well. Thus we have multidimensional slope and intercept heterogeneity, the environment to which MO-OLS is suited. The MO-OLS estimates show no evidence of bias, and the estimated MO-OLS standard errors are very accurate for both sample sizes. In contrast, the MG-OLS estimator is upward biased due to correlation between the regressor and the time-varying intercept, and its estimated standard error badly underestimates the true variance of the estimator.

Scenario 4 increases the variance of the slope heterogeneity, which has the effect of increasing the empirical standard deviation of both estimators. Again, the MO-OLS standard error proposed here appears to be very accurate. The MG-OLS standard error badly underestimates its empirical standard deviation.

[1]When $N, T = 500$, the empirical standard deviation of MO-OLS is 0.0084 while the mean estimated standard error is 0.0101. This demonstrates how the upward bias in the standard error decreases with sample size.

Table 2. Monte Carlo simulations of the MO-OLS standard error.

	N,T = 50			N,T = 150		
	Mean	Std. Dev.	Est. S.E.	Mean	Std. Dev.	Est. S.E.
Scenario 1:						
MO-OLS	1.003	0.086	0.106	1.000	0.026	0.034
MG-OLS	1.003	0.075	0.074	1.000	0.023	0.024
Scenario 2:						
MO-OLS	0.996	0.110	0.123	1.001	0.049	0.052
MG-OLS	0.996	0.102	0.102	1.001	0.047	0.047
Scenario 3:						
MO-OLS	1.004	0.129	0.126	0.999	0.066	0.066
MG-OLS	1.451	0.221	0.094	1.450	0.126	0.047
Scenario 4:						
MO-OLS	1.009	0.212	0.213	0.999	0.117	0.118
MG-OLS	1.458	0.423	0.161	1.456	0.243	0.085
Scenario 5:						
MO-OLS	1.010	0.213	0.208	1.003	0.120	0.117
MG-OLS	3.170	0.434	0.200	3.156	0.256	0.112

Note: Monte Carlo replications were set to 1000 for this study.

Finally, Scenario 5 adds correlation between x_{it} and the slope heterogeneity by setting $\alpha = 1$. Despite this complication, MO-OLS performs just as well as in Scenario 4. In contrast, the bias of MG-OLS increases very sharply.

In summary, the simulations show that in environments that conform to the multidimensional slope heterogeneity framework, the MO-OLS estimator shows little evidence of bias, and our estimate of the standard error works well even in small samples.

Manuscript received 16 February, 2021; final version accepted 17 February, 2021.