The Stata Journal (2007)
7, Number 3, pp. 281–312

Robust standard errors for panel regressions with cross-sectional dependence

Daniel Hoechle
Department of Finance
University of Basel
Basel, Switzerland
daniel.hoechle@unibas.ch

Abstract. I present a new Stata program, `xtscc`, that estimates pooled ordinary least-squares/weighted least-squares regression and fixed-effects (within) regression models with Driscoll and Kraay (*Review of Economics and Statistics* 80: 549–560) standard errors. By running Monte Carlo simulations, I compare the finite-sample properties of the cross-sectional dependence–consistent Driscoll–Kraay estimator with the properties of other, more commonly used covariance matrix estimators that do not account for cross-sectional dependence. The results indicate that Driscoll–Kraay standard errors are well calibrated when cross-sectional dependence is present. However, erroneously ignoring cross-sectional correlation in the estimation of panel models can lead to severely biased statistical results. I illustrate the `xtscc` program by considering an application from empirical finance. Thereby, I also propose a Hausman-type test for fixed effects that is robust to general forms of cross-sectional and temporal dependence.

Keywords: st0128, xtscc, robust standard errors, nonparametric covariance estimation

1 Introduction

In social sciences, and particularly in economics, analyzing large-scale microeconometric panel datasets has become common. Compared with purely cross-sectional data, panels are attractive since they often contain far more information than single cross-sections and thus allow for an increased precision in estimation. Unfortunately, however, actual information of microeconometric panels is often overstated since microeconometric data are likely to exhibit all sorts of cross-sectional and temporal dependencies. In the words of Cameron and Trivedi (2005, 702), "NT correlated observations have less information than NT independent observations." Therefore, erroneously ignoring possible correlation of regression disturbances over time and between subjects can lead to biased statistical inference. To ensure validity of the statistical results, most recent studies that include a regression on panel data therefore adjust the standard errors of the coefficient estimates for possible dependence in the residuals. However, according to Petersen (2007), many recently published articles in leading finance journals still fail to adjust the standard errors appropriately. Furthermore, although most empirical studies now provide standard error estimates that are heteroskedasticity- and autocorrelation consistent, cross-sectional or "spatial" dependence is still largely ignored.

 st0128

However, assuming that the disturbances of a panel model are cross-sectionally independent is often inappropriate. Although it might be difficult to convincingly argue why country- or state-level data should be spatially uncorrelated, many studies on social learning, herd behavior, and neighborhood effects clearly indicate that microeconometric panel datasets are likely to exhibit complex patterns of mutual dependence between the cross-sectional units (e.g., individuals or firms).[1] Furthermore, because social norms and psychological behavior patterns typically enter panel regressions as unobservable common factors, complex forms of spatial and temporal dependence may even arise when the cross-sectional units have been randomly and independently sampled.

Provided that the unobservable common factors are uncorrelated with the explanatory variables, the coefficient estimates from standard panel estimators—e.g., fixed-effects (FE) estimator, random-effects (RE) estimator, or pooled ordinary least-squares (OLS) estimation—are still consistent (but inefficient). However, standard error estimates of commonly applied covariance matrix estimation techniques—e.g., OLS, White, and Rogers or clustered standard errors—are biased, and hence statistical inference based on such standard errors is invalid. Fortunately, Driscoll and Kraay (1998) propose a nonparametric covariance matrix estimator that produces heteroskedasticity- and autocorrelation-consistent standard errors that are robust to general forms of spatial and temporal dependence.

Stata has long provided the option to estimate standard errors that are robust to certain violations of the underlying econometric model. This article aims to contribute to this tradition by providing a Stata implementation of Driscoll and Kraay's (1998) covariance matrix estimator for use with pooled OLS estimation and FE regression. In contrast to Driscoll and Kraay's original contribution that considers only balanced panels, I adjust their estimator for use with unbalanced panels and use Monte Carlo simulations to investigate the adjusted estimator's finite sample performance in case of medium- and large-scale (microeconometric) panels. Consistent with Driscoll and Kraay's original finding for small balanced panels, the Monte Carlo experiments reveal that erroneously ignoring spatial correlation in panel regressions typically leads to overly optimistic (anticonservative) standard error estimates, irrespective of whether a panel is balanced. Although Driscoll and Kraay standard errors tend also to be slightly optimistic, their small-sample properties are considerably better than those of the alternative covariance estimators when cross-sectional dependence is present.

The rest of this article is organized as follows. In the next section, I motivate why Driscoll and Kraay's covariance matrix estimator is a valuable supplement to Stata's existing capabilities. Section 3 describes the `xtscc` program that produces Driscoll and Kraay standard errors for coefficients estimated by pooled OLS/weighted least-squares (WLS) regression and FE (within) regression. Section 4 provides the formulas as they are implemented in the `xtscc` program. In section 5, I present the setup and the results of Monte Carlo experiments that compare the finite-sample properties of the Driscoll–Kraay estimator with those of other more commonly used covariance matrix

1. See Trueman (1994), Welch (2000), Feng and Seasholes (2004), and the survey article by Hirshleifer and Teoh (2003).

estimation techniques when the cross-sectional units are spatially dependent. Section 6 considers an empirical example from financial economics and demonstrates practical use of the `xtscc` program. Furthermore, by extending the line of arguments proposed by Wooldridge (2002, 290), I show how one can apply the `xtscc` program to perform a Hausman test for FE that is robust to general forms of cross-sectional and temporal dependence. Section 7 concludes.

2 Motivation for the Driscoll–Kraay estimator

To ensure valid statistical inference when some of the underlying regression model's assumptions are violated, relying on robust standard errors is common. Probably the most popular of these alternative covariance matrix estimators has been developed by Huber (1967), Eicker (1967), and White (1980). Provided that the residuals are independently distributed, standard errors that are obtained by aid of this estimator are consistent even if the residuals are heteroskedastic. In Stata 9, heteroskedasticity-consistent or "White" standard errors are obtained by choosing option `vce(robust)`, which is available for most estimation commands.

Extending the work of White (1980, 1984) and Huber (1967), Arellano (1987), Froot (1989), and Rogers (1993) show that it is possible to somewhat relax the assumption of independently distributed residuals. Their generalized estimator produces consistent standard errors if the residuals are correlated within but uncorrelated between clusters. Stata's estimation commands with option `vce(robust)` also contain a `cluster()` option, that allows the computation of so-called Rogers or clustered standard errors.[2]

Another approach to obtain heteroskedasticity- and autocorrelation (up to some lag)-consistent standard errors was developed by Newey and West (1987). Their generalized method of moments–based covariance matrix estimator is an extension of White's estimator, as it can be shown that the Newey–West estimator with lag length zero is identical to the White estimator. Although Newey–West standard errors have initially been proposed for use with time-series data only, panel versions are available. In Stata, Newey–West standard errors for panel datasets are obtained by choosing option `force` of the `newey` command.

Although all these techniques of estimating the covariance matrix are robust to certain violations of the regression model assumptions, they do not consider cross-sectional correlation. However, because of social norms and psychological behavior patterns, spatial dependence can be a problematic feature of any microeconometric panel dataset even if the cross-sectional units (e.g., individuals or firms) have been randomly selected. Therefore, assuming that the residuals of a panel model are correlated within but uncorrelated between groups of individuals often imposes an artificial and inappropriate constraint on empirical models. Assuming that the residuals are correlated both within groups as well as between groups would often be more natural.

2. If the panel identifier (e.g., individuals, firms, or countries) is the `cluster()` variable, then Rogers standard errors are heteroskedasticity and autocorrelation consistent.

In an early attempt to account for heteroskedasticity as well as for temporal and spatial dependence in the residuals of time-series cross-section models, Parks (1967) proposes a feasible generalized least-squares (FGLS)–based algorithm that Kmenta (1986) made popular. Unfortunately, however, the Parks–Kmenta method, which is implemented in Stata's `xtgls` command with option `panels(correlated)`, is typically inappropriate for use with medium- and large-scale microeconometric panels for at least two reasons. First, this method is infeasible if the panel's time dimension, T, is smaller than its cross-sectional dimension, N, which is almost always the case for microeconometric panels.[3] Second, Beck and Katz (1995) show that the Parks–Kmenta method tends to produce unacceptably small standard error estimates.

To mitigate the problems of the Parks–Kmenta method, Beck and Katz (1995) suggest relying on OLS coefficient estimates with panel-corrected standard errors (PCSEs). In Stata, pooled OLS regressions with PCSEs can be estimated with the `xtpcse` command. Beck and Katz (1995) convincingly demonstrate that their large-T asymptotics–based standard errors, which correct for contemporaneous correlation between the subjects, perform well in small panels. Nevertheless, it has to be expected that the finite-sample properties of the PCSE estimator are rather poor when the panel's cross-sectional dimension N is large compared to the time dimension T. Beck and Katz's (1995) PCSE method estimates the full $N \times N$ cross-sectional covariance matrix, and this estimate will be rather imprecise if the ratio T/N is small.

Therefore, when working with medium- and large-scale microeconometric panels, it seems tempting to implement parametric corrections for spatial dependence. However, with large-N asymptotics, such corrections require strong assumptions about the corrections' form because the number of cross-sectional correlations grows with rate N^2, whereas the number of observations increases only by rate N. To maintain the model's feasibility, empirical researchers therefore often presume that the cross-sectional correlations are the same for every pair of cross-sectional units such that the introduction of time dummies purges the spatial dependence. However, constraining the cross-sectional correlation matrix is prone to misspecification, and hence implementing nonparametric corrections for the cross-sectional dependence is desirable.

By relying on large-T asymptotics, Driscoll and Kraay (1998) demonstrate that the standard nonparametric time-series covariance matrix estimator can be modified such that it is robust to general forms of cross-sectional as well as temporal dependence. Driscoll and Kraay's approach loosely applies a Newey-West–type correction to the sequence of cross-sectional averages of the moment conditions. Adjusting the standard error estimates in this way guarantees that the covariance matrix estimator is consistent, independently of the cross-sectional dimension N (i.e., also for $N \rightarrow \infty$). Therefore, Driscoll and Kraay's approach eliminates the deficiencies of other large-T–consistent covariance matrix estimators such as the Parks–Kmenta and the PCSE approach, which typically become inappropriate when the cross-sectional dimension N of a microeconometric panel gets large.

3. The Parks–Kmenta and other large-T asymptotics–based covariance matrix estimators become infeasible with large N relative to T because obtaining a nonsingular estimate of the $N \times N$ matrix of cross-sectional covariances when $T < N$ is impossible. See Beck and Katz (1995) for details.

Table 1 gives an overview for selected Stata commands and options that produce robust standard error estimates for linear panel models.

Table 1: Selection of Stata commands and options that produce robust standard error estimates for linear panel models

Command	Option	SE estimates are robust to disturbances that are	Notes
`reg`, `xtreg`	`vce(robust)`	heteroskedastic	
`reg`, `xtreg`	`cluster()`	heteroskedastic and autocorrelated	
`xtregar`		autocorrelated with AR(1)[a]	
`newey`		heteroskedastic and autocorrelated of type MA(q)[b]	
`xtgls`	`panels()`, `corr()`	heteroskedastic, contemporaneously cross-sectionally correlated, and autocorrelated of type AR(1)	$N < T$ required for feasibility; tends to produce optimistic SE estimates
`xtpcse`	`correlation()`	heteroskedastic, contemporaneously cross-sectionally correlated, and autocorrelated of type AR(1)	large-scale panel regressions with `xtpcse` take a lot of time
`xtscc`		heteroskedastic, autocorrelated with MA(q), and cross-sectionally dependent	

[a] AR(1) refers to first-order autoregression.
[b] MA(q) denotes autocorrelation of the moving average type with lag length q.

3 The xtscc program

3.1 Syntax

`xtscc` *depvar* [*varlist*] [*if*] [*in*] [*weight*] [`, lag(`#`) fe pooled level(`#`)`]

3.2 Description

`xtscc` produces Driscoll and Kraay (1998) standard errors for coefficients estimated by pooled OLS/WLS and FE (within) regression. *depvar* is the dependent variable and *varlist* is an optional list of explanatory variables.

The error structure is assumed to be heteroskedastic, autocorrelated up to some lag and possibly correlated between the groups (panels). These standard errors are robust to general forms of cross-sectional (spatial) and temporal dependence when the time dimension becomes large. Because this nonparametric technique of estimating standard errors places no restrictions on the limiting behavior of the number of panels, the size of the cross-sectional dimension in finite samples does not constitute a constraint on feasibility—even if the number of panels is much larger than T. Nevertheless, because the estimator is based on an asymptotic theory, one should be somewhat cautious with applying this estimator to panels that contain a large cross-section but only a short time dimension.

The `xtscc` program is suitable for use with both balanced and unbalanced panels. Furthermore, it can handle missing values.

3.3 Options

`lag(#)` specifies the maximum lag to be considered in the autocorrelation structure. By default, a lag length of $m(T) = \text{floor}[4(T/100)^{2/9}]$ is assumed (see sec. 4.4).

`fe` performs FE (within) regression with Driscoll and Kraay standard errors. These standard errors are heteroskedasticity consistent and robust to general forms of cross-sectional (spatial) and temporal dependence when the time dimension becomes large. If the residuals are assumed to be heteroskedastic only, use `xtreg, fe vce(robust)`. When the standard errors should be heteroskedasticity- and autocorrelation consistent, use `xtreg, fe cluster()`. Weights are not allowed if option `fe` is chosen.

`pooled` is the default option for `xtscc`. It performs pooled OLS/WLS regression with Driscoll and Kraay standard errors. These standard errors are heteroskedasticity consistent and robust to general forms of cross-sectional (spatial) and temporal dependence when the time dimension becomes large. If the residuals are assumed to be heteroskedastic only, use `regress, vce(robust)`. When the standard errors should be heteroskedasticity- and autocorrelation consistent, use either `regress, cluster()` or `newey, lag(#) force`. Analytic weights are allowed for use with option `pooled`; see [U] **11.1.6 weight** and [U] **20.16 Weighted estimation**.

`level(#)` specifies the confidence level, as a percentage, for confidence intervals. The default is `level(95)` or as set by `set level`; see [U] **23.5 Specifying the width of confidence intervals**.

3.4 Remarks

The main procedure of `xtscc` is implemented in Mata and is based in parts on Driscoll and Kraay's original GAUSS program, which can be downloaded from John Driscoll's home page (http://www.johncdriscoll.net).

The `xtscc` program includes several functions from Ben Jann's `moremata` package.

4 Panel models with Driscoll and Kraay standard errors

Although Driscoll and Kraay's (1998) covariance matrix estimator is perfectly general and by no means limited to the use with linear panel models, I restrict the presentation of the estimator to the case implemented in the `xtscc` program, i.e., to linear regression. In contrast to Driscoll and Kraay's original formulation, the estimator below is adjusted for use with both balanced and unbalanced panel datasets.[4]

When option `fe` is chosen or if analytic weights are provided along with the `pooled` option, the `xtscc` program first transforms the variables in a way that allows for estimation by OLS. For FE estimation, the corresponding transform is the within transformation, and for WLS estimation the transform applied is the WLS transform. I describe both transforms below.

4.1 Driscoll and Kraay standard errors for pooled OLS estimation

Consider the linear regression model

$$y_{it} = \mathbf{x}'_{it}\theta + \varepsilon_{it}, \qquad i = 1, \ldots, N, \quad t = 1, \ldots, T$$

where the dependent variable y_{it} is a scalar, $\mathbf{x}_{it}$ is a $(K+1) \times 1$ vector of independent variables whose first element is 1, and θ is a $(K+1) \times 1$ vector of unknown coefficients. i denotes the cross-sectional units (individuals) and t denotes time. Stacking all observations as follows is common:

$$\mathbf{y} = [y_{1t_{11}} \ \ldots \ y_{1T_1} \ y_{2t_{21}} \ \ldots \ y_{NT_N}]' \quad \text{and} \quad \mathbf{X} = [\mathbf{x}_{1t_{11}} \ \ldots \ \mathbf{x}_{1T_1} \ \mathbf{x}_{2t_{21}} \ \ldots \ \mathbf{x}_{NT_N}]'$$

This formula allows the panel to be unbalanced since for individual i only a subset $t_{i1}, \ \ldots \ , T_i$ with $1 \leq t_{i1} \leq T_i \leq T$ of all T observations may be available. It is assumed that the regressors $\mathbf{x}_{it}$ are uncorrelated with the scalar disturbance term ε_{is} for all s, t (strong exogeneity). However, the disturbances ε_{it} themselves are allowed to be autocorrelated, heteroskedastic, and cross-sectionally dependent. Under these presumptions θ can consistently be estimated by OLS regression, which yields

$$\widehat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Driscoll and Kraay standard errors for the coefficient estimates are then obtained as the square roots of the diagonal elements of the asymptotic (robust) covariance matrix,

$$V(\widehat{\theta}) = (\mathbf{X}'\mathbf{X})^{-1}\widehat{\mathbf{S}}_T(\mathbf{X}'\mathbf{X})^{-1}$$

where $\widehat{\mathbf{S}}_T$ is defined as in Newey and West (1987):

$$\widehat{\mathbf{S}}_T = \widehat{\mathbf{\Omega}}_0 + \sum_{j=1}^{m(T)} w(j,m)[\widehat{\mathbf{\Omega}}_j + \widehat{\mathbf{\Omega}}'_j] \tag{1}$$

4. For details on the regularity conditions under which Driscoll and Kraay standard errors are consistent, see Driscoll and Kraay (1998) and Newey and West (1987).

In (1), $m(T)$ denotes the lag length up to which the residuals may be autocorrelated and the modified Bartlett weights,

$$w(j,m) = 1 - j/\{m(T)+1\}$$

ensure positive semidefiniteness of $\widehat{\mathbf{S}}_T$ and smooth the sample autocovariance function such that higher-order lags receive less weight. The $(K+1)\times(K+1)$ matrix $\widehat{\boldsymbol{\Omega}}_j$ is defined as

$$\widehat{\Omega}_j = \sum_{t=j+1}^{T} \mathbf{h}_t(\widehat{\theta})\mathbf{h}_{t-j}(\widehat{\theta})' \quad \text{with} \quad \mathbf{h}_t(\widehat{\theta}) = \sum_{i=1}^{N(t)} \mathbf{h}_{it}(\widehat{\theta}) \tag{2}$$

In (2), the sum of the individual time t moment conditions $\mathbf{h}_{it}(\widehat{\theta})$ runs from 1 to $N(t)$, where N is allowed to vary with t. This small adjustment to Driscoll and Kraay's (1998) original estimator suffices to make their estimator ready for use with unbalanced panels. For pooled OLS estimation, the individual orthogonality conditions $\mathbf{h}_{it}(\widehat{\theta})$ in (2) are the $(K+1)\times 1$ dimensional moment conditions of the linear regression model; i.e.,

$$\mathbf{h}_{it}(\widehat{\theta}) = \mathbf{x}_{it}\widehat{\varepsilon}_{it} = \mathbf{x}_{it}(y_{it} - \mathbf{x}_{it}'\widehat{\theta})$$

From (1) and (2), it follows that Driscoll and Kraay's covariance matrix estimator equals the heteroskedasticity- and autocorrelation-consistent covariance matrix estimator of Newey and West (1987) applied to the time series of cross-sectional averages of the $\mathbf{h}_{it}(\widehat{\theta})$.[5] By relying on cross-sectional averages, standard errors estimated by this approach are consistent independently of the panel's cross-sectional dimension N. Driscoll and Kraay (1998) show that this consistency result holds even for the limiting case where $N \to \infty$. Furthermore, estimating the covariance matrix with this approach yields standard errors that are robust to general forms of cross-sectional and temporal dependence.

4.2 Fixed-effects regression with Driscoll and Kraay standard errors

The `xtscc` program's option `fe` estimates FE (within) regression models with Driscoll and Kraay standard errors. The respective FE estimator is implemented in two steps. In the first step all model variables $z_{it} \in \{y_{it}, \mathbf{x}_{it}\}$ are within-transformed as follows (see [XT] **xtreg**):

$$\widetilde{z}_{it} = z_{it} - \overline{z}_i + \overline{\overline{z}}, \quad \text{where} \quad \overline{z}_i = T_i^{-1}\sum_{t=t_{i1}}^{T_i} z_{it} \quad \text{and} \quad \overline{\overline{z}} = \left(\sum T_i\right)^{-1}\sum_i\sum_t z_{it}$$

Since we recognize that the within-estimator corresponds to the OLS estimator of

$$\widetilde{y}_{it} = \widetilde{\mathbf{x}}_{it}'\theta + \widetilde{\varepsilon}_{it} \tag{3}$$

the second step then estimates the transformed regression model in (3) by pooled OLS estimation with Driscoll–Kraay standard errors (see sec. 4.1).

5. Although this representation of Driscoll and Kraay's covariance matrix estimator emphasizes that the estimator belongs to the robust group of covariance matrix estimators, the exposition in Driscoll and Kraay (1998) clarifies that their estimator indeed applies a Newey-West–type correction to the sequence of cross-sectional averages of the moment conditions.

4.3 WLS regression with Driscoll and Kraay standard errors

As for the FE estimator, WLS regression with Driscoll and Kraay standard errors is also performed in two steps. The first step applies the WLS transform $\widetilde{z}_{it} = \sqrt{w_{it}} z_{it}$ to all model variables including the constant (i.e., $z_{it} \in \{y_{it}, \mathbf{x}_{it}\}$), and the second step then estimates the transformed model in (4) by pooled OLS estimation (see [R] **regress** and Verbeek [2004, 84]) with Driscoll and Kraay standard errors:

$$\widetilde{y}_{it} = \widetilde{\mathbf{x}}'_{it}\theta + \widetilde{\varepsilon}_{it} \tag{4}$$

4.4 Note on lag length selection

In (1), $m(T)$ denotes the lag length up to which the residuals may be autocorrelated. Strictly, by constraining the residuals to be autocorrelated up to some lag $m(T)$, only moving-average (MA) processes of the residuals are considered. Fortunately, this is not necessarily a problem since autoregressive (AR) processes normally can be well approximated by finite-order MA processes. However, for using modified Bartlett weights (see above), Newey and West (1987) have shown that their estimator is consistent if the number of lags included in the estimation of the covariance matrix, $m(T)$, increases with T but at a rate slower than $T^{1/4}$. Therefore, selecting an $m(T)$ that is close to the maximum lag length (i.e., $m(T) = T - 1$) is not advisable, even if one is convinced that the residuals follow an AR process.

To assist the researcher by choosing $m(T)$, Andrews (1991), Newey and West (1994), and others have developed what are known as *plug-in* estimators. Plug-in estimators are automized procedures that deliver the optimum number of lags according to an asymptotic mean squared error criterion. Hence, the lag length $m(T)$ that is selected by a plug-in estimator depends on the data at hand. Unfortunately, however, no such procedure is available in official Stata right now.

Therefore, the `xtscc` program uses a simple rule for selecting $m(T)$ when no `lag(#)` option is specified. The heuristic applied is taken from the first step of Newey and West's (1994) plug-in procedure and sets

$$m(T) = \text{floor}[4(T/100)^{2/9}]$$

However, choosing the lag length like this is not necessarily optimal because this choice is essentially independent from the underlying data. In fact, this simple rule of selecting the lag length tends to choose an $m(T)$ that might often be too small.

5 Monte Carlo evidence

By theory, the coefficient estimate of a 95% confidence interval should contain the true coefficient value in 95 of 100 cases. The coverage rate measures how well this assumption is met in practice. For example, if an econometric estimator is perfectly calibrated, then the coverage rate of the 95% confidence interval should be close to the nominal value, i.e., close to 0.95. However, when coverage rates and hence standard error estimates are biased, statistical tests (such as the t test) lose their validity. Therefore, coverage rates are an important measure for assessing whether statistical inference is valid under certain circumstances.

Although the coverage rates of OLS standard errors are perfectly calibrated when all OLS assumptions are met, we know little about how well standard error estimates perform when the residuals and the explanatory variables of large-scale microeconometric panels are cross-sectionally and temporally dependent. Although there are both studies that address the consequences of spatial and temporal dependence explicitly[6] and studies that consider medium- and large-scale panels,[7] I know of no analysis that investigates the small-sample properties of standard error estimates for large-scale panel datasets with spatially dependent cross-sections. But as others have argued before, assuming that the subjects (e.g., individuals or firms) of medium- and large-scale microeconometric panels are independent of each other might often be equivocal in practice because of things like social norms, herd behavior, and neighborhood effects.

The Monte Carlo simulations presented here consider both large-scale panels and intricate forms of cross-sectional and temporal dependence. By comparing the coverage rates from several techniques of estimating (robust) standard errors for linear panel models, I can replicate and extend Driscoll and Kraay's original finding that cross-sectional dependence can lead to severely biased standard error estimates if it is not accounted for appropriately. Even though coverage rates of Driscoll and Kraay standard errors are typically below their nominal value, Driscoll and Kraay standard errors have considerably better small-sample properties than those of commonly applied alternative techniques for estimating standard errors when cross-sectional dependence is present. This result holds, irrespective of whether a panel dataset is balanced.

5.1 Specification

Without loss of generality, the Monte Carlo experiments are based on fitting the following bivariate regression model:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{5}$$

In (5), it is assumed that the independent variable x_{it} is uncorrelated with the disturbance term ε_{it}; i.e., $\text{Corr}(x_{it}, \varepsilon_{it}) = 0$. To introduce cross-sectional and temporal dependence, both the explanatory variable x_{it} and the disturbance term ε_{it} contain three components: an individual specific long-run mean ($\overline{x}_i$, $\overline{\varepsilon}_i$), an autocorrelated common factor (g_t, f_t), and an idiosyncratic forcing term (ω_{it}, ϑ_{it}). Accordingly, x_{it} and ε_{it} are specified as follows:

$$x_{it} = \overline{x}_i + \theta_i g_t + \omega_{it} \quad \text{and} \quad \varepsilon_{it} = \overline{\varepsilon}_i + \lambda_i f_t + \vartheta_{it} \tag{6}$$

6. See Driscoll and Kraay (1998) and Beck and Katz (1995).
7. See Bertrand, Duflo, and Mullainathan (2004) and Petersen (2007).

The common factors g_t and f_t in (6) are constructed as AR(1) processes:

$$g_t = \gamma g_{t-1} + w_t \quad \text{and} \quad f_t = \rho f_{t-1} + v_t \tag{7}$$

For simplicity, but again without loss of generality, it is assumed that the within-variance of x_{it}, ε_{it}, g_t, and f_t is one. Together with the condition that the forcing terms ω_{it}, ϑ_{it}, w_t, and v_t are independently and identically distributed (i.i.d.), it follows that

$$\omega_{it} \overset{iid}{\sim} N(0, 1-\theta_i^2\rho)\ ,\ w_t \overset{iid}{\sim} N(0, 1-\gamma^2)\ ,\ \vartheta_{it} \overset{iid}{\sim} N(0, 1-\lambda_i^2)\ ,\ v_t \overset{iid}{\sim} N(0, 1-\rho^2)$$

If we consider these distributional assumptions about the forcing terms, some algebra yields that for realized values of $\overline{x}_i$ and θ_i, the correlation between x_{it} and $x_{j,t-s}$ is given by

$$\text{Corr}(x_{it}, x_{j,t-s}) = \begin{cases} 1 & \text{if } i = j \text{ and } s = 0 \\ \theta_i\theta_j\gamma^s & \text{otherwise} \end{cases}$$

Similarly, for the correlation between ε_{it} and $\varepsilon_{j,t-s}$ it follows that

$$\text{Corr}(\varepsilon_{it}, \varepsilon_{j,t-s}) = \begin{cases} 1 & \text{if } i = j \text{ and } s = 0 \\ \lambda_i\lambda_j\rho^s & \text{otherwise} \end{cases}$$

To complete the specification of the Monte Carlo experiments, it is assumed that both the subject-specific FE ($\overline{x}_i$, $\overline{\varepsilon}_i$) and the idiosyncratic factor sensitivities (θ_i, λ_i) are uniformly distributed:

$$\overline{x}_i \sim U[-a, +a]\ ,\ \overline{\varepsilon}_i \sim U[-b, +b]\ ,\ \theta_i \sim U[\tau_1, \tau_2]\ ,\ \lambda_i \sim U[\iota_1, \iota_2] \tag{8}$$

5.2 Parameter settings (scenarios)

Because the parameters a and b in (8) are irrelevant for the correlations *between* subjects, they are arbitrarily fixed to $a = 1.5$ and $b = 0.6$ in all Monte Carlo experiments.[8] Accordingly, the total variances (i.e., within plus between variance) of x_{it} and e_{it} are $\sigma_x^2 = 1 + a^2/3 = 1.75$ and $\sigma_\varepsilon^2 = 1 + b^2/3 = 1.12$, respectively. By contrast, parameter values for τ_1, τ_2, ι_1, and ι_2 are altered in the simulations because they directly affect the degree of spatial dependence. I consider six different scenarios:

1. $\tau_1 = \tau_2 = \iota_1 = \iota_2 = 0$. This is the reference case where all assumptions of the FE (within) regression model are perfectly met. In this scenario, x_{it} and ε_{it} both contain an individual specific fixed effect, but they are independently distributed between subjects and across time. By denoting with $r(p, q)$ the average or expected correlation between p and q, it follows immediately that here we have $r(x_{it}, x_{j,t-s}) = r(\varepsilon_{it}, \varepsilon_{j,t-s}) = 0$ (for $i \neq j$ or $s \neq 0$).

2. $\tau_1 = \iota_1 = 0$ and $\tau_2 = \iota_2 = \sqrt{1/2}$. Here the expected contemporaneous between-subject correlations are given by $r(x_{it}, x_{jt}) = r(\varepsilon_{it}, \varepsilon_{jt}) = 0.125$ (for $i \neq j$).

8. For a detailed discussion on the consequences of changes in the size of subject-specific FE for statistical inference, see Petersen (2007).

3. $\tau_1 = \iota_1 = 0$ and $\tau_2 = \iota_2 = 1$. This yields $r(x_{it}, x_{jt}) = r(\varepsilon_{it}, \varepsilon_{jt}) = 0.25$ (for $i \neq j$).

4. $\tau_1 = \iota_1 = 0.6$ and $\iota_2 = \tau_2 = 1$. Here the expected contemporaneous between-subject correlations are high: $r(x_{it}, x_{jt}) = r(\varepsilon_{it}, \varepsilon_{jt}) = 0.64$ (for $i \neq j$).

5. $\tau_1 = 0.6$, $\tau_2 = 1$, $\iota_1 = 0$, and $\iota_2 = \sqrt{1/2}$. This results in $r(x_{it}, x_{jt}) = 0.64$ and $r(\varepsilon_{it}, \varepsilon_{jt}) = 0.125$ (for $i \neq j$).

6. $\tau_1 = 0$, $\tau_2 = \sqrt{1/2}$, $\iota_1 = 0.6$, and $\iota_2 = 1$. Here the independent variable is only weakly correlated [$r(x_{it}, x_{jt}) = 0.125$ for $i \neq j$] between subjects, but the residuals are highly dependent [$r(\varepsilon_{it}, \varepsilon_{j,t}) = 0.64$].

These six scenarios are simulated for three levels of autocorrelation, where for brevity I consider only the case $\rho = \gamma$. The autocorrelation parameters ρ and γ are set to 0, 0.25, and 0.5. Finally, y_{it} is generated according to (5) with parameters α and β arbitrarily set to 0.1 and 0.5.

5.3 Results

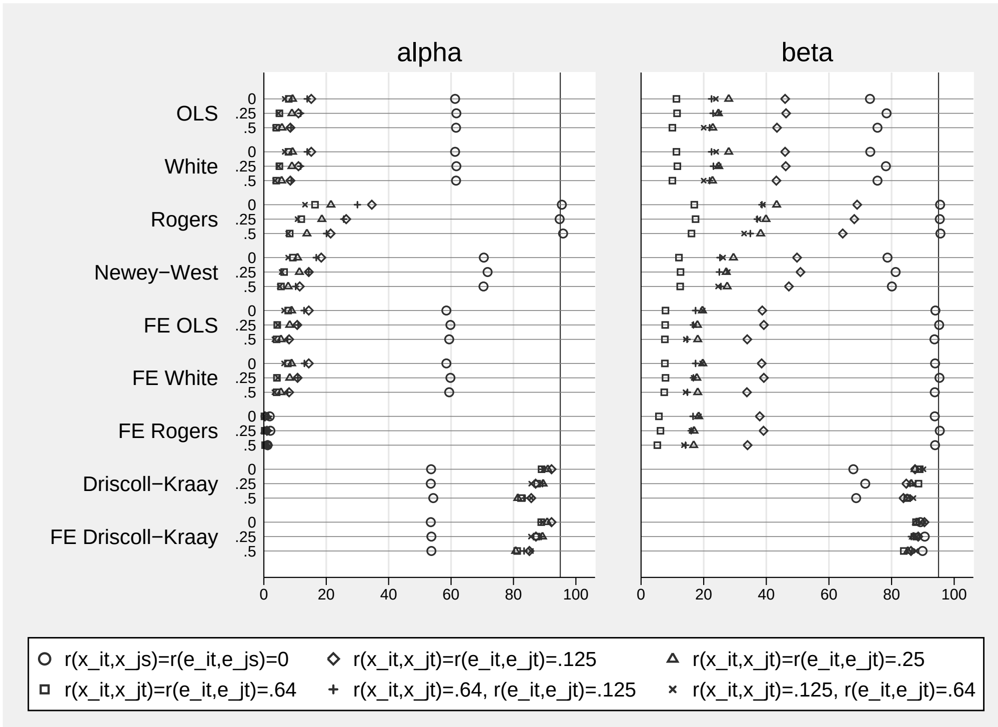

Figure 1: Coverage rates of 95% confidence intervals: comparison of different techniques for estimating standard errors. Monte Carlo simulation with 1,000 runs per parameter setting for a balanced panel with $N = 1{,}000$ subjects and $T = 40$ observations per subject. The total number of observations in the panel regressions is $NT = 40{,}000$, and the y-axis labels 0, .25, and .5 denote the values of the autocorrelation parameters ρ and γ ($\rho = \gamma$).

Reference simulation: Medium-sized microeconometric panel with quarterly data

In the first simulation, I consider a medium-sized microeconometric panel with $N = 1{,}000$ subjects and time dimension $T_{\max} = 40$ as it is typically encountered in corporate finance studies with quarterly data. For all parameter settings, a Monte Carlo simulation with 1,000 replications is run for both balanced and unbalanced panels. To generate the datasets for the unbalanced panel simulations, I assume that the panel starts with a full cross-section of $N = 1{,}000$ subjects that are labeled by a running number ranging from 1 to 1,000. Then, from $t = 2$ on, only the subjects i with $i > \text{floor}\{N(t-1)/(T_{\max}-1)\}$ remain in the panel. Hence, whereas the datasets in the balanced panel simulations contain 40,000 observations, those of the unbalanced panel simulations comprise only 20,018 observations.

In summary, the Monte Carlo simulations for each of the 18 parameter settings—i.e., six scenarios, three levels of autocorrelation—defined in section 5.2 proceed as follows:

1. Generation of a panel dataset with $N = 1{,}000$ subjects and $T_{\max} = 40$ periods as specified above.
2. Estimation of the regression model in (5) by pooled OLS and FE regression. For pooled OLS estimation, five covariance matrix estimators are considered. For the FE regression, four techniques of obtaining standard errors are applied.
3. After having replicated steps (1) and (2) 1,000 times, the coverage rates of the 95% confidence intervals for all nine standard error estimates are gathered. This is achieved by obtaining the fraction of times that the nominal 95% confidence interval for $\widehat{\alpha}$ ($\widehat{\beta}$) contains the true coefficient value of $\alpha = 0.1$ ($\beta = 0.5$).

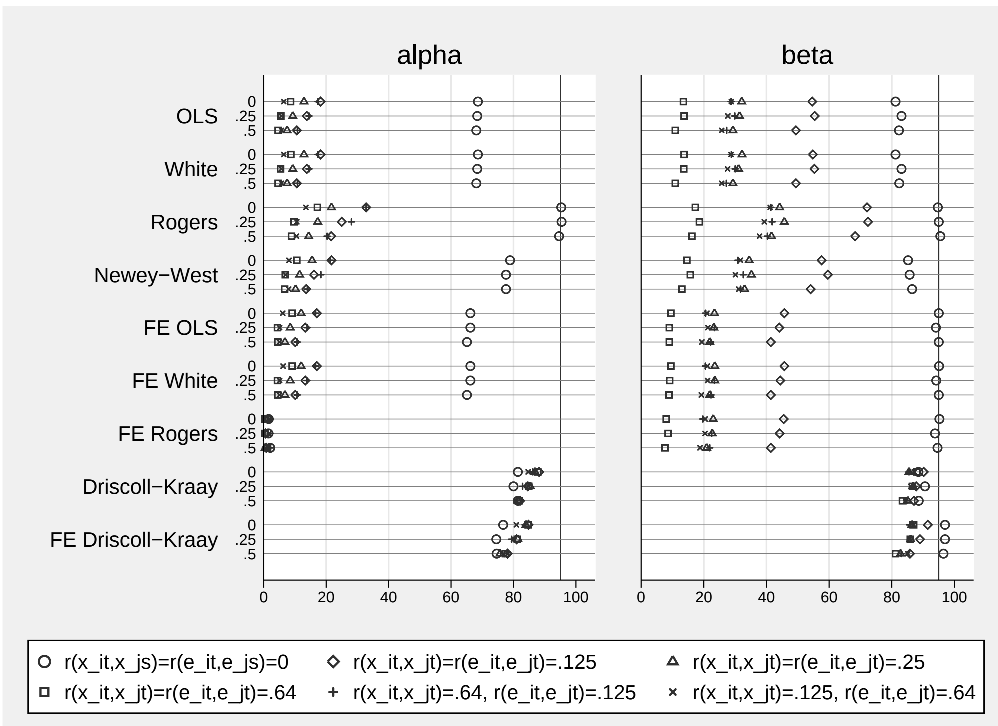

Figure 2: Coverage rates of 95% confidence intervals: comparison of different techniques for estimating standard errors. Monte Carlo simulation with 1,000 runs per parameter setting for an unbalanced panel with $N = 1{,}000$ subjects and at most $T = 40$ observations per subject. The total number of observations in the panel regressions is 20,018, and the y-axis labels 0, .25, and .5 denote the values of the autocorrelation parameters ρ and γ ($\rho = \gamma$).

Figure 1 contains the results of the balanced panel simulation. Interestingly, although the reference case of scenario 1 perfectly meets the assumptions of the FE (within) regression model, pooled OLS estimation delivers coverage rates for the intercept term $\widehat{\alpha}$ that are not worse than those of the FE regressions. On the contrary, Rogers standard errors obtained from pooled OLS regression are the single SE estimates for which the coverage rates of $\widehat{\alpha}$ correspond to their nominal value.

For the estimates for the slope coefficient $\widehat{\beta}$, figure 1 reveals that here all the standard error estimates obtained from FE regression are perfectly calibrated under the parameter settings of scenario 1 (i.e., $\tau_1 = \tau_2 = \iota_1 = \iota_2 = 0$). Considering that scenario 1 perfectly obeys the FE regression model, these simulation results are perfectly in line with the theoretical properties of the FE estimator. Surprisingly, though, not only are FE standard errors appropriate under scenario 1 but also Rogers standard errors for pooled OLS estimation. Finally, and consistently with Driscoll and Kraay's (1998) original findings, figure 1 reveals that the coverage rates of Driscoll and Kraay standard errors are slightly worse than those of the other more commonly used covariance matrix estimators when the residuals are spatially uncorrelated (i.e., for scenario 1).

However, the results change substantially when cross-sectional dependence is present. For OLS, White, Rogers, and Newey–West standard errors' cross-sectional correlation leads to coverage rates that are far below their nominal value, irrespective of whether regression model (5) is estimated by pooled OLS or FE regression. Even worse, although the true model contains individual-specific FE, the coverage rates of the within regressions are actually lower than those of the pooled OLS estimation. Interestingly, Rogers standard errors for pooled OLS are again comparably well calibrated. However, they also tend to be overly optimistic when the cross-sectional units are spatially dependent.

Figure 1 also indicates that the coverage rates of OLS, White, Rogers, and Newey–West standard errors are negatively related to the level of cross-sectional dependence. The more spatially correlated the subjects are, the more severely upward biased will be the t values of linear panel models fitted with OLS, White, Rogers, and Newey–West standard errors. Furthermore, a comparison of the results for scenarios (2) and (5) suggests that an increase in the cross-sectional dependence of the explanatory variable x_{it} exacerbates underestimation of the standard errors and correspondingly lowers coverage rates further.

When we look at the consequences of temporal dependence, figure 1 shows that autocorrelation tends to worsen coverage rates. However, appropriately assessing the effect of serial correlation for the coverage rates is somewhat difficult, as the simulation presented here considers only comparably low levels of autocorrelation, the highest average or expected autocorrelation coefficient being equal to

$$r(\varepsilon_{i,t}, \varepsilon_{i,t-1}) = \gamma_{\max} \cdot E(\lambda_i)^2 = 0.5 \cdot 0.8^2 = 0.32$$

Nevertheless, the figure indicates that the (additional) effect of autocorrelation for the coverage rates of coefficient estimates is relatively small when cross-sectional dependence is present.

Finally, from figure 1 we see that Driscoll and Kraay standard errors tend to be slightly optimistic, too. However, when spatial dependence is present, then Driscoll–Kraay standard errors are much better calibrated (and thus far more robust) than OLS, White, Rogers, and Newey–West standard errors. Furthermore, and in contrast to the aforementioned estimators, the coverage rates of Driscoll–Kraay standard errors are almost invariant to changes in the level of cross-sectional and temporal correlation.

A comparison of figures 1 and 2 reveals that the results of the unbalanced panel simulation are qualitatively similar to those of the balanced panel simulation. Hence, the slight adjustment of Driscoll and Kraay's (1998) original estimator implemented in the `xtscc` command seems to work well in practice.

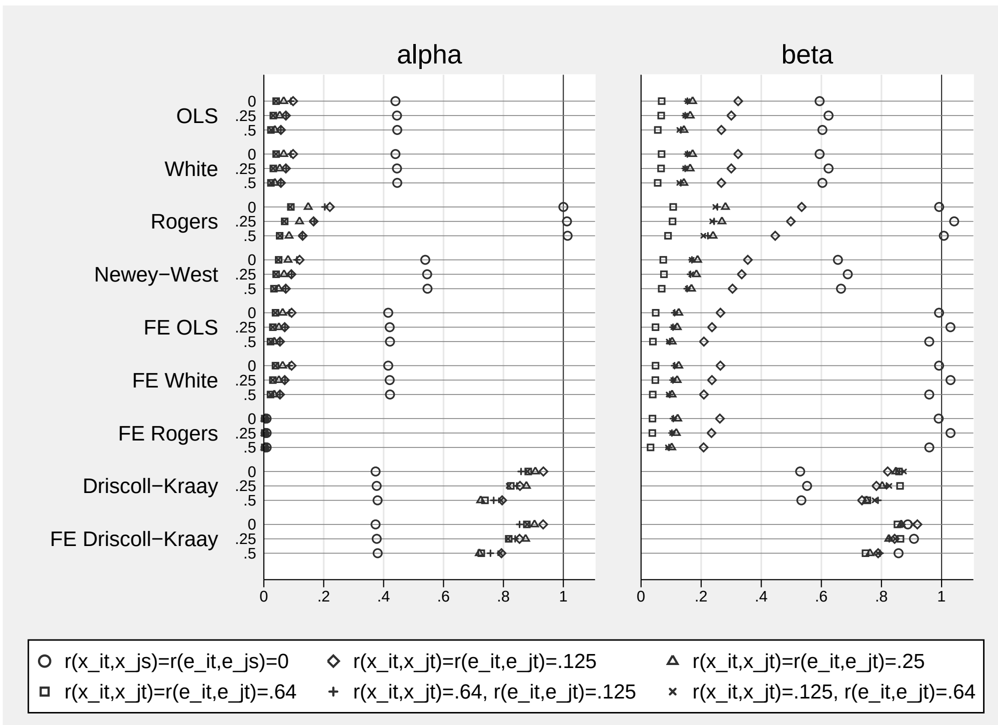

Figure 3: Ratio of estimated to true standard deviation: Monte Carlo simulation with 1,000 runs per parameter setting for a balanced panel with $N = 1{,}000$ subjects and $T = 40$ observations per subject. The total number of observations in the panel regressions is $NT = 40{,}000$, and the y-axis labels 0, .25, and .5 denote the values of the autocorrelation parameters ρ and γ ($\rho = \gamma$).

Figure 3 contains a complementary representation of the results presented in figure 1. Here, for each covariance matrix estimator considered in the analysis, the average standard error estimate from the simulation is divided by the standard deviation of the coefficient estimates. The standard deviation of the estimated coefficients is the true standard error of the regression. Therefore, for a covariance matrix estimator to be unbiased, this ratio should be close to one. Consistent with the findings from above, figure 3 shows that Rogers standard errors for pooled OLS are perfectly calibrated when no cross-sectional correlation is present. However, OLS, White, Rogers, and Newey–West standard errors worsen when spatial correlation increases. Contrary to this, calibration of the Driscoll and Kraay covariance matrix estimator is largely independent from cross-sectional dependence. Since the results of the unbalanced panel simulation turn out to be qualitatively similar to those presented in figure 3, for brevity I do not show them here.

Alternative simulations: Large-scale microeconometric panel with annual data

The results of the reference simulation discussed in the last section suggest that the small-sample properties of Driscoll–Kraay standard errors outperform those of other (more) commonly used covariance matrix estimators when cross-sectional dependence is present. However, by considering that Driscoll and Kraay's (1998) nonparametric covariance matrix estimator relies on large-T asymptotics, one might argue that specifying $T = 40$ in the reference simulation is clearly in favor of the Driscoll–Kraay estimator.

(*Continued on next page*)

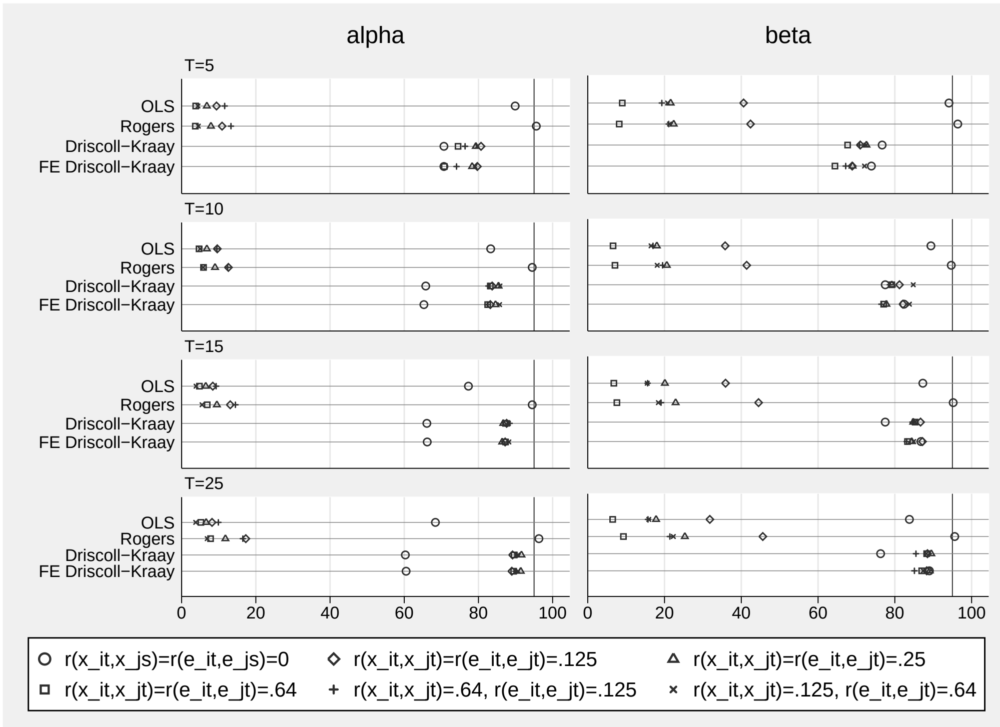

Figure 4: Coverage rates of 95% confidence intervals: comparison of different techniques for estimating standard errors of linear panel models. Monte Carlo simulations with 1,000 replications per parameter setting for balanced panels with $N = 2{,}500$ subjects and temporally uncorrelated common factors f_t and g_t (i.e., ρ=γ=0).

As a robustness check and to obtain a more comprehensive picture about the small-sample performance of Driscoll–Kraay standard errors, I therefore perform a set of four more simulations. Specifically, I consider a large-scale microeconometric panel containing $N = 2{,}500$ subjects whose time dimension amounts to $T = 5$, 10, 15, and 25 periods.

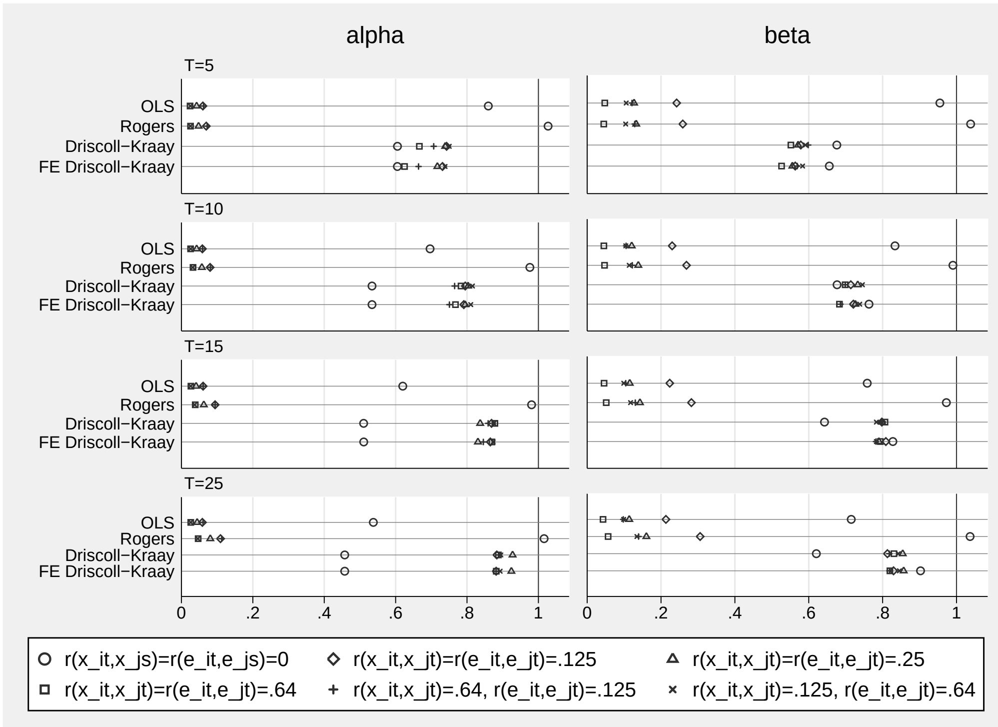

Figure 5: Ratio of estimated to true standard deviation: comparison of different techniques for estimating standard errors of linear panel models. Monte Carlo simulations with 1,000 replications per parameter setting for balanced panels with $N = 2{,}500$ subjects and temporally uncorrelated common factors f_t and g_t (i.e., ρ=γ=0).

Although being somewhat superior when there is no spatial dependence, coverage rates of OLS and Rogers standard errors in figure 4 are clearly dominated by those of the Driscoll–Kraay estimator when cross-sectional correlation is present. Moreover, figure 5 indicates that OLS and Rogers standard errors for pooled OLS tend to severely overstate actual information inherent in the dataset when the subjects are mutually dependent. Interestingly, both these results hold, irrespective of the panel's time dimension T, and they are particularly pronounced when the degree of cross-sectional dependence is high.[9]

Finally, figures 4 and 5 also demonstrate the consequences of the Driscoll and Kraay (1998) estimator being based on large-T asymptotics: the longer the time dimension T of a panel is, the better calibrated are the Driscoll–Kraay standard errors.

9. For brevity, figures 4 and 5 depict only the results for a representative subset of the covariance matrix estimators considered in the simulations. However, the omitted results are qualitatively similar to those for OLS and Rogers standard errors.

6 Example: bid–ask spread of stocks

Here I consider an empirical example from financial economics. The dataset used in the application is by no means special in that cross-sectional dependence is particularly pronounced. Rather, the dataset considered here is just an ordinary small-scale microeconometric panel, as it might be used in any empirical study. This exercise shows that choosing different techniques for obtaining standard error estimates can have substantial consequences for statistical inference. Furthermore, I demonstrate how the `xtscc` program can be used to perform a Hausman test for FE that is robust to general forms of spatial and temporal dependence. In the last part of the example, I show how to test whether the residuals of a panel model are cross-sectionally dependent.

6.1 Introduction

The bid–ask spread is the difference between the asking price for which an investor can buy a financial asset and the (normally lower) bidding price for which the asset can be sold. The bid–ask spread of stocks has long played an important role in financial economics. It therefore constitutes a major component of the transaction costs of equity trades (Keim and Madhavan 1998) and has become a popular measure for a stock's liquidity in empirical finance studies.[10]

According to Glosten (1987), the bid–ask spread depends on several determinants, the most important being the degree of information asymmetries between market participants. Put simply, his theoretical model states that the more pronounced information asymmetries between market participants are, the wider should be the bid–ask spread. In this application, I want to investigate whether typical measures for information asymmetries between market participants (e.g., firm size) can explain parts of the differences in quoted bid–ask spreads as suggested by Glosten's (1987) model.

I analyze a panel of 219 European mid- and large-cap stocks that have been *randomly* selected from the Morgan Stanley Capital International (MSCI) Europe constituents list as of December 31, 2000. The data are month-end figures from Thomson Financial Datastream, and the sample period ranges from December 2000 to December 2005 (61 months).

6.2 Description of the data

`BidAskSpread.dta` comprises an unbalanced panel whose subjects (i.e., the stocks) are identified by variable `ID` and whose time dimension is set by variable `TDate`. The quoted bid–ask spread, `BA`, serves as the dependent variable. Per Roll (1984), who argues that percent bid–ask spreads may be more easily interpreted than absolute ones, variable `BA` is defined in relative terms as follows:

10. Campbell, Lo, and MacKinlay (1997, 99) define liquidity of stocks as "the ability to buy or sell significant quantities of a security quickly, anonymously, and with relatively little price impact."

$$\text{BA}_{it} = 100 \cdot \frac{\text{Ask}_{it} - \text{Bid}_{it}}{0.5(\text{Ask}_{it} + \text{Bid}_{it})} \tag{9}$$

In (9), Bid_{it} and Ask_{it} denote the last bid and ask prices of stock i in month t.

Variable `TRMS` contains the monthly return of the MSCI Europe total return index in U.S. dollars (as a percentage), and variable `TRMS2` is its squared value. `TRMS2` constitutes a simple proxy for the stock market risk and hence reflects uncertainty about future economic prospects. The `Size` variable comprises the stocks' size decile. A value of 1 (10) indicates that the U.S. dollar market capitalization of a stock was among the smallest (largest) 10% of the sample stocks in a given month. Finally, variable `aVol` measures the stocks' abnormal trading volume, which is defined as follows:

$$\text{aVol}_{it} = 100 \cdot \left\{ \ln(\text{Vol}_{it}) - \frac{1}{T_i} \sum_t \ln(\text{Vol}_{it}) \right\}$$

Here Vol_{it} and T_i denote the number (in thousands) of stocks i being traded on the last trading day of month t and the total number of nonmissing observations for stock i, respectively.

The following Stata output lists an arbitrary excerpt of six consecutive observations from `BidAskSpread.dta`:

```
. use BidAskSpread
. list ID TDate BA-Size in 70/75, sep(0) noobs
```

ID	TDate	BA	TRMS	TRMS2	aVol	Size
ABB LTD.	2001:08	0.244	-2.578	6.648	-88.977	8
ABB LTD.	2001:09	0.526	-9.978	99.559	-50.142	8
ABB LTD.	2001:10	0.297	3.171	10.058	4.515	8
ABB LTD.	2001:11	0.363	4.016	16.128	-33.736	8
ABB LTD.	2001:12	.	2.562	6.565	-73.165	8
ABB LTD.	2002:01	0.440	-5.215	27.200	-27.169	8

❑ **Technical note**

These data contain all the typical characteristics for microeconometric panels. Although the dataset starts as a full panel, 27 of 219 stocks leave the sample early. In addition to being unbalanced, the `BidAskSpread` panel contains gaps. For instance, variable `BA` is missing for all the stocks on March 29, 2002.

❑

6.3 Regression specification and formulation of the hypothesis

To investigate whether the information differences between market participants can partially explain cross-sectional differences in quoted bid–ask spreads, I fit the following linear regression model:

$$\texttt{BA}_{it} = \alpha + \beta_{\text{aVol}} \cdot \texttt{aVol}_{it} + \beta_{\text{Size}} \cdot \texttt{Size}_{it} + \beta_{\text{TRMS2}} \cdot \texttt{TRMS2}_{it} + \beta_{\text{TRMS}} \cdot \texttt{TRMS}_{it} + \epsilon_{it} \quad (10)$$

Here $i = 1, \ldots, 219$ denotes the stocks and $t = 491, \ldots, 551$ is the month in Stata's time-series format.

Glosten's (1987) model predicts that the degree of asymmetric information between market participants should be positively related to the bid–ask spread. In the finance literature, it is generally believed that paid prices of frequently traded stocks contain more information than those of rarely traded ones. Accordingly, asymmetric information between market participants is assumed to be smaller for liquid than for nonliquid stocks, which leads to the hypothesis that frequently traded stocks should have tighter bid–ask spreads than illiquid ones.

If this conjecture is correct, we would expect that estimating regression model (10) yields $\beta_{\text{aVol}} < 0$ because stock prices contain more information when abnormal trading volume is high than when it is low. Furthermore, similar reasoning leads to the expectation that the coefficient estimate for the `Size` variable is also negative since small stocks tend to be less frequently traded than large stocks.

However, in addition to being negative, the coefficient estimate for β_{Size} should be highly significant. Besides Roll (1984), who finds that firm size is closely related to the stocks' *effective* bid–ask spread, many studies in empirical finance find evidence for fundamental return differences between small and large stocks.[11]

Since the volatility of stock market returns is closely related to the uncertainty about future economic prospects, bid–ask spreads are expected to be positively correlated with stock market risk. Hence, β_{TRMS2} should be positive. Finally, for variable `TRMS` no such information story or another compelling economic argument is immediate. Accordingly, whether the coefficient estimate for β_{TRMS} should be positive or negative is indefinite on an ex ante basis.

6.4 Pooled OLS estimation

Fitting the regression model in (10) is likely to produce residuals that are positively correlated over time. Furthermore, cross-sectional dependence cannot be completely ruled out because of possibly available common factors that are not considered in the analysis. Therefore, I follow the suggestion from section 5.3 and fit regression model (10) by pooled OLS with Driscoll–Kraay standard errors. Somewhat arbitrarily, a lag length of 8 months is chosen. However, the results turn out to be robust to changes in the selected lag length.

11. See Banz (1981) and Fama and French (1992, 1993).

```
. xtscc BA aVol Size TRMS2 TRMS, lag(8)

Regression with Driscoll-Kraay standard errors   Number of obs     =     11775
Method: Pooled OLS                               Number of groups  =       219
Group variable (i): ID                           F(  4,    218)    =    142.84
maximum lag: 8                                   Prob > F          =    0.0000
                                                 R-squared         =    0.0290
                                                 Root MSE          =    2.6984

------------------------------------------------------------------------------
             |             Drisc/Kraay
          BA |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        aVol |  -.0017793   .0010938    -1.63   0.105    -.0039351    .0003764
        Size |   -.151868   .0102688   -14.79   0.000    -.1721068   -.1316291
       TRMS2 |   .0033298   .0008826     3.77   0.000     .0015902    .0050694
        TRMS |   -.001836   .0052329    -0.35   0.726    -.0121496    .0084777
       _cons |   1.459139   .1354202    10.77   0.000     1.192238    1.726039
------------------------------------------------------------------------------
```

The regression results confirm the hypothesis about the signs of the coefficient estimates. Furthermore, and consistent with my conjecture from above, β_{Size} is not only negative but also highly significant.

It is interesting to compare the results of pooled OLS estimation with Driscoll–Kraay standard errors with those of alternative (more) commonly applied standard error estimates. Table 2 shows that statistical inference indeed depends substantially on the choice of the covariance matrix estimator. This effect can probably best be seen from variable `aVol`. Although OLS standard errors lead to the conclusion that β_{aVol} is highly significant at the 1% level, Driscoll and Kraay standard errors indicate that β_{aVol} is insignificant even at the 10% level. However, Driscoll and Kraay standard errors need not necessarily be more conservative than those of other covariance estimators, as can easily be inferred from the t values of β_{Size}.

(*Continued on next page*)

Table 2: Comparison of standard error estimates for pooled OLS estimation

SE	OLS	White	Rogers	Newey–West	Driscoll–Kraay
aVol	-0.0018^{***}	-0.0018^{**}	-0.0018^{*}	-0.0018^{*}	-0.0018
	(-4.006)	(-2.043)	(-1.831)	(-1.760)	(-1.627)
Size	-0.1519^{***}	-0.1519^{***}	-0.1519^{***}	-0.1519^{***}	-0.1519^{***}
	(-17.412)	(-12.496)	(-6.756)	(-10.717)	(-14.789)
TRMS2	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0033^{***}
	(5.295)	(5.520)	(5.495)	(5.582)	(3.773)
TRMS	-0.0018	-0.0018	-0.0018	-0.0018	-0.0018
	(-0.370)	(-0.353)	(-0.381)	(-0.340)	(-0.351)
Const.	1.4591^{***}	1.4591^{***}	1.4591^{***}	1.4591^{***}	1.4591^{***}
	(25.266)	(18.067)	(9.172)	(14.883)	(10.775)
No. of obs.	11,775	11,775	11,775	11,775	11,775
No. of clusters			219		219
R^2	0.029	0.029	0.029	0.029	0.029

NOTE: This table provides the coefficient estimates from the regression model in (10) estimated by pooled OLS. The t stats (in parentheses) are based on standard error estimates obtained from the covariance matrix estimators in the column headings. The dataset contains monthly data from December 2000 to December 2005 for a panel of 219 stocks that have been randomly selected from the MSCI Europe constituents list as of December 31, 2000. The dependent variable in the regression is the relative bid–ask spread `BA`. `aVol` is the abnormal trading volume, `Size` contains the stock's size decile, `TRMS` denotes the monthly return as a percentage of the MSCI Europe total return index, and `TRMS2` is the square of it. *, **, and *** imply statistical significance at the 10%, 5%, and 1% level, respectively.

Comparing the results for variable `TRMS2` is of particular interest. Although being significant at the 1% level, the t stat obtained from Driscoll–Kraay standard errors is markedly lower than that of the other covariance matrix estimators considered in table 2. This finding is perfectly in line with the Monte Carlo evidence presented above, as in the presence of cross-sectional dependence coverage rates of OLS, White, Rogers, and Newey–West standard errors are low when an explanatory variable is highly correlated between subjects. Being a common factor, variable `TRMS2` is perfectly positively correlated between the firms. Therefore, coverage rates of OLS, White, Rogers, and Newey–West standard errors are expected to be particularly low when spatial correlation is present. As a result, the comparably low t stat of the Driscoll–Kraay estimator for variable `TRMS2` indicates that cross-sectional dependence might indeed be present here. Unfortunately, however, this conjecture cannot be formally tested because no adequate testing procedure for cross-sectional dependence in the residuals of pooled OLS regressions is available in Stata right now. Therefore, I must defer a formal test for spatial dependence in the regression residuals to section 6.7. There, I perform Pesaran's (2004) cross-sectional dependence (CD) test on the residuals of the regression model in (10) being estimated by FE regression.

6.5 Robust Hausman test for FE

If the pooled OLS model in (10) is correctly specified and the covariance between ϵ_{it} and the explanatory variables is zero, then either $N \rightarrow \infty$ or $T \rightarrow \infty$ is sufficient for consistency. However, pooled OLS regression yields inconsistent coefficient estimates when the true model is the FE model; i.e.,

$$\texttt{BA}_{it} = \alpha_i + \mathbf{x}'_{it}\beta + e_{it} \tag{11}$$

with $\mathrm{Cov}(\alpha_i, \mathbf{x}_{it}) \neq 0$ and $i = 1, \ldots, 219$. Under the assumption that the unobservable individual effects α_i are time invariant but correlated with the explanatory variables $\mathbf{x}_{it}$, the regression model in (11) can be consistently estimated by FE or within regression.

To test for the presence of subject-specific FE, performing a Hausman test is common. The null hypothesis of the Hausman test states that the RE model is valid, i.e., that $E(\alpha_i + e_{it}|\mathbf{x}_{it}) = 0$. Here I explain how the `xtscc` program can be used to perform a Hausman test that is heteroskedasticity consistent and robust to general forms of spatial and temporal dependence. The exposition starts with the standard Hausman test as implemented in Stata's `hausman` command. Then Wooldridge's (2002, 288ff.) suggestion on how to perform a panel-robust version of the Hausman test is adapted to form a test that is also consistent if cross-sectional dependence is present.

Standard Hausman test as implemented in Stata

Although pooled OLS regression yields consistent coefficient estimates when the RE model is true [i.e., $E(\alpha_i + e_{it}|\mathbf{x}_{it}) = 0$], its coefficient estimates are inefficient under the null hypothesis of the Hausman test. Therefore, pooled OLS regression should not be used when testing for FE. Because FGLS estimation is both consistent and efficient under the null hypothesis of the Hausman test, comparing the coefficient estimates obtained from FGLS with those of the FE estimator is more appropriate.[12] For numerical reasons, Wooldridge (2002, 290) recommends performing the Hausman test for FE with either the FE or the random-effects estimates of σ_e^2. Thanks to the `hausman` command's option `sigmamore`, Stata makes performing a standard Hausman test in the way suggested by Wooldridge (2002) easy:

```
. qui xtreg BA aVol Size TRMS2 TRMS, re       // FGLS estimation
. estimates store REgls
. qui xtreg BA aVol Size TRMS2 TRMS, fe       // within regression
. estimates store FE
```

12. However, the FGLS estimator is no longer fully efficient under the null when α_i or e_{it} is not i.i.d. In this likely case, the standard Hausman test becomes invalid and a more general testing procedure is required.

```
. hausman FE REgls, sigmamore        // see Wooldridge (2002, 290) for details
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      FE          REgls        Difference          S.E.
-------------+----------------------------------------------------------------
        aVol |   -.0017974    -.0017916       -5.81e-06        .0000128
        Size |   -.1875486    -.1603143       -.0272343        .0337314
       TRMS2 |    .0031042     .0031757       -.0000715        .0000239
        TRMS |   -.0014581     -.001634         .000176        .0001959
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       11.64
                Prob>chi2 =      0.0203
```

Provided that the Hausman test applied here is valid (which it probably is not), the null hypothesis of no FE is rejected at the 5% level of significance. Therefore, the standard Hausman test leads to the conclusion that pooled OLS estimation is likely to produce inconsistent coefficient estimates for the regression model in (10). As a result, the regression model in (10) should be estimated by FE (within) regression.

Alternative formulation of the Hausman test and robust inference

In his seminal work on specification tests in econometrics, Hausman (1978) showed that performing a Wald test of $\gamma = \mathbf{0}$ in the auxiliary OLS regression

$$\mathrm{BA}_{it} - \widehat{\lambda}\overline{\mathrm{BA}}_i = (1-\widehat{\lambda})\mu + (\mathbf{x}_{1it} - \widehat{\lambda}\overline{\mathbf{x}}_{1i})'\beta_1 + (\mathbf{x}_{1it} - \overline{\mathbf{x}}_{1i})'\gamma + v_{it} \tag{12}$$

is asymptotically equivalent to the chi-squared test conducted above. In (12), $\mathbf{x}_{1it}$ denotes the time-varying regressors, $\overline{\mathbf{x}}_{1i}$ are the time-demeaned regressors, and $\widehat{\lambda} = 1 - \sigma_e/\sqrt{\sigma_e^2 + T\sigma_\alpha^2}$. For $\gamma = \mathbf{0}$, expression (12) reduces to the two-step representation of the RE estimator. As a result, the null hypothesis of this alternative test (i.e., $\gamma = \mathbf{0}$) states that the RE model is appropriate.[13] Although this alternative formulation of the Hausman test does not necessarily have better finite-sample properties than those of the standard Hausman test implemented in Stata's `hausman` command, it has the advantage of being computationally more stable in finite samples because it never encounters problems with non–positive definite matrices.

When α_i or e_{it} is not i.i.d., the RE estimator is not fully efficient under the null hypothesis of $E(\alpha_i + e_{it}|\mathbf{x}_{it}) = 0$. As a result, estimating the augmented regression in (12) with OLS standard errors or running Stata's `hausman` test leads to invalid statistical inference. Unfortunately, however, α_i or e_{it} is probably not i.i.d., as heteroskedasticity and other forms of temporal and cross-sectional dependency are often encountered in microeconometric panel datasets. To ensure valid statistical inference for the Hausman test when α_i or e_{it} is non-i.i.d., Wooldridge (2002, 288ff.) therefore proposes estimat-

13. See Cameron and Trivedi (2005, 717ff.) for details.

ing the auxiliary regression in (12) with panel-robust standard errors. In Stata, the respective analysis can be performed as follows:

```
. qui xtreg BA aVol Size TRMS2 TRMS, re
. scalar lambda_hat = 1 - sqrt(e(sigma_e)^2/(e(g_avg)*e(sigma_u)^2+e(sigma_e)^2))
. gen in_sample = e(sample)
. sort ID TDate
. qui foreach var of varlist BA aVol Size TRMS2 TRMS   {
.     by ID: egen `var'_bar = mean(`var') if in_sample
.     gen `var'_re = `var' - lambda_hat*`var'_bar if in_sample  // GLS-transform
.     gen `var'_fe = `var' - `var'_bar if in_sample             // within-transform
.     }
. * Wooldridge's auxiliary regression for the panel-robust Hausman test:
. reg BA_re aVol_re Size_re TRMS2_re TRMS_re aVol_fe Size_fe TRMS2_fe
> TRMS_fe if in_sample, cluster(ID)
  (output omitted)
. * Test of the null-hypothesis ``gamma==0'':
. test aVol_fe Size_fe TRMS2_fe TRMS_fe
 ( 1)  aVol_fe = 0
 ( 2)  Size_fe = 0
 ( 3)  TRMS2_fe = 0
 ( 4)  TRMS_fe = 0
       F(  4,   218) =    2.40
            Prob > F =    0.0510
```

Here the null hypothesis of no FE has to be rejected at the 10% level. Because of the marginal rejection of the null hypothesis, the regression model in (10) should be estimated by FE regression to ensure consistency of the results. However, even though this alternative specification of the Hausman test is more robust than the one presented above, it is still based on the assumption that $\mathrm{Cov}(e_{it}, e_{js}) = 0$ for $i \neq j$. Therefore, statistical inference will be invalid if cross-sectional dependence is present, which is likely for microeconometric panel regressions.

To perform a Hausman test that is robust to general forms of spatial and temporal dependence and that should be suitable for most microeconometric applications, I adapt Wooldridge's suggestion and fit the auxiliary regression in (12) with Driscoll and Kraay standard errors:

```
. xtscc BA_re aVol_re Size_re TRMS2_re TRMS_re aVol_fe Size_fe TRMS2_fe TRMS_fe
> if in_sample, lag(8)
  (output omitted)
. test aVol_fe Size_fe TRMS2_fe TRMS_fe
 ( 1)  aVol_fe = 0
 ( 2)  Size_fe = 0
 ( 3)  TRMS2_fe = 0
 ( 4)  TRMS_fe = 0
       F(  4,   218) =    1.65
            Prob > F =    0.1632
```

The F stat from the test of $\gamma = \mathbf{0}$ is much smaller than that of the panel-robust Hausman test encountered before, and the null hypothesis of $E(\alpha_i + e_{it}|\mathbf{x}_{it}) = 0$ can no longer be rejected at any standard level of significance. Thus, after fully accounting for cross-sectional and temporal dependence, the Hausman test indicates that the coefficient estimates from pooled OLS estimation should be consistent.

If the average cross-sectional dependence of a microeconometric panel is positive (negative) on average, then the spatial correlation–robust Hausman test suggested here is less (more) likely to reject the null hypothesis than the versions of the Hausman test described before.

6.6 FE estimation

Although the spatial correlation consistent version of the Hausman test indicates that the coefficient estimates from pooled OLS estimation should be consistent, I nevertheless estimate regression model (10) by FE regression. Table 3 compares the results from different techniques of obtaining standard error estimates for the FE estimator.

Table 3: Comparison of standard error estimates for FE regression

SE	FE	White	Rogers	Driscoll–Kraay
aVol	-0.0018^{***}	-0.0018^{**}	-0.0018^{*}	-0.0018^{**}
	(-4.161)	(-2.166)	(-1.852)	(-2.057)
Size	-0.1875^{***}	-0.1875^{***}	-0.1875^{***}	-0.1875^{***}
	(-4.994)	(-4.883)	(-4.186)	(-6.977)
TRMS2	0.0031^{***}	0.0031^{***}	0.0031^{***}	0.0031^{***}
	(5.072)	(5.717)	(5.370)	(3.835)
TRMS	-0.0015	-0.0015	-0.0015	-0.0015
	(-0.302)	(-0.300)	(-0.311)	(-0.279)
Const.	1.6670^{***}	1.6670^{***}	1.6670^{***}	1.6670^{***}
	(7.750)	(7.452)	(6.737)	(7.965)
No. of obs.	11,775	11,775	11,775	11,775
No. of stocks	219	219	219	219
Overall R^2	0.029	0.029	0.029	0.029

NOTE: This table provides the coefficient estimates from the regression model in (10) fitted by FE (within) regression. The t stats (in parentheses) are based on standard error estimates obtained from the covariance matrix estimators in the column headings. The dataset contains monthly data from December 2000 to December 2005 for a panel of 219 stocks that have been randomly selected from the MSCI Europe constituents list as of December 31, 2000. The dependent variable in the regression is the relative bid–ask spread `BA`. `aVol` is the abnormal trading volume, `Size` contains the stock's size decile, `TRMS` denotes the monthly return as a percentage of the MSCI Europe total return index, and `TRMS2` is the square root of `TRMS`. *, **, and *** imply statistical significance at the 10%, 5%, and 1% level, respectively.

With the exception of the t values for the `size` variable, which are markedly smaller, both the coefficient estimates and the t values are similar to those of the pooled OLS estimation above. However, the t value of variable `aVol`, which was insignificant for the pooled OLS estimation with Driscoll and Kraay standard errors, now indicates significance at the 5% level.

6.7 Testing for cross-sectional dependence

Tables 2 and 3 indicate that standard error estimates depend substantially on the choice of the covariance matrix estimator. But which standard error estimates are consistent for the regression model in (10)? The Monte Carlo evidence presented in section 5 indicates that the calibration of Driscoll–Kraay standard errors is worse than that of, say, Rogers standard errors if the subjects are spatially uncorrelated. However, Driscoll and Kraay standard errors are much more appropriate when cross-sectional dependence is present. To test whether the residuals from an FE estimation of regression model (10) are spatially independent, I perform Pesaran's (2004) CD test.[14] The null hypothesis of the CD test states that the residuals are cross-sectionally uncorrelated. Correspondingly, the test's alternative hypothesis presumes that spatial dependence is present. Thanks to Rafael De Hoyos and Vasilis Sarafidis, who implemented Pesaran's CD test in their `xtcsd` command, the CD test is readily available in Stata.[15] Because `xtcsd` is implemented as a postestimation command for `xtreg`, an FE (or RE) regression model with OLS standard errors must be estimated before calling the `xtcsd` program:

```
. qui xtreg BA aVol Size TRMS2 TRMS, fe
. xtcsd, pesaran abs

Pesaran's test of cross sectional independence =    94.455, Pr = 0.0000

Average absolute value of the off-diagonal elements =     0.160
```

From the output of the `xtcsd` command, one can see that estimating (10) with FE produces regression residuals that are cross-sectionally dependent. On average, the (absolute) correlation between the residuals of two stocks is 0.16. Therefore, it comes as no surprise that Pesaran's CD test rejects the null hypothesis of spatial independence at any standard level of significance. Regression (10) should therefore be estimated with Driscoll–Kraay standard errors since they are robust to general forms of cross-sectional and temporal dependence.

7 Conclusion

The `xtscc` program presented here produces Driscoll and Kraay (1998) standard errors for linear panel models. Besides being heteroskedasticity consistent, these standard

14. Pesaran's CD test is suitable for panels with N and T tending to infinity in any order.
15. See DeHoyos and Sarafidis (2006) for more details about `xtcsd`. Besides Pesaran's CD test, the `xtcsd` program can also perform the cross-sectional independence tests suggested by Friedman (1937) and Frees (1995). However, only Pesaran's CD test is adequate for use with unbalanced panels.

error estimates are robust to general forms of cross-sectional and temporal dependence. In contrast to Driscoll and Kraay's (1998) original covariance matrix estimator, which is for use with balanced panels only, the `xtscc` program works with both balanced and unbalanced panels.

Cross-sectional dependence constitutes a problem for many (microeconometric) panel datasets, as it can arise even when the subjects are randomly sampled. The reasons for spatial correlation in the disturbances of panel models are manifold. Typically, it arises because social norms, psychological behavior patterns, and herd behavior cannot be quantitatively measured and thus enter panel regressions as unobserved common factors.

The Monte Carlo experiments considered here indicate that the choice of the covariance matrix estimator is crucial for the validity of the statistical results. OLS, White, Rogers, and Newey–West standard errors are therefore well calibrated when the residuals of a panel regression are homoskedastic as well as spatially and temporally independent. However, when the residuals are cross-sectionally correlated, then the aforementioned covariance matrix estimators lead to severely downward-biased standard error estimates for both pooled OLS and FE (within) regression. By contrast, Driscoll–Kraay standard errors are well calibrated when the regression residuals are cross-sectionally dependent, but they are slightly less adequate than, say, Rogers standard errors when spatial dependence is absent.

To ensure that statistical inference is valid, testing whether the residuals of a linear panel model are cross-sectionally dependent is therefore important. If they are, then statistical inference should be based on the Driscoll–Kraay estimator. However, when the residuals are believed to be spatially uncorrelated, then Rogers standard errors are preferred. Although no testing procedure for cross-sectional dependence in the residuals of pooled OLS regression models is currently available in Stata, DeHoyos and Sarafidis (2006) implemented Pesaran's (2004) CD test for the FE and the RE estimator in their `xtcsd` command.

8 Acknowledgments

I thank Christopher F. Baum, David Drukker, William Gould, Matthias Grueninger, Ben Jann, Christian Kleiber, David Rey, and Markus Schmid for helpful comments and discussions.

9 References

Andrews, D. W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59: 817–858.

Arellano, M. 1987. Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics* 49: 431–434.

Banz, R. W. 1981. Relationship between return and market value of common stocks. *Journal of Financial Economics* 9: 3–18.

Beck, N., and J. N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review* 89: 634–647.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119: 249–275.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay. 1997. *Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

DeHoyos, R. E., and V. Sarafidis. 2006. Testing for cross-sectional dependence in panel-data models. *Stata Journal* 6: 482–496.

Driscoll, J., and A. C. Kraay. 1998. Consistent covariance matrix estimation with spatially dependent data. *Review of Economics and Statistics* 80: 549–560.

Eicker, F. 1967. Limit theorems for regressions with unequal and dependent errors. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, ed. L. LeCam and J. Neyman, 59–82. Berkeley, CA: University of California Press.

Fama, E. F., and K. R. French. 1992. Cross-section of expected stock returns. *Journal of Finance* 47: 427–465.

———. 1993. Common risk factors in returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56.

Feng, L., and M. S. Seasholes. 2004. Correlated trading and location. *Journal of Finance* 59: 2117–2144.

Frees, E. W. 1995. Assessing cross-sectional correlation in panel data. *Journal of Econometrics* 69: 393–414.

Friedman, M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* 32: 675–701.

Froot, K. A. 1989. Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis* 24: 333–355.

Glosten, L. R. 1987. Components of the bid–ask spread and the statistical properties of transaction prices. *Journal of Finance* 42: 1293–1307.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271.

Hirshleifer, D., and S. H. Teoh. 2003. Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management* 9: 25–66.

Huber, P. J. 1967. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 221–233. Berkeley, CA: University of California Press.

Keim, D., and A. Madhavan. 1998. Cost of institutional equity trades. *Financial Analysts Journal* 54: 50–69.

Kmenta, J. 1986. *Elements of Econometrics*. 2nd ed. New York: Macmillan.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

———. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61: 631–653.

Parks, R. 1967. Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association* 62: 500–509.

Pesaran, M. H. 2004. General diagnostic tests for cross section dependence in panels. University of Cambridge, Faculty of Economics, Cambridge Working Papers in Economics No. 0435.

Petersen, M. A. 2007. Estimating standard errors in finance panel datasets: Comparing approaches. *Review of Financial Studies*. Forthcoming.

Rogers, W. H. 1993. sg17: Regression standard errors in clustered samples. *Stata Technical Bulletin* 13: 19-23. Reprinted in *Stata Technical Bulletin Reprints*, vol. 3, pp. 88–94. College Station, TX: Stata Press.

Roll, R. 1984. Simple implicit measure of the effective bid–ask spread in an efficient market. *Journal of Finance* 39: 1127–1140.

Trueman, B. 1994. Analyst forecasts and herding behavior. *Review of Financial Studies* 7: 97–124.

Verbeek, M. 2004. *A Guide to Modern Econometrics*. 2nd ed. West Sussex, UK: Wiley.

Welch, I. 2000. Herding among security analysts. *Journal of Financial Economics* 58: 369–396.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48: 817–838.

———. 1984. *Asymptotic Theory for Econometricians*. Orlando, FL: Academic Press.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

About the author

Daniel Hoechle is a Ph.D. student at the Department of Finance, University of Basel, Basel, Switzerland.