Heteroskedasticity, Autocorrelation, and Spatial Correlation Robust Inference in Linear Panel Models with Fixed-Effects

Tim Vogelsang

Department of Economics, Michigan State University

June 19, 2009

Background I

- Linear panel models estimated by pooled OLS with fixed-effects and possibly time period fixed-effects are widely used in practice.
- Reliable inference requires the use of robust standard errors.
  - Heteroskedasticity in the cross-section and serial correlation in the time dimension are easily handled using the well known "clustered" standard errors (Arellano, 1987, OxBul).
  - Many empirical users weren't aware of these standard errors when the paper by Bertrand, Duflo and Mullainathan (2004, QJE) was published (see their Table 1) even though Wooldridge's (2002) graduate book covers them.

Background II

- Clustered standard errors have been well studied.
  - Asymptotic theory with T fixed, n large indicates that t-tests which use clustered standard errors have $N(0,1)$ asymptotic distributions under the assumption of random sampling in the cross-section (or no correlation across individuals).
  - Rules out spatial correlation in the cross-section but serial correlation can be general including nonstationary.

Background III

- Hansen (2007, JoE) provides a thorough theoretical analysis of clustered standard errors.
  - With n fixed and T is large, t-tests using clustered standard errors are approximately distributed $\sqrt{\frac{n}{n-1}}t_{n-1}$.
  - This result holds when there is no spatial correlation or heterogeneity in the cross-section and there is weak dependence over time.
- Hansen's assumptions require strict exogeneity in the time dimension when fixed-effect estimation is used. His results do not include the case where time period dummies are also included in the model.

Standard Errors Robust to Spatial Correlation

- If something is known about the form of spatial correlation, it might be possible to group the cross-section individuals into clusters.
- Observations within a cluster are correlated.
- Clusters are "roughly" uncorrelated with each other (see Wooldridge, 2003, AER Papers and Proceedings).
- Bester, Conley and Hansen (2008) provide an asymptotic analysis of spatially clustered standard errors where the number of clusters is held fixed and the number of observations per clusters is large.
- Given a distance measure related to the strength of the spatial correlation, Spatial-HAC standard errors of Conley (1999) can be used.
- Bester, Conley, Hansen and Vogelsang (2008) develop fixed-b asymptotic results for Spatial-HAC based tests.

What if Nothing is Known About the Spatial Correlation?

- Options when T is not too small and there is weak time dependence:
  - Fama-Macbeth (1973, JPE): Estimate the model for each cross-section generating a time series of parameter estimates. Average these estimates over time, use HAC robust standard errors. Problem: estimates are inconsistent if fixed-effects are correlated with regressors.
  - Aggregate data within each cross-section to form a time series regression, use HAC standard errors. With intercept included, exact invariance to fixed-effects is obtained. Potentially throws away a lot of information: less efficient estimation.
  - Driscoll and Kraay (1998, Restat): Use pooled OLS with standard errors that handle spatial correlation. Very similar to time series HAC standard errors. Their results do not allow fixed-effect dummies. Goncalves (2009) allows dummies.
  - The Driscoll and Kraay standard errors require a choice of kernel and bandwidth. They use consistency of the standard errors - does not capture affect of kernel/bandwidth choice on the test. Fixed-b asymptotics can capture those choices.

Contributions of This Paper I

1. Develop a fixed-b asymptotic theory for heteroskedasticity, autocorrelation, spatial correlation (HACSC) robust standard errors in a linear panel model.
2. Provide results for models with fixed-effects and possibly time period dummies as well.
3. For completeness results are provided for generalizations of the usual "clustered" standard errors. Extend some results of Hansen (2007) to allow weak exogeneity over time.
4. The analysis is carried out using "fixed-n, large-T" asymptotics and "large-n, large-T" asymptotics.
5. The fixed-b asymptotic distribution of the HACSC robust tests are the same whether fixed-effects are included in the model or both fixed-effects and time periods dummies are included. The results are the same whether n is fixed or is allowed to grow with T.

Contributions of This Paper II

6. The exogeneity assumptions and form of spatial correlation required to obtain the asymptotic results depend on what dummies are included and whether n is fixed or large.
7. Monte Carlo simulations indicate the fixed-b approximation (fixed-b critical values) works well when the strength of the serial correlation is not too strong relative to the time dimension sample size, T. The value of n does not play a critical role in the accuracy of the approximation but does matter for power.

The Model I

- Consider a standard linear panel model

$$y_{it} = x_{it}'\beta + a_i + u_{it} \tag{1}$$

  where x_{it} is a $k \times 1$ vector of regressors.
- It is assumed that x_{it} is uncorrelated with u_{it} but a_i can be correlated with x_{it}.
- One might also want to allow time effects

$$y_{it} = x_{it}'\beta + a_i + f_t + u_{it}. \tag{2}$$

- The parameters of interest are β.

The Model II

- Consider the pooled OLS fixed-effects estimator of β

$$\widehat{\beta} = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{x}_{it}' \right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{y}_{it}$$

where $\tilde{y}_{it} = \ddot{y}_{it} = y_{it} - \overline{y}_i$ and $\tilde{x}_{it} = \ddot{x}_{it} = x_{it} - \overline{x}_i$ in model (1) and $\tilde{y}_{it} = \ddot{y}_{it} - \overline{\ddot{y}}_t$ and $\tilde{x}_{it} = \ddot{x}_{it} - \overline{\ddot{x}}_t$ in model (2) with $\overline{\ddot{y}}_t = n^{-1} \sum_{i=1}^{n} \ddot{y}_{it}$ and $\overline{\ddot{x}}_t = n^{-1} \sum_{i=1}^{n} \ddot{x}_{it}$.

The Model III

- Plugging in for $\tilde{y}_{it}$ gives

$$\widehat{\beta} - \beta = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{x}_{it}' \right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{u}_{it}.$$

or

$$\begin{aligned} \sqrt{T} \left(\widehat{\beta} - \beta \right) &= \left(\sum_{i=1}^{n} T^{-1} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{x}_{it}' \right)^{-1} \sum_{i=1}^{n} T^{-1/2} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{u}_{it}, \\ &= \left(\sum_{i=1}^{n} T^{-1} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{x}_{it}' \right)^{-1} \sum_{i=1}^{n} T^{-1/2} \sum_{t=1}^{T} \tilde{v}_{it}, \end{aligned}$$

where $\tilde{v}_{it} = \tilde{x}_{it} \tilde{u}_{it}$.

The Model IV

- The form of the standard errors depends on

$$var\left(\sum_{i=1}^{n} T^{-1/2}\sum_{t=1}^{T}\tilde{v}_{it}\right).$$

- Standard errors would be taken from the matrix

$$T^{-1}\left(\sum_{i=1}^{n} T^{-1}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}\widehat{var}\left(\sum_{i=1}^{n} T^{-1/2}\sum_{t=1}^{T}\tilde{v}_{it}\right)\left(\sum_{i=1}^{n} T^{-1}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}$$

Standard Errors: No Cross-section Correlation I

- Suppose $\tilde{v}_{it}$ is uncorrelated with $\tilde{v}_{jt}$ for $i \neq j$ (no spatial correlation), then

$$var\left(\sum_{i=1}^{n} T^{-1/2}\sum_{t=1}^{T}\tilde{v}_{it}\right) = \sum_{i=1}^{n} var\left(T^{-1/2}\sum_{t=1}^{T}\tilde{v}_{it}\right).$$

- Assume $\tilde{v}_{it}$ is covariance stationary. From a time series perspective for a given individual, i,

$$\lim_{T\to\infty} var\left(T^{-1/2}\sum_{t=1}^{T}\tilde{v}_{it}\right) = \Gamma_0^{(i)} + \sum_{j=1}^{\infty}\left(\Gamma_j^{(i)} + \Gamma_j^{(i)\prime}\right) \equiv \Omega_i,$$

$$\Gamma_j^{(i)} = E\left[v_{it}v_{it-j}'\right], \quad v_{it} = x_{it}u_{it}.$$

Standard Errors: No Cross-section Correlation II

- Each Ω_i can be estimated using a nonparametric kernel HAC estimator of the form

$$\widehat{\Omega}_i = \widehat{\Gamma}_0^{(i)} + \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\left(\widehat{\Gamma}_j^{(i)} + \widehat{\Gamma}_j^{(i)\prime}\right),$$

where

$$\widehat{\Gamma}_j^{(i)} = T^{-1} \sum_{t=j+1}^{T} \widehat{v}_{it}\widehat{v}_{it-j}', \qquad \widehat{v}_{it} = \tilde{x}_{it}\widehat{u}_{it} = \widehat{v}_{it} = \tilde{x}_{it}(\tilde{y}_{it} - \tilde{x}_{it}'\widehat{\beta}).$$

- $k(x)$ is a kernel weighting function and M is the bandwidth. A larger M puts more weight on higher order sample autocovariances.
- Adding the $\widehat{\Omega}_i$ across individuals leads to the estimator:

$$\sum_{i=1}^{n} \widehat{\Omega}_i.$$

Standard Errors: No Cross-section Correlation III

- A special case is the kernel $k(x) = 1_{|x|\leq 1}$. When $M = T$, full weight is put on the sample autocovariances. This gives the usual Arellano (1987) clustered standard errors because it can be shown that in this case

$$\widehat{\Omega}_i = \widehat{\Gamma}_0^{(i)} + \sum_{j=1}^{T-1} \left(\widehat{\Gamma}_j^{(i)} + \widehat{\Gamma}_j^{(i)\prime} \right) = T^{-1} \widehat{S}_{iT} \widehat{S}_{iT}', \qquad \widehat{S}_{iT} = \sum_{t=1}^{T} \widehat{v}_{it}.$$

- It is well known that for fixed-T and large-n, the clustered standard errors are consistent for general forms of serial correlation as long as there is no spatial correlation in the model.

Standard Errors: Allowing Cross-section Correlation I

- Now suppose we want to allow $cov(\tilde{v}_{it}, \tilde{v}_{js}) \neq 0$ for $i \neq j$. This is a very general form of spatial correlation including spatial correlation over time because we can have $t \neq s$.
- Now write
$$var\left(\sum_{i=1}^{n} T^{-1/2} \sum_{t=1}^{T} \tilde{v}_{it}\right) = var\left(T^{-1/2} \sum_{t=1}^{T} \left(\sum_{i=1}^{n} \tilde{v}_{it}\right)\right).$$
- For fixed-n the quantity
$$\overline{\tilde{v}}_t = \sum_{i=1}^{n} \tilde{v}_{it}$$
is simply a vector of stationary time series and the spatial correlation has been aggregated by cross-section averaging.

Standard Errors: Allowing Cross-section Correlation II

- For a given value of n, from the time series perspective we have

$$\lim_{T\to\infty} var\left(T^{-1/2}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}\tilde{v}_{it}\right)\right) = \lim_{T\to\infty} var\left(T^{-1/2}\sum_{t=1}^{T}\bar{\tilde{v}}_{t}\right)$$

$$= \overline{\Omega} = \overline{\Gamma}_0 + \sum_{j=1}^{\infty}\left(\overline{\Gamma}_j + \overline{\Gamma}_j'\right),$$

$$\overline{\Gamma}_j = E\left[\overline{v}_t \overline{v}_{t-j}'\right], \quad \overline{v}_t = \sum_{i=1}^{n} v_{it}$$

and $\overline{v}_t$ will be covariance stationary (across t) if v_{it} is covariance stationary for all i.

Standard Errors: Allowing Cross-section Correlation III

- $\overline{\Omega}$ can be estimated using a nonparametric kernel HAC estimator of the form

$$\widehat{\overline{\Omega}} = \widehat{\overline{\Gamma}}_0 + \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\left(\widehat{\overline{\Gamma}}_j + \widehat{\overline{\Gamma}}_j'\right),$$

$$\widehat{\overline{\Gamma}}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{\overline{v}}_t \widehat{\overline{v}}_{t-j}', \qquad \widehat{\overline{v}}_t = \sum_{i=1}^{n} \widehat{v}_{it} = \sum_{i=1}^{n} \tilde{x}_{it} \widehat{u}_{it}.$$

- Driscoll and Kraay (1998) established the consistency of $\widehat{\overline{\Omega}}$ under large-T and large-n asymptotics for the Bartlett kernel using standard bandwidth rates for M but their results do not apply when fixed-effects dummies (or time period dummies) are included in the model.

Standard Errors: Allowing Cross-section Correlation IV

- Putting full weight on all the sample autocovariances is not an option here because then we would have

$$\widehat{\widehat{\Omega}} = \widehat{\widehat{\Gamma}}_0 + \sum_{j=1}^{T-1} \left(\widehat{\widehat{\Gamma}}_j + \widehat{\widehat{\Gamma}}_j' \right) = T^{-1} \widehat{\widehat{S}}_T \widehat{\widehat{S}}_T' = 0,$$

using

$$\widehat{\widehat{S}}_T = \sum_{t=1}^{T} \widehat{\widehat{v}}_t = \sum_{t=1}^{T} \sum_{i=1}^{n} \widehat{v}_{it} = \sum_{t=1}^{T} \sum_{i=1}^{n} \tilde{x}_{it} \widehat{u}_{it} = 0,$$

Summary of Robust Standard Errors in a Linear Panel I

- Form of the standard errors depends on what can be assumed about the spatial correlation and serial correlation.
- If there is no spatial correlation in the cross-section, use the cross-section average of individual-by-individual HAC estimators.
  - Putting full weight on the sample autocovariances leads to the usual clustered standard errors.
  - These standard errors are robust to general forms of serial correlation and heteroskedasticity.

Summary of Robust Standard Errors in a Linear Panel II

- If there is spatial correlation in the cross-section of unknown form and we have weak time dependence, use a HAC estimator of the cross-section average.
  - These standard errors are robust to general forms of spatial correlation, stationary serial correlation and heteroskedasticity.
  - Cannot put full weight on all the sample autocovariances because this gives an estimator that is identically zero for any data set.
  - Must use a kernel and choose a bandwidth.
  - Consistency results by Driscoll and Kraay (1998) only provide a crude approximation that does not depend on the kernel or bandwidth.

Robust Test Statistics I

- Consider testing the set of linear hypothesis

$$H_0 : R\beta = r,$$

  where R is a $q \times k$ matrix of know constants with full rank and $q \leq k$ and r is a $q \times 1$ vector of known constants.
- Focus on $q = 1$ and t-statistics ($q > 1$ and Wald statistics are in the paper).

Robust Test Statistics II

$$t_{clus(n)} = \frac{(R\widehat{\beta} - r)}{\sqrt{R\widehat{V}_{clus(n)}R'}}$$

$$\widehat{V}_{clus(n)} = T\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1}\left(\sum_{i=1}^{n}T^{-1}\widehat{S}_{iT}\widehat{S}_{iT}'\right)\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{x}_{it}\widetilde{x}_{it}'\right)^{-1}$$

$$t_{ave(n)} = \frac{(R\widehat{\beta} - r)}{\sqrt{R\widehat{V}_{ave(n)}R'}}$$

$$\widehat{V}_{ave(n)} = T\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}\left(\sum_{i=1}^{n}\widehat{\Omega}_{i}\right)\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}$$

Robust Test Statistics III

$$t_{HACSC} = \frac{(R\widehat{\beta} - r)}{\sqrt{R\widehat{V}_{HACSC}R'}}$$

$$\widehat{V}_{HACSC} = T\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}\widehat{\widehat{\Omega}}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}$$

Basic Ideas Behind Fixed-b Asymptotics I

- The HAC estimators are of the form

$$\widehat{\Omega} = \widehat{\Gamma}_0 + \sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\left(\widehat{\Gamma}_j + \widehat{\Gamma}_j'\right)$$

$$\widehat{\Omega}_{Bart} = \widehat{\Gamma}_0 + \sum_{j=1}^{M-1} \left(1 - \frac{j}{M}\right)\left(\widehat{\Gamma}_j + \widehat{\Gamma}_j'\right)$$

- The choice of $k(x)$ and M affect $\widehat{\Omega}$ and an empirical researcher can fiddle with these choices and move $\widehat{\Omega}$ around.
- Problem: the typical asymptotic approximation (consistency) doesn't reflect these choices.
- By assuming $M = bT$ where $b \in (0, 1]$ is held fixed as T increases, an asymptotic theory can be built for $\widehat{\Omega}$ that reflects the choice of kernel and M.

Basic Ideas Behind Fixed-b Asymptotics II

- This changes the asymptotic distribution of t statistics from $N(0,1)$ to something like a t random variable: the fixed-b asymptotic distribution of t will be a ratio of a $N(0,1)$ and the square root of a "chi-square" type random variable. The "chi-square" type random variable will depend on $k(x)$ and b.
- In practice one would use critical values from this nonstandard random variable for $b = M/T$.
- Using a different value of M for a given value of T, changes "b" and changes the critical value of the t-test.

Some Definitions I

Let $B_h(r)$ denote a generic $h \times 1$ vector of stochastic processes. Let the random matrix, $P(b, B_h)$, be defined as follows for $b \in (0, 1]$.

- Case (iii): if $k(x)$ is the Bartlett kernel, $k(x) = 1 - |x|$ for $|x| \leq 1$ and $k(x) = 0$ for $|x| \geq 1$,

$$P(b, B_h) = \frac{2}{b}\int_0^1 B_h(r)B_h(r)'dr$$

$$-\frac{1}{b}\int_0^{1-b}\left(B_h(r+b)B_h(r)' + B_h(r)B_h(r+b)'\right)dr$$

$$-\frac{1}{b}\int_{1-b}^{1}\left(B_h(1)B_h(r)' + B_h(r)B_h(1)'\right)dr + B_h(1)B_h(1)'.$$

- Let I_k denote a $k \times k$ identity matrix.
- Let ι denote an $n \times 1$ vector of ones. Let e_i denote an $n \times 1$ vector with i^{th} element equal to one an zeros otherwise, i.e.

$$e_i = (0, 0, ...0, 1, 0, ..., 0)'.$$

Large–T, Fixed-n Assumptions I

Sufficient assumptions when individual fixed-effects dummies are in model:

1. $p\lim T^{-1}\sum_{t=1}^{[rT]}\tilde{x}_{it}\tilde{x}_{it}' = rQ_i$ for $r \in (0,1]$ and Q_i^{-1} exists.
2. $E(u_{it}|x_{it}) = 0$.
3. $T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \Lambda W(r)$, where $W(r)$ is an $nk \times 1$ vector of standard Wiener processes and $\Lambda\Lambda'$ is the $nk \times nk$ long run variance matrix of the mean zero $nk \times 1$ vector

$$v_t = (v_t^{11\prime}, v_t^{22\prime}, ..., v_t^{nn\prime})',$$

where v_t^{ii} is a $k \times 1$ vector defined as

$$v_t^{ii} = (x_{it} - \mu_i)u_{it}, \qquad \mu_i = E(x_{it}).$$

Theoretical Results: Individual Dummies Only I

- (Theorem) Suppose the model is estimated with individual dummies but no time period dummies. Let $\tilde{B}_k^i(r)$ denote a $k \times 1$ vector of stochastic processes defined as

$$\tilde{B}_k^i(r) = B_k^i(r) - rQ_i \left(\sum_{j=1}^{n} Q_j \right)^{-1} \sum_{j=1}^{n} B_k^j(1),$$

  where $B_k^i(r) = A_i \Lambda W(r)$ and $A_i = (e_i' \otimes I_k)(e_i' \otimes I_{nk})$. Let $W_q^*(r)$ denote a $q \times 1$ vector of independent standard Wiener processes and define $\tilde{W}_q^*(r) = W_q^*(r) - rW_q^*(1)$ to be a $q \times 1$ vector of standard Brownian bridges.

Theoretical Results: Individual Dummies Only II

Under fixed-b asymptotics with n fixed and $T \to \infty$:

$$\sqrt{T}(\widehat{\beta} - \beta) \Rightarrow \left(\sum_{i=1}^{n} Q_i \right)^{-1} \sum_{i=1}^{n} B_k^i(1),$$

$$t_{clus(n)} \Rightarrow \frac{R \left(\sum_{i=1}^{n} Q_i \right)^{-1} \sum\limits_{i=1}^{n} B_k^i(1)}{\sqrt{R \left(\sum_{i=1}^{n} Q_i \right)^{-1} \sum\limits_{i=1}^{n} \tilde{B}_k^i(1) \tilde{B}_k^i(1)' \left(\sum_{i=1}^{n} Q_i \right)^{-1} R'}},$$

$$t_{ave(n)} \Rightarrow \frac{R \left(\sum_{i=1}^{n} Q_i \right)^{-1} \sum\limits_{i=1}^{n} B_k^i(1)}{\sqrt{R \left(\sum_{i=1}^{n} Q_i \right)^{-1} \sum\limits_{i=1}^{n} P(b, \tilde{B}_k^i) \left(\sum_{i=1}^{n} Q_i \right)^{-1} R'}},$$

Theoretical Results: Individual Dummies Only III

$$t_{HACSC} \Rightarrow \frac{W_1^*(1)}{\sqrt{P(b, \tilde{W}_1^*)}}, \quad t_{HACSC}^{BCH} \Rightarrow \frac{W_1^*(1)}{\sqrt{P(\lambda_0, \lambda_1 .., \lambda_G, \tilde{W}_1^*)}}.$$

- (Corollary) The asymptotic results for the *HACSC* statistics are unchanged when individual fixed-effects dummies are not included in the model and x_t contains an intercept.

Theoretical Results: Individual Dummies Only IV

- (Corollary) Suppose there is random sampling in the cross-section. The results simplify as

$$t_{clus} \Rightarrow \sqrt{\frac{n}{n-1}} t_{n-1},$$

$$t_{ave(n)} \Rightarrow n\overline{W}_q(1) / \sqrt{\sum_{i=1}^{n} P_1(b, \widehat{W}_q^i(r))},$$

where $W_q^i(r)$ are $q \times 1$ vectors of independent Wiener processes that are independent of each other and $\widehat{W}_q^i(r) = W_q^i(r) - r\overline{W}_q(1)$ and $\overline{W}_q(1) = \frac{1}{n}\sum_{j=1}^{n} W_q^j(1)$.

Including Time Period Dummies Requires Stronger Assumptions I

1. $E(u_{it}|x_{jt}) = 0$ for all $i, j = 1, 2, ..., n$.
2. $T^{-1/2} \sum_{t=1}^{[rT]} v_t^{ex} \Rightarrow \Lambda^{ex} W^{ex}(r)$, where $W^{ex}(r)$ is an $n^2 k \times 1$ vector of standard Wiener processes and $\Lambda^{ex} \Lambda^{ex\prime}$ is the $n^2 k \times n^2 k$ long run variance matrix of the mean zero $n^2 k \times 1$ vector

$$v_t^{ex} = (v_1^{1\prime}, v_t^{2\prime}, ..., v_t^{n\prime})'$$

where

$$v_t^j = (v_t^{1j\prime}, v_t^{2j\prime}, ..., v_t^{nj\prime})'$$

and

$$v_t^{ij} = (x_{it} - \mu_i) u_{jt}.$$

- The "*ex*" superscript denotes an extended vector that includes vectors v_t^{ij} with $i \neq j$.

Theoretical Results: Individual and Time Period Dummies I

- (Theorem) Suppose the model is estimated with individual and time period dummies. Under fixed-b asymptotics, for n fixed and $T \to \infty$:
  1. The limits of the $HACSC$ statistics are the same as before.
  2. The limits of $\sqrt{T}(\widehat{\beta} - \beta)$ and the $clus(n)$ and $ave(n)$ statistics take the same form as before except the definition of $B_k^i(r)$ changed to be

$$B_k^i(r) = A_i^{ex} \Lambda^{ex} W^{ex}(r),$$

where A_i^{ex} is a nonstochastic matrix given by

$$A_i^{ex} = A_{ii} - \frac{1}{n} \sum_{j=1}^{n} A_{ij}, \text{ where } A_{ij} = \left[(e_i' \otimes I_k) - \frac{1}{n} (\iota' \otimes I_k) \right] (e_j' \otimes I_{nk}).$$

Theoretical Results: Individual and Time Period Dummies II

- (Corollary) Suppose there is random sampling in the cross-section. The asymptotic distributions simplify as follows:

$$t_{clus(n)} \Rightarrow \sqrt{\frac{n}{n-2}} t_{n-1} \equiv \sqrt{\frac{(n-1)}{(n-2)} \frac{n}{(n-1)}} t_{n-1},$$

$$t_{ave(n)} \Rightarrow n\overline{\tilde{W}}_1^*(1) / \sqrt{\sum_{i=1}^{n} P_1(b, \widehat{\tilde{W}}_1^{*ii})},$$

where

$$\widehat{\tilde{W}}_q^{*ii}(r) = \tilde{W}_q^{*ii}(r) - r\overline{\tilde{W}}_q^*(1), \quad \overline{\tilde{W}}_q^*(1) = \frac{1}{n} \sum_{i=1}^{n} \tilde{W}_q^{*ii}(1),$$

Theoretical Results: Individual and Time Period Dummies III

with

$$\tilde{W}_q^{*ii}(r) = \tilde{W}_q^{ii}(r) - \frac{1}{n}\sum_{l=1}^{n}\tilde{W}_q^{li}(r),$$

$$\tilde{W}_q^{ij}(r) = W_q^{ij}(r) - \frac{1}{n}\sum_{l=1}^{n}W_q^{il}(r),$$

and $W_q^{ij}(r)$ are $q \times 1$ vectors of independent Wiener processes that are independent of each other.

Large–T, Large-n Assumptions

- To obtain results in the large-T, large-n case use theory of random fields.
- A random field is a stochastic process indexed by a vector and includes pure time series as a special case when the index is a scalar. Panel data is naturally indexed by the vector (i, t).

Sufficient Assumptions With Individual Dummies

1. $p\lim \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{[rT]} \tilde{x}_{it}\tilde{x}_{it}' = rQ, \quad Q^{-1}$ exists.
2. $E(u_{it}|x_{is}) = 0$ for all t, s.
3. $\frac{1}{\sqrt{nT}} \sum_{i=1}^{n} \sum_{t=1}^{[rT]} (x_{it} - \mu_i)u_{it} \Rightarrow \Lambda W_k(1, r)$, where $\mu_i = E(x_{it})$ and $W_k(1, r)$ is a $k \times 1$ vector of standard Brownian sheets.
4. The process $(x_{it} - \mu_i)u_{is}$ is a mean zero vector of three dimensional stationary random fields indexed by i, t, s.

Theoretical Results

- (Theorem) Suppose the model is estimated with individual fixed-effects but no time period dummies are included. As $n, T \to \infty$

$$\sqrt{nT}(\widehat{\beta} - \beta) \Rightarrow Q^{-1}\Lambda W_k(1,1) \sim N\left(\mathbf{0}, Q^{-1}\Lambda\Lambda' Q^{-1}\right),$$

$$t_{HACSC} \Rightarrow \frac{W_1^*(1)}{\sqrt{P(b, \tilde{W}_1^*)}}.$$

- The limit of t_{HACSC} is the same as when n is fixed.

Stronger Assumptions With Individual and Time Period Dummies

1. $E(u_{it}|x_{js}) = 0$ for all i, j and t, s.
2. The process $(x_{it} - \mu_i)u_{js}$ is a mean zero vector of four dimensional stationary random fields indexed by i, j, t, s.

Theoretical Results

- (Theorem) Suppose the model is estimated with individual fixed-effects and time period dummies. As $n, T \to \infty$

$$\sqrt{nT}(\widehat{\beta} - \beta) \Rightarrow Q^{-1}\Lambda W_k(1,1) \sim N\left(\mathbf{0}, Q^{-1}\Lambda\Lambda' Q^{-1}\right),$$

$$t_{HACSC} \Rightarrow \frac{W_1^*(1)}{\sqrt{P(b, \tilde{W}_1^*)}}.$$

- The asymptotic distribution of t_{HACSC} is the same as when only individual fixed-effects dummies are included.

Summary of Asymptotic Results I

- The usual clustered standard errors are not valid when there is spatial correlation in the model.
- The *HACSC* robust standard errors have the "standard" fixed-b asymptotic distributions as derived by Kiefer and Vogelsang (2005) for time series models.
    - This result holds whether individual dummies and/or time period dummies are included.
    - This result holds for large-T and both fixed-n and large-n.

Summary of Asymptotic Results II

- But .... the type of dummies included and the treatment of n affect the exogeneity assumptions needed and affect assumptions about the stationarity of the spatial correlation:
  - Large-T, fixed-n, individual dummies only: weak exogeneity in cross-section and over time, spatial correlation unconstrained
  - Large-T, fixed-n, individual and time period dummies: strict exogeneity in cross-section, weak exogeneity over time, spatial correlation unconstrained
  - Large-T, large-n, individual dummies only: weak exogeneity in cross-section, strict exogeneity over time and spatial correlation stationary
  - Large-T, large-n, individual and time period dummies: strict exogeneity in both directions and spatial correlation stationary
- Fixed-b critical values and p-values for the HACSC statistics based on the Bartlett kernel are easily computed using the numerical approach given in Appendix B of the paper.

Finite Sample Performance I

- Assess the asymptotic approximations using a simple but relevant Monte Carlo simulation.
- The DGP:

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \beta_4 x_{4it} + u_{it},$$

where

$$u_{it} = \rho u_{i,t-1} + \lambda \eta_t + \varepsilon_{it}, \quad u_{i0} = 0, \quad \varepsilon_{it} \sim N(0,1), \quad \eta_t \sim N(0,1),$$

$$\begin{aligned} cov(\varepsilon_{it}, \varepsilon_{js}) &= 0, \ t \neq s, \\ cov(\eta_t, \eta_s) &= 0 \ for \ t \neq s, \\ cov(\eta_t, \varepsilon_{js}) &= 0 \ \forall t, s, j. \end{aligned}$$

Finite Sample Performance II

For $l = 1, 2, 3, 4$

$$x_{lit} = \rho x_{li,t-1} + \lambda \upsilon_{lt} + e_{lit}, \quad x_{li0} = 0, \quad e_{lit} \sim N(0,1), \quad \upsilon_{lt} \sim N(0,1),$$

$$cov(e_{it}, e_{js}) = 0 \text{ for } i \neq j \text{ and/or } t \neq s,$$
$$cov(\upsilon_{lt}, \upsilon_{ls}) = 0 \text{ for } t \neq s,$$
$$cov(\upsilon_{lt}, e_{ljs}) = 0 \ \forall t, s, j,$$

- The idiosyncratic errors, ε_{it}, are modeled in the cross-section as spatial MA(2) errors on a square grid with MA parameters θ and θ^2.
- When $\lambda \neq 0$, there is equi-spatial correlation, ρ_{spat}, between individuals in the cross-section in both the regressors and the idiosyncratic error. The value of that correlation is $\rho_{spat} = \lambda^2/(1+\lambda^2)$.
- When $\theta = 0$ and $\lambda = 0$, there is no spatial correlation in the model.

Finite Sample Performance III

- The null hypothesis is $H_0 : \beta_1 = 0$. Without loss of generality set $\beta_2 = 0$, $\beta_3 = 0$, $\beta_4 = 0$ and $a_i = 0 \ \forall i$.
- Results are given for $n = 10, 50, 250$ and $T = 10, 50, 250$.
- 2,000 replications were used in all cases. The nominal level is 0.05.
- Results are reported for the Bartlett kernel for a selection of bandwidths.

Table 2: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
10	10	.0	.134	.167	.223	.311	.390	.110	.091	.086	.087	.082	.083	.063
		.3	.167	.206	.259	.356	.414	.112	.120	.112	.106	.102	.102	.071
		.6	.238	.270	.339	.420	.496	.128	.194	.167	.163	.149	.149	.081
		.9	.363	.368	.436	.501	.563	.150	.311	.256	.234	.219	.218	.106
	50	.0	.097	.123	.199	.281	.352	.096	.068	.055	.050	.045	.050	.061
		.3	.103	.139	.209	.286	.352	.094	.068	.064	.060	.060	.061	.061
		.6	.141	.178	.244	.323	.400	.095	.100	.090	.086	.081	.086	.063
		.9	.311	.313	.385	.460	.535	.115	.253	.190	.173	.168	.165	.082
	250	.0	.082	.116	.178	.270	.349	.083	.049	.049	.049	.047	.047	.051
		.3	.084	.111	.174	.271	.337	.079	.053	.050	.045	.047	.048	.049
		.6	.087	.119	.186	.268	.332	.081	.060	.053	.046	.044	.047	.048
		.9	.140	.170	.225	.318	.394	.082	.096	.086	.079	.078	.077	.056

Table 2: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
50	10	.0	.120	.156	.225	.312	.382	.059	.088	.077	.072	.073	.074	.056
		.3	.155	.176	.256	.341	.411	.067	.111	.108	.100	.094	.096	.059
		.6	.231	.242	.326	.409	.491	.071	.175	.153	.147	.133	.134	.062
		.9	.361	.366	.435	.499	.560	.068	.302	.254	.234	.210	.209	.062
	50	.0	.090	.123	.190	.282	.344	.059	.056	.058	.058	.058	.060	.053
		.3	.099	.128	.191	.274	.347	.055	.061	.060	.056	.059	.053	.050
		.6	.119	.145	.205	.298	.372	.055	.081	.075	.063	.064	.064	.048
		.9	.284	.281	.340	.430	.489	.059	.231	.171	.162	.150	.152	.053
	250	.0	.082	.110	.168	.259	.335	.057	.047	.051	.055	.050	.051	.048
		.3	.081	.113	.172	.263	.331	.057	.056	.049	.050	.050	.051	.051
		.6	.093	.126	.190	.284	.352	.057	.057	.057	.062	.059	.057	.051
		.9	.142	.173	.240	.339	.401	.056	.104	.085	.080	.079	.080	.052

Table 2: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
250	10	.0	.109	.144	.201	.280	.353	.045	.070	.062	.058	.061	.059	.043
		.3	.129	.163	.233	.320	.393	.048	.091	.075	.081	.077	.080	.045
		.6	.209	.236	.306	.391	.463	.052	.165	.144	.128	.128	.133	.049
		.9	.339	.335	.405	.473	.532	.058	.286	.245	.218	.202	.199	.055
	50	.0	.088	.121	.176	.265	.340	.048	.055	.052	.053	.051	.054	.046
		.3	.103	.138	.202	.272	.337	.058	.064	.060	.061	.058	.060	.055
		.6	.140	.169	.238	.326	.382	.059	.097	.084	.082	.079	.081	.055
		.9	.306	.303	.372	.446	.510	.044	.239	.178	.171	.173	.171	.041
	250	.0	.073	.104	.172	.262	.340	.044	.048	.042	.041	.041	.039	.043
		.3	.084	.115	.168	.254	.327	.051	.052	.049	.043	.044	.046	.044
		.6	.092	.120	.175	.261	.336	.048	.060	.058	.057	.053	.055	.047
		.9	.130	.157	.231	.302	.373	.043	.094	.080	.077	.076	.080	.041

Table 3: Empirical Null Rejection Probabilities, 5% level, Bartlett Kernel, MA(2) Spatial correlation in cross-section, $\lambda = 0$, $\theta = 0.5$. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
9	10	.0	.413	.413	.421	.428	.430	.423	.154	.133	.118	.118	.119
		.3	.425	.426	.426	.420	.425	.408	.172	.154	.136	.136	.137
		.6	.469	.467	.463	.451	.449	.414	.229	.198	.187	.185	.186
		.9	.530	.516	.508	.491	.473	.411	.321	.288	.260	.240	.241
	50	.0	.382	.389	.401	.422	.433	.456	.070	.065	.059	.064	.065
		.3	.393	.401	.408	.416	.426	.441	.081	.084	.081	.074	.079
		.6	.434	.430	.434	.443	.446	.437	.120	.102	.099	.100	.095
		.9	.535	.501	.485	.466	.460	.423	.272	.227	.209	.197	.200
	250	.0	.383	.393	.406	.417	.434	.469	.056	.051	.042	.049	.049
		.3	.398	.406	.421	.434	.446	.478	.058	.052	.049	.049	.049
		.6	.411	.416	.436	.450	.458	.475	.064	.063	.063	.060	.062
		.9	.462	.452	.449	.455	.459	.471	.122	.118	.107	.106	.108

Table 3: Empirical Null Rejection Probabilities, 5% level, Bartlett Kernel, MA(2) Spatial correlation in cross-section, $\lambda = 0$, $\theta = 0.5$. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
49	10	.0	.526	.526	.523	.525	.525	.523	.110	.097	.093	.092	.091
		.3	.537	.537	.533	.529	.527	.517	.133	.120	.108	.112	.114
		.6	.584	.569	.561	.547	.538	.513	.196	.164	.162	.158	.162
		.9	.637	.602	.585	.566	.548	.498	.297	.246	.230	.215	.212
	50	.0	.469	.469	.474	.480	.485	.500	.061	.064	.060	.061	.065
		.3	.489	.489	.493	.495	.494	.498	.072	.070	.065	.069	.070
		.6	.522	.513	.510	.506	.504	.502	.103	.095	.090	.088	.094
		.9	.631	.601	.581	.554	.542	.508	.270	.219	.190	.184	.186
	250	.0	.466	.469	.473	.480	.484	.496	.043	.045	.045	.048	.049
		.3	.480	.484	.485	.492	.493	.503	.060	.056	.051	.050	.051
		.6	.482	.482	.481	.484	.490	.501	.069	.061	.059	.058	.064
		.9	.538	.526	.518	.513	.516	.506	.124	.107	.092	.091	.096

Table 3: Empirical Null Rejection Probabilities, 5% level, Bartlett Kernel, MA(2) Spatial correlation in cross-section, $\lambda = 0$, $\theta = 0.5$. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
256	10	.0	.546	.544	.544	.543	.538	.533	.085	.081	.072	.070	.074
		.3	.564	.562	.554	.540	.530	.512	.109	.099	.098	.094	.097
		.6	.624	.609	.598	.574	.555	.520	.171	.143	.135	.131	.132
		.9	.688	.657	.636	.608	.587	.525	.301	.244	.216	.205	.199
	50	.0	.525	.523	.523	.523	.521	.518	.056	.050	.051	.056	.053
		.3	.532	.532	.528	.526	.525	.522	.061	.064	.065	.062	.065
		.6	.566	.554	.543	.536	.532	.519	.087	.079	.074	.072	.077
		.9	.679	.644	.625	.594	.577	.533	.251	.183	.175	.168	.168
	250	.0	.505	.506	.509	.508	.510	.506	.047	.046	.049	.043	.047
		.3	.500	.501	.502	.503	.504	.503	.050	.048	.050	.049	.048
		.6	.510	.506	.505	.505	.505	.507	.057	.053	.050	.048	.051
		.9	.559	.542	.533	.524	.518	.506	.093	.081	.076	.073	.073

Table 9: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$). Fixed-b, F Critical Values.

			$H_0 : \beta_1 = 0, \beta_2 = 0$						$H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0.$					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
9	10	.0	.217	.186	.170	.169	.174	.591	.370	.322	.328	.314	.309	.771
		.3	.254	.225	.205	.204	.205	.578	.413	.369	.356	.354	.354	.770
		.6	.334	.289	.268	.257	.264	.582	.521	.458	.441	.441	.438	.755
		.9	.466	.406	.362	.361	.358	.580	.704	.624	.594	.601	.602	.744
	50	.0	.086	.073	.065	.064	.061	.623	.088	.084	.085	.086	.082	.778
		.3	.095	.083	.076	.075	.075	.618	.114	.104	.097	.099	.100	.786
		.6	.154	.129	.112	.111	.111	.620	.204	.167	.183	.182	.178	.762
		.9	.405	.331	.289	.291	.291	.606	.581	.490	.491	.484	.487	.755
	250	.0	.058	.050	.057	.055	.055	.654	.051	.052	.052	.049	.048	.795
		.3	.065	.063	.062	.060	.061	.665	.059	.053	.057	.053	.052	.777
		.6	.077	.074	.070	.071	.072	.661	.077	.073	.076	.071	.069	.786
		.9	.168	.146	.145	.138	.139	.644	.234	.206	.210	.206	.203	.775

Table 9: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$). Fixed-b, F Critical Values.

			$H_0: \beta_1 = 0, \beta_2 = 0$						$H_0: \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0.$					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
49	10	.0	.143	.126	.123	.115	.122	.717	.257	.227	.233	.222	.217	.907
		.3	.182	.160	.145	.150	.152	.706	.292	.261	.260	.257	.250	.911
		.6	.278	.232	.220	.217	.219	.706	.455	.394	.383	.380	.375	.908
		.9	.453	.366	.339	.331	.329	.727	.693	.590	.582	.586	.583	.910
	50	.0	.061	.062	.063	.057	.059	.699	.070	.069	.060	.064	.060	.886
		.3	.082	.081	.077	.076	.074	.705	.087	.077	.082	.076	.077	.891
		.6	.135	.112	.107	.102	.106	.708	.163	.144	.143	.141	.135	.900
		.9	.365	.280	.258	.257	.258	.712	.551	.446	.445	.445	.444	.896
	250	.0	.051	.047	.046	.050	.048	.699	.053	.053	.057	.058	.058	.894
		.3	.057	.052	.051	.051	.049	.694	.064	.059	.060	.057	.060	.896
		.6	.067	.064	.059	.061	.062	.700	.069	.060	.063	.064	.064	.887
		.9	.152	.124	.118	.112	.117	.717	.181	.160	.155	.154	.160	.897

Table 9: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$). Fixed-b, F Critical Values.

			$H_0 : \beta_1 = 0, \beta_2 = 0$						$H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0.$					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
256	10	.0	.110	.098	.095	.089	.087	.711	.192	.170	.168	.154	.154	.897
		.3	.146	.136	.121	.115	.118	.723	.234	.213	.207	.198	.196	.900
		.6	.253	.210	.196	.192	.187	.723	.400	.320	.324	.317	.306	.900
		.9	.453	.353	.320	.311	.311	.739	.658	.557	.538	.546	.540	.912
	50	.0	.057	.057	.053	.055	.053	.730	.062	.060	.059	.060	.059	.917
		.3	.073	.072	.067	.065	.067	.736	.083	.089	.080	.077	.078	.914
		.6	.121	.106	.098	.105	.102	.738	.155	.144	.142	.142	.141	.916
		.9	.344	.247	.231	.237	.236	.716	.513	.408	.410	.411	.405	.907
	250	.0	.052	.051	.053	.051	.051	.718	.051	.049	.054	.053	.049	.913
		.3	.057	.059	.058	.059	.059	.713	.058	.055	.053	.058	.053	.909
		.6	.068	.058	.061	.063	.065	.717	.072	.067	.072	.076	.070	.894
		.9	.122	.102	.099	.092	.095	.712	.170	.154	.152	.152	.150	.908

Summary of the Simulations I

- The $N(0,1)$ approximation is not too horrible for $t_{ave(n)}$ and $t_{clus(n)}$ and would be better for larger n.
- The $N(0,1)$ approximation is often quite bad for t_{HACSC} and becomes progressively worse as the bandwidth is increased.
- The fixed-b approximation is much better overall than the $N(0,1)$ approximation especially for t_{HACSC}. The improvement of fixed-b over the chi-square approximation for $Wald_{HACSC}$ can be quite substantial.
- The stronger the serial correlation (closer ρ is to 1), the larger T needs to be for the asymptotic approximations to be accurate. The value of n does not play a big role.
- When there is spatial correlation, $t_{ave(n)}$ and $t_{clus(n)}$ substantially over-reject. The problem becomes more acute as n increases.

Summary of the Simulations II

- t_{HACSC} handles the spatial correlation well if T isn't too small relative to the strength of the serial correlation. Using a larger bandwidth reduces the over-rejection problem. The stronger the serial correlation, the larger T needs to be to avoid over-rejection problems (standard finding in pure time series settings). The value of n does not play a large role for t_{HACSC}.
- When there is no spatial correlation, $t_{ave(n)}$ and t_{clus} are more powerful than t_{HACSC}. Power of $t_{ave(n)}$ and t_{HACSC} decreases as the bandwidth increases. Power of t_{clus} is lower than $t_{ave(n)}$ regardless of the bandwidth used for $t_{ave(n)}$.
- When there is spatial correlation, power of t_{HACSC} decreases as the bandwidth increases.

Conclusions

- The popular Arellano (1987) clustered standard errors are not valid when there is spatial correlation.
- *HACSC* robust standard errors used in conjunction with fixed-b critical values handle spatial correlation pretty well in practice provided the time dimension sample size is large enough relative to the strength of the serial correlation.
- Empirical researchers should use *HACSC* robust standard errors and fixed-b critical values whenever spatial correlation might be present. They should keep in mind the exogeneity assumptions that are needed and if the cross-section sample size is large (whatever that means) the spatial correlation needs to be stationary. Finally, if the serial correlation is too strong relative to the time dimension sample size, the tests will not be reliable.

Computation of Fixed-b Critical Values and p-values: Bartlett Kernel I

- Let $cv_{\alpha}(b)$ denote the right tail fixed-b asymptotic critical value for a given statistic where $b = M/T$.
- Let z_{α} denote the right tail critical value for the standard asymptotic distribution of the statistic.
- Consider the polynomial

$$cv_{\alpha}(b) = z_{\alpha} + \delta(b) + \gamma(b) z_{\alpha} + \theta(b) z_{\alpha}^2$$

where

$$\begin{aligned} \delta(b) &= \delta_1 b + \delta_2 b^2 + \delta_3 b^3 \\ \gamma(b) &= \gamma_1 b + \gamma_2 b^2 + \gamma_3 b^3 \\ \theta(b) &= \theta_1 b + \theta_2 b^2 + \theta_3 b^3. \end{aligned}$$

Computation of Fixed-b Critical Values and p-values: Bartlett Kernel II

- By construction $cv_\alpha(0) = z_\alpha$ so that the fixed-b critical values simplify to the standard critical values as b approaches zero.
- By plugging in the formulas for the b polynomials, rewrite $cv_\alpha(b)$ as a regression-type relationship in terms of polynomials of b, z_α and their interactions as

$$\begin{aligned} cv_\alpha(b) &= z_\alpha + \delta_1 b + \delta_2 b^2 + \delta_3 b^3 + \gamma_1 b z_\alpha + \gamma_2 b^2 z_\alpha + \gamma_3 b^3 z_\alpha \\ &\quad + \theta_1 b z_\alpha^2 + \theta_2 b^2 z_\alpha^2 + \theta_3 b^3 z_\alpha^2. \end{aligned}$$

- Estimate this regression with OLS using simulated asymptotic critical values over a grid of b values.
- Because the relationship between $cv_\alpha(b)$ and z_α is quadratic, it is easy to obtain a p-value function by inverting the formula for $cv_\alpha(b)$.

Computation of Fixed-b Critical Values and p-values: Bartlett Kernel III

- The inverse functional relationship is given by

$$z\left(cv(b)\right) = \frac{-\left(1+\gamma(b)\right)+\sqrt{\left(1+\gamma(b)\right)^2+4\theta(b)\left(cv(b)-\delta(b)\right)}}{2\theta(b)}.$$

- Let Z denote the standard asymptotic random variable, i.e. χ_q^2. The fixed-b asymptotic p-value is given by

$$P(W_q^*(1)'P(b,\widetilde{W}_q^*)^{-1}W_q^*(1) > x)$$

or equivalently

$$P(Z > z(x)).$$

- The paper provides a table of the estimated coefficients for $cv_\alpha(b)$ for the t-statistic and the *Wald* statistic for testing up to 50 restrictions. The fit of the OLS regressions was very close to 1.0 in all cases.

Time Clustered HAC Estimators I

- Bester, Conley and Hansen (2008): An interesting twist on kernel HAC estimators.
- They analyze covariance matrix estimators in general situations where the data can be divided into clusters that are asymptotically independent.
- In the time dimension weak dependence allows the data to be clustered in the following general manner:
  - Divide the time dimension into G contiguous (non-overlapping) exhaustive groups (time clusters) corresponding to the time periods $\{1, ..., [\lambda_1 T]\}$, $\{[\lambda_1 T] + 1, ..., [\lambda_2 T]\}$, ..., $\{[\lambda_{G-1} T] + 1, ..., T\}$ where $0 < \lambda_1 < \lambda_2 < ... < \lambda_{G-1} < 1$.
  - Set $\lambda_0 = 0$ and $\lambda_G = 1$. Let $K_{lj} = 1$ if time periods l and j are in the same cluster and let $K_{lj} = 0$ otherwise.

Time Clustered HAC Estimators II

- This leads to the HAC estimators

$$\widehat{\Omega}_i = T^{-1} \sum_{l=1}^{T} \sum_{j=1}^{T} K_{lj} \widehat{v}_{il} \widehat{v}_{ij}' = T^{-1} \sum_{g=1}^{G} \left[\left(\sum_{t=[\lambda_{g-1}T]+1}^{[\lambda_g T]} \widehat{v}_{it} \right) \left(\sum_{t=[\lambda_{g-1}T]+1}^{[\lambda_g T]} \widehat{v}_{it}' \right) \right],$$

$$\widehat{\widehat{\Omega}} = T^{-1} \sum_{l=1}^{T} \sum_{j=1}^{T} K_{lj} \widehat{\widehat{v}}_l \widehat{\widehat{v}}_j' = T^{-1} \sum_{g=1}^{G} \left[\left(\sum_{t=[\lambda_{g-1}T]+1}^{[\lambda_g T]} \widehat{\widehat{v}}_t \right) \left(\sum_{t=[\lambda_{g-1}T]+1}^{[\lambda_g T]} \widehat{\widehat{v}}_t' \right) \right].$$

Time Clustered HAC Estimators III

- These estimators can be viewed as a generalizations of the variance covariance matrix estimator used by Gotze and Kunsch (1996) in the block bootstrap.

Choosing the Bandwidth

- How should the bandwidth be chosen in practice?
- There is a trade-off between accuracy of the null approximation (type I error) and power (type II error).
- The panel literature has been willing, implicitly, to trade power for more accuracy of the null approximation given the nearly exclusive use of t_{clus}.
- Why not use $M = T$ when using t_{HACSC}? If n is relatively large, power is probably less of an issue than the over-rejection problem and $M = T$ is best for controlling the over-rejection problem caused by strong serial correlation.

Empirical Application: Divorce Rates I

- Re-examine some of the empirical findings of Wolfers (2006, AER) on the relationship between unilateral divorce laws and divorce rates.
- The focus is on column 1 of panel A from Table 4 in Wolfers (2006).
- Panel of state-level annual data from 1956-1988 was used to estimate the dynamic change over time of divorce rates as a function of the time since a state adopted no-fault unilateral divorce laws.
- The dependent variable is the divorce rate per 1000 persons for a given state in a given year.
- The regressors are dummy variables for the number of years the law change had been in effect for a given year in a given state.
- The model includes state and time fixed-effects.

Empirical Application: Divorce Rates II

- The OLS estimates (using state population weights) indicate that divorce rates rose soon after the unilateral divorce laws were passed but within a decade, divorce rates had fallen and continued to fall over time.
- According to OLS standard errors, the estimates are fairly precise and all coefficients are statistically significant at the 5% level.
- Robust standard errors can be implemented in Stata using the cluster command (cluster on state to compute the Arellano standard errors) and the xtscc command to compute the *HACSC* standard errors using the Bartlett kernel.

Table 10: Empirical Application: Divorce Rates, U.S. State Level Annual Data 1956-1988, Dependent Variable is Divorce Rate per 1,000 Persons per Year

Panel A: Estimates and Standard Errors (includes State and Year fixed-effects)

Law Change in Effect for:	OLS Estimate	OLS se	$clus(n)$ se	*HACSC* Robust se $M=3$	$M=7$	$M=17$	$M=33$
1-2 years	.267	.085	.188	.144	.114	.075	.054
3-4 years	.210	.085	.159	.090	.079	.046	.035
5-6 years	.164	.085	.171	.060	.057	.028	.024
7-8 years	.158	.084	.174	.043	.039	.021	.016
9-10 years	-.121	.084	.163	.039	.043	.037	.026
11-12 years	-.324	.083	.180	.043	.045	.037	.024
13-14 years	-.461	.084	.199	.054	.051	.043	.028
15+ years	-.507	.080	.233	.042	.047	.037	.029

Table 10: Empirical Application: Divorce Rates, U.S. State Level Annual Data 1956-1988, Dependent Variable is Divorce Rate per 1,000 Persons per Year

Panel B: 95% Confidence Intervals with *HACSC* Robust se

	$N(0,1)$ Critical Values				Fixed-*b* Critical Values			
	$M=3$	$M=7$	$M=17$	$M=33$	$M=3$	$M=7$	$M=17$	$M=33$
1-2 years	-.02, .55	.04, .49	.12, .41	.16, .37	-.05, .59	-.02, .56	.003, .53	.01, .52
3-4 years	.03, .39	.06, .36	.12, .30	.14, .28	.01, .41	.01, .41	.05, .37	.04, .38
5-6 years	.05, .28	.05, .28	.11, .22	.12, .21	.03, .30	.02, .31	.07, .26	.05, .28
7-8 years	.07, .24	.08, .23	.12, .20	.13, .19	.06, .25	.06, .26	.08, .23	.08, .23
9-10 years	-.20, -.04	-.20, -.04	-.19, -.05	-.17, -.07	-.21, -.03	-.23, -.01	-.25, .01	-.25, .003
11-12 years	-.41, -.24	-.41, -.24	-.40, -.25	-.37, -.28	-.42, -.23	-.44, -.21	-.45, -.19	-.44, -.21
13-14 years	-.57, -.36	-.56, -.36	.54, -.38	-.52, -.41	-.58, -.34	-.59, -.33	-.61, -.31	-.59, -.33
15+ years	-.59, -.42	-.60, -.41	-.58, -.43	-.56, -.45	-.60, -.41	-.63, -.39	-.63, -.38	-.65, -.37

Table 10: Empirical Application: Divorce Rates, U.S. State Level Annual Data 1956-1988, Dependent Variable is Divorce Rate per 1,000 Persons per Year

Panel C: 95% Confidence Interval Lengths

	Using $N(0,1)$ Critical Values				Using Fixed-b Critical Values			
	$M=3$	$M=7$	$M=17$	$M=33$	$M=3$	$M=7$	$M=17$	$M=33$
1-2 years	.564	.447	.294	.212	.644	.582	.527	.515
3-4 years	.353	.310	.180	.137	.402	.403	.323	.334
5-6 years	.235	.223	.110	.094	.268	.291	.197	.229
7-8 years	.169	.152	.082	.063	.192	.199	.148	.153
9-10 years	.153	.169	.145	.102	.174	.220	.260	.248
11-12 years	.169	.176	.145	.094	.192	.230	.260	.229
13-14 years	.212	.200	.169	.110	.241	.260	.302	.267
15+ years	.164	.184	.145	.114	.188	.240	.260	.277

Table 1: Empirical Null Rejection Probabilities, 5% level, $t_{ave(n)}$ (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			$t_{ave(n)}$, values of b					$t_{clus(n)}$	$t_{ave(n)}$, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
10	10	.0	.084	.085	.090	.091	.096	.110	.083	.079	.075	.071	.068	.063
		.3	.112	.117	.120	.121	.122	.112	.110	.110	.106	.097	.091	.071
		.6	.178	.164	.162	.156	.154	.128	.174	.158	.146	.129	.115	.081
		.9	.277	.241	.225	.206	.196	.150	.275	.226	.209	.181	.162	.106
	50	.0	.067	.067	.070	.077	.085	.096	.066	.063	.060	.059	.058	.061
		.3	.077	.076	.080	.086	.090	.094	.076	.071	.068	.065	.066	.061
		.6	.102	.095	.089	.094	.096	.095	.101	.090	.081	.078	.076	.063
		.9	.234	.185	.173	.158	.146	.115	.232	.180	.156	.132	.117	.082
	250	.0	.054	.056	.059	.066	.068	.083	.052	.052	.048	.049	.048	.051
		.3	.048	.048	.053	.063	.070	.079	.047	.045	.046	.047	.045	.049
		.6	.061	.061	.066	.072	.074	.081	.060	.054	.050	.048	.048	.048
		.9	.097	.088	.082	.083	.090	.082	.095	.084	.074	.066	.060	.056

Table 1: Empirical Null Rejection Probabilities, 5% level, $t_{ave(n)}$ (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			$t_{ave(n)}$, values of b					$t_{clus(n)}$	$t_{ave(n)}$, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
50	10	.0	.074	.071	.069	.069	.069	.059	.070	.070	.065	.064	.059	.056
		.3	.094	.094	.089	.084	.080	.067	.091	.088	.086	.077	.072	.059
		.6	.158	.137	.123	.103	.093	.071	.154	.133	.117	.098	.086	.062
		.9	.258	.209	.177	.149	.118	.068	.255	.204	.172	.137	.106	.062
	50	.0	.051	.052	.053	.054	.057	.059	.051	.050	.048	.048	.049	.053
		.3	.063	.060	.059	.057	.055	.055	.061	.059	.055	.050	.049	.050
		.6	.083	.076	.070	.065	.060	.055	.082	.073	.064	.060	.057	.048
		.9	.202	.147	.127	.103	.089	.059	.199	.146	.118	.094	.081	.053
	250	.0	.052	.051	.053	.053	.053	.057	.049	.050	.050	.050	.050	.048
		.3	.052	.051	.053	.055	.056	.057	.051	.049	.049	.049	.051	.051
		.6	.057	.055	.053	.055	.057	.057	.055	.053	.050	.050	.051	.051
		.9	.095	.077	.074	.067	.066	.056	.090	.075	.066	.062	.058	.052

Table 1: Empirical Null Rejection Probabilities, 5% level, $t_{ave(n)}$ (Bartlett Kernel) and $t_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Two-Tailed Test of $H_0 : \beta_1 = 0$.

			$N(0,1)$ Critical Values						Fixed-b, t Critical Values					
			$t_{ave(n)}$, values of b					$t_{clus(n)}$	$t_{ave(n)}$, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
250	10	.0	.059	.057	.053	.051	.051	.045	.058	.055	.051	.049	.048	.043
		.3	.076	.072	.068	.059	.055	.048	.073	.071	.066	.057	.053	.045
		.6	.141	.117	.106	.089	.074	.052	.137	.115	.105	.086	.072	.049
		.9	.218	.169	.150	.128	.097	.058	.217	.167	.150	.124	.095	.055
	50	.0	.052	.051	.049	.051	.049	.048	.049	.049	.049	.048	.048	.046
		.3	.066	.065	.063	.063	.060	.058	.066	.064	.062	.060	.059	.055
		.6	.090	.079	.072	.067	.066	.059	.089	.077	.071	.066	.066	.055
		.9	.202	.148	.125	.097	.081	.044	.200	.146	.124	.095	.079	.041
	250	.0	.045	.044	.044	.044	.044	.044	.043	.044	.043	.043	.043	.043
		.3	.050	.050	.049	.050	.050	.051	.049	.049	.049	.049	.047	.044
		.6	.057	.053	.051	.050	.051	.048	.055	.052	.050	.049	.048	.047
		.9	.086	.071	.063	.056	.052	.043	.084	.070	.062	.053	.051	.041

Table 4: Size-adjusted Power, 5% level, Bartlett Kernel
No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$).
Two-Tailed Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
10	10	.0	.153	.141	.150	.149	.148	.131	.117	.130	.123	.106	.115
		.3	.122	.125	.136	.130	.127	.116	.103	.099	.110	.099	.100
		.6	.110	.111	.109	.106	.103	.114	.104	.102	.102	.098	.099
		.9	.089	.087	.086	.086	.087	.084	.091	.098	.077	.079	.086
	50	.0	.578	.577	.577	.553	.550	.498	.510	.499	.460	.449	.447
		.3	.513	.502	.484	.462	.459	.405	.452	.393	.368	.375	.371
		.6	.312	.300	.289	.274	.260	.264	.289	.266	.249	.250	.243
		.9	.137	.131	.129	.122	.111	.107	.112	.101	.086	.091	.099
	250	.0	.998	.998	.998	.998	.997	.993	.995	.990	.966	.956	.956
		.3	.997	.997	.995	.994	.993	.981	.993	.983	.955	.916	.909
		.6	.933	.927	.918	.906	.897	.854	.907	.849	.822	.761	.758
		.9	.374	.367	.368	.355	.343	.304	.311	.305	.256	.265	.263

Table 4: Size-adjusted Power, 5% level, Bartlett Kernel
No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$).
Two-Tailed Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
50	10	.0	.535	.544	.536	.533	.523	.501	.423	.424	.396	.386	.398
		.3	.471	.468	.461	.463	.465	.457	.392	.360	.350	.311	.323
		.6	.342	.333	.335	.336	.331	.334	.292	.284	.251	.245	.251
		.9	.235	.239	.229	.226	.218	.211	.227	.209	.192	.199	.201
	50	.0	1.00	1.00	1.00	1.00	1.00	1.00	.998	.990	.960	.940	.928
		.3	.994	.993	.993	.993	.993	.992	.988	.967	.936	.915	.912
		.6	.941	.936	.937	.934	.929	.929	.894	.851	.775	.782	.780
		.9	.450	.459	.456	.456	.460	.451	.423	.378	.333	.321	.320
	250	.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.999	.998
		.9	.952	.953	.953	.953	.950	.948	.917	.899	.824	.786	.793

Table 4: Size-adjusted Power, 5% level, Bartlett Kernel
No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$).
Two-Tailed Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			$t_{ave(n)}$, values of b					$t_{clus(n)}$	t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0
250	10	.0	.998	.998	.998	.998	.998	.998	.991	.977	.949	.907	.912
		.3	.996	.996	.995	.995	.995	.996	.967	.955	.922	.882	.876
		.6	.959	.960	.959	.958	.959	.958	.890	.872	.777	.777	.777
		.9	.831	.830	.282	.828	.829	.833	.743	.710	.687	.630	.617
	50	.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.997	.997
		.9	.988	.988	.988	.988	.988	.987	.960	.941	.882	.830	.829
	250	.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		.9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.998

Table 5: Size-adjusted Power, 5% level, Bartlett Kernel
MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$)
Two-Tail Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0
9	10	.0	.076	.070	.062	.058	.059
		.3	.071	.066	.059	.067	.069
		.6	.070	.059	.067	.074	.073
		.9	.062	.067	.059	.059	.060
	50	.0	.165	.157	.145	.120	.123
		.3	.147	.133	.141	.136	.138
		.6	.128	.123	.108	.106	.106
		.9	.070	.073	.071	.075	.076
	250	.0	.505	.500	.464	.426	.413
		.3	.453	.441	.396	.404	.397
		.6	.318	.275	.258	.269	.268
		.9	.131	.113	.102	.106	.111

Table 5: Size-adjusted Power, 5% level, Bartlett Kernel
MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$)
Two-Tail Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0
49	10	.0	.097	.095	.089	.091	.095
		.3	.104	.090	.088	.090	.089
		.6	.088	.083	.083	.080	.086
		.9	.085	.091	.076	.074	.079
	50	.0	.414	.374	.328	.324	.321
		.3	.355	.317	.300	.275	.281
		.6	.242	.218	.181	.189	.188
		.9	.123	.117	.111	.096	.102
	250	.0	.971	.953	.915	.880	.873
		.3	.932	.898	.866	.832	.826
		.6	.721	.680	.639	.591	.597
		.9	.250	.235	.229	.212	.202

Table 5: Size-adjusted Power, 5% level, Bartlett Kernel
MA(2) Spatial correlation in cross-section ($\lambda = 0$, $\theta = 0.5$)
Two-Tail Test of $H_0 : \beta_1 = 0$. Alternative value is $\beta_1 = 0.1$.

			t_{HACSC}, values of b				
n	T	ρ	0.1	0.2	0.4	0.7	1.0
256	10	.0	.323	.286	.246	.255	.253
		.3	.294	.267	.229	.242	.243
		.6	.237	.204	.191	.186	.185
		.9	.168	.172	.168	.158	.157
	50	.0	.937	.906	.862	.800	.814
		.3	.875	.838	.779	.729	.744
		.6	.717	.688	.600	.587	.596
		.9	.275	.246	.216	.206	.226
	250	.0	1.00	1.00	1.00	.998	.998
		.3	1.00	1.00	.999	.996	.995
		.6	.999	.996	.986	.974	.971
		.9	.764	.696	.606	.585	.596

Table 6: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, Two Group Case: Within Group Correlation 0.75, Between Group Correlation 0.25. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			No Time Dummies						With Time Dummies					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
10	10	.0	.154	.142	.125	.124	.123	.344	.161	.142	.124	.124	.126	.333
		.3	.172	.153	.133	.140	.141	.350	.188	.166	.152	.143	.140	.330
		.6	.222	.201	.180	.175	.170	.354	.235	.204	.194	.185	.187	.313
		.9	.301	.270	.243	.223	.226	.346	.316	.280	.243	.234	.225	.287
	50	.0	.057	.055	.059	.058	.061	.331	.071	.061	.059	.058	.061	.372
		.3	.072	.070	.067	.066	.069	.353	.080	.076	.070	.070	.070	.367
		.6	.111	.096	.095	.094	.095	.361	.138	.115	.104	.103	.107	.348
		.9	.277	.223	.204	.193	.198	.360	.293	.243	.210	.203	.203	.322
	250	.0	.051	.051	.050	.046	.048	.358	.053	.049	.049	.052	.052	.373
		.3	.055	.057	.052	.049	.050	.349	.052	.050	.050	.050	.052	.372
		.6	.063	.066	.061	.058	.061	.346	.062	.055	.059	.058	.061	.352
		.9	.128	.114	.091	.096	.095	.358	.130	.112	.102	.100	.106	.357

Table 6: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, Two Group Case: Within Group Correlation 0.75, Between Group Correlation 0.25. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			No Time Dummies						With Time Dummies					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
50	10	.0	.156	.143	.133	.131	.131	.683	.157	.141	.123	.125	.127	.616
		.3	.181	.166	.155	.148	.152	.706	.177	.166	.154	.147	.152	.608
		.6	.245	.222	.211	.196	.200	.693	.235	.212	.189	.181	.188	.573
		.9	.331	.291	.268	.245	.245	.642	.321	.286	.257	.237	.235	.514
	50	.0	.067	.060	.059	.060	.060	.664	.061	.059	.060	.061	.064	.663
		.3	.074	.069	.068	.070	.071	.669	.076	.067	.071	.071	.073	.647
		.6	.108	.099	.088	.086	.089	.666	.107	.098	.088	.088	.085	.638
		.9	.281	.220	.200	.196	.201	.674	.272	.222	.194	.178	.179	.566
	250	.0	.047	.043	.040	.039	.042	.671	.049	.054	.053	.054	.051	.662
		.3	.055	.057	.051	.052	.053	.650	.060	.054	.060	.061	.062	.661
		.6	.069	.061	.063	.065	.064	.664	.075	.068	.069	.070	.071	.665
		.9	.129	.113	.099	.099	.104	.680	.142	.117	.113	.110	.109	.680

Table 6: Empirical Null Rejection Probabilities, 5% level, t_{HACSC} (Bartlett Kernel) and $t_{clus(n)}$, Two Group Case: Within Group Correlation 0.75, Between Group Correlation 0.25. Two-Tailed Test of $H_0 : \beta_1 = 0$. Fixed-b, t Critical Values

			No Time Dummies						With Time Dummies					
			t_{HACSC}, values of b					$t_{clus(n)}$	t_{HACSC}, values of b					$t_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
250	10	.0	.135	.120	.114	.103	.106	.846	.162	.140	.130	.135	.135	.826
		.3	.147	.133	.131	.122	.121	.851	.182	.156	.149	.141	.148	.806
		.6	.207	.185	.174	.161	.164	.841	.237	.210	.195	.191	.191	.768
		.9	.295	.257	.235	.227	.224	.835	.314	.276	.238	.228	.228	.730
	50	.0	.061	.059	.056	.056	.057	.862	.070	.067	.060	.059	.059	.830
		.3	.077	.071	.062	.062	.065	.855	.080	.074	.071	.070	.074	.838
		.6	.117	.104	.091	.095	.095	.863	.117	.100	.093	.091	.097	.825
		.9	.290	.240	.216	.204	.206	.847	.296	.241	.212	.205	.206	.793
	250	.0	.051	.053	.046	.049	.048	.849	.044	.049	.048	.045	.049	.853
		.3	.052	.048	.044	.045	.049	.854	.058	.054	.051	.056	.056	.844
		.6	.061	.056	.056	.056	.055	.853	.061	.060	.057	.058	.062	.831
		.9	.127	.112	.094	.090	.090	.868	.124	.105	.089	.090	.091	.821

Table 7: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0$.

			χ^2_2 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
10	10	.0	.198	.264	.392	.533	.625	.173	.116	.104	.107	.097	.092	.060
		.3	.248	.314	.451	.586	.670	.189	.142	.133	.135	.129	.129	.060
		.6	.372	.424	.560	.678	.751	.219	.258	.216	.207	.200	.197	.080
		.9	.558	.568	.673	.774	.830	.243	.449	.360	.322	.329	.331	.106
	50	.0	.141	.199	.322	.469	.569	.173	.068	.063	.059	.066	.059	.052
		.3	.147	.219	.351	.499	.594	.180	.078	.066	.071	.074	.071	.060
		.6	.209	.281	.407	.561	.647	.176	.124	.108	.108	.103	.107	.069
		.9	.469	.489	.618	.730	.793	.212	.376	.263	.256	.248	.253	.072
	250	.0	.115	.181	.314	.477	.580	.159	.050	.044	.046	.051	.052	.052
		.3	.122	.188	.323	.464	.572	.160	.057	.056	.059	.057	.056	.052
		.6	.142	.203	.341	.485	.582	.151	.074	.069	.064	.063	.065	.050
		.9	.227	.291	.421	.563	.655	.172	.146	.120	.114	.108	.110	.070

Table 7: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0$.

			χ_2^2 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
50	10	.0	.188	.260	.391	.532	.618	.075	.111	.100	.098	.091	.089	.059
		.3	.247	.315	.442	.578	.662	.083	.146	.141	.134	.129	.129	.067
		.6	.385	.428	.560	.680	.750	.087	.275	.216	.209	.208	.208	.068
		.9	.564	.579	.681	.761	.817	.096	.459	.381	.342	.338	.344	.073
	50	.0	.120	.204	.331	.474	.568	.072	.053	.054	.053	.056	.056	.052
		.3	.145	.212	.343	.508	.601	.070	.073	.062	.059	.070	.067	.056
		.6	.195	.259	.403	.556	.651	.076	.114	.102	.096	.092	.095	.059
		.9	.463	.472	.597	.712	.777	.086	.343	.260	.240	.236	.236	.066
	250	.0	.106	.176	.313	.465	.575	.065	.051	.048	.045	.047	.043	.044
		.3	.111	.175	.314	.467	.572	.066	.055	.055	.054	.055	.054	.050
		.6	.132	.204	.338	.502	.591	.074	.069	.063	.065	.068	.070	.053
		.9	.211	.278	.408	.543	.643	.071	.130	.107	.106	.104	.103	.055

Table 7: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0$.

			χ^2_2 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
250	10	.0	.183	.238	.632	.507	.598	.043	.102	.094	.085	.088	.087	.040
		.3	.217	.285	.409	.552	.643	.046	.129	.113	.101	.105	.103	.043
		.6	.333	.385	.526	.646	.729	.049	.234	.191	.184	.172	.172	.048
		.9	.544	.562	.658	.754	.815	.052	.452	.360	.319	.319	.319	.049
	50	.0	.131	.201	.316	.467	.565	.052	.059	.060	.060	.058	.059	.049
		.3	.148	.216	.342	.494	.580	.052	.069	.064	.066	.066	.066	.048
		.6	.205	.275	.406	.559	.636	.054	.111	.101	.106	.097	.100	.049
		.9	.470	.482	.605	.714	.774	.057	.350	.260	.248	.237	.242	.052
	250	.0	.098	.164	.306	.453	.564	.047	.043	.043	.042	.041	.042	.043
		.3	.105	.176	.314	.480	.570	.042	.053	.047	.045	.048	.044	.041
		.6	.127	.203	.344	.489	.588	.049	.068	.066	.058	.056	.056	.046
		.9	.232	.308	.449	.575	.666	.055	.133	.119	.118	.120	.122	.051

Table 8: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$.

			χ^2_4 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
10	10	.0	.417	.544	.744	.855	.904	.361	.207	.195	.188	.175	.171	.070
		.3	.472	.597	.784	.878	.920	.375	.267	.225	.238	.219	.219	.075
		.6	.641	.728	.868	.938	.962	.414	.413	.354	.360	.343	.336	.086
		.9	.836	.873	.939	.975	.981	.466	.680	.573	.562	.565	.563	.110
	50	.0	.223	.399	.641	.806	.875	.349	.070	.067	.067	.070	.070	.053
		.3	.253	.425	.654	.814	.886	.351	.085	.088	.087	.086	.081	.059
		.6	.374	.529	.732	.856	.916	.358	.158	.140	.148	.141	.138	.058
		.9	.745	.796	.911	.961	.978	.422	.525	.433	.436	.435	.425	.092
	250	.0	.195	.364	.636	.789	.868	.331	.053	.054	.059	.053	.056	.055
		.3	.211	.379	.627	.800	.870	.341	.053	.063	.065	.060	.064	.043
		.6	.240	.415	.654	.818	.889	.330	.066	.063	.071	.067	.068	.048
		.9	.406	.551	.763	.877	.923	.343	.193	.174	.171	.168	.169	.060

Table 8: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$.

			χ^2_4 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
50	10	.0	.392	.536	.736	.853	.908	.096	.187	.181	.176	.171	.162	.051
		.3	.463	.603	.771	.887	.930	.101	.249	.225	.228	.207	.205	.060
		.6	.646	.738	.882	.943	.966	.105	.412	.347	.344	.344	.337	.065
		.9	.848	.876	.939	.971	.984	.119	.678	.563	.552	.573	.559	.069
	50	.0	.207	.376	.625	.781	.855	.087	.058	.055	.051	.050	.053	.047
		.3	.246	.416	.653	.811	.873	.089	.084	.077	.073	.071	.069	.051
		.6	.353	.517	.733	.861	.919	.096	.149	.135	.142	.136	.132	.058
		.9	.730	.789	.901	.955	.973	.109	.511	.407	.426	.415	.413	.065
	250	.0	.186	.369	.614	.796	.871	.084	.044	.042	.050	.047	.042	.042
		.3	.192	.362	.636	.796	.867	.083	.053	.050	.056	.057	.055	.048
		.6	.228	.391	.641	.808	.870	.085	.069	.061	.062	.062	.059	.051
		.9	.385	.539	.746	.879	.920	.095	.163	.153	.163	.154	.155	.058

Table 8: Empirical Null Rejection Probabilities, 5% level, $Wald_{HACSC}$ (Bartlett Kernel) and $Wald_{clus(n)}$, No spatial correlation in cross-section ($\lambda = 0$, $\theta = 0$). Testing $H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$.

			χ^2_4 Critical Values						Fixed-b, F Critical Values					
			$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$	$Wald_{HACSC}$, values of b					$Wald_{clus(n)}$
n	T	ρ	0.1	0.2	0.4	0.7	1.0		0.1	0.2	0.4	0.7	1.0	
250	10	.0	.383	.519	.740	.857	.912	.057	.184	.167	.170	.165	.157	.049
		.3	.431	.577	.767	.885	.927	.064	.232	.201	.204	.199	.193	.057
		.6	.612	.716	.864	.932	.967	.066	.398	.325	.328	.325	.316	.057
		.9	.837	.866	.944	.974	.983	.068	.654	.545	.545	.552	.541	.060
	50	.0	.196	.635	.625	.792	.858	.047	.059	.051	.057	.053	.054	.043
		.3	.226	.395	.629	.809	.873	.051	.066	.064	.069	.067	.068	.045
		.6	.346	.502	.733	.860	.909	.054	.133	.119	.128	.123	.118	.051
		.9	.730	.783	.906	.955	.974	.060	.510	.408	.414	.418	.420	.051
	250	.0	.191	.366	.631	.796	.871	.049	.046	.051	.047	.046	.043	.044
		.3	.201	.372	.630	.799	.860	.054	.055	.051	.053	.051	.049	.046
		.6	.229	.393	.643	.804	.877	.063	.069	.060	.062	.063	.059	.057
		.9	.399	.562	.761	.871	.928	.064	.172	.145	.148	.146	.147	.058