Spatial panel econometrics

Paul Elhorst

June 15-16, 2021

SEAI, Rome

Faculty of Economics and Business
University of Groningen
the Netherlands
Email: j.p.elhorst@rug.nl

Accompanying papers

Elhorst J.P. (2021) Spatial Panel Models and Common Factors. In: Fischer M.M., Nijkamp P. (eds) Handbook of Regional Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-60723-7_86

Elhorst J.P. (2021) The general nesting spatial econometric model for spatial panels with common factors: Further raising the bar. Presented at the Spatial Econometric World Conference, May 27, 2021, Tokyo (online).

+

Zip-file with Stata and excel-files.

The general nesting spatial (GNS) econometric model for spatial panels with common factors (CF) reads as

$$Y_t = \boldsymbol{\tau Y_{t-1}} + \boldsymbol{\rho W Y_t} + \boldsymbol{\eta W Y_{t-1}} + X_t\beta + \boldsymbol{W X_t \theta} + \sum_r \boldsymbol{\Gamma_r^T f_{rt}} + u_t, \quad u_t = \boldsymbol{\lambda W u_t} + \varepsilon_t$$

Spatial lags are in red: 1+*K*+1=*K*+2 in total (*K* is number of *X* variables)

Dynamic effects are in green: 2 in total

Common factors are in blue: # parameters depend on type of CF, they are reported below.

Spatial econometric model

Linear regression model (**Y=Xβ+ε**) extended to include

Endogenous interaction effect (1): **ρWY**

- Dependent variable y of unit A ↔ Dependent variable y of unit B
- Y denotes an **N**×1 vector consisting of one observation on the dependent variable for every unit in the sample (i=1,…,**N**)
- **W** is an **N**×**N** nonnegative matrix describing the arrangement of the units in the sample

Exogenous interaction effects (K): **WXθ**

- Independent variable x of unit A → Dependent variable y of unit B
- X denotes an N×**K** matrix of exogenous explanatory variables

Interaction effect among **error** terms (1): **λWu**

- Error term u of unit A ↔ Error term u of unit B

Dynamics: Y_{t-1} and WY_{t-1}

Habit persistence. It takes time to change behavior.

Korniotis (2010): Internal and external habit persistence.

Models without WY_t

Anselin et al. (2008): time-space recursive spatial econometric model. Suitable to explain spatial diffusion phenomena. Think of the rise and spread of the Covid-19 virus on a daily/weekly basis.

LeSage and Pace (2009, ch. 7): spatiotemporal (partial adjustment) model. High temporal dependence and low spatial dependence might nonetheless imply a long-run equilibrium with high spatial dependence.

Fogli and Veldkamp (2011): Information diffusion can change preferences (female labor force participation), but people require time to gather information, creating a delay in the decision-making process, and hence spatial dependence takes time to manifest itself.

Dynamic spatial panel data model with FE, RE or CF

$Y_t = \boldsymbol{\tau Y_{t-1}} + \boldsymbol{\rho W Y_t} + \boldsymbol{\eta W Y_{t-1}} + X_t\beta + \boldsymbol{W X_t \theta}$ + error terms

Short-term effects in dynamic model (ignore τ and η) /

Long-term effects in static model: $\boldsymbol{Y_{t-1}}$ and $\boldsymbol{WY_{t-1}}$ are not included

$$\left[\frac{\partial E(Y)}{\partial x_{1k}} \quad \cdots \quad \frac{\partial E(Y)}{\partial x_{Nk}}\right]_t = (I-\rho W)^{-1}[\beta_k I_N + \theta_k W].$$

Long-term (set $Y_{t-1}=Y_t=Y^*$ and $WY_{t-1}=WY_t=WY^*$)

$$\left[\frac{\partial E(Y)}{\partial x_{1k}} \quad \cdots \quad \frac{\partial E(Y)}{\partial x_{Nk}}\right] = [(1-\tau)I-(\rho+\eta)W]^{-1}[\beta_k I_N + \theta_k W].$$

Direct effect: Mean diagonal element N by N matrix

Indirect effect: Mean row sum of off-diagonal elements N by N matrix

Note: Error terms (FE,RE,CF) drop out due to taking expectations

Note: t-values of direct and indirect effects are bootstrapped or determined by the delta method.

Important facts in empirical research

Indirect effect=spatial spillover effect=main focus of spatial economists/spatial econometricians

Generally, it is hard to find significant spatial spillover effects since they depend on so many parameters (3 short term, 5 long term); many empirical studies do not recognize this.

It is likely that the magnitude of the direct effects is greater than that of the indirect effects. If not, explain!

Always verify whether $\rho<1$ (static), or $\tau+\rho+\eta<1$ (dynamic)

Spatial econometric models with different combinations of spatial interaction effects and the flexibility regarding spatial spillovers

Type of model	Spatial interaction effects	#	**Flexibility spillovers**
OLS, Ordinary least squares model	-	0	Zero by construction
SAR, Spatial autoregressive model	WY	1	Constant ratios
SEM, Spatial error model	Wu	1	Zero by construction
SLX, Spatial lag of X model	WX	K	Fully flexible
SAC, Spatial autoregressive combined model (SARAR)	WY, Wu	2	Constant ratios
SDM, Spatial Durbin model	WY, WX	K+1	Fully flexible
SDEM, Spatial Durbin error model	WX, Wu	K+1	Fully flexible
GNS, General nesting spatial model	WY, WX, Wu	K+2	Fully flexible

Direct and spillover effects corresponding to different model specifications
Short term effects dynamic model/long term effects static model

Model	Direct effect	Spillover effect
OLS / SEM (*Wu*)	β_k	0
SAR (*WY*)/ SAC (*WY*, *Wu*) *	Average diagonal element of $(\mathbf{I}-\rho\mathbf{W})^{-1}\beta_k$	Average row sum of off-diagonal elements of $(\mathbf{I}-\rho\mathbf{W})^{-1}\beta_k$
SLX / SDEM (*WX* / *Wu*)	β_k	θ_k
SDM / GNS (*WY*+*WX*/*Wu*)	Average diagonal element of $(\mathbf{I}-\rho\mathbf{W})^{-1}[\beta_k+\mathbf{W}\theta_k]$	Average row sum of off-diagonal elements of $(\mathbf{I}-\rho\mathbf{W})^{-1}[\beta_k+\mathbf{W}\theta_k]$

* **Ratio between the spillover effect and the direct effect in the SAR/SAC model is the same for every explanatory variable**.

Global versus local spillover effects

Indirect effects that occur if ρ=0 (of WY) and $\theta \neq 0$ are known as ***local spillover effects***

$$\begin{bmatrix} \frac{\partial E(y_1)}{\partial x_{1k}} & . & \frac{\partial E(y_1)}{\partial x_{Nk}} \\ . & . & . \\ \frac{\partial E(y_N)}{\partial x_{1k}} & . & \frac{\partial E(y_N)}{\partial x_{Nk}} \end{bmatrix} = \begin{bmatrix} \beta_k & w_{12}\theta_k & . & w_{1N}\theta_k \\ w_{21}\theta_k & \beta_k & . & w_{2N}\theta_k \\ . & . & . & . \\ w_{N1}\theta_k & w_{N2}\theta_k & . & \beta_k \end{bmatrix} = \beta_k I_N + \theta_k W.$$

This is local because the indirect effects only fall on spatial units for which the elements of W are non-zero. Local spillovers go together with dense(r) W matrix.

Indirect effects that occur if ρ≠0 (of WY) and θ=0 are known as ***global spillover effects***

$$\begin{bmatrix} \frac{\partial E(y_1)}{\partial x_{1k}} & \cdot & \frac{\partial E(y_1)}{\partial x_{Nk}} \\ \cdot & \cdot & \cdot \\ \frac{\partial E(y_N)}{\partial x_{1k}} & \cdot & \frac{\partial E(y_N)}{\partial x_{Nk}} \end{bmatrix} = (I-\rho W)^{-1}\beta_k = (\mathrm{I}+\rho \mathrm{W}+\rho^2 W^2+\rho^3 W^3+\ldots)\beta_k$$

This is global because the indirect effects fall on all units; even if W contains many zero elements, $(I-\rho W)^{-1}$ will not.

Global spillovers tend to go together with sparse(r) W matrix; due to the higher-order terms $\rho^g W^g$ ($g>1$) locations farther away are reached anyway even if they are not directly connected.

$$\sum_r \boldsymbol{\Gamma}_r^T \boldsymbol{f}_{rt} = \boldsymbol{\mu} + \boldsymbol{\alpha}_t \boldsymbol{\iota}_N$$

$\boldsymbol{\mu}$: vector of spatial fixed or random effects (don't confuse spatial fixed effects with spatial lags)

$\boldsymbol{\alpha}_t$: time period fixed or random effects (t=1,…,T)

$$Y_t=\tau Y_{t-1}+\rho WY_t+\eta WY_{t-1}+X_t\beta+WX_t\theta+\mu+\alpha_t\iota_N+\varepsilon_t$$

The standard reasoning behind spatial specific effects $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N)'$ is that they control for all space-specific time-invariant variables whose omission could bias the estimates in a typical cross-sectional study (Baltagi, 2005).

The spatial specific effects may be treated as **fixed effects** or as **random effects**. In the fixed effects model, a dummy variable is introduced for each spatial unit, while in the random effects model, μ_i (i=1,…,N) is treated as a random variable that is independently and identically distributed with zero mean and variance σ_μ. Furthermore, it is assumed that the random variables μ_i and ε_{it} are independent of each other.

The same for $\boldsymbol{\alpha}_t$.

The random effects model is quite popular among spatial econometricians/practitioners, which can be explained by three reasons.

(1) It may be considered as a **compromise solution** to the all or nothing way of utilizing the cross-sectional component of the data. Panel data models with controls for spatial fixed effects only utilize the time-series component of the data, whereas these models without such controls employ both time-series and cross-sectional components. The parameter ϕ in random effects models, which can take values on the interval [0,1], may be used to estimate the weight that may be attached to the cross-sectional component of the data. If this weight equals 0, the random effects model reduces to the fixed effects model; if it goes to 1, it converges to its counterpart without controls for spatial fixed effects.

(2) The random effects model avoids the **loss of degrees of freedom** incurred in the fixed effects model associated with a relatively large N. Besides, the spatial fixed effects can only be estimated consistently when T is sufficiently large, because the number of observations available for the estimation of each μ_i is T. Importantly, the inconsistency of μ_i is not transmitted to the estimator of the slope coefficients $\boldsymbol{\beta}$, since it is not a function of the estimated μ_i.

In other words, the incidental parameters problem does not matter when $\boldsymbol{\beta}$ are the coefficients of interest and the spatial fixed effects μ_i are not, which is the case in most empirical studies.

(3) It avoids the problem that the coefficients of **time-invariant variables** or variables that only vary a little cannot be estimated.

Despite its popularity, the question whether the random effects model is also an appropriate specification is often left unanswered. **Three conditions** should be satisfied before the random effects model may be implemented:

(1) The number of units should potentially be able to go to infinity.
(2) The units of observation should be representative of a larger population.
(3) The traditional assumption of zero correlation between the random effects μ_i and the explanatory variables needs to be made, which in general is particularly restrictive.

These conditions do not tend to be satisfied in spatial research.

There are two types of asymptotics that are commonly used in the context of spatial observations:

(a) The **'infill'** asymptotic structure, where the sampling region remains bounded as $N \to \infty$. In this case more units of information come from observations taken from between those already observed; and

(b) The **'increasing domain'** asymptotic structure, where the sampling region grows as $N \to \infty$. In this case there is a minimum distance separating any two spatial units for all N.

According to Lahiri (2003), there are also two types of sampling designs:

(a) The stochastic design where the spatial units are **randomly drawn**; and

(b) The **fixed design** where the spatial units lie on a nonrandom field, possibly irregularly spaced.

The spatial econometric literature mainly focuses on increasing domain asymptotics under the fixed sample design (Cressie 1993, p. 100; Griffith and Lagona 1998; Lahiri 2003).

Although the number of spatial units under the fixed sample design can potentially go to infinity, this design is incompatible with the increasing domain asymptotic structure. If there is a minimum distance separating spatial units and the researcher wants to collect data for a certain type of spatial units within a particular study area, there will be an upper bound on the number of spatial units. Furthermore, when data on all spatial units within a study area are collected it is questionable whether they are still representative of a larger population.

For a given set of regions, such as all counties of a state or all regions in a country, the population may be said '*to be sampled exhaustively*' (Nerlove and Balestra 1996, p. 4), and '*the individual spatial units have characteristics that actually set them apart from a larger population*' (Anselin 1988, p. 51). In other words, if the data happen to be a random sample of the population, unconditional inference about the population necessitates estimation with random effects. If, however, the objective is limited to making conditional inferences about the sample, then fixed effects should be specified.

In spatial research there is a prominent reason why investigators generally do not draw a limited sample of units from a particular study area, but rather work with cross-sectional or space-time data of adjacent spatial units located in unbroken study areas. **This is because otherwise the spatial weight matrix cannot consistently be specified and the impact of spatial interaction effects cannot be consistently estimated.** Only when neighboring units are also part of the sample, it is possible to measure the impact of these neighboring units. In other words, this type of research just requires that the data covers the whole population, since it would break down when having a random sample of the population.

In conclusion, we can say that the fixed effects model is generally more appropriate than the random effects model since spatial econometricians tend to work with space-time data of adjacent spatial units located in unbroken study areas, such as all counties of a state or all regions in a country.

The same applied to time. Researchers tend to work with consecutive time spans; otherwise the impact of dynamic effects cannot be consistently be estimated.

To test the assumption of zero correlation between the random effects μ_i and the explanatory variables, the **Hausman specification test** might be used (Baltagi 2005, pp. 66-68). The hypothesis being tested is H_0: h=0, where

$$h = d^T [\mathrm{var}(d)]^{-1} d$$

$$d = \hat{\beta}_{FE} - \hat{\beta}_{RE}$$

$$\mathrm{var}(d) = \hat{\sigma}^2_{RE} (X^{\bullet T} X^{\bullet})^{-1} - \hat{\sigma}^2_{FE} (X^{*T} X^{*})^{-1}$$

This test statistic has a chi-squared distribution with K degrees of freedom (the number of explanatory variables in the model, excluding the constant term). Hausman's specification test can also be used when the model is extended to include spatial lags (**WY,WX**). Then the coefficients of these spatial lags should also be included in this comparison test (Lee and Yu, 2012).

Empirical illustration: Cigarette Demand in the US

Baltagi and Li (2004) estimate a demand model for cigarettes based on a panel from 46 U.S. states (N=46)

$$\log(C_{it}) = \alpha + \beta_1 \log(P_{it}) + \beta_2 \log(Y_{it}) + \mu_i \text{(optional)} + \lambda_t \text{(optional)} + \varepsilon_{it} ,$$

where C_{it} is real per capita sales of cigarettes by persons of smoking age (14 years and older). This is measured in packs of cigarettes per capita. P_{it} is the average retail price of a pack of cigarettes measured in real terms. Y_{it} is real per capita disposable income. Whereas Baltagi and Li (2004) use the first 25 years for estimation to reserve data for out of sample forecasts, we use the full data set covering the period 1963-1992 (T=30). Details on data sources are given in Baltagi and Levin (1986, 1992) and Baltagi et al. (2000). They also give reasons to assume the state-specific effects (μ_i) and time-specific effects (λ_t) fixed, in which case one includes state dummy variables and time dummies for each year.

We have reasons to believe that spatial interaction effects need to be included in this model!

BOOTLEGGING

- The main motivation to extend the basic model to include spatial interaction effects is the so-called **bootlegging effect**; consumers are expected to purchase cigarettes in nearby states, legally or illegally (smuggling), if there is a price advantage.
- This smuggling behavior is a result of significant price variation in cigarettes across US states and partly due to the disparities in state cigarette tax rates. Baltagi and Levin (1986, 1992) incorporate the minimum real price of cigarettes in any neighboring state as a proxy for the bootlegging effect.
- A limitation is that this proxy does not account for cross-border shopping that may take place between other states than the minimum-price neighboring state (Baltagi and Levin, 1986). This can be due to smuggling taking place over longer distances by trucks since cigarettes can be stored and are easy to transport (Baltagi and Levin, 1992) or due to geographically large states where cross-border shopping may occur in different neighboring states.
- To take this into account, other studies have extended the model to explicitly incorporate spatial interaction effects. However, while the specification originally adopted by Baltagi and Levin (1992) resembles the SLX model but then with only one exogenous interaction effect (price), applied spatial econometric studies have either included: (i) endogenous interaction effects, (ii) interaction effects among the error terms or (iii) a combination of endogenous and exogenous interaction effects.

STEP 1: Empirical illustration: Cigarette Demand in the US

Estimation results of cigarette demand using panel data models without spatial interaction effects

Determinants	(1)	(2)	(3)	(4)
	Pooled OLS	Spatial fixed effects	Time-period fixed effects	Spatial and time-period fixed effects
Log(P)	-0.859 (-25.16)	-0.702 (-38.88)	-1.205 (-22.66)	-1.035 (-25.63)
Log(Y)	0.268 (10.85)	-0.011 (-0.66)	0.565 (18.66)	0.529 (11.67)
Intercept	3.485 (30.75)			
R^2	0.321	0.853	0.440	0.896
LogL	370.3	1425.2	503.9	1661.7
LM spatial lag	66.47	136.43	44.04	46.90
LM spatial error	153.04	255.72	62.86	54.65
robust LM spatial lag	58.26	29.51	0.33	1.16
robust LM spatial error	144.84	148.80	19.15	8.91

LR test spatial fixed effects: (2315.7, with 45 degrees of freedom [df], $p < 0.01$)

LR test time-period fixed effects: (473.1, 29 df, $p < 0.01$)

(robust) LM test (critical value 3.84): error model

One of the main questions is which model best describes the data. One of the criterions that may be used for this purpose is the likelihood ratio (LR) test based on the log-likelihood function values of the different models. The LR test is based on minus two times the difference between the value of the log-likelihood function in the restricted model and the value of the log-likelihood function of the unrestricted model: $-2*(\log L_{restricted} - \log L_{unrestricted})$. This test statistic has a chi-squared distribution with degrees of freedom equal to the number of restriction imposed.

To investigate the (null) hypothesis that the spatial fixed effects are jointly insignificant, one may perform a likelihood ratio (LR) test. The results (2315.7, with 45 degrees of freedom [df], $p < 0.01$) indicate that this hypothesis must be rejected. Similarly, the hypothesis that the time-period fixed effects are jointly insignificant must be rejected (473.1, 29 df, $p < 0.01$). These test results justify the extension of the model with spatial and time-period fixed effects.

Robust LM tests point to spatial error model. Critical value for 1 degree of freedom is 3.84. However, this result might be misleading since variables $\boldsymbol{WX_t}$ have not been controlled for. See Corrado and Fingleton (2012): $\boldsymbol{WY_t}$ variable picks up omitted $\boldsymbol{WX_t}$ variables or nonlinearities in the X variables.

STEP 2: Focus on SDM (Limitation: SDEM left aside)

To test the hypothesis whether the spatial Durbin model can be simplified to the spatial error model, H_0: $\theta+\lambda\beta=0$, one may perform a Wald or LR test. The results reported in the second column using the Wald test (8.18, with 2 degrees of freedom [df], p=0.017) or using the LR test (8.28, 2 df, p=0.016) indicate that this hypothesis must be rejected. Similarly, the hypothesis that the spatial Durbin model can be simplified to the spatial lag model, H_0: $\theta=0$, must be rejected (Wald test: 17.96, 2 df, p=0.000; LR test: 15.80, 2 df, p=0.000). This implies that both the spatial error model and the spatial lag model must be rejected in favor of the spatial Durbin model.

Estimation results cigarette demand: Spatial Durbin model specification with spatial and time-period specific effects

Determinants	(1)		(2)	
	Spatial and time-period **fixed effects** bias-corrected*		**Random spatial effects**, Fixed time-period effects	
W*Log(C)	0.264	(8.25)	0.224	(6.82)
Log(P)	-1.001	(-24.36)	-1.007	(-24.91)
Log(Y)	0.603	(10.27)	0.593	(10.71)
W*Log(P)	0.093	(1.13)	0.066	(0.81)
W*Log(Y)	-0.314	(-3.93)	-0.271	(-3.55)
Phi			0.087	(6.81)
σ^2		0.005		0.005
(Pseudo) R^2		0.902		0.880
(Pseudo) Corrected R^2		0.400		0.317
LogL		1691.4		1555.5
Wald test spatial lag		17.96 (p=0.000)		13.90 (p=0.001)
LR test spatial lag		15.80 (p=0.000)		14.48 (p=0.000)
Wald test spatial error		8.18 (p=0.017)		7.38 (p=0.025)
LR test spatial error		8.28 (p=0.016)		7.27 (p=0.026)

Corrected R^2 is R^2 without the contribution of fixed effects (**double-check this in Stata**)
Bias-correction based on Lee and Yu (2010).

Hausman test for random effects instead of fixed effects

The last column in the Table above reports the parameter estimates if we treat c_i as a random variable rather than a set of fixed effects. Hausman's specification test can be used to test the random effects model against the fixed effects model. The results (30.61, 5 df, $p<0.01$) indicate that the random effects model must be rejected.

Another way to test the random effects model against the fixed effects model is to estimate the parameter "phi" (ϕ^2 in Baltagi, 2005) using ML, which measures the weight attached to the cross-sectional component of the data and which can take values on the interval [0,1]. If this parameter equals 0, the random effects model converges to its fixed effects counterpart; if it goes to 1, it converges to a model without any controls for spatial specific effects. We find phi=0.087, with t-value of 6.81, which just as Hausman's specification test indicates that the fixed and random effects models are significantly different from each other.

Direct and indirect effects cigarette demand: Spatial Durbin model with spatial and time-period specific effects

Direct effect Log(P)	-1.013 (-24.73)	-1.012 (-23.93)	-1.018 (-24.64)	-1.018 (-25.03)
Indirect effect Log(P)	**-0.220** (-2.26)	**-0.215** (-2.12)	**-0.199** (-2.28)	**-0.195** (-2.19)
Total effect Log(P)	-1.232 (-11.31)	-1.228 (-11.26)	-1.217 (-12.43)	-1.213 (-12.21)
Direct effect Log(Y)	0.594 (10.45)	0.594 (10.67)	0.586 (10.68)	0.583 (10.53)
Indirect effect Log(Y)	-0.197 (-2.15)	-0.196 (-2.18)	-0.169 (-2.03)	-0.171 (-2.06)
Total effect Log(Y)	0.397 (4.61)	0.398 (4.62)	0.417 (5.45)	0.412 (5.37)

Notes: t-values in parentheses. Direct and indirect effects estimates:

Left column $(I-\lambda W)^{-1}$ computed every draw, right column $(I-\lambda W)^{-1}$ approximated.

HOWEVER: No evidence of bootlegging or substitution effect!

We found evidence in favour of spatial Durbin model with fixed effects, but not of the bootlegging effect. **What happens if we add dynamics?**

Estimation results of cigarette demand using different model specifications

Determinants	(1)		(2)		
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}		
Intercept					
$Log(C)_{-1}$			**0.865**	**(65.04)**	**INTERNAL HABIT PERSISTENCE: Significant**
W*Log(C)	0.264	(8.25)	0.076	(2.00)	
$W*Log(C)_{-1}$			**-0.015**	**(-0.29)**	**EXTERNAL HABIT PERSISTENCE: Insignificant**
Log(P)	-1.001	(-24.36)	-0.266	(-13.19)	
Log(Y)	0.603	(10.27)	0.100	(4.16)	
W*Log(P)	0.093	(1.13)	**0.170**	**(3.66)**	
W*Log(Y)	-0.314	(-3.93)	-0.022	(-0.87)	
R^2		0.902		0.977	
LogL		1691.4		2623.3	

Notes: t-values in parentheses

Determinants	(1)		(2)	
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}	
$Log(C)_{-1}$			0.865	(65.04)
W*Log(C)	0.264	(8.25)	0.076	(2.00)
$W*Log(C)_{-1}$			-0.015	(-0.29)
Log(P)	-1.001	(-24.36)	-0.266	(-13.19)
Log(Y)	0.603	(10.27)	0.100	(4.16)
W*Log(P)	0.093	(1.13)	0.170	(3.66)
W*Log(Y)	-0.314	(-3.93)	-0.022	(-0.87)
LogL		**1691.4**		**2623.3**

To investigate whether the extension of the non-dynamic model to the dynamic spatial panel data model increases the explanatory power of the model, one may test whether the coefficients of the variables Y_{t-1} and WY_{t-1} are jointly significant using an LR-test. The outcome of this test (2×(2623.3-1691.4)=1863.8 with 2 df) evidently justifies the extension of the model with dynamic effects.

Conclusion: Dynamic spatial Durbin model outperforms its non-dynamic counterpart.

Effects estimates of cigarette demand using different model specifications

Determinants	(1)		(2)	
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}	
Short-term direct effect Log(P)			-0.262	(-11.48)
Short-term indirect effect Log(P)			0.160	(3.49)
Short-term direct effect Log(Y)			0.099	(3.36)
Short-term indirect effect Log(Y)			-0.018	(-0.45)
Long-term direct effect Log(P)	**-1.013**	**(-24.73)**	-1.931	(-9.59)
Long-term indirect effect Log(P)	-0.220	(-2.26)	0.610	(0.98)
Long-term direct effect Log(Y)	**0.594**	**(10.45)**	0.770	(3.55)
Long-term indirect effect Log(Y)	-0.197	(-2.15)	0.345	(0.48)

Notes: t-values in parentheses

- A static (non-dynamic) spatial Durbin model cannot be used to calculate short-term effect estimates of the explanatory variables.
- The direct effects estimates of the two explanatory variables in (1) are significantly different from zero and have the expected signs. Higher prices restrain people from smoking, while higher income levels have a positive effect on cigarette demand. The price elasticity amounts to -1.013 and the income elasticity to 0.594 in a non-dynamic model.

Effects estimates of cigarette demand using different model specifications

Determinants	(1)		(2)	
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}	
Short-term direct effect Log(P)			-0.262	(-11.48)
Short-term indirect effect Log(P)			0.160	(3.49)
Short-term direct effect Log(Y)			0.099	(3.36)
Short-term indirect effect Log(Y)			-0.018	(-0.45)
Long-term direct effect Log(P)	-1.013	(-24.73)	-1.931	(-9.59)
Long-term indirect effect Log(P)	**-0.220**	**(-2.26)**	0.610	(0.98)
Long-term direct effect Log(Y)	0.594	(10.45)	0.770	(3.55)
Long-term indirect effect Log(Y)	**-0.197**	**(-2.15)**	0.345	(0.48)

Notes: t-values in parentheses

The spatial spillover effects of both variables in the non-dynamic spatial Durbin model are negative and significant. Own-state price increases will restrain people not only from buying cigarettes in their own state, but to a limited extent also from buying cigarettes in neighboring states (elasticity -0.220). By contrast, whereas an income increase has a positive effects on cigarette consumption in the own state, it has a negative effect in neighboring states.

The negative price spillover effects is not consistent with Baltagi and Levin (1992), who found that price increases in a particular state —due to tax increases meant to reduce cigarette smoking and to limit the exposure of non-smokers to cigarette smoke— encourage consumers in that state to search for cheaper cigarettes in neighboring states.

However, whereas Baltagi and Levin's (1992) model is dynamic, it is not spatial; and whereas our model so far contains spatial interaction effects, it is not (yet) dynamic.

Effects estimates of cigarette demand using different model specifications

Determinants	(1)		(2)	
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}	
Short-term direct effect Log(P)			-0.262	(-11.48)
Short-term indirect effect Log(P)			**0.160**	**(3.49)**
Short-term direct effect Log(Y)			0.099	(3.36)
Short-term indirect effect Log(Y)			**-0.018**	**(-0.45)**
Long-term direct effect Log(P)	-1.013	(-24.73)	-1.931	(-9.59)
Long-term indirect effect Log(P)	-0.220	(-2.26)	**0.610**	**(0.98)**
Long-term direct effect Log(Y)	0.594	(10.45)	0.770	(3.55)
Long-term indirect effect Log(Y)	-0.197	(-2.15)	**0.345**	**(0.48)**

Notes: t-values in parentheses

The short-term spatial spillover effect of a price increase turns out to be positive; the elasticity amounts to 0.160 and is highly significant (t-value 3.49). Although greater and again positive, we do NOT find empirical evidence that the long-term spatial spillover effect is also significant. A similar result is found by Debarsy et al. (2011).

The spatial spillover effect of an income increase is not significant either. A similar result is found by Debarsy et al. (2011).

Effects estimates of cigarette demand using different model specifications

Determinants	(1)		(2)	
	Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with lag WY_{t-1}	
Short-term direct effect Log(P)			**-0.262**	**(-11.48)**
Short-term indirect effect Log(P)			0.160	(3.49)
Short-term direct effect Log(Y)			**0.099**	**(3.36)**
Short-term indirect effect Log(Y)			-0.018	(-0.45)
Long-term direct effect Log(P)	**-1.013**	**(-24.73)**	**-1.931**	**(-9.59)**
Long-term indirect effect Log(P)	-0.220	(-2.26)	0.610	(0.98)
Long-term direct effect Log(Y)	**0.594**	**(10.45)**	**0.770**	**(3.55)**
Long-term indirect effect Log(Y)	-0.197	(-2.15)	0.345	(0.48)

Notes: t-values in parentheses

Consistent with microeconomic theory, **the short-term direct effects** appear to be substantially smaller than **the long-term direct effects**; -0.262 versus -1.931 for the price variable and 0.099 versus 0.770 for the income variable.

The long-term direct effects in the dynamic spatial Durbin model, on their turn, appear to be greater (in absolute value) than their counterparts in the non-dynamic spatial Durbin model; -1.931 versus -1.013 for the price variable and 0.770 versus 0.594 for the income variable. Apparently, the non-dynamic model underestimates the long-term effects.

Conclusion

If a dynamic rather than a non-dynamic spatial econometric model is adopted and W is specified as binary contiguity matrix, empirical evidence is found in favor of the bootlegging effect. All other results also make sense from an economic-theoretical point of view.

On the other hand, questions that remain are the theoretical motivation in favor of adding a spatial lag W*Log(C), whether global spillovers make sense (SDM rather than SDEM), and whether is W correctly specified.

See SLX paper (Halleck Vega and Elhorst, JRS, 2015) and Bayesian comparison test of LeSage (2014, SDM vs. SDEM) for alternative approaches.

Common factors

The **dynamic spatial Durbin model** with spatial specific and time-period specific effects reads as

$$Y_t = \tau Y_{t-1} + \rho W Y_t + \eta W Y_{t-1} + \alpha \iota_N + X_t \beta + W X_t \theta + \mu + \alpha_t \iota_N + u_t.$$

where $\mu = (\mu_1, \dots, \mu_N)^T$. The spatial and time-period specific effects may be treated as fixed effects or as random effects. In the fixed effects model, a dummy variable is introduced for each spatial unit and for each time period (except one to avoid perfect multicollinearity), while in the random effects model, μ_i and α_t are treated as random variables that are independently and identically distributed with zero mean and variance σ_μ^2 and σ_α^2, respectively. Furthermore, it is assumed that the random variables μ_i, α_t, and ε_t are independent of each other.

$$Y_t = \boldsymbol{\tau Y_{t-1}} + \boldsymbol{\rho W Y_t} + \boldsymbol{\eta W Y_{t-1}} + X_t\beta + \boldsymbol{W X_t \theta} + \sum_r \boldsymbol{\Gamma_r^T f_{rt}} + u_t, \quad u_t = \boldsymbol{\lambda W u_t} + \varepsilon_t$$

Option 1 for $\sum_r \boldsymbol{\Gamma_r^T f_{rt}}$.

If two factors are considered, $f_{1t} = (1, \dots, 1)^T$ and $f_{2t} = (\alpha_1, \dots, \alpha_T)^T$, and the parameter restrictions $\Gamma_1^T = (\mu_1, \dots, \mu_N)$ and $\Gamma_2^T = (1, \dots, 1)$ are imposed, the dynamic spatial Durbin model is obtained with **Spatial and time-period fixed effects = two common factors**.

Formally, the spatial fixed effects represent one common factor (f_{1t}) which is constant over time but with heterogenous coefficients (Γ_1). The time-period fixed effects represent another common factor of length *T* (f_{2t}) which changes over time but which has homogenous coefficients (Γ_2). The total number of common factor parameters to be estimated in this setting amounts to *N+T-1*.

Demeaning variables by time dummies

$y_{it} \coloneqq y_{it} - \bar{y}_{.t}$ (expression for individual observations)

In vector notation

$$Y_t \coloneqq Y_t - \begin{bmatrix} \frac{1}{N} & \cdots & \frac{1}{N} \\ \vdots & \ddots & \vdots \\ \frac{1}{N} & \cdots & \frac{1}{N} \end{bmatrix} \begin{bmatrix} y_{1t} \\ \vdots \\ y_{Nt} \end{bmatrix},$$

and **the same for Y_{t-1} and all X-variables.**

Conclusion: Including time dummies as a common factor has the effect that all variables are demeaned, i.e., are taken in deviation of their corresponding cross-sectional averages with a **W** matrix in which all elements (including the diagonal elements) are $1/N$. These demeaned variables, including Y_t, may be treated as being exogenous provided that N goes to infinity, based on assumption 5 of Pesaran (2006). If time dummies are not included, ρ of $\boldsymbol{WY_t}$ will be overestimated.

Option 2: Keep the cross-sectional fixed effects, but replace the time dummies by cross-sectional averages (CSAs): $\bar{Y}_t = \frac{1}{N}\sum_{i=1}^{N} Y_{it}$, $\bar{Y}_{t-1} = \frac{1}{N}\sum_{i=1}^{N} Y_{it-1}$, and $\bar{X}_{kt} = \frac{1}{N}\sum_{i=1}^{N} X_{ikt}$ (k=1,..,K).

Objection to time period fixed effects: each time dummy has the same homogeneous impact on all observations in period t, while it is likely that, for example, business cycle effects hit one unit harder than another unit. Total number of common factor parameters to be estimated when accounting for heterogeneity by CSAs increases to $N+(2+K)*N$.

Since the numbers of parameters to be estimated increases rapidly with the number of common factors, most empirical studies try to keep the number of cross-sectional averages to a minimum. Often controlling for $\bar{Y}_t$ and $\bar{Y}_{t-1}$ only already effectively filters out the common time trends in the data (Cicarelli and Elhorst, 2018).

Pesaran (2006, assumption 5 and remark 3): CSAs may be treated as exogenous explanatory variables since the contribution of each unit to the CSAs at a particular point in time goes to zero if N goes to infinity.

Link with cyclical sensitivity literature

Thirlwall (1966) and Brechling (1967) demonstrate that regional unemployment rates tend to move in tandem with the national unemployment rate, but within the common rises and falls over time, the extent to which a region's rate responds to changes in the national rate can be quite heterogeneous.

This implies that heterogeneity is considered in both the old cyclical sensitivity literature and in the modern CSA literature and thus that common factors can be embedded in the economic-theoretical literature on cyclical sensitivity.

I use Matlab routines to estimate models with CF, but I will give an example on slide 40 how to use the xlsme command in Stata (Belotti et al. 2017) if you want to add cross-sectional averages.

Option 3: Principal components, in which case the *Γ* parameters represent the factor loadings of the principal components.

Shi and Lee (2017): Develop QML estimator for the dynamic GNS model with CF specified as principal components. This estimator does not require any specification of the distribution function of the disturbance term (explains the Q in QML). The coefficients estimates are bias-corrected for the Nickell bias and the impact of this bias on the other coefficients in the equation.

For this purpose, a Matlab routine called SFactors has been developed, which the first author (Shi) made available at his web site www.w-shi.net. I extended this with the calculation of R2 and the log-likelihood function value and posted this at spatial-panels.com.

A potential disadvantage of principal components is that they are often difficult to interpret, especially if they are compared with cross-sectional averages.

Every principal component requires the estimation of 2N additional parameters.

The command in Stata running a dynamic spatial panel data model with spatial fixed effects and common factor $\bar{Y}_t$ reads as: xsmle Y X $\bar{Y}_{1t}$ … $\bar{Y}_{Nt}$ wmat(W) model(sdm) durbin(X) dlag(3) fe type(ind) effects nsim(1000)

The option type(ind) controls for spatial fixed effects and the variable list $\bar{Y}_{1t}$ … $\bar{Y}_{Nt}$ controls for the cross-sectional average of the depedent variable with N unit-specific coefficients. Time dummies should not be included to avoid (near) perfect multicollinearity. The data structure to read in the cross-sectional averages takes the following form:

Unit	Time	Unit 1	Unit 2	…	Unit N
1	1	$\bar{Y}_1$	0	…	0
2	1	0	$\bar{Y}_1$	…	0
⋮	⋮	⋮	⋮		⋮
N	1	0	0		$\bar{Y}_1$
⋮	⋮	⋮	⋮		⋮
1	T	$\bar{Y}_T$	0	…	0
2	T	0	$\bar{Y}_T$	…	0
⋮	⋮	⋮	⋮		⋮
N	T	0	0	…	$\bar{Y}_T$

Testing for common factors

Cross-sectional dependence test of Pesaran (2015) in Econometric Reviews

The CD test uses the correlation coefficients between the time-series for each panel unit, which for N units results in N x (N-1) correlations between unit *r* and all other units, for *r*=1 to *N*-1. Denoting these estimated correlation coefficients between the time-series of two units *r* and *j* as $\hat{\rho}_{rj}$, the Pesaran (2015, eq.10) **CD test** is defined as $\mathrm{CD} = \sqrt{2T/N(N-1)}\sum_{r=1}^{N-1}\sum_{j=r+1}^{N}\hat{\rho}_{rj}$, where *T* is the number of observations on each unit over the observation period. This test statistic has the limiting *N*(0,1) distribution as *N* and *T* go to infinity first. This implies that the critical values of this two-sided test are -1.96 and 1.96 at the five percent significance level.

Two null hypotheses can be tested: *H_0: cross-sectional independence, H_1: cross-sectional dependence* (Theorem 2), ***H_0: weak cross-sectional dependence (α<1/2), H_1: strong cross-sectional dependence (α>1/2)*** **(Theorem 3).**

Exponent α of Bailey et al. (2016) in Journal of Applied Econometrics

To investigate the strength of the found cross-sectional dependence, one can compute the **exponent α of Bailey et al. (2016).** This statistic can take values on the interval (0,1] and measures the rate at which the variance of the cross-sectional averages tends to zero; $\alpha \leq 1/2$ points to weak cross-sectional dependence only and $\alpha = 1$ to strong cross-sectional dependence. Values in between indicate moderate to strong cross-sectional dependence and require additional research to discriminate between weak and strong cross-sectional dependence.

Both the CD-test and the exponent α are available in Stata: xtcd2 and xtcse2 (see do-file computer lab).

Testing for common-factors: CD-test and exponent α-estimator

Elhorst, J.P., Gross, M., Tereanu, E. (2021) Cross-sectional dependence and spillovers in space and time: where spatial econometrics and Global VAR models meet. *Journal of Economic Surveys* 35(1):192-226

Interplay between cross-section dependence, CF, weight structure and estimation

Note: α can be estimated consistently only for $1/2 < \alpha \leq 1$. Use Pesaran's CD test to find out whether α is smaller or greater than ½.

A	Cross section dependence	Weight structure	Estimation
0<α<0.5	weak	sparse: local, mutually dominant units	ML/IV/GMM
0.5<α<0.75	moderate	still quite sparse	
0.75<α<1	quite strong	Dense (GVAR literature)	OLS sufficient
1	strong	CS averages or PC (no weights involved)	

Practical guide suggested by Elhorst, Gross and Tereanu (2021):

1. **Assess the degree of strong cross-sectional dependence in the raw data** – Compute the CD-test of Pesaran (2004, 2015) and the corresponding exponent α of BHP (2016). A nonsignificant CD-test result or a significant CD-test result with a value of α significantly smaller than 3/4 indicates that the data are weakly dependent or moderately dependent. Then a spatial econometric model without CF suffices. By contrast, a significant CD-test and a value of α not significantly smaller than 1 suggests the presence of CF.
2. **Assess the degree of cross-sectional dependence in the residuals from step 1**. Apply the CD-test on the "de-factored" observations from step 1 in case a common factor model has been chosen. Failure to reject the null indicates possibly remaining weak cross-sectional dependence. The appropriate method would then be a common factor model with a sparse connectivity matrix W estimated by means of ML/IV/GMM.

Raw data cigarette consumption: CD=176.145, α=1.004424.

Estimation results of cigarette demand using different model specifications

Determinants	(1)		(2)		(3)		(4)		(5)	
	Non-dynamic spatial Durbin model no fixed effects		Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with cross-sectional averages		Dynamic spatial Durbin model with principal components	
Intercept	2.631	(15.82)								
Log(C)$_{-1}$					0.865	(65.04)	0.812	(59.71)	0.873	(72.27)
WLog(C)	0.337	(11.09)	0.264	(8.25)	0.076	(2.00)	0.069	(1.99)	0.001	(0.02)
WLog(C)$_{-1}$					−0.015	(−0.29)	0.093	(2.55)	0.080	(1.67)
Log(P)	−1.251	(−21.80)	−1.001	(−24.36)	−0.266	(−13.19)	−0.319	(−14.12)	−0.245	(−11.22)
Log(Y)	0.554	(14.96)	0.603	(10.27)	0.100	(4.16)	0.175	(5.10)	0.141	(4.75)
WLog(P)	0.780	(11.15)	0.093	(1.13)	0.170	(3.66)	0.304	(11.22)	0.260	(6.75)
WLog(Y)	−0.444	(11.09)	−0.314	(−3.93)	−0.022	(−0.87)	−0.172	(−4.93)	−0.027	(−0.70)
Werror									−0.013	(−0.19)
R^2	0.435		0.902		0.977		0.914		0.977	
LogL	475.5		1691.4		2623.3		2699.1		3078.5	
CD-test	25.16		−2.28		0.29		−2.93		−3.08	

Notes: t-values in parentheses

Effects estimates of cigarette demand using different model specifications

Determinants	(1)		(2)		(3)		(4)		(5)	
	Non-dynamic spatial Durbin model no fixed effects		Non-dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with fixed effects		Dynamic spatial Durbin model with cross-sectional averages		Dynamic spatial Durbin model with principal components	
Short-term direct effect Log(P)					−0.262	(−11.48)	−0.312	(−14.65)	−0.246	(−10.44)
Short-term indirect effect Log(P)					0.160	(3.49)	0.295	(11.39)	0.260	(6.04)
Short-term direct effect Log(Y)					0.099	(3.36)	0.172	(5.13)	0.143	(4.90)
Short-term indirect effect Log(Y)					−0.018	(−0.45)	−0.169	(−4.96)	−0.030	(−0.75)
Long-term direct effect Log(P)	−1.216	(−23.39)	−1.013	(−24.73)	−1.931	(−9.59)	−1.634	(−1.74)	−1.670	(−3.85)
Long-term indirect effect Log(P)	0.508	(7.27)	−0.220	(−2.26)	0.610	(0.98)	0.254	(0.01)	2.498	(0.27)
Long-term direct effect Log(Y)	0.530	(15.48)	0.594	(10.45)	0.770	(3.55)	0.899	(1.51)	1.384	(0.99)
Long-term indirect effect Log(Y)	−0.366	(−7.47)	−0.197	(−2.15)	0.345	(0.48)	−0.008	(−0.00)	2.680	(0.07)

Notes: t-values in parentheses

Empirical Results CF

To find out which set of common factors is able to filter out common factors most effectively, the cross-sectional dependence (CD) test developed by Pesaran (2015) may be used.

The conclusion from three empirical studies — Cicarelli and Elhorst (2018), Elhorst et al. (2020) and Elhorst (2021) — is that the best option (1, 2 or 3) to control for common time trends might differ from one empirical study to another.

Recommendation for empirical researchers:

-Always include WX variables to end up with flexible spillovers.

-If WY is included, motivate it from an economic-theoretical point of view.

-It is possible to estimate a dynamic GNS with principal components in Matlab, and a dynamic SDM with cross-sectional averages in Matlab or Stata.

Economic-theoretical underpinnings of $\boldsymbol{WY_t}$

-Anselin (2006): Conceptualization of strategic interaction or a spatial reaction function, $y_i = R(y_{\_i}, x_i)$, where $y_{\_i}$ reflects decisions by other agents. See for strategic interaction among local governments (Wildasin, 1988; Besley and Case, 1995; Brueckner, 2003, 2006; Allers and Elhorst, 2011).

-Pinkse et al. (2002) and LeSage et al. (2017): When one petrol station decreases its price, geographically nearby service stations need to follow in order not to lose market share.

-Hanson (2005): augmented market-potential function derived from Krugman's model of economic geography, reflecting the impact of scale economies and transport costs, explaining wages.

-Behrens et al. (2012): a quantity-based structural gravity equation system in which both trade flows and error terms are cross-sectionally correlated.

-Blonigen et al (2007): foreign direct investments (FDI).

-Xu and Lee (2019): SAR can be regarded as a model on the Nash equilibrium of a static complete information game with a linear-quadratic utility function.

Economic-theoretical underpinnings of $\boldsymbol{WX_t}$ variables

LeSage and Pace (2009): Several motivations, though mostly statistical.

Ertur and Koch (2007): SD model of GDP per capita growth (initial income level, savings rate, population growth rate).

Yesilyurt and Elhorst (2017): SD model of military expenditures as a ratio of GDP.

Firmino Costa da Silva et al. (2017): dynamic SD and GNS model of a spatially augmented population growth model.

Heijnen and Elhorst (2018): SD diffusion model of waste disposal taxes across municipalities.

Xu and Lee (2019): game-theoretical model can be extended with $\boldsymbol{WX_t}$ variables.

Conclusion

Dynamic spatial econometric models for spatial panels with common factors (CF) are the most advanced models currently available for empirical research.

I encourage more scholars to work with these models (CF) in their empirical research.

At the same time, I should warn you that these models are difficult model to work with since the estimation results produced by this model are often quite puzzling, especially in the beginning.

These advanced models require extensive research experience in spatial econometrics and sufficient economic-theoretical knowledge of the problem at hand. Often the results are not immediately in line with initial expectations, but after thinking them over and debating them with other researchers, progress towards an acceptable model specification can be made step by step.