J Geograph Syst (1999) 1:91–106

Journal of Geographical Systems

The econometric estimation and testing of DARP models

Emilio Casetti[1,2], **Ayse Can**[3]

[1] Department of Geography, Ohio State University, Columbus, OH 43210, USA
[2] Department of Economics, Odense University, DK-5230 Odense, Denmark
[3] Fannie Mae Foundation, 4000 Wisconsin Avenue NW, Washington, DC 20016-2804, USA
(e-mail: ayse_can@fanniemaefoundation.org)

Received: 4 March 1998/Accepted: 23 December 1998

Abstract. DARP, acronym for Drift Analysis of Regression Parameters, originated as a heuristic technique for the investigation of parametric drift in any arbitrary 'expansion space', geographic or otherwise. DARP was intended as an exploratory tool useful to aid with the formal specification of parametric drift. In this paper, the DARP technique is reformulated in terms of 'DARP models', and the estimation and testing of these models by GLS, FGLS, and ML are discussed. The ML estimation of a spatial DARP model is demonstrated using empirical data.

Key words: Darp models, expansion method, parametric drift, heteroskedasticity

JEL classification: C12, C13, C51, C52

1. Introduction

DARP (Casetti 1982) stands for Drift Analysis of Regression Parameters. It was proposed as a heuristic technique useful in the specification of models constructed by the Expansion Method. This introductory section outlines briefly the Expansion Method, DARP, and their relation to each other. Subsequent sections define the 'DARP models' and discuss their estimation by Generalized Least Squares (GLS), by Feasible Generalized Least Squares (FGLS), and by Maximum Likelihood (ML). An example of Maximum Likelihood estimation of a geographical DARP model and a concluding section cap the paper.

The Expansion Method (Casetti 1972, 1997) is both an approach to the construction of mathematical models suited to formalize and investigate parametric drift, and a research philosophy intimating that the variation of interesting relations across relevant contexts is a more fruitful object of enquiry than any quest for invariant 'laws'. It involves taking an initial model with at least some of its parameters in letter form and redefining at least some

Comments and criticisms by two anonymous reviewers are gratefully acknowledged.

of these letter parameters into functions of expansion variables. The expanded model thus obtained is capable of portraying the drift of the initial model in the space spanned by the expansion variables.

A wide class of expanded models can be written in terms of a linear initial model

$$Y = X\beta + \varepsilon \tag{1}$$

and expansion equation(s)

$$\beta = f(Z, \mathfrak{p}) + \eta \tag{2}$$

which yield the expanded model

$$Y = Xf(Z, \mathfrak{p}) + X\eta + \varepsilon, \tag{3}$$

where Y is a dependent variable, X is a matrix of independent variables, Z is a matrix of expansion variables, β and ρ are vectors of parameters, and ε and η are vectors of error terms. Given an initial model, a specification of $f(,)$, and the probability distributions of ε and η, the terminal model (3) can be estimated by whatever technique is appropriate, and the hypothesis that the initial model drifts can be tested.

In many research situations, the initial model is a well established formulation such as, for instance, a demand function, while its expansion spaces and expansion equation(s) are dealt with only implicitly in the pertinent substantive literatures. In these situations the question can arise of whether the initial model holds with different parameters across an expansion space, but we neither know that this drift does in fact occur, nor what forms it takes. Circumstances such as these call for trying and testing a number of specifications of $f(,)$ and of η. Clues as to which values β tends to have at selected points in the expansion space can help these searches, much in the same way that scatter diagrams can aid in the specification of econometric models. DARP was proposed as a heuristic device that can provide these clues.

DARP was designed to produce estimates of the initial model at locations in the expansion space referred to as 'reference points'. To this effect, DARP defines observation specific weights that are a decreasing function of the distances in Z space between each observation and a reference point. These weights 'deemphasize' the observations to an extent depending upon their distance from the reference point, so that in the estimation of an initial model, the observations more distant from the reference point will matter less then the observations close to it. Hence, DARP consists in the weighted estimation of the initial model's parameters using weights that are specific to a reference point. Repeating this weighted estimation for, say, a grid of reference points will yield parameter estimates that are potential clues about the drift of the initial model in the expansion space.

The plots or maps of the DARP estimates of the initial model's parameters versus the Z variable(s) can suggest specifications of the expansion equation(s) and of the terminal model that can be then subjected to statistical/econometric estimation and testing. One useful application of DARP in the spatial sciences is when the expansion space is spanned by the geographical coordinates of the observations as in Casetti and Jones (1983).

The recently proposed 'Geographically Weighted Regression' (Brunsdon et al. 1996; Fotheringham et al. 1997) is a geographical implementation of

DARP, complemented by a scheme for parameter estimation and testing. It is not clear, however, whether and to what extent, within the context of this scheme, the small and/or large sample properties of the estimator proposed have been established, and the statistical/econometric bases for inferences about the population parameters addressed. Of course, both of these are necessary to move away from a heuristic frame of reference.

DARP is related to the literatures on Kernel and Nearest Neighbor regressions (Cleveland 1979; Cleveland and Devlin 1988; Eubank 1988; Hardle 1990; Hastie and Tibshirani 1990; Muller 1988). However, the center of gravity of these techniques is on the often heuristic investigation of a model's drift in its predictor space for the purpose of obtaining a better specification of the model.

The Jacknifed DARP (JDARP for short) is a complement of DARP (Casetti 1983). While DARP deflates the observations distant from a reference point, JDARP deflates the observations close to it. If influential outliers are located in Z space in proximity to the reference point, JDARP reduces their impact on the estimation results. Thus JDARP results can suggest the removal of the observations located in some region(s) in Z space. Both DARP and JDARP yield estimates of the initial model at any arbitrary point in the expansion space. However, only the DARP estimates are useful in the investigation of parametric drift.

The DARP and JDARP techniques can be thought of as information filters centered upon a reference point and capable of weakening the information content of the observations distant from, or close to, it. However, they are reminiscent of the filters discussed in the Fourier analysis literatures, rather than of the ones designed to alleviate unwanted characteristics of the data (Getis 1995).

2. The DARP models

In this paper DARP and JDARP are recast in terms of DARP models. A DARP model is generated by expanding the variance of the error term of an initial model into a monotonic function of the distance between the observations and a reference point in an expansion space. Let us illustrate this definition using a standard linear econometric model as initial model, geographical space as the expansion space, and a geographical location as reference point. Let

$$Y = X\beta + \varepsilon \tag{4}$$

$$E(\varepsilon\varepsilon') = \sigma^2 I \tag{5}$$

be a standard econometric model. Here and throughout this paper it is assumed that the elements of ε are normally distributed. Assume that a sample of N observations is given, that two geographical coordinates identify each observation as a point in geographic space, and that a reference point is also specified in terms of its coordinates in geographic space. The reference point may or may not coincide with the location of an observation.

Denote by h a mononically increasing function of distance from a reference point, and by h_i the function's value generated by the distance between the ith observation and the reference point. 'Expand' the parameter σ^2 of (5) by the expansion equation $\sigma^2 = g(h, \gamma)$, so that

$$\sigma_i^2 = g(h_i, \gamma) \tag{6}$$

where γ is a vector of parameters, and σ_i^2 denotes the variance of the error term for the ith observation. Equations (4) and (6) define a DARP model.

In general, a DARP model is an econometric model in which the variance of the error terms has been expanded into a monotonic function of the observations' distance from a reference point in an expansion space Z. The obvious implication of this definition is that for the same specification of (4) and (6) we can have as many DARP models as there are points in the Z space: that is, infinitely many.

The specification of (6) employed in this paper is

$$\sigma_i^2 = \exp(\gamma_0 + \gamma_1 h_i) \tag{7}$$

where $\sigma_0^2 = \exp(\gamma_0)$, and h_i is the square of the distance between the ith observation and the reference point.

In matrix notation, the covariance of this class of DARP models can be written in the two equivalent formulations

$$E(\varepsilon\varepsilon') = \Phi = \sigma_0^2 \Psi \tag{8}$$

where Φ is a diagonal matrix in which the ith diagonal element is $\exp(\gamma_0 + \gamma_1 h_i)$, and Ψ is also a diagonal matrix with the ith diagonal element equal to $\exp(\gamma_1 h_i)$.

Let us focus briefly upon the significance of the parameter γ_1. A positive γ_1 deemphasizes the observations distant from the reference point, and consequently produces estimates of (4) that are 'local' with respect to it. A negative γ_1 deemphasizes the observations closer to the reference point, thus treating them as quasi outliers with a smaller role in the estimation of (4). If $\gamma_1 = 0$ the distance in Z space between the observations and the reference point does not affect the variance of the error terms, and $\sigma_i^2 = \sigma_0^2$ for all i's.

A DARP analysis involves the estimation of a set of DARP models sharing an initial model, an expansion space, and variance expansions, but associated with a set of distinct reference points. In the case considered here the variance expansions are specified by (7). The γ_1's in (7) can be assumed or estimated. A DARP analysis in which the same positive value of γ_1 is assumed for every DARP model in it generates local estimates of the initial model that can suggest the occurrence of parameter drift. If the assumed γ_1 has the same negative value in all the DARP models, the analysis implements a search for differential performance of the initial model across the expansion space, and its results can suggest the occurrence of 'performance' drift.

It should be noted that a DARP analysis in which the value of γ_1 is assumed is quite legitimate. The straightforward estimation of a standard econometric model by OLS is equivalent to the estimation of a DARP model assuming that $\gamma_1 = 0$. If it is legitimate to assume that $\gamma_1 = 0$, it must be also legitimate to assume that γ_1 has any other arbitrary value, provided that there is a rationale for it. The search for evidence of parametric or performance drift is one such rationale.

If instead γ_1 is estimated, it is the analysis itself that determines whether and where parameter drift or performance drift occurs, and these conditions may very well coexist in different regions in the expansion space.

Summing up, the DARP models focussed upon here are defined by (4) and (8). Three approaches to the estimation of these models are discussed in the sections that follow.

3. The estimation of DARP models

The DARP models are a special class of heteroskedastic models. The DARP models with the expansion specified by (7) are a special type of multiplicative heteroskedasticity models (Harvey 1976). In the section that follows we discuss briefly and in generalities the estimation of the DARP models defined by (4) and (8) by Generalized Least Square (GLS), by Feasible Generalized Least Square (FGLS), and by Maximum Likelihood (ML). Discussions of the GLS, FGLS, and ML estimators in general and with special regard to the estimation of models with multiplicative heteroskedasticity can be found, for instance, in Fomby (1988), Harvey (1990), and Judge et al. (1985, 1988).

In the paragraphs that follow, estimators will be denoted by the roman counterparts of the Greek letters indicating the corresponding parameters, and the GLS, FGLS, and ML estimators will be denoted by GS, FS, and ML subscripts, which are the first and last letters in their respective acronyms. For instance, b_{FS} and b_{iFS} indicate respectively the FGLS estimators of the parameter vector β and of the parameter β_i.

3.1 GLS estimation

If the value of γ_1 is either known or assumed, Ψ is also known, and the GLS estimators of β, of its covariance, and of σ_0^2 are respectively

$$b_{GS} = (X'\Psi^{-1}X)^{-1}X'\Psi^{-1}Y, \tag{9}$$

$$var(b_{GS}) = s_{GS}^2(X'\Psi^{-1}X)^{-1}, \tag{10}$$

$$s_{GS}^2 = \frac{(Y - Xb_{GS})'\Psi^{-1}(Y - Xb_{GS})}{N - K}, \tag{11}$$

where K stands for the number of columns in X. Notably, b_{GS} is BLUE, s_{GS}^2 is an unbiased estimator of σ_0^2, and since the elements of ε are assumed to be normally distributed, the t and F statistics calculated using (9), (10), and (11) can be used to test hypotheses as in the case of a standard linear model estimated by OLS.

The matrix Ψ^{-1} that appears in (9), (10), and (11) is not needed to calculate the GLS estimators, since a weighted regression using a vector of weights, w, with elements $w_i = \exp(-\gamma_1 h_i)$ would yield the GLS estimators and their statistics. In fact Ψ^{-1} is a diagonal matrix with these weights in its principal diagonal.

If the assumed γ_1 is positive, the GLS estimates of β obtained are potential indicators of the parametric drift of the initial model. If it is negative, these estimates are indicators of performance drift of initial model in Z space. Hypothesis testing and confidence intervals can be used to identify regions in Z space in which parametric and/or performance drift materialize. Consequently, the estimation of DARP models for a set of reference points and for a range of values of γ_1 can indeed produce information useful, in the aggregate, to specify suitable expansion equations for the β parameters, and also to decide whether subsets of the sample should be either removed, or dealt with in separate analyses.

The GLS estimation of a DARP model presupposes that the parameter γ_1 is known or assumed. To estimate γ_1 as well as β, Feasible Generalized Least Square (FGLS) or Maximum Likelihood (ML) can be used. Let us consider the FGLS estimation first.

3.2 FGLS estimation

The FGLS estimation of a DARP model involves obtaining an estimate of γ_1, for instance by LS, and then using it as if it were the known γ_1 parameter in GLS estimation. Specifically, the steps required are as follows. The OLS estimator of β, b, and the residuals $e = Y - Xb$ are obtained. Using the elements of e, the vector $q' = (\ln(e_1^2), \ldots, \ln(e_N^2))$ is constructed and LS is applied to $q = \gamma_0 + \gamma_1 h + v$, where $v' = ((\ln(e_1^2/\sigma_1^2), \ldots, \ln(e_N^2/\sigma_N^2))$, to obtain the estimator $\hat{\gamma}_1$ of γ_1. The elements of v have non zero expectations and covariances and are heteroskedastic. However, it can be shown that they are asymptotically homoskedastic and with zero covariances, and also that while $\hat{\gamma}_0$ is biased with a known bias, $\hat{\gamma}_1$ is unbiased. Which implies that $\hat{\gamma}_1$ is a consistent estimator of γ_1 but $\hat{\gamma}_0$ is not a consistent estimator of γ_0. Here only $\hat{\gamma}_1$ is needed. By replacing $\hat{\gamma}_1$ for γ_1 in (8) the estimator $\hat{\Psi}$ of Ψ is obtained.

The FGLS estimator of β is

$$b_{FS} = (X'\hat{\Psi}^{-1})^{-1} X'\hat{\Psi}^{-1} Y \tag{12}$$

The small sample properties of b_{FS} are unknown. It is known, however, that b_{FS} is asymptotically normally distributed, with mean β, and with a covariance matrix that is consistently estimated by

$$s_{FS}^2 (X'\hat{\Psi}^{-1} X)^{-1} \tag{13}$$

where

$$s_{FS}^2 = \frac{(Y - Xb_{FS})'\hat{\Psi}^{-1}(Y - Xb_{FS})}{N - K} \tag{14}$$

is a consistent estimator of σ_0^2. Hypothesis testing and confidence intervals can be based on these results, under the presupposition that the sample used is large 'enough'. For a full discussion of the FGLS estimation of multiplicative hetheroskedasticity models cfr Judge et al. (1988 p. 367 ff) or Harvey (1976 p. 462).

The statistic $(\hat{\gamma}_1)^2/4.9348\Sigma h_i$ has an asymptotic chi-square distribution with one degree of freedom, and it can be used to test the null hypothesis that $\gamma_1 = 0$ (Judge et al. 1988 p. 370).

3.3 ML estimation

When the probability distribution of the error terms in a model is known, the model's likelihood can be expressed as a function of the parameters to be estimated, and the ML estimates are the parameter values that maximize the model's likelihood for the data given. Here the error terms in the vector ε are normally distributed and independent. The ML estimation of a DARP model specified by (4) and (8) involves expressing the model's likelihood as a function of β and γ given Y, X, and h. The logarithmic transformation of the likelihood function, that is conventionally used since it attains a maximum for the same parameter values that maximize the likelihood function is here

$$L(\beta, \gamma | Y, X, H) = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\ln|\Phi| - \frac{1}{2}(Y - X\beta)'\Phi^{-1}(Y - X\beta), \quad (15)$$

where the N by 2 matrix H can be defined in terms of its ith row $H_i = (1, h_i)$. In order to maximize L with respect to β and γ the partial derivatives $\partial L/\partial\beta$ and $\partial L/\partial\gamma$ are obtained and set to zero. The resulting system of non linear equations is solved numerically to obtain the ML estimators of β and γ. These estimators can be calculated iteratively by the method of scoring. The ML estimators b_{ML} and c_{ML} are consistent, and are asymptotically independent and normally distributed with asymptotic covariance matrices

$$(X'\Phi^{-1}X)^{-1} \quad (16)$$

and

$$2(H'H)^{-1}. \quad (17)$$

Testing and confidence intervals of the elements of b_{ML} and c_{ML} can be based on the covariance matrices (16) and (17), with the estimated Φ constructed from the c_{ML} vector at convergence of the scoring algorithm. However, the test of the null hypothesis that $\gamma_1 = 0$ can be also based on the Wald statistic $(c_{1ML})^2/(2\Sigma h_i)$ that has a chi-square distribution with one degree of freedom. For a discussion of the ML estimation of multiplicative heteroskedasticity models cfr Harvey (1976; 1991 p. 98 ff and 134 ff), Greene (1997 p. 565 ff), and Fomby et al. (1984 p. 183 ff).

It is useful to comment on the comparative advantages of the FGLS and ML estimators. Both b_{FS} and b_{ML} are consistent, and both have an asymptotic covariance matrix equal to $(X'\Phi^{-1}X)^{-1}$, which means that they are equally asymptotically efficient. The FGLS and the ML estimators of γ_1 are also both consistent, but their asymptotic variances are respectively $4.9348\Sigma h_i$ and $2\Sigma h_i$. This implies that c_{1ML} is asymptotically more efficient than c_{1FS}. An important implication of this difference is that the ML estimates of both β and γ will converge faster to the true parameter values than their FGLS counterparts, and consequently will perform better with any sample of less than infinite size.

While the ML estimators are preferable to their FGLS counterparts, the latter represent a workable alternative if the iterations required to obtain the ML estimators fail to converge. However, the analyst who wishes to pursue this course of action might prefer the estimation alternative suggested by

Harvey (1976) whereby $\hat{\gamma}_1$ is replaced by an estimator of γ_1 produced by one iteration of the scoring algorithm (Fomby et al. 1984 p. 185 ff).

The demonstration that follows employs the ML estimation. The empirical analyses in the demonstration were carried out using the command for the estimation of multiplicative heteroskedasticity models (HREG) in the econometric package LIMDEP ver. 7.

4. A demonstration

In the section that follows, the ML estimation of DARP models is demonstrated using the 'Expenditures Data Set' published in Pindyck and Rubinfeld (1991, Table 6.2, pp. 155–156). The observations in this data set are the 48 conterminous states in the US. Its variables are: total state and local government expenditures, EXP; state income, STINC; federal grants to the states, AID; and state population, POP. EXP, STINC, and AID are in millions of dollars, POP is in million persons. These data were complemented by the spatial coordinates of the states' centroids.

The analyses reported here were carried out on the per capita variables PCEXP, PCINC, and PCAID, obtained by dividing EXP, STINC and AID by POP. The DARP models estimated are

$$\text{PCEXP} = \beta_0 + \beta_1 \text{PCINC} + \beta_2 \text{PCAID} + \varepsilon \tag{18}$$

$$E(\varepsilon\varepsilon') = \Phi, \tag{19}$$

where Φ is a diagonal matrix with the ith element in its principal diagonal equal to $\exp(\gamma_0 + \gamma_1 h_i)$; h_i is the squared distance between the centroid of the ith state and a reference point. Since the states' centroids were selected as reference points, 48 DARP models, one per observation, were estimated by ML. The estimation itself was not computationally burdensome. The number of iterations required ranged from a minimum of 7 to a maximum of 28.

Selected estimation results are shown in Table 1. Each line in the table reports results pertaining to the DARP model for the State identified by the code in column 1. Columns 2, 3, and 4 give the ML estimates of β, the estimate of γ_1 is in column 5, and the one's and zero's in column 6 indicate whether the null hypothesis that $\gamma_1 = 0$ is rejected or not. The test is based on the asymptotic normality of the ML estimators and uses the covariance estimated at the point of convergence of the scoring algorithm. Since all the estimates of the β's are significant no inferentials concerning them are reported. Also unreported are the estimates of γ_0, that is not relevant to the discussion carried out here.

Special attention need be given to the estimates of γ_1 that tell us whether in the expansion space there are regions characterized by parametric and/or performance drift. Of the 48 estimates of γ_1 32 are negative, and 16 are positive. Of the 32 negative estimates 25 are significant; of the 15 positive only 2 are significant.

The spatial patterns of the estimated γ_1's are portrayed in Fig. 1. The triangular symbols in Fig. 1 show where the positive/negative significant/nonsignificant estimates of γ_1 are located. If we draw on the map a line from

Table 1. DARP analyses: ML estimates

Code (1)	b_{OML} (2)	b_{1ML} (3)	b_{2ML} (4)	c_{1ML} (5)	sigf(c_{1ML}) (6)
ME	−0.445960	0.215698	1.66422	−0.0015336	1
NH	−0.442991	0.215347	1.65960	−0.0016317	1
VT	−0.444211	0.215516	1.66223	−0.0017184	1
MA	−0.440019	0.214913	1.65531	−0.0016027	1
RI	−0.438757	0.214711	1.65356	−0.0015635	1
CT	−0.438822	0.214715	1.65449	−0.0016406	1
NY	−0.442875	0.215269	1.66331	−0.0019222	1
NJ	−0.436029	0.214167	1.65417	−0.0017213	1
PA	−0.438293	0.214350	1.66293	−0.0020209	1
OH	−0.439127	0.213141	1.69493	−0.0025267	1
IN	−0.440121	0.210904	1.74618	−0.0027968	1
IL	−0.441915	0.207530	1.82067	−0.0028293	1
MI	−0.457044	0.214898	1.74229	−0.0032334	1
WI	−0.467845	0.211356	1.86440	−0.0037897	1
MN	−0.470856	0.206048	1.97285	−0.0032203	1
IA	−0.447715	0.203882	1.91002	−0.0023521	0
MO	−0.432826	0.202781	1.85661	−0.0014320	0
ND	−0.445783	0.203349	1.88993	−0.0008193	0
SD	−0.433592	0.203038	1.83573	−0.0002410	0
NE	−0.425948	0.203234	1.79010	0.0003021	0
KS	−0.424419	0.203675	1.76777	0.0006764	0
DE	−0.433665	0.213646	1.65443	−0.0017281	1
MD	−0.433535	0.213515	1.65726	−0.0018140	1
VA	−0.430367	0.212435	1.66561	−0.0018865	1
WV	−0.433320	0.212718	1.67493	−0.0021295	1
NC	−0.425973	0.211103	1.67164	−0.0017601	1
SC	−0.422023	0.209292	1.68906	−0.0016814	1
GA	−0.418809	0.206684	1.72433	−0.0015724	1
FL	−0.414232	0.204783	1.73540	−0.0011214	1
KY	−0.431103	0.209481	1.73117	−0.0023451	1
TN	−0.425734	0.206785	1.75733	−0.0020328	1
AL	−0.419470	0.203916	1.77754	−0.0013680	0
MS	−0.424605	0.202538	1.81556	−0.0006987	0
AR	−0.429179	0.202860	1.82019	−0.0002074	0
LA	−0.433929	0.204033	1.80134	0.0006179	0
OK	−0.426022	0.204316	1.75443	0.0010312	0
TX	−0.429902	0.205969	1.72677	0.0015175	0
MT	−0.417643	0.203324	1.75082	0.0004603	0
ID	−0.412726	0.204453	1.70476	0.0006978	0
WY	−0.415740	0.203911	1.72643	0.0007331	0
CO	−0.417252	0.204744	1.71113	0.0010132	0
NM	−0.422239	0.206326	1.69736	0.0012233	0
AZ	−0.421913	0.207162	1.68518	0.0010388	1
UT	−0.415680	0.205566	1.69166	0.0009231	0
NV	−0.415489	0.206118	1.68294	0.0008242	0
WA	−0.411624	0.204267	1.70623	0.0005707	0
OR	−0.412181	0.205169	1.69005	0.0006647	0
CA	−0.417403	0.206908	1.67697	0.0007931	1

North Dakota to Florida all the States to the right hand side of the line have negative and mostly significant c_{1ML}'s, while to the left of the line the c_{1ML}'s are mostly positive and non significant.

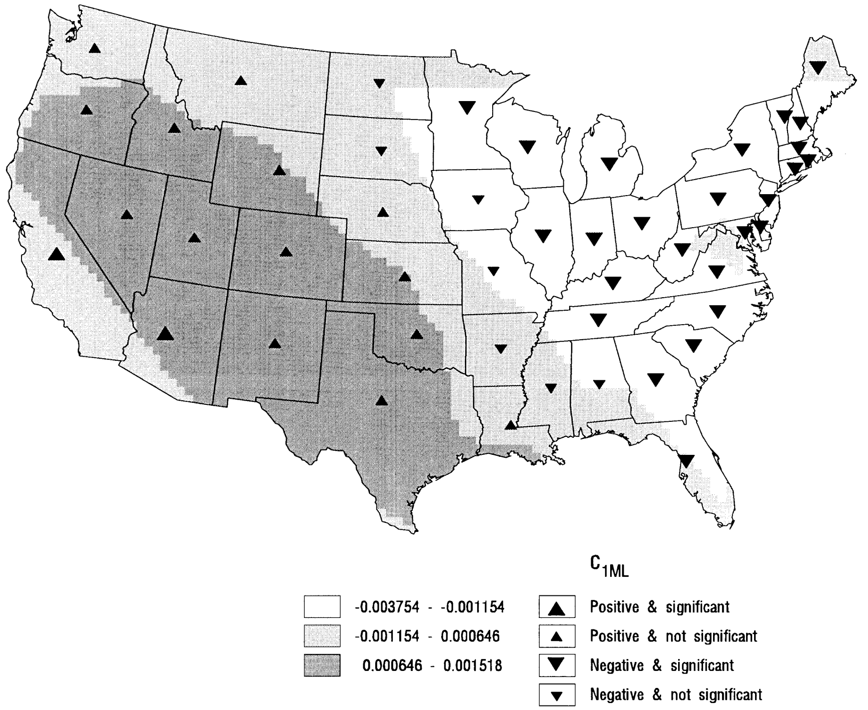

Fig. 1. Spatial distribution of c_{1ML}

Specifically, the spatial pattern of the estimated γ_1's suggest that the parameters of the initial model (18) may drift across the geographical space south of North Dakota to Florida line, and that the model performs comparatively less well north of this line. However, the possibility that parametric drift does occur in this region also cannot be ruled out. DARP estimates with $c_{1ML} < 0$ signal a weak performance of the initial model at and near the reference points involved, but do not imply the absence of parametric drift. Conversely, DARP estimates with $c_{1ML} > 0$ as in the case of the region south of the North Dakota to Florida line suggest the possible occurrence of parametric drift, but do not imply the absence of spatial variation in the performance of (18) across this region. Both performance drift and parameter drift can drive the DARP estimates. When the first prevails, $c_{1ML} < 0$; when the second prevails, $c_{1ML} > 0$.

A DARP analysis can uncover the possible tendency of the initial model's parameters to have different values across the expansion space or portions of it ('parametric drift'), and the possible occurrence of regions in the expansion space in which the initial model fits less well or not at all ('performance drift'). These two 'types' of results are qualitatively different. The simultaneous occurrence of both within the same analysis, as in Table 1 and Fig. 1, can be conceptualized in terms of two complementary constrained estimations. In one of these the DARP model is estimated subject to the constraint that $\gamma_1 \leq 0$, while in the other the same model is subject to the constraint that $\gamma_1 \geq 0$.

The $\gamma_1 \geq 0$ constraint is here taken to mean that whenever $c_{1ML} < 0$ we replace the estimate of β obtained, with the one based on assuming that $\gamma_1 = 0$. This $\gamma_1 \geq 0$ constraint is appropriate when we investigate the possible occurrence of parametric drift, since it results in estimates that do not address the performance drift. Instead, when the variation in the initial model's performance is focussed upon, the complementary $\gamma_1 \leq 0$ constraint is appropriate. This constraint yield estimates that do not address the possible occurrence of parametric drift, and is operationalized by replacing, whenever $c_{1ML} > 0$, the estimates of β obtained with the ones based on assuming that $\gamma_1 = 0$. Only the constrained ML estimates of β based on the $\gamma_1 \geq 0$ constraint are dealt with in this demonstration. These estimates will be denoted by a D in their subscripts. For instance, b_D and b_{iD} indicate the constrained ML estimates of β and β_i based on the $\gamma_1 \geq 0$ constraint.

The ML estimation of (18), and (19) for $\gamma_1 = 0$ is identical to the ML estimation of (18) with the stipulation that $E(\varepsilon\varepsilon') = \sigma^2 I$. The results from such estimation are

$$\underset{}{\text{PCEXP}} = \underset{(-4.29)}{-0.42967} + \underset{(10.81)}{0.20307}\ \text{PCINC} + \underset{(8.21)}{1.8145}\ \text{PCAID} + e \quad R^2 = 0.76. \tag{20}$$

The t values are in parenthesis under their respective coefficients. The b_D resulting from the constrained ML estimation of (18) and (19) based on the $\gamma_1 \geq 0$ constraint can be obtained by replacing the b_{ML} in the lines of Table 1 in which $c_{1ML} < 0$ by the corresponding estimates in (20). This b_D documents the possible occurrence of parametric drift.

In order to display the parametric drift of the β's across the expansion space considered, b_{0D}, b_{1D}, and b_{2D} were mapped. Figures 2, 3, and 4, do indeed suggest that the β's drift in geographical space. In a demonstration such as the one presented here these comments will suffice. In a substantive research, though, these results would have to be investigated more closely, and the analyst might want to experiment with the respecification of the initial model and/or with the segmentation of the data and/or with the expansion of some or all the initial model's parameters.

The segmentation of the data set could involve either the removal of some observations, or the partitioning of the data set into subsets to be analyzed separately. The most obvious expansions suggested by the DARP results considered are 'trend surfaces expansions', that involve redefining some or all the β's as low degree polynomials in the observations' coordinates. However, the trend surface expansions can prove difficult to work with because of the very substantial multicollinearity that they tend to engender. Alternative, more promising expansions could consist in redefining the β's into functions of an index reflecting the spatial patterns revealed by the maps. The fact that the states to the right of the North Dakota to Florida line tend to be older, smaller, and more closely spaced than the ones to the left of the line, can perhaps suggest suitable indices and can also point to the substantive rationales behind the parametric drift observed.

Let us note that our demonstration of a DARP analysis was successful. The ML estimation of the demonstration models converged to a result in a few iterations at all the reference points considered. Also, the estimation results are readily interpretable and substantively intriguing. They suggest the

Fig. 2. Spatial distribution of b_{0D}

occurrence of both parametric and performance instability in geographic space. Overall, the demonstration indicates, albeit in a preliminary fashion, that ML is a workable approach to the estimation of DARP models.

5. Conclusions

Let us place in perspective the nature, scope, and utility of the DARP analyses in general, and of the spatial DARP analyses in particular. The 'DARP models' discussed in this paper are generated by 'expanding' the variance parameter of an initial econometric model into a monotonic function of distance from reference points in an expansion space. Thus, they qualify as 'expansion models'. The estimates of DARP models associated with a collection of reference points constitute a DARP analysis, and can be regarded as a nonparametric expansion of the initial model involved.

The DARP analyses are applicable to investigating the parametric and performance drift of any conceivable initial model across any conceivable expansion spaces. These include the spaces spanned by some or all the predictor variables, or by variables that do not appear among the predictor variables, or by a mix of both. Consequently, the scope of the DARP analyses is wider than that of the Kernel and Nearest Neighbor regressions that focus upon a model's instability across its predictor space. In this respect these techniques represent a special case within the frame of reference encompassing DARP and the Expansion Method.

Fig. 3. Spatial distribution of b_{1D}

Let us sketch briefly some possible outcomes of a DARP analysis. If for none of the models in a DARP analysis the null hypothesis of no drift can be rejected, the analysis suggests that the initial model is stable in the expansion space considered. The 'opposite' outcome occurs when we have a successful DARP analysis, namely, when at least some of the estimated DARP models identify regions in the expansion space in which parametric and/or performance instabilities exist. A successful DARP analysis opens the door to a variety of responses.

Some responses might start from a respecification of the initial model. Others would retain the same initial model, but take the drift conditions uncovered as the starting point of a sequence involving data segmentations and/or expansions followed by new DARP analyses. The sequences involved would terminate when the models arrived at are stable over the data from which they are estimated.

However, another response to a successful DARP analysis may consist in the simple recognition that parametric drift and/or performance drift exist. Such recognition could be followed by the attempt to use pertinent bodies of qualitative and quantitative knowledge to explain, interpret, and possibly theorize, this finding. Scholarly outputs bundling the description of a DARP analysis with graphs, maps and narratives can constitute an implementation of this response. There seems to be no a priori rationale suggesting that a particular type of response to a successful DARP analysis is best under all circumstances.

Next, let us consider the 'spatial' DARP analyses. A spatial DARP model

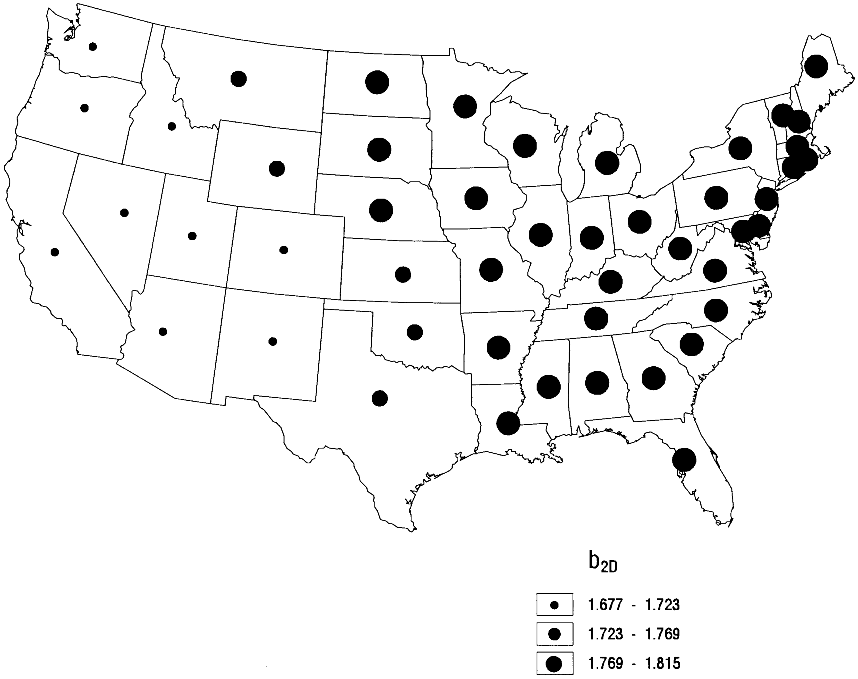

Fig. 4. Spatial distribution of b_{2D}

is generated by expanding the variance parameter of an initial econometric model, spatial or otherwise, into a monotonic function of the observations' distance from a reference location in a geographic expansion space. A spatial DARP analysis is the set of estimates of spatial DARP models associated with a set of reference locations in a geographic space. Questions concerning, for example, the drift of a production function across a region, or the drift of a spatial interaction model across a continent can be readily addressed via spatial DARP analyses. The results obtained constitute non parametric spatial expansions, and may lead to the subsequent specification and testing of parametric spatial expansions based on geographical coordinates (or transformations thereof), or on indices of some phase of 'spatial differentiation'.

Within the context of a spatial DARP the reference points are locations that do not necessarily coincide with the locations of the observations. This disconnect between reference points and observations is potentially very useful. Consider for example a DARP analysis in which the observations are the 3100 plus counties in the United States. Using as reference points a few hundred locations at the intersections of a grid rather than county specific locations can render much easier to obtain and map the results of the analysis. In other circumstances it may be convenient to locate equally spaced reference points on a transect, or on a route such as an interstate highway, in order to investigate the possible parametric or performance drift of an initial model as we move along the transect or route.

It is useful to touch briefly upon the interface between spatial DARP and related parametric expansions, on one hand, and some central themes of spatial econometrics (Anselin 1988). Suppose that a model is estimated from spatially referenced observations, and that subsequent testing indicates a significant spatial autocorrelation in its residuals. One typical response calls for testing, and possibly modifying, the econometric specification of the model. Spatially autoregressive terms and spatially autocorrelated error terms are a likely object of this specification search, and may eventually produce a successful model, purged of any significant spatial autocorrelation in its residuals. What characterizes this response is that it does not require a respecification of the substantive segment of the original model.

An alternative response to positive spatial autocorrelation tests might consist in a specification search focussed upon the substantive segment of the original model. Spatial DARP analyses and subsequent parametric expansions inspired by their results are one possible avenue to implement it. The DARP analyses are oriented toward seeking and addressing the structural instability that constitutes a substantive shortcoming of the initial model. Any expanded model arrived at by this process can be subjected to further DARP analyses in order to determine whether both the instabilities originally observed and the spatial autocorrelation have in fact disappeared.

However, initial models including spatially autoregressive terms and/or spatially autocorrelated error terms can be the specification an analyst starts from. In this situation also, suitable spatial DARP analyses and subsequent parametric expansions could be used to investigate and address the possible structural instability of the initial formulation.

Summing up, the searches centering on DARP and/or on parametric expansions, and those centering on spatial econometric respecifications are distinct responses to the perverse outcome of diagnostic tests. Which one of these two types of searches is appropriate when, constitutes an interesting and important question. Possibly, the circumstances specific to a given project and the judgement of the project's scholars, should be relied upon for the guidance to address it.

References

Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht

Brundson C, Fotheringham AS, Charlton ME (1996) Geographically weighted regregression: A method for exploring spatial nonstationarity. Geographical Analysis 28:281–298

Casetti E (1972) Generating models by the expansion method: Applications to geographic research. Geographical Analysis 4:81–91

Casetti E (1982) Drift analysis of regression parameters: An application to the investigation of fertility development relations. Modeling and Simulation 13 (Part 3):961–966

Casetti E (1983) Jacknifed DARP: An application to non linear dynamics. Modeling and Simulation 14 (Part 3):751–756

Casetti E (1997) The expansion method, mathematical modeling, and spatial econometrics. International Regional Science Review 20:9–33

Casetti E, Jones III JP (1983) Regional shifts in the manufacturing productivity Response to Output Growth: Sunbelt Versus Snowbelt. Economic Geography 4:285–301

Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74:829–836

Cleveland WS, Devlin SJ (1988) Locally weighted regression: An approach to regression analysis by local fitting. Journal of the American Statistical Association 83:596–610

Eubank RL (1988) Spline smoothing and nonparametric regression. Marcel Dekker, New York

Fomby TB, Hill CR, Johnson SR (1988) Advanced econometric methods. Springer, Berlin Heidelberg New York

Fotheringham AS, Charlton M, Brunsdon ME (1997) Measuring spatial variations in relationships with geographically weighted regression. In Fischer MM, Getis A (eds) Recent developments in spatial analysis. Springer, Berlin Heidelberg New York, pp 60–82

Getis A (1995) Spatial filtering in a regression framework: Examples using data on urban crime, regional inequality, and government expenditures. In: Anselin L, Feorax RJGM (eds) New Directions in Spatial Econometrics. Springer, Berlin Heidelberg New York, Chap 8, pp 172–185

Greene WH (1997) Econometric analysis. Prentice-Hall, Englewood Cliffs, NJ

Hardle W (1990) Applied nonparametric regression. Cambridge University Press, New York

Harvey AC (1976) Estimating regression models with multiplicative heteroscedasticity. Econometrica 44:461–465

Harvey A (1990) The econometric analysis of time series, 2nd edn. MIT Press, Cambridge, MA

Hastie TJ, Tibshirani RJ (1990) Generalized additive models. Chapman and Hall, London

Judge GG, Griffiths WE, Lutkepohl H, Lee TC (1985) The theory and practice of econometrics, 2nd edn. Wiley, New York

Judge GG, Hill RC, Griffiths WE, Lutkepohl H, Lee TC (1988) Introduction to the theory and practice of econometrics, 2nd edn. Wiley, New York

Muller HG (1988) Nonparametric regression analysis of longitudinal data. Springer, Berlin Heidelberg New York

Pindyck RS, Rubinfeld DL (1991) Econometric models and economic forecasts, 3rd edn. McGraw Hill, New York