Environment and Planning A, 1984, volume 16, pages 529-543

Spatial flows and spatial patterns†

A S Fotheringham
Department of Geography, Indiana University, Bloomington, IN 47405, USA
Received 28 March 1983; in revised form 24 May 1983

Abstract. The misspecification of gravity spatial interaction models has recently been described by the author. The bias in parameter estimates that results from such misspecification appears to produce the 'map pattern effect' or 'spatial structure bias' in estimated distance-decay parameters. A further aspect of the misspecification bias in gravity parameter estimates is explored here. The severity of the bias is shown to vary in a predictable manner with variations in spatial structure. In particular, the bias is shown to be dependent upon the pattern of accessibility that exists within a spatial system. The relevant aspects of this pattern are discussed for intraurban and interurban (or interregional) flow matrices. It is shown that from an examination of the spatial structure of centres in a spatial system it is possible, a priori, to identify whether significant bias will arise in the calibration of a gravity model. Certain configurations of centres are shown theoretically to produce maximal bias. The author thus answers the question, "why do gravity parameter estimates appear to be biased in some systems but not in others?"

1 Introduction

In spatial interaction modelling, distance-decay can be defined as the rate at which the volume of interaction between centres decreases as the distance between the centres increases, ceteris paribus. An estimated distance-decay parameter obtained in the calibration of a gravity interaction model is assumed to measure this rate and, hence, is usually given a behavioural interpretation. The parameter estimate is traditionally assumed to measure the perceived deterrence of distance to interaction. However, I have recently shown (Fotheringham, 1983a; 1983b) that gravity interaction models are misspecifications of reality because they fail to model accurately relationships between destinations, and this misspecification produces bias in parameter estimates. Relationships between destinations may arise either from forces of competition or from forces of agglomeration. If the former are present, then, as the proximity of a destination to other destinations increases, interaction to that destination decreases, ceteris paribus. If the latter are present, then, as the proximity of a destination to other destinations increases, interaction to that destination increases, ceteris paribus. Spatial interaction models that fail to model competition or agglomeration forces accurately are henceforth described as misspecified, meaning that the models are misspecifications of reality. The bias in parameter estimates that results from such misspecification is related to the spatial arrangement of centres in the system under investigation so that such estimates cannot be given any behavioural interpretation; nor can they be expected to remain constant over time or space, an assumption often made in forecasting.

Two questions arise concerning the stability and regularity of the bias in gravity model parameters. One question concerns the relationship between model format and parameter bias and arises because of differences in the patterns of parameter estimates that are observed when different forms of the gravity model are calibrated. I have shown (Fotheringham, 1983a), however, such that differences are to be expected since different degrees of misspecification exist between gravity models: even

† This research was undertaken with the assistance of a grant (number SES 82-08339) from the National Science Foundation.

when competition or agglomeration effects occur, attraction-constrained and doubly-constrained gravity models can be correctly specified under certain conditions, whereas total-flow-constrained and production-constrained gravity models are always misspecified. The other question, which as yet remains unanswered, concerns differences in the size and direction of parameter bias over space. Why do such differences exist even when the same model is employed? For example, whereas in most systems, the bias in gravity-model distance-decay parameters produces a relationship between estimated distance-decay parameters and origin accessibility such that accessible origins have less-negative parameter estimates than inaccessible origins (Fotheringham, 1981), in at least one instance, the obverse has been noted (Hervitz, 1983)[(1)]. It is the purpose of this paper to explain why such variation is possible and to demonstrate that this variation in no way invalidates the gravity-model misspecification hypothesis. In so doing, it is shown that the strength of the relationship between parameter estimates and origin accessibility can be estimated a priori for any spatial system and that the exact spatial structure bias in gravity parameter estimates can be identified. This is undertaken for the two gravity models in which spatial structure bias is most common, the total-flow-constrained and production-constrained models. The first type of gravity model is used, à la Curry et al (1975) and Cliff et al (1974), solely because its tractable format allows the exposition of concepts that would otherwise become extremely obfuscated in more complex model formulations. It is not meant to give the impression that such a model should be used for serious predictive purposes.

The investigation here is important in aiding the comprehension of gravity-model misspecification and the concomitant bias in gravity parameter estimates. In turn, the understanding of gravity parameter bias is essential to the accurate interpretation of gravity parameter estimates and in aiding the formulation of more accurate spatial interaction models. These points are discussed more fully elsewhere (Fotheringham, 1983a).

2 Misspecification bias in gravity parameters

Consider the general form of the gravity interaction model:

$$I_{ij} = \mathrm{f}(O_i, D_j, S_{ij}, \xi_{ij}) , \tag{1}$$

where I_{ij} represents the interaction volume between origin i and destination j; O_i represents the propulsiveness of i; D_j represents the attractiveness of j in terms of its size or mass; S_{ij} represents the separation of i and j; ξ_{ij} represents a random error component. S_{ij} is here measured by distance, so that a distance-decay parameter is estimated in the calibration of the model.

In an analysis of constrained gravity models, I have recently demonstrated (Fotheringham, 1983a; 1983b) that the general form of equation (1) results in misspecified interaction models. Such models fail to account for relationships between destinations; a more correctly specified interaction model has the general form:

$$I_{ij} = \mathrm{f}(O_i, D_j, S_{ij}, A_{ij}, \mu_{ij}) , \tag{2}$$

where A_{ij} represents the relationship between destination j and all other destinations

(1) A strong relationship between estimated distance-decay parameters and origin accessibility is a symptom of gravity-model misspecification. Although it might be tempting to relate such misspecification to a relationship between interaction and origin accessibility, no evidence has been provided to support such a thesis. The initial ideas concerning gravity-model misspecification are couched in terms of destination choice (Fotheringham, 1983b) and no similar theory appears valid to explain why differences in flow volumes from the same origin would be related to the accessibility of the origin. That origin accessibility does not account for the bias in parameter estimates described here is clearly seen in a production-constrained gravity model where parameter estimates can be strongly biased and yet the model contains an origin accessibility variable (the balancing factor).

that are possible receivers of interactions from origin i; and μ_{ij} is a random error component. The A_{ij} variable can be measured by the accessibility of destination j to all other possible destinations available to origin i. If accessibility is measured as a potential (Rich, 1980), then A_{ij} can be defined as:

$$A_{ij} \equiv \sum_{\substack{k \\ k \neq i, k \neq j}} \frac{D_k}{S_{jk}} , \tag{3}$$

where the summation is over all possible alternative destinations available to origin i and hence excludes both i and j. The rationale for the addition of this variable to the gravity-model format was presented in an earlier paper (Fotheringham, 1983a) where it is shown that the general format of equation (2) generates a set of spatial interaction models termed competing destinations models. If there is a negative parameter estimate for the accessibility variable in such models, then competition forces between destinations are dominant (as a destination becomes more accessible to other destinations, interaction to that destination decreases, ceteris paribus). If the accessibility variable has a positive parameter estimate, then agglomeration forces are dominant (as a destination becomes more accessible to other destinations, interaction to that destination increases, ceteris paribus). An analysis of two types of interaction models generated from equations (1) and (2) is now presented; this analysis allows the spatial structure bias in gravity parameter estimates to be identified exactly. The concentration of this paper is on origin-specific parameter estimates since the effects of spatial structure upon such estimates are easier to appreciate: ceteris paribus, gravity parameter estimates will vary between origins since different origins face different spatial arrangements of destinations. A discussion is also given of the spatial structure bias in system-wide parameter estimates, that is, for 'average' parameter estimates obtained for a whole system.

2.1 *Spatial structure bias in origin-specific total-flow-constrained gravity models*

If the general measure of spatial separation employed so far, S_{ij}, is replaced by a specific measure, say d_{ij}^{β}, where d_{ij} represents distance between i and j, and β is a distance-decay parameter, then an origin-specific total-flow-constrained gravity model derived from equation (1) can be written as:

$$I_{ij} = \alpha_i D_j^{\gamma_i} d_{ij}^{\beta_i} \xi_{ij} , \tag{4}$$

and the equivalent competing destinations model written as:

$$I_{ij} = \alpha_i D_j^{\gamma_i} d_{ij}^{\beta_i} A_{ij}^{\delta_i} \mu_{ij} . \tag{5}$$

In both models, it is assumed that the estimate of α_i is chosen so that the total predicted interaction volume is equal to the known total. It is assumed that the estimates of the other parameters are obtained by ordinary least squares regression after first taking logarithms so that equations (4) and (5) can be rewritten as:

$$I_{ij}^* = \alpha_i^* + \gamma_i D_j^* + \beta_i d_{ij}^* + \xi_{ij}^* , \tag{6}$$

and

$$I_{ij}^* = \alpha_i^* + \gamma_i D_j^* + \beta_i d_{ij}^* + \delta_i A_{ij}^* + \mu_{ij}^* , \tag{7}$$

respectively[2]. The symbol '*' denotes a logarithmic value. If equation (7) represents the 'true' model, then the error term associated with the model given in

(2) When the models are calibrated by regression, the total flow constraint has to be imposed a posteriori since, as Heien (1968) and Haworth and Vincent (1979) note, the logarithmic transformation produces a downward bias in the estimated constant. The other parameter estimates which are the focus of this paper are unaffected by the logarithmic transformation.

equation (6) is:

$$\xi^*_{ij} = \delta_i A^*_{ij} + \mu^*_{ij} . \tag{8}$$

Whence, it is well known (inter alia, Hanushek and Jackson, 1977, pages 79–86) that the expected estimates of the gravity-model parameters from equation (6) are:

$$\mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \left[\frac{r(A^*_{ij}, D^*_j) - r(d^*_{ij}, D^*_j) r(A^*_{ij}, d^*_{ij})}{1 - r^2(d^*_{ij}, D^*_j)} \frac{S(A^*_{ij})}{S(D^*_j)} \right] , \tag{9}$$

and

$$\mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \left[\frac{r(A^*_{ij}, d^*_{ij}) - r(d^*_{ij}, D^*_j) r(A^*_{ij}, D^*_j)}{1 - r^2(d^*_{ij}, D^*_j)} \frac{S(A^*_{ij})}{S(d^*_{ij})} \right] , \tag{10}$$

where $r(X, Y)$ represents the correlation coefficient between variables X and Y; $r^2(X, Y)$ represents the square of the correlation coefficient; $S(X)$ represents the standard deviation of X; and where '$\hat{}$' denotes a predicted value. The right-hand terms of equations (9) and (10) thus represent the bias inherent in mass and distance parameter estimates obtained in the calibration of the gravity model in equation (6). To derive an interpretation of this bias in terms of spatial structure, initially assume that there is no linear relationship between D^*_j and d^*_{ij} (this assumption is relaxed later). Then equations (9) and (10) can be rewritten as:

$$\mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \left[r(A^*_{ij}, D^*_j) \frac{S(A^*_{ij})}{S(D^*_j)} \right] , \tag{11}$$

and

$$\mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \left[r(A^*_{ij}, d^*_{ij}) \frac{S(A^*_{ij})}{S(d^*_{ij})} \right] . \tag{12}$$

With the equality,

$$r(X, Y) = \frac{C(X, Y)}{S(X) S(Y)} , \tag{13}$$

where $C(X, Y)$ is the covariance between X and Y, equations (11) and (12) can be rewritten as:

$$\mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \left[\frac{C(A^*_{ij}, D^*_j)}{S^2(D^*_j)} \right] , \tag{14}$$

and

$$\mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \left[\frac{C(A^*_{ij}, d^*_{ij})}{S^2(d^*_{ij})} \right] . \tag{15}$$

Then, since $C(Y, X)/S^2(X)$ represents the estimated slope parameter obtained in the regression of Y on X,

$$\mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \hat{\tau}_i , \tag{16}$$

and

$$\mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \hat{\pi}_i , \tag{17}$$

where $\hat{\tau}_i$ represents the estimated slope parameter obtained in regressing the logarithm of destination accessibility on the logarithm of destination size; and $\hat{\pi}_i$ represents the estimated slope parameter obtained in regressing the logarithm of destination accessibility on the logarithm of distance.

To interpret the bias in the gravity parameter estimates, consider the estimate of the distance-decay parameter (the interpretation of the bias in the mass parameter estimate follows exactly). The misspecification bias or spatial structure effect in $\hat{\beta}_i$ is given by the term $\delta_i \hat{\pi}_i$, which can be interpreted as the indirect elasticity of interaction with respect to distance; δ_i is the elasticity of interaction from origin i with respect to the accessibility of a destination; $\hat{\pi}_i$ is the estimated elasticity of the accessibility of a destination with respect to the distance between the destination and origin i. The estimated distance-decay parameter from equation (6) is biased, since it includes this indirect effect. Assuming that there is no linear relationship between D_j^* and d_{ij}^*, if $\delta_i = 0$ and/or $\hat{\pi}_i = 0$, there is no bias in the parameter estimate. The first case indicates that there is no log-linear relationship between the accessibility of a destination and interaction to that destination: the second case indicates that there is no log-linear relationship between the distance from an origin to a destination and destination accessibility. If $\delta_i = 0$, the total-flow-constrained gravity model is correctly specified. If $\hat{\pi}_i = 0$, and $\delta_i \neq 0$, however, destination accessibility is a relevant explanatory variable of interaction patterns, but there is no linear relationship between the logarithm of destination accessibility and the logarithm of distance. Consequently, in such a situation, the gravity distance-decay parameter estimate would be unbiased, but more accurate forecasts of interaction patterns would be obtained from the competing destinations model than from the gravity model since the former contains an additional relevant variable.

If the assumption that there is no linear relationship between d_{ij}^* and D_j^* is relaxed, another source of parameter bias exists. From equation (10), if $\delta_i \neq 0$, and $r(A_{ij}^*, d_{ij}^*) = 0$, then $\mathrm{E}(\hat{\beta}_i) \neq \beta_i$ if $r(d_{ij}^*, D_j^*) \neq 0$, and $r(A_{ij}^*, D_j^*) \neq 0$. That is, the estimated gravity distance-decay parameter can still be biased if a linear relationship exists between d_{ij}^* and D_j^* and between A_{ij}^* and D_j^*. This bias in the gravity distance-decay parameter results from an 'indirect' relationship between accessibility and distance via the mass variable. This indirect relationship is independent of the direct relationship between distance and accessibility since it exists even when $r(A_{ij}^*, d_{ij}^*) = 0$. Thus, there are three major determinants of bias in a gravity distance-decay parameter estimate. In order of their probable importance, they are:

1 the relationship between destination accessibility and interaction (δ_i);
2 the 'direct' relationship between destination accessibility and distance $[r(A_{ij}^*, d_{ij}^*)]$;
3 the 'indirect' relationship between destination accessibility and distance $[r(d_{ij}^*, D_j^*)$ and $r(A_{ij}^*, D_j^*)]$.

To assess the direction of the bias in gravity distance-decay parameter estimates and to demonstrate how this bias can produce particular spatial patterns of origin-specific estimates, assume that the third determinant of parameter bias is relatively unimportant and that the bias can be represented by equation (17). The direction of the parameter bias for a particular origin thus depends on the values of δ_i and $\hat{\pi}_i$, and it is variations in these values that produce different spatial patterns of estimated distance-decay parameters as described by figure 1. Assume that, for any interaction system, δ_i has the same sign for all i; that is, in any interaction system, similar relationships exist between the destinations available to each origin, and either competition forces or agglomeration forces are completely dominant. This appears reasonable, since it is unlikely that forces of competition could exist among the destinations available to one origin when forces of agglomeration exist among the destinations available to another origin. For ease of explanation, assume initially that competition forces are dominant ($\delta_i < 0$, $\forall i$) and that δ_i is constant over space, or does not vary systematically (there is no reason to suspect otherwise). Then, for accessible origins, $\hat{\pi}_i$ will be negative (as the distance between the origin and the destination increases, the destination becomes less accessible), and $\delta_i \hat{\pi}_i > 0$ so that $\mathrm{E}(\hat{\beta}_i) > \beta_i$. Conversely, for inaccessible origins, $\hat{\pi}_i$ will be positive (as the distance between the origin and destination increases, the destination becomes more accessible) and $\delta_i \hat{\pi}_i < 0$ so that $\mathrm{E}(\hat{\beta}_i) < \beta_i$. Since $\beta_i < 0$, $\hat{\beta}_i$ will be more negative for inaccessible origins than for accessible origins even if people's perception of distance as a deterrent to interaction is constant over space. This explains the commonly found pattern of gravity origin-specific distance-decay parameters [see Fotheringham (1981) for a review]. The finding that accessible origins tend to have less-negative parameter estimates than inaccessible origins probably has little or nothing to do with variations in interaction behaviour, but everything to do with variations in the spatial arrangement of destinations faced by each origin. An extreme example of this bias occurs when $\delta_i \hat{\pi}_i > |\beta_i|$, since then $\hat{\beta}_i$ will be positive. This could only occur for very accessible origins where $\hat{\pi}_i$ is large and negative; this theoretical finding is also supported by empirical evidence (Fotheringham, 1981).

If agglomeration forces, rather than competition forces, were present (that is, $\delta_i > 0$, $\forall i$), the above reasoning could be followed to demonstrate that the pattern of origin-specific distance-decay parameter estimates would be one in which accessible origins have more-negative parameter estimates than do inaccessible origins [Hervitz's results (1983)]. Since this pattern has rarely been seen, a tentative conclusion can be reached that forces of competition between destinations appear to be more common than forces of agglomeration. This result needs to be examined, however, for different types of interaction and for interactions at different scales.

The nature of the misspecification bias identified above also applies to other distance functions that may be used in interaction modelling. For example, if an exponential distance function is employed, the same procedure as outlined in equations (4)–(17) can be followed to identify the bias in gravity-model distance-decay parameter estimates. The bias would then be $\delta_i \hat{\omega}_i$, where δ_i has the same interpretation as above, but $\hat{\omega}_i$ relates logarithmic accessibility to distance rather than to logarithmic distance. A similar pattern of bias results, although it is likely that there will be some minor variations caused by differences in $\hat{\pi}_i$ and $\hat{\omega}_i$.

It is also clear from equations (4)–(16) that estimates of mass parameters obtained in the calibration of the gravity model also contain a potential misspecification bias. If it is assumed again, for the sake of exposition, that there is no linear relationship between D_j^* and d_{ij}^*, the bias in the gravity mass parameter given in equation (16) is

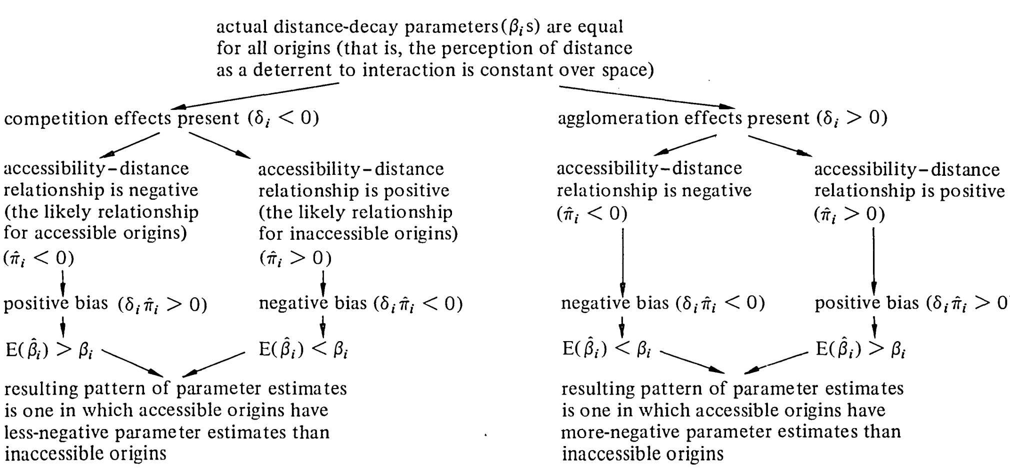


Figure 1. Different types of bias in gravity parameter estimates; the case of origin-specific distance-decay parameter estimates.

$\delta_i \hat{\tau}_i$, where $\hat{\tau}_i$ represents the estimated elasticity of destination accessibility with respect to destination size (the estimated slope parameter obtained in the simple regression of A_{ij}^* on D_j^*). One can hypothesise that $\hat{\tau}_i$ will generally be closer to zero than will $\hat{\pi}_i$ since strong relationships between accessibility and size are probably less likely than strong relationships between accessibility and distance. Although accessibility patterns often exhibit fairly strong spatial trends, it is quite likely that in a spatial system there will be both small and large centres in accessible locations and there will be both small and large centres in inaccessible locations. Consequently, in general one might expect mass parameter estimates obtained in the calibration of an origin-specific gravity model to exhibit less spatial bias than equivalent distance-decay parameter estimates. The remainder of this paper concentrates on the misspecification bias in estimated distance-decay parameters, but equivalent procedures can be followed to examine bias in estimated mass parameters.

2.2 *Spatial structure bias in origin-specific production-constrained gravity models*
Define $p_{ij} \equiv I_{ij} \Big/ \sum_j I_{ij}$, and then an origin-specific production-constrained gravity model derived from equation (1) can be written as:

$$p_{ij} = \frac{D_j^{\gamma_i} d_{ij}^{\beta_i}}{\sum_j D_j^{\gamma_i} d_{ij}^{\beta_i}} \xi_{ij} \,. \tag{18}$$

The equivalent form of the competing destinations model derived from equation (2) is:

$$p_{ij} = \frac{D_j^{\gamma_i} d_{ij}^{\beta_i} A_{ij}^{\delta_i}}{\sum_j D_j^{\gamma_i} d_{ij}^{\beta_i} A_{ij}^{\delta_i}} \mu_{ij} \,. \tag{19}$$

Nakanishi and Cooper (1974) have shown that equations (18) and (19) can be represented in a linear format by employing the following steps: (1) take the product over j of both sides of the equation; (2) take the nth root of both sides (where n is the number of destinations); and (3) divide both sides into p_{ij}. The equations are then;

$$\frac{p_{ij}}{n\left(\prod_{j=1}^{n} p_{ij}\right)^{1/2}} = \frac{D_j^{\gamma_i} d_{ij}^{\beta_i}}{n\left(\prod_{j=1}^{n} D_j^{\gamma_i} d_{ij}^{\beta_i}\right)^{1/2}} \,, \tag{20}$$

and

$$\frac{p_{ij}}{n\left(\prod_{j=1}^{n} p_{ij}\right)^{1/2}} = \frac{D_j^{\gamma_i} d_{ij}^{\beta_i} A_{ij}^{\delta_i}}{n\left(\prod_{j=1}^{n} D_j^{\gamma_i} d_{ij}^{\beta_i} A_{ij}^{\delta_i}\right)^{1/2}} \,, \tag{21}$$

which on transforming become

$$\frac{p_{ij}}{n\left(\prod_{j=1}^{n} p_{ij}\right)^{1/2}} = \left[\frac{D_j}{n\left(\prod_{j=1}^{n} D_j\right)^{1/2}}\right]^{\gamma} \left[\frac{d_{ij}}{n\left(\prod_{j=1}^{n} d_{ij}\right)^{1/2}}\right]^{\beta} , \tag{22}$$

and

$$\frac{p_{ij}}{n\left(\prod_{j=1}^{n} p_{ij}\right)^{1/2}} = \left[\frac{D_j}{n\left(\prod_{j=1}^{n} D_j\right)^{1/2}}\right]^{\gamma} \left[\frac{d_{ij}}{n\left(\prod_{j=1}^{n} d_{ij}\right)^{1/2}}\right]^{\beta} \left[\frac{A_{ij}}{n\left(\prod_{j=1}^{n} A_{ij}\right)^{1/2}}\right]^{\delta} . \tag{23}$$

Since $n\left(\prod_{j=1}^{n} x_j\right)^{1/2}$ is the geometric mean of x, the equations can be written, after taking the logarithms, as:

$$\left(\frac{p_{ij}}{\overline{p_i}}\right)^* = \gamma_i\left(\frac{D_j}{\overline{D}}\right)^* + \beta_i\left(\frac{d_{ij}}{\overline{d_i}}\right)^* + \left(\frac{\xi_{ij}}{\overline{\xi_j}}\right)^* , \tag{24}$$

and

$$\frac{p_{ij}}{\overline{p_i}}^* = \gamma_i\left(\frac{D_j}{\overline{D}}\right)^* + \beta_i\left(\frac{d_{ij}}{\overline{d_i}}\right)^* + \delta_i\left(\frac{A_{ij}}{\overline{A_i}}\right)^* + \left(\frac{\mu_{ij}}{\overline{\mu_i}}\right)^* , \tag{25}$$

where the symbol '*' denotes a natural logarithm and '—' denotes a geometric mean. Nakanishi and Cooper (1974) suggest calibrating equations (24) and (25) by generalized least squares regression if sampling errors are likely to be large. However, in terms of obtaining accurate parameter estimates, Stetzer (1976) provides evidence that ordinary least squares regression is probably an equally adequate calibration technique in most instances[3].

With the procedure outlined for the total-flow-constrained models, it is clear that, if equation (25) represents reality but equation (24) is calibrated, the error term $(\xi_{ij}/\overline{\xi_i})^*$ can be represented as:

$$\left(\frac{\xi_{ij}}{\overline{\xi_i}}\right)^* = \delta_i\left(\frac{A_{ij}}{\overline{A_i}}\right)^* + \left(\frac{\mu_{ij}}{\overline{\mu_i}}\right)^* , \tag{26}$$

whence, the expected values of the parameter estimates from equation (24) are:

$$\begin{aligned} \mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \Bigg\langle & r\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right] \Bigg/ \left\{1 - r^2\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right]\right\} \\ & - r\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right] r\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*, \left(\frac{d_{ij}}{\overline{d_i}}\right)^*\right] \Bigg/ \left\{1 - r^2\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right]\right\}\Bigg\rangle \\ & \times S\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*\right] \Bigg/ S\left[\left(\frac{D_j}{\overline{D}}\right)^*\right] , \end{aligned} \tag{27}$$

and

$$\begin{aligned} \mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \Bigg\langle & r\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*, \left(\frac{d_{ij}}{\overline{d_i}}\right)^*\right] \Bigg/ \left\{1 - r^2\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right]\right\} \\ & - r\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right] r\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right] \Bigg/ \left\{1 - r^2\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*, \left(\frac{D_j}{\overline{D}}\right)^*\right]\right\}\Bigg\rangle \\ & \times S\left[\left(\frac{A_{ij}}{\overline{A_i}}\right)^*\right] \Bigg/ S\left[\left(\frac{d_{ij}}{\overline{d_i}}\right)^*\right] . \end{aligned} \tag{28}$$

Again, for ease of exposition, assume that there is no relationship between the size and distance variables, that is, between $(D_j/\overline{D})^*$ and $(d_{ij}/\overline{d_i})^*$. Then, the expected values of the parameter estimates can be written as:

$$\mathrm{E}(\hat{\gamma}_i) = \gamma_i + \delta_i \hat{\phi}_i , \tag{29}$$

and

$$\mathrm{E}(\hat{\beta}_i) = \beta_i + \delta_i \hat{\lambda}_i , \tag{30}$$

[3] Stetzer (1976) also provides evidence that the calibration of these linear equations by ordinary least squares regression produces more-accurate parameter estimates than the more traditional maximum likelihood estimation procedure used to calibrate production-constrained interaction models. However, the trend of the misspecification bias in the estimates derived from both procedures is likely to be very similar, and the linear format of equations (24) and (25) allows the identification of this bias.

where $\hat{\phi}_i$ represents the estimated slope coefficient obtained in the regression of $(A_{ij}/\overline{A_i})^*$ on $(D_j/\overline{D})^*$, and $\hat{\lambda}_i$ represents the estimated slope coefficient obtained in the regression of $(A_{ij}/\overline{A_i})^*$ on $(d_{ij}/\overline{d_i})^*$. The biases in the two types of gravity model, total-flow-constrained and production-constrained, are therefore very similar, although there will be some slight differences due to differences between $\hat{\tau}_i$ and $\hat{\phi}_i$ and between $\hat{\pi}_i$ and $\hat{\lambda}_i$. For instance, in terms of the distance-decay parameter, the bias in the total-flow-constrained gravity model is a function of the linear relationship between A_{ij}^* and d_{ij}^*, $\hat{\pi}_i$, whereas the bias in the production-constrained gravity model is a function of the linear relationship between $(A_{ij}/\overline{A_i})^*$ and $(d_{ij}/\overline{d_i})^*$, $\hat{\lambda}_i$. Given the likely similarity in these relationships, however, the pattern of bias in the parameter estimates derived from the two gravity models will probably be very similar.

An alternative method of linearising the production-constrained interaction models is given by Baxter (1979). By use of Baxter's methodology, equations (18) and (19) can be represented as:

$$\left(\frac{p_{ij}}{p_{ik}}\right)^* = \gamma_i\left(\frac{D_j}{D_k}\right)^* + \beta_i\left(\frac{d_{ij}}{d_{ik}}\right)^* + \left(\frac{\xi_{ij}}{\xi_{ik}}\right)^*, \tag{31}$$

and

$$\left(\frac{p_{ij}}{p_{ik}}\right)^* = \gamma_i\left(\frac{D_j}{D_k}\right)^* + \beta_i\left(\frac{d_{ij}}{d_{ik}}\right)^* + \delta_i\left(\frac{A_{ij}}{A_{ik}}\right)^* + \left(\frac{\mu_{ij}}{\mu_{ik}}\right)^*, \tag{32}$$

respectively, where k represents an arbitrarily selected destination. By use of the above procedure to derive the bias in the gravity parameter estimates, it can be shown that the bias in each estimate is the same as that derived from the Nakanishi and Cooper linearisation.

2.3 *Spatial structure bias in 'system-wide' parameter estimates*

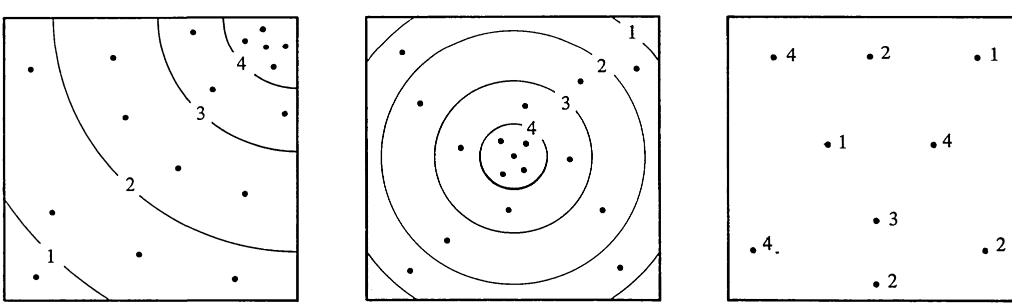
Although 'spatial structure bias' and 'map pattern effect' are terms normally used in connection with origin-specific parameter estimates, system-wide gravity parameter estimates also contain a potential misspecification bias. For example, with the above methodology, the bias in a system-wide distance-decay parameter derived from a total-flow-constrained gravity model is $\delta\hat{\pi}$, the overall indirect elasticity of interaction with respect to distance. System-wide biases, however, are likely to have smaller absolute values (and hence, will be less serious) than most of the origin-specific biases, since the system-wide bias can be considered as an 'average' of the origin-specific biases. As described above, the biases for accessible and inaccessible origins are likely to have different signs and in the system-wide biases these opposite effects will tend to cancel each other out. This point is expanded below, where the seriousness of misspecification bias in various types of spatial systems is considered. For simplicity, the discussion concentrates on one type of distance-decay parameter bias, that present in total-flow-constrained gravity models; but the conclusions reached apply in a similar manner to production-constrained gravity models.

3 Misspecification bias in various spatial systems

It is useful here to consider in what types of spatial systems the spatial structure bias in gravity distance-decay parameters will be at a maximum and in what types of spatial systems the bias will be at a minimum. There are different answers to this question depending upon whether origin-specific distance-decay parameters are estimated or whether one distance-decay parameter is estimated for the whole system. The situation in which origin-specific estimates are obtained is discussed initially.

Equation (17) demonstrates that, if there is no linear relationship between d_{ij}^* and D_j^*, the bias in an origin-specific distance-decay parameter estimated in the calibration of a total-flow-constrained gravity model is $\delta_i\hat{\pi}_i$. This term is interpreted as the indirect elasticity of interaction with respect to distance from origin i; for each origin it is a product of the elasticity of interaction with respect to the accessibility of a destination (δ_i) and the estimated elasticity of the accessibility of a destination with respect to the distance between the destination and the origin ($\hat{\pi}_i$). The parameter δ_i is a behavioural parameter and is assumed constant over space. For simplicity, it is again assumed that $\delta_i < 0$, for all i. Consequently, variations in the severity of the bias, $\delta_i \hat{\pi}_i$, can be ascribed to variations in $\hat{\pi}_i$, which is a structural parameter determined solely by the sizes and configurations of centres within a spatial system. When $\hat{\pi}_i = 0$, there is no bias in the gravity parameter estimate; when $\hat{\pi}_i > 0$, there is a negative bias, and when $\hat{\pi}_i < 0$, there is a positive bias. The extent to which $\hat{\pi}_i$ varies within a spatial system depends upon the distribution of accessibility within the system. Consider the three distributions of accessibility in figure 2[(4)]. In figure 2(a), accessibility increases regularly in one direction. In such a system, $\hat{\pi}_i$ will be highly negative for the accessible origins because, as distance from such origins increases, accessibility decreases. Conversely, $\hat{\pi}_i$ will be highly positive for the inaccessible origins because, as distance from such origins increases, accessibility increases. Consequently, the spatial structure bias in the estimated distance-decay parameters from such a system will be significantly troublesome for most origins. The empirical evidence presented in an earlier paper (Fotheringham, 1983a) is derived from a spatial system that has the characteristics of figure 2(a)—accessibility in the United States of America generally increases from the Southwest to the Northeast.

In figure 2(b), accessibility decreases as distance from a central point in the system increases, and it represents the situation that would be present in most cities. In such a system, $\hat{\pi}_i$ will be highly negative for accessible origins for the same reason as above.



strong spatial structure bias in $\{\hat{\beta}_i\}$ — weak bias in $\hat{\beta}$

moderate spatial structure bias in $\{\hat{\beta}_i\}$ — strong bias in $\hat{\beta}$

no spatial structure bias in $\{\hat{\beta}_i\}$ — no bias in $\hat{\beta}$

Figure 2. Accessibility in spatial systems: (a) strong linear distribution of accessibility; (b) strong centralised distribution of accessibility; (c) random distribution of accessibility. Iso-lines represent the general pattern of accessibility present in each system; higher numbers represent greater accessibility.

(4) Accessibility is defined as a function of distance and size. Although information on distance is given in each of the diagrams in figure 2, no information is provided on the sizes of the centres. Thus, depending upon variations in size, any pattern or trend in accessibility is possible for any arrangement of centres. In figure 2(c), for example, although the centres are regularly spaced, a fairly random pattern of accessibility values has been described. Boundary effects, which might otherwise have produced a centralised pattern of accessibility, are offset by variations in size. Note that the accessibility being described in figure 2 is the accessibility of destination j to all other possible destinations. Since the definition of such accessibility excludes the accessibility of a destination to itself, no local maxima in accessibility, produced by dividing by the distance from a centre to itself, will result and hence the generalised patterns in figure 2 are reasonably accurate.

However, the parameter will be close to 0 for inaccessible origins, since there is little or no relationship between distance from such origins and accessibility: as distance from an inaccessible origin increases, accessibility increases until the centre of the system is reached and then decreases thereafter. Consequently, the spatial structure bias in the estimated distance-decay parameters from such a system is likely to be significantly troublesome for only the most accessible origins.

In figure 2(c) there is no pattern to the accessibility of centres and $\hat{\pi}_i$ will be close to 0 for all origins. Consequently, the spatial structure bias in the estimated distance-decay parameters from such a system will be minimal for all origins. However, such a system will be rarely seen in reality since there is usually a clustering of centres around one or more locations and this clustering produces regularities in the distribution of accessibility.

Consider now the situation where system-wide parameter estimates are obtained from the gravity model. For example, β rather than β_i is estimated. With the procedure outlined above, the bias in β will be $\delta\pi$, where δ represents the overall relationship between I_{ij}^* and A_{ij}^*, and $\hat{\pi}$ represents the overall relationship between A_{ij}^* and d_{ij}^*. Assume that δ, the behavioural parameter, is a constant in the three spatial systems and is negative. Then consider the behaviour of $\hat{\pi}$ in each of the systems. As already mentioned, in figure 2(a), $\hat{\pi}_i$ will be highly negative for accessible origins and highly positive for inaccessible origins. These relationships will tend to counteract each other and $\hat{\pi}$ will be close to 0. Thus, in figure 2(a) there will be only a weak bias in $\hat{\beta}$. In figure 2(b) the overall relationship between A_{ij}^* and d_{ij}^* will be negative since this relationship exists for the accessible origins and no relationship exists for the inaccessible origins. Consequently, $|\hat{\pi}|$ will be larger, and the spatial structure bias in $\hat{\beta}$ will be greater in figure 2(b) than in figure 2(a). In figure 2(c) there will be little or no overall relationship between A_{ij}^* and d_{ij}^* and hence there will be little or no spatial structure bias in $\hat{\beta}$.

It is important to note that the above discussion is concerned solely with the effect of model misspecification on gravity-model parameter estimates and it is not concerned with the accuracy of model forecasts. It is shown that for some types of spatial systems the bias in $\hat{\beta}$ or in $\hat{\beta}_i$ will be very small. However, even in such systems the gravity model still fails to account for the level of competition between destinations and the addition of such a variable will produce more accurate predictions of flows.

4 Empirical evidence

To demonstrate some of the preceding theoretical arguments the production-constrained models in equations (24) and (25) were calibrated by ordinary least squares regression using a matrix of airline passenger flows in 1970 between the twenty-five largest cities in the United States of America. All flows over distances of less than 160 miles were ignored in the calibrations and the destination accessibility variable was defined as the accessibility of a destination to the ninety-eight largest alternative destinations available to the residents of origin i[5]. The results of the model calibrations, presented in table 1, demonstrate the bias in gravity distance-decay and mass parameter estimates. It is clear that the bias in the gravity distance-decay

[5] Eliminating all interactions over distances of less than 160 miles from the analysis removes a problem peculiar to airline passenger flow data: few people travel very short distances by aircraft and the inclusion of such data in the model calibrations would bias the results. The accessibility variable is defined as the accessibility of a destination to all other potential destinations available to origin i regardless of whether or not such destinations appear in a sample matrix. The ninety-eight largest alternative destinations were assumed to be a reasonable surrogate for all such destinations. More information on these points and on the data in general is given in a previous paper (Fotheringham, 1983a).

Table 1. Calibration results for equations (24) and (25) (in text) based on a 25 × 25 matrix of airline passenger flows (the standard errors of the parameter estimates are given in brackets).

City	$\hat{\beta}_i$			$\hat{\gamma}_i$			$\hat{\delta}_i$	R^2	
	G	CD	diff	G	CD	diff		G	CD
Atlanta	−0.7624	−1.3714	0.6090	0.7399	1.0513	−0.3114	−0.9350	0.534	0.664
	(0.2117)	(0.2698)		(0.1633)	(0.1727)		(0.3090)		
Baltimore	−0.2500	−0.8905	0.6405	0.8099	0.7884	0.0215	−1.1858	0.584	0.614
	(0.1669)	(0.4356)		(0.1771)	(0.1710)		(0.7496)		
Boston	−0.4806	−1.2943	0.8137	1.0716	0.9713	0.1003	−1.2469	0.756	0.801
	(0.1742)	(0.3722)		(0.1801)	(0.1677)		(0.5171)		
Chicago	−0.1220	−0.3899	0.2679	0.7583	0.8723	−0.1140	−0.4857	0.626	0.673
	(0.1271)	(0.1784)		(0.1207)	(0.1263)		(0.2413)		
Cincinnati	−0.4209	−0.5014	0.0805	0.8584	0.8816	−0.0232	−0.1424	0.701	0.688
	(0.1408)	(0.2694)		(0.1376)	(0.1552)		(0.4034)		
Cleveland	−0.3078	−0.8023	0.4945	0.9991	1.0464	−0.0473	−0.7843	0.731	0.734
	(0.1572)	(0.4772)		(0.1491)	(0.1545)		(0.7150)		
Dallas-Fort Worth	−0.8937	−0.9299	0.0362	0.7806	0.9632	−0.1826	−0.7322	0.506	0.729
	(0.2365)	(0.1753)		(0.1673)	(0.1310)		(0.1712)		
Denver	−1.4774	−1.1078	−0.3696	1.0718	1.1622	−0.0904	−0.6436	0.741	0.829
	(0.2861)	(0.2562)		(0.1428)	(0.1190)		(0.1874)		
Detroit	−0.1365	−0.3096	0.1731	0.9178	0.9417	−0.0239	−0.2937	0.650	0.634
	(0.1537)	(0.3327)		(0.1506)	(0.1584)		(0.4984)		
Houston	−1.0893	−1.1203	0.0310	0.9274	1.0765	−0.1491	−0.6085	0.628	0.746
	(0.2211)	(0.1830)		(0.1612)	(0.1408)		(0.1856)		
Kansas City	−0.5890	−0.8280	0.2390	0.7908	1.0218	−0.2310	−0.8697	0.531	0.797
	(0.2233)	(0.1337)		(0.1601)	(0.1139)		(0.1631)		
Los Angeles	−1.0418	−0.7320	−0.3098	1.0790	1.2104	−0.1314	−0.5994	0.817	0.865
	(0.1443)	(0.1634)		(0.1183)	(0.1112)		(0.2060)		
Memphis	−0.8403	−1.1522	0.3119	0.7987	1.0490	−0.2503	−0.7635	0.567	0.716
	(0.2070)	(0.1904)		(0.1683)	(0.1543)		(0.2205)		
Miami	−1.1665	−1.1136	−0.0529	1.4108	1.3740	0.0368	0.1266	0.908	0.906
	(0.1303)	(0.1497)		(0.1006)	(0.1132)		(0.1703)		
Minneapolis	−0.1756	−0.7449	0.5693	0.9510	1.1708	−0.2198	−1.0625	0.489	0.705
	(0.3221)	(0.2823)		(0.1975)	(0.1596)		(0.2626)		
New Orleans	−1.0317	−1.2424	0.2107	0.9461	1.1811	−0.2350	−0.7704	0.566	0.725
	(0.2471)	(0.2049)		(0.1860)	(0.1615)		(0.2121)		
New York	−0.2984	−1.0292	0.7308	1.0315	1.0566	−0.0251	−1.3715	0.554	0.624
	(0.1641)	(0.3687)		(0.2246)	(0.2066)		(0.6316)		
Philadelphia	−0.4726	−1.4939	1.0213	0.9391	0.9959	−0.0568	−1.7390	0.695	0.805
	(0.1520)	(0.3289)		(0.1682)	(0.1355)		(0.5204)		
Phoenix	−1.3154	−1.1546	−0.1608	1.1207	1.1621	−0.0414	−0.2205	0.833	0.829
	(0.1932)	(0.2878)		(0.1306)	(0.1428)		(0.2900)		
Pittsburgh	−0.3211	−0.5263	0.2052	1.0808	1.0917	−0.0109	−0.3604	0.734	0.722
	(0.1510)	(0.6463)		(0.1655)	(0.1726)		(1.1022)		
St Louis	−0.4006	−0.6706	0.2700	0.7403	0.9127	−0.1724	−0.6039	0.660	0.789
	(0.1362)	(0.1296)		(0.1135)	(0.1008)		(0.1625)		
San Francisco	−0.8862	−0.5369	−0.3493	1.0007	1.1192	−0.1185	−0.5426	0.800	0.825
	(0.1635)	(0.2324)		(0.1196)	(0.1268)		(0.2717)		
Seattle	−1.7037	−1.4418	−0.2619	0.9857	1.0495	−0.0638	−0.3261	0.880	0.888
	(0.1966)	(0.2521)		(0.0998)	(0.1045)		(0.2064)		
Tampa	−1.2540	−1.2585	0.0045	1.1508	1.1544	−0.0036	−0.0120	0.830	0.822
	(0.1581)	(0.1799)		(0.1290)	(0.1464)		(0.2113)		
Washington	−0.2791	−1.0561	0.7770	1.0271	1.0319	−0.0048	−1.4183	0.768	0.808
	(0.1365)	(0.3694)		(0.1428)	(0.1298)		(0.6351)		

Notes:
$\hat{\beta}_i$, $\hat{\gamma}_i$, and $\hat{\delta}_i$ denote estimates of the distance-decay parameter, mass parameter, and accessibility parameter, respectively; G denotes an estimated parameter derived from the calibration of a gravity model; CD denotes an estimated parameter derived from the calibration of a competing destinations model. diff represents the difference in equivalent gravity and competing destinations model parameter estimates; R^2 is the adjusted coefficient of determination.

parameter estimates (denoted by the difference column) is strongly related to origin accessibility, or more precisely, to the linear correlation between the accessibility and distance variables, which in turn is strongly related to the value of $\hat{\lambda}_i$ in equation (30). The bias in the gravity estimates of $\hat{\gamma}_i$ and $\hat{\beta}_i$ can be obtained either by subtracting the competing destinations parameter estimate (which is assumed to be unbiased) from the gravity estimate or by calculating the right-hand terms in equations (27) and (28). The data presented in tables 1 and 2 demonstrate that the two methods are equivalent.

Since the correlation coefficient between the distance and size variables is generally low for each of the origins, the estimate of the bias in the gravity distance-decay parameter given by equation (30) is a reasonable one for this data set. With $\hat{\delta}_i$ negative for twenty-four of the twenty-five origins, the bias in $\hat{\beta}_i$ will generally be positive whenever $\hat{\lambda}_i$ is negative (that is, for accessible origins) and negative whenever $\hat{\lambda}_i$ is positive (that is, for inaccessible origins). The difference column indicates that the bias in the gravity distance-decay parameter estimate is negative only for the most inaccessible origins in the data set: Denver, Los Angeles, Miami, Phoenix, San Francisco, and Seattle. Consequently, although most of the origins in the data set are accessible and have competing destinations parameter estimates which are more negative than the corresponding gravity estimates, these inaccessible origins have

Table 2. Correlation coefficients and standard deviations associated with the calibration results presented in table 1.

City	$r(A,d)$	$r(d,D)$	$r(A,D)$	$S(A)$	$S(d)$	$S(D)$
Atlanta	−0.6243	0.1937	0.3360	0.4873	0.5359	0.6946
Baltimore	−0.9353	−0.3402	0.2918	0.4285	0.7500	0.7068
Boston	−0.9112	−0.4558	0.3243	0.4849	0.7159	0.6925
Chicago	−0.7085	−0.0054	0.3206	0.4851	0.6215	0.6543
Cincinnati	−0.8373	−0.1741	0.3739	0.4782	0.6744	0.6900
Cleveland	−0.9466	−0.2522	0.3258	0.4627	0.6773	0.7138
Dallas-Forth Worth	0.1058	0.4343	0.3379	0.4839	0.4938	0.6980
Denver	0.4646	0.2539	0.3073	0.4767	0.3429	0.6870
Detroit	−0.8880	−0.2342	0.3226	0.4689	0.6838	0.6978
Houston	0.0946	0.4187	0.3318	0.4829	0.5089	0.6978
Kansas City	−0.2206	0.1538	0.3323	0.4871	0.4933	0.6881
Los Angeles	0.6914	0.3088	0.4924	0.4645	0.5359	0.6536
Memphis	−0.3569	0.2171	0.3496	0.4874	0.5417	0.6663
Miami	−0.3660	0.2372	0.3089	0.4771	0.5316	0.6879
Minneapolis	−0.4932	−0.0729	0.3313	0.4861	0.4278	0.6977
New Orleans	−0.1309	0.3686	0.3209	0.4852	0.5116	0.6795
New York	−0.9169	−0.2137	0.2177	0.4657	0.7970	0.5824
Philadelphia	−0.9309	−0.1590	0.1932	0.4255	0.6692	0.6046
Phoenix	0.7118	0.0303	0.2899	0.4732	0.4566	0.6755
Pittsburgh	−0.9738	−0.3251	0.3582	0.4642	0.7817	0.7135
St Louis	−0.4753	0.1441	0.3323	0.4869	0.5818	0.6980
San Francisco	0.7218	0.0637	0.3692	0.4684	0.5100	0.6962
Seattle	0.6290	0.0108	0.3071	0.4506	0,3511	0.6918
Tampa	−0.3222	0.2486	0.3141	0.4830	0.5562	0.6815
Washington	−0.9469	−0.3029	0.2918	0.4285	0.7393	0.7068

Notes:
$r(A,d)$, $r(d,D)$, and $r(A,D)$ represent the correlation coefficients between the accessibility and distance variables, the distance and size variables, and the accessibility and size variables, respectively. The form of these variables is that given in equations (29) and (30) in the text.
$S(A)$, $S(d)$, and $S(D)$ represent the standard deviations of the accessibility, distance, and size variables, respectively. The form of these variables is that given in equations (29) and (30) in the text.

competing destinations parameter estimates which are less negative than their gravity equivalents. Hence, the variation of the gravity estimates is much larger than that of the competing destinations estimates (the coefficient of variation associated with the gravity estimates is 0.6331 compared with 0.3454 for the competing destinations estimates). The results in tables 1 and 2 also demonstrate why there is little variation in the gravity mass parameter estimates (the coefficient of variation is 0.1605): there is probably little variation in the population parameters since the competing destinations estimates exhibit little variation; and the spatial structure bias in the gravity mass parameter estimates is small since the correlation between the accessibility and size variables is generally very low. In terms of equation (29), the values of $\hat{\phi}_i$ are all very small.

The results yield further support for the argument that destination accessibility is a relevant explanatory variable of spatial flows. For nineteen of the twenty-five origins, the addition of the destination accessibility variable produces an increase in the adjusted R^2 statistic associated with the regressions and the estimated accessibility parameter, $\hat{\delta}_i$, is significantly different from 0 at the 95% confidence level for fourteen of the twenty-five origins[6].

Last, the results in tables 1 and 2 suggest several interesting research problems. For instance, in the data set employed here the accessibility parameter estimates are not particularly stable over space (the coefficient of variation is 0.6377) and it would be interesting to investigate the stability of such parameter estimates in other spatial systems and to investigate the causes of instability where it exists. It is possible that such instability results from peculiarities associated with particular origins. For instance, the very low negative estimate for Tampa and the slightly positive estimate for Miami reported in table 1 probably arise from the strong vacation-ties that exist between Florida and the cities in the eastern half of the United States of America. For Tampa and Miami, the destination accessibility variable probably separates the accessible cities in the eastern half of the country from the inaccessible cities in the western half of the country; the larger volumes of interaction between Florida and the accessible cities being reflected in the low negative and positive accessibility parameter estimates. Thus, it may be that an analysis of estimated accessibility parameters could yield interesting information on patterns of connectivity that exist within spatial systems. Another research problem prompted by the variation between the accessibility parameter estimates concerns the ability of the destination accessibility measure used here to account accurately for the possible competition or agglomeration effects present. Are there superior measures that would further reduce the variation in $\hat{\beta}_i$ and further increase the ability of the model to replicate flow patterns?

5 Conclusions

In earlier papers (Fotheringham, 1983a; 1983b) it was demonstrated that gravity models are misspecified and that this misspecification is the likely cause of a spatial structure bias in estimated distance-decay parameters. A more precise definition of this bias is derived here; it allows the identification of the types of spatial systems in which the bias will be at a maximum. It is shown that the strength and direction of the bias depend primarily upon the relationships between interaction and destination accessibility and between destination accessibility and distance. When origin-specific parameter estimates are obtained in the calibration of a gravity model, the bias in the set of estimates is likely to be at a maximum in spatial systems where the accessibility of centres regularly increases in one direction within the system. An example of such a

[6] The number of origins having significant accessibility parameter estimates is probably misleadingly low, since in some cases the variances of the individual estimates are inflated because of high degrees of multicollinearity between the accessibility and distance variables.

system is the distribution of US cities where accessibility generally increases in a Southwest to Northeast direction. The bias in a single gravity distance-decay parameter estimated for the whole system is likely to be greatest in systems where accessibility generally decreases outwards from a central point. Most cities will exhibit this pattern of accessibility.

References

Baxter M J, 1979, "The application of logit regression analysis to production-constrained gravity models" *Journal of Regional Science* **19** 171-177

Cliff A D, Martin R L, Ord J K, 1974, "Evaluating the friction of distance parameters in gravity models" *Regional Studies* **8** 281-286

Curry L, Griffith D A, Sheppard E S, 1975, "Those gravity parameters again" *Regional Studies* **9** 289-296

Fotheringham A S, 1981, "Spatial structure and distance-decay parameters" *Annals of the Association of American Geographers* **71** 425-436

Fotheringham A S, 1983a, "A new set of spatial-interaction models: the theory of competing destinations" *Environment and Planning A* **15** 15-36

Fotheringham A S, 1983b, "Some theoretical aspects of destination choice and their relevance to production-constrained gravity models" *Environment and Planning A* **15** 1121-1132

Hanushek E A, Jackson J E, 1977 *Statistical Methods for Social Scientists* (Academic Press, New York)

Haworth J M, Vincent P J, 1979, "The stochastic disturbance specification and its implications for log-linear regression" *Environment and Planning A* **11** 781-790

Heien P M, 1968, "A note on log-linear regression" *Journal of the American Statistical Association* **63** 1034-1038

Hervitz H, 1983 *An Analysis of the Determinants of Interregional Migration Within the Framework of a Multinomial Logit Model* PhD thesis, Department of Economics, Indiana University, Bloomington, IN

Nakanishi M, Cooper L G, 1974, "Parameter estimation for a multiplicative competitive interaction model—least squares approach" *Journal of Marketing Research* **11** 303-311

Rich D C, 1980 *Concepts and Techniques in Modern Geography 26. Potential Models in Human Geography* (Geo-Abstracts, Norwich)

Stetzer F, 1976, "Parameter estimation for the constrained gravity model: a comparison of six methods" *Environment and Planning A* **8** 673-683