Journal of the Royal Statistical Society
SERIES A Statistics in Society

*J. R. Statist. Soc.* A (2011)
**174**, *Part* 4, *pp.* 1007–1027

# New Orleans business recovery in the aftermath of Hurricane Katrina

James P. LeSage,
*Texas State University—San Marcos, USA*

R. Kelley Pace and Nina Lam,
*Louisiana State University, Baton Rouge, USA*

Richard Campanella
*Tulane University, New Orleans, USA*

and Xingjian Liu
*Texas State University—San Marcos, USA*

[Received April 2010. Revised January 2011]

**Summary.** We analyse decisions to reopen in the aftermath of Hurricane Katrina made by business establishments on major business thoroughfares in New Orleans by using a spatial probit methodology. Our approach allows for interdependence between decisions to reopen by one establishment and those of its neighbours. There is a large literature on the role that is played by spatial dependence in firm location decisions, and we find evidence of strong dependence in decisions by firms to reopen in the aftermath of a natural disaster such as Hurricane Katrina. This interdependence has important statistical implications for how we analyse business recovery after disasters, as well as government aid programmes.

*Keywords*: Spatial auto-regressive model; Spatial spillovers

## 1. Introduction

Empirical observations on how businesses respond after a major catastrophe are rare, especially for a catastrophe as great as Hurricane Katrina, which hit New Orleans, Louisiana, on August 29th, 2005. We analyse the reopened status of all firms on three major business thoroughfares in New Orleans (Carrollton Avenue, St Claude Avenue and Magazine Street) during the periods from 0 to 3 months, 0 to 6 months and 0 to 12 months after the hurricane. In the aftermath of a disaster individual firms must make decisions about investing in repairs that are necessary to restore business operations. The outcome of the decision is likely to depend on decisions that are made by neighbouring firms. The conventional probit model that is used to analyse binary decision outcomes attempts to explain (cross-sectional) variation in the set of firms' decisions to reopen as a function of a set of explanatory variables pertaining to each firm, but it ignores decisions that are made by neighbouring firms.

*Address for correspondence*: James P. LeSage, McCoy College of Business Administration, Department of Finance and Economics, Texas State University—San Marcos, San Marcos, TX 78666, USA.
E-mail: jlesage@spatial-econometrics.com


0964–1998/11/1741007

This seems unrealistic in that an individual firm's decision to reopen is likely to be influenced by the decisions of neighbouring establishments. In the case of retail firms traffic generated by neighbouring establishments may be an important factor in generating spatial spillover business, i.e. customer traffic on a street may depend on the number of neighbouring establishments in operation. Patrons of a restaurant may also patronize neighbouring entertainment venues, art galleries or retail shopping establishments. Spatial spillover business can arise from neighbouring establishments that offer competing or complementary products and services. For example, neighbouring restaurants (competing businesses) on the same street may generate spatial spillover business since this attracts diners to the area, and neighbouring retail shops (complementary businesses) may also serve to attract potential spatial spillover patrons for restaurants.

This raises the question—would a single firm on a street decide to reopen knowing that all neighbouring firms on the street had decided not to reopen? Such an extreme case makes it clear that a role is played by the reopening decisions of neighbouring establishments when we consider a single firm's decision to reopen. The specific economic mechanism at work here is that some part of the unobserved net profitability associated with the decision to reopen derives from spatial spillover business reflected by the traffic that is generated from products and services offered by neighbouring establishments.

It seems reasonable that many of the motivations for firm location decisions set forth in location theories still hold in the aftermath of a disaster. For that reason, we enumerate some of the theoretical underpinnings in the firm location literature below.

Literature on street level retail location (in non-disaster situations) has long emphasized the importance of spatial interaction in firm level decisions. For example, Hotelling (1929) showed how location decisions as well as those of rival firms play an important role in maximization of profit and can lead to firms locating to near each other. Reilly (1931) set forth a 'law of retail gravitation', drawing on an analogy with Newton's gravitational law as it related to retail shopping behaviour and store location decisions. More recently, Fujita and Smith (1990) and Hinloopena and van Marrewijk (1999), as well as Brown (1989, 1994), focused on explaining clustering behaviour that is often observed in retail outlets within shopping districts. Of course, spatial clustering is consistent with the statistical concept of spatial dependence. Lee and Pace (2005) fitted a retail gravity model using a chain of stores and found spatial dependence among consumers as well as among stores.

There is also a demand side approach that focuses on street level consumer behaviours and cognitions. It ranges from the analysis of consumer interchanges between retail outlets (Nelson, 1958), to the analysis of Reilly (1931), gravity or frictional effects of distance (McGoldrick and Thompson, 1992) and to mental mapping of retail locations (Golledge and Timmermans, 1990). These analyses rely on several generalizations of consumer behaviours that can be used to motivate spatial interactions between retail outlets from another perspective. Nelson's 'rule of retail compatibility' (Nelson, 1958) states that compatible businesses that are close selling related types of goods benefit from customer flows and interchanges drawn by each other. Similarly, the 'theory of cumulative attraction' (Nelson, 1958; Hunt and Crompton, 2008) argues that even competitive businesses (those selling the same products) would cluster in an effort to cater to consumers' desire for comparative shopping before making a purchase.

A supply side approach has empirically investigated retailers' spatial strategies, finding that microscale decisions on location often take into account the location of competitive or complementary establishments (Berman and Evans, 1991; Hernandez and Bennison, 2000).

In summary, all firm location theories and models arrive at the conclusion that choices of consumers (or patrons) and decisions on location by retailers are influenced by extant patterns of retail activity. The decision of a business establishment to reopen after a natural disaster such as Hurricane Katrina could be viewed as similar to the original location decision made by the firm when it entered business. If the surrounding environment was important for the original decision, it should be equally influential in the decision to reopen in the same or another location after the disaster. This provides a strong motivation for adopting an empirical modelling method that allows one firm's decision to reopen to depend on similar decisions made by neighbouring firms.

In Section 2 we set forth our spatial auto-regressive (SAR) variant of the conventional probit regression model that allows for interdependence between firms' decisions regarding reopening. Interpreting the coefficient estimates from this type of regression model is also discussed. Section 3 discusses our sample data and presents estimates from the SAR probit model applied to this data set.

The data and the programs that were used to analyse them can be obtained from

`http://www.blackwellpublishing.com/rss`

## 2. Spatial auto-regressive probit model

We let the $n \times 1$ vector $y$ be a 0–1 binary vector reflecting the closed–open status of the $n$ firms at some point in time (say 3 months after Hurricane Katrina). A conventional probit model would attempt to explain variation in the binary vector $y$ by using an $n \times k$ matrix of firm-specific explanatory variables $X$ (where a single variable is represented by $x_v$ for $v = 1, \ldots, k$) and associated $k \times 1$ vector of parameters $\beta$, under the assumption that each observed decision is independent of all others.

LeSage and Pace (2009) set forth an SAR variant of the conventional probit model that takes the form shown in expression (1), where $y^*$ represents an $n \times 1$ vector reflecting the latent unobservable utility or profit that is associated with the closed–open status of firms:

$$y^* = \rho W y^* + X\beta + \varepsilon, \qquad \varepsilon \sim N(0, I_n). \tag{1}$$

The spatial lag of the latent dependent variable $Wy^*$ involves the $n \times n$ spatial weight matrix $W$ that contains elements consisting of either $1/m$ or 0, where $m$ is some number of nearest neighbours. If observation or firm $j$ represents one of the $m$ nearest neighbouring establishments to firm $i$, then the $(i, j)$th element of $W$ contains the value $1/m$. All elements in the $i$th row of the matrix $W$ that are not associated with neighbouring observations take values of 0. By construction, $W$ is row stochastic (non-negative and each row sums to 1). This results in the $n \times 1$ vector $Wy^*$ consisting of an average of the $m$ neighbouring firms' utility or profit, creating a mechanism for modelling interdependence in firms' decisions to reopen in the aftermath of a disaster. The scalar parameter $\rho$ measures the strength of dependence, with a value of 0 indicating independence. Clearly, a conventional non-spatial probit model emerges when $\rho = 0$.

The Bayesian approach to modelling binary limited dependent variables treats the binary 0–1 observations in $y$ as indicators of latent, unobserved $y^*$ (net) profit in the closed–open states, with the unobservable profit underlying the observed choice outcomes. For example, in our case where the binary dependent variable reflects the closed–open state of the firm, the decision to reopen would be made when the net profit from the open *versus* closed state is positive. As already motivated, there are theoretical reasons why the profits of an establishment depend on those of neighbours, especially in the street level retail environment that we are considering. The Bayesian estimation approach to these models is to replace the unobserved latent profit with *parameters* that are estimated. For the case of an SAR probit model, given estimates of the $n \times 1$ vector of missing or unobserved (parameter) values that we denote as $y^*$, we can proceed to estimate the remaining model parameters $\beta$ and $\rho$ by sampling from the same conditional distributions that are used in the continuous dependent variable Bayesian SAR models (see chapter 10 in LeSage and Pace (2009)).

More formally, the choice depends on the difference in profits: $\pi_{1i} - \pi_{0i}$, $i = 1, \ldots, n$, associated with observed 0–1 choice indicators, where $\pi_{1i}$ represents profits (of firm $i$) in the open state and $\pi_{0i}$ in the closed state. The probit model assumes that this difference, $y_i^* = \pi_{1i} - \pi_{0i}$, follows a normal distribution. We do not observe $y_i^*$, only the choices made, which are reflected in

$$y_i = \begin{cases} 1, & \text{if } y_i^* \geqslant 0, \\ 0, & \text{if } y_i^* < 0. \end{cases}$$

If the vector of latent profits $y^*$ were known, we would also know $y$, which led Albert and Chib (1993) to conclude that $p(\beta, \sigma_\varepsilon^2 | y^*) = p(\beta, \sigma_\varepsilon^2 | y^*, y)$. The insight here is that, if we view $y^*$ as an additional set of parameters to be estimated, then the (joint) conditional posterior distribution for the model parameters $\beta$ and $\sigma_\varepsilon^2$ (conditioning on both $y^*$ and $y$) takes the same form as a Bayesian regression problem involving a continuous dependent variable rather than the problem involving the discrete-valued vector $y$. This approach was used by LeSage and Pace (2009) to implement a Bayesian Markov chain Monte Carlo estimation procedure for the SAR model (1).

### 2.1. Interpreting the model estimates

Interpreting the way in which changes in the explanatory variables in the matrix $X$ impact the probability of a firm's reopening in the SAR probit model requires some care. Expressions (2) make it clear that the probability (of a 0–1 event outcome) is a non-linear function $F(\cdot)$ (the multivariate probability rule) of a function $(I_n + \rho W + \rho^2 W^2 + \ldots) X\beta$ of the explanatory variables in the model $X$:

$$\left. \begin{aligned} y^* &= \rho W y^* + X\beta + \varepsilon, \\ y^* &= S(\rho) X\beta + S(\rho)\varepsilon, \\ X\beta &= \sum_{v=1}^{k} x_v \beta_v, \\ S(\rho) &= (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \ldots \\ \Pr(y=1) &= F\{S(\rho) X\beta\}. \end{aligned} \right\} \tag{2}$$

Since the interpretation of the spatial probit model builds on the interpretation of changes of independent variables on dependent observations over space as well as the well-known non-linear transformations due to the probit model, to simplify the exposition we first consider the simpler case of a non-probit SAR model shown in expression (3), where $z$ denotes a continuous $n \times 1$ dependent variable vector:

$$\left. \begin{aligned} z &= \alpha \iota_n + \rho W z + X\beta + \varepsilon, \\ \partial z / \partial x_v' &= (I_n - \rho W)^{-1} I_n \beta_v, \qquad v = 1, \ldots, k, \\ &= S(\rho) \beta_v. \end{aligned} \right\} \tag{3}$$

LeSage and Pace (2009) proposed the use of an average of the diagonal elements from the $n \times n$ matrix, $\partial z / \partial x_v'$, to produce a scalar summary of the *direct effects*, which are derived from the own partial derivatives, $\partial z_i / \partial x_{v,i}$.

They also used an average of the (cumulated) off-diagonal elements from the $n \times n$ matrix, $\partial z/\partial x'_v$ $(i \neq j)$, to produce a scalar summary of the (cumulative) *indirect effects* that are associated with the cross-partial derivatives, $\partial z_i/\partial x_{v,j}$. This scalar summary measure cumulates the spatial spillovers falling on neighbouring establishments as well as neighbours to these neighbours, and so on.

When we allow for dependence among observations or firms, changes in the explanatory variables that are associated with one firm, say firm $i$, will influence the dependent variable value of firm $i$ as well as other firms, say $j$. The SAR model collapses to an independence model when the scalar spatial dependence parameter $\rho$ takes a value of 0. In this case, the cross-partial derivatives are all 0.

For the case of spatial dependence, the (non-zero) cross-partials represent what are commonly thought of as *spatial spillover* effects. Changes in the value of a single observation or firm $i$ explanatory variable can (potentially) influence all $n-1$ other observations or firms. This is true for all $i = 1, \ldots, n$ values of the $v$th explanatory variable, leading to the $n \times n$ matrix of own- and cross-partial derivatives. This motivates the need for scalar summary measures that average across the sample of observations similar in spirit to the way that we interpret conventional least squares regression estimates. An important point is that the scalar summary measure of *indirect effects* cumulates the spatial spillovers falling on all other observations, but the magnitude of impact will be greatest for nearby neighbours and will decline in magnitude for higher order neighbours.

The sum of the two effects (direct and indirect) represents the (cumulative) *total effect* that is associated with a change in an observation for that explanatory variable.

Turning to the more complicated case of the SAR probit model, consider the effect on observation or firm $i$ arising from a change in a variable $x_v$ (say the depth of the flood) at firm $j$ location, a single cross-derivative:

$$\eta = S(\rho)X\beta = E(y^*), \tag{4}$$

$$\begin{aligned} \frac{\partial \Pr(y_i = 1)}{\partial x_{v,j}} &= \left( \frac{\partial F(\eta)}{\partial \eta} \Big| \eta_i \right) S_{ij}(\rho)\beta_v, \\ \frac{\partial \Pr(y_i = 1)}{\partial x_{v,j}} &= \mathrm{pdf}(\eta_i)\ S_{ij}(\rho)\beta_v. \end{aligned} \tag{5}$$

We note that, if $\rho = 0$ so that $S(\rho) = I_n$ (and $S_{ij}(\rho) = 0$), we arrive at the standard probit result where changes in $x$-values of neighbouring firm $j$ have no influence on firm $i$'s decisions.

We now set up a matrix version of these own- and cross-partial derivatives. Let $d(\cdot)$ represent the $n \times 1$ vector on the diagonal of a diagonal matrix $D(\cdot)$ where the non-diagonal elements are 0s. By construction, $D(\cdot)$ is symmetric. The $n \times 1$ vector $d\{f(\eta)\}$ contains the probability density function evaluated at the predictions for each observation and associated $n \times n$ diagonal matrix $D\{f(\eta)\}$ which has $d\{f(\eta)\}$ on the diagonal:

$$f(\eta) = \mathrm{pdf}(\eta),$$

$$\frac{\partial \Pr(y = 1)}{\partial x'_v} = D\{f(\eta)\}S(\rho)I_n\beta_v.$$

We can now expand the expression into component matrices:

$$S(\rho) = I_n + \rho W + \rho^2 W^2 + \ldots,$$

$$\frac{\partial \Pr(y = 1)}{\partial x'_v} = [D\{f(\eta)\} + \rho\ D\{f(\eta)\}W + \rho^2\ D\{f(\eta)\}W^2 + \ldots]\beta_v.$$

From these expressions, the $n \times 1$ vector of (cumulative) total effects can be written as

$$\begin{aligned}\left(\frac{\partial \Pr(y=1)}{\partial x_v'}\right)\iota_n &= [D\{f(\eta)\}\iota_n + \rho\, D\{f(\eta)\}W\iota_n + \rho^2\, D\{f(\eta)\}W^2\iota_n + \ldots]\beta_v \\ &= [D\{f(\eta)\}\iota_n + \rho\, D\{f(\eta)\}\iota_n + \rho^2\, D\{f(\eta)\}\iota_n + \ldots]\beta_v \\ &= (D\{f(\eta)\}\iota_n)(1-\rho)^{-1}\beta_v \\ &= (d\{f(\eta)\})(1-\rho)^{-1}\beta_v.\end{aligned}$$

As a scalar summary measure of *average total effect*, we can use an average of the vector of (cumulative) total effects, specifically

$$n^{-1}(d\{f(\eta\})'\iota_n(1-\rho)^{-1}\beta_v. \tag{6}$$

To summarize the *average direct effect* we use

$$\frac{1}{n}\mathrm{tr}\left\{\frac{\partial \Pr(y=1)}{\partial x_v'}\right\} = (\mathrm{tr}[D\{f(\eta)\}] + \rho\,\mathrm{tr}[D\{f(\eta)\}W] + \rho^2\,\mathrm{tr}[D\{f(\eta)\}W^2] + \ldots)\,\frac{\beta_v}{n}$$

and for the (cumulative) *average indirect effect* we use the difference: average total effect − average direct effect. We note that there are several approaches to computing $\mathrm{tr}[D\{f(\eta)\}W^j]$ efficiently (LeSage and Pace (2009), chapter 4).

### 2.2. Simple illustration

To provide a concrete illustration, we consider the case of seven firms located in a line along (one side of) a street, and we use a simple spatial weight matrix that identifies a single left and right neighbour to each observation. Let $y$ be a 0–1 binary vector, and

$$\begin{aligned}\Pr(y=1) &= F\{S(\rho)X\beta\}, \\ S(\rho) &= (I_n - \rho W)^{-1}.\end{aligned}$$

We generated the vectors of flood depth and firm size. Subsequently, we calculated the probability of reopening by using a model based on two explanatory variables, flood depth and firm size, using a value of the spatial dependence parameter $\rho$ equal to 0.8, and assumed that the parameters that are associated with flood depth and size equalled −0.25 and 0.5:

$$\Pr(\text{reopening}) = F\{S(\rho=0.8)(-0.25\ \text{flood depth} + 0.5\ \text{firm size})\}. \tag{7}$$

The resulting values shown in Table 1 illustrate that, using the conventional practice of interpreting predicted probabilities less than 0.5 as implying that $y=0$ and greater than 0.5 as $y=1$, the model perfectly predicts the pattern of observed 0–1 values.

To illustrate the effect of changing a single observation on the probabilities we increased the flood depth at the location of firm 3 from 20 to 60, *ceteris paribus*. Table 2 shows the original predicted probabilities $\Pr(y=1)$, the new probabilities $\Pr(y'=1)$ and the change in probabilities or predictions that is implied by the SAR probit model $\Pr(y'=1) - \Pr(y'=1)$.

The first point to note is that the change in flood depth at the location of firm 3 leads to a decrease in the probability that this firm will reopen by 14.14%. However, neighbouring firms 4 and 5 are also impacted by the higher level of flooding at firm 3, leading to a lower probability that these firms will reopen. Specifically, the probability that the immediately neighbouring firm 4 will reopen decreases by 9.28%, and that for the neighbour to this neighbouring firm number 5 by 1.9%. These represent an illustration of spatial spillover effects that arise when we allow for interdependence of other firms' decisions on firm 3's decision to open or to close.

**Table 1.** Illustration based on $n = 7$ firms

| *Firm* | *y-value* | *Pr(y = 1)* | *Flood depth* | *Firm size* |
|---|---|---|---|---|
| Observation 1 | 0 | 0.0036 | 40 | 1 |
| Observation 2 | 0 | 0.0231 | 30 | 2 |
| Observation 3 | 0 | 0.1964 | 20 | 3 |
| Observation 4 | 1 | 0.5131 | 25 | 4 |
| Observation 5 | 1 | 0.8569 | 20 | 4 |
| Observation 6 | 1 | 0.9907 | 20 | 8 |
| Observation 7 | 1 | 0.9968 | 20 | 8 |

**Table 2.** Effect of changing a single observation

| *Firm* | *y-value* | *Pr(y = 1)* | *Pr(y′ = 1)* | *Pr(y′ = 1)− Pr(y′ = 1)* | *Flood depth* |
|---|---|---|---|---|---|
| Observation 1 | 0 | 0.0036 | 0.0098 | 0.0061 | 40 |
| Observation 2 | 0 | 0.0231 | 0.0241 | 0.0010 | 30 |
| Observation 3 | 0 | 0.1964 | 0.0550 | −0.1414 | 60 |
| Observation 4 | 1 | 0.5131 | 0.4203 | −0.0928 | 25 |
| Observation 5 | 1 | 0.8569 | 0.8375 | −0.0194 | 20 |
| Observation 6 | 1 | 0.9907 | 0.9904 | −0.0003 | 20 |
| Observation 7 | 1 | 0.9968 | 0.9968 | 0.0000 | 20 |

The direct effect arising from the change in observation 3's flood depth is −0.1414, whereas the cumulative indirect effect is −0.1054, with the (cumulative) total effect being −0.2468, the sum of all non-zero changes.

We also note that the change in flood depth at firm 3's location leads the model to predict that firm 4 would not reopen, since the probability of reopening has fallen from 0.51 to 0.42 as a result of increased flooding at the neighbouring location. This directly lowers the chances that firm 3 will reopen by 14% and indirectly influences the reopening decision of firm 4, its neighbour (as well as firm 5, the neighbour to its neighbour).

Interdependence in this model means that increasing flood depth for observation 3 reduced its own chance of opening and those of most other observations. However, the chance of opening for observations 1 and 2 actually rose (very slightly). This occurs because increasing the flood depth for observation 3 which has closed status means that observations 1 and 2 which share the same closed status have relatively less flooding compared with observation 3 in the original scenario. Therefore, the new estimated probabilities rise slightly for observations 1 and 2.

The effects that are presented in Table 2 represent the effect of changing only a single observation, number 3, whereas the scalar summary measures based on the expressions for own- and cross-partial derivatives (which were described earlier) would average over changes in all observations (for each explanatory variable in the model).

### *2.3. Considerations in interpreting spatial probit models*

We discuss some points regarding interpretation of the scalar summary measures for direct, indirect and total effects estimates that are associated with the SAR probit model.

A minor issue is that there is a need to determine measures of dispersion for the mean scalar summary magnitudes described, since these are required for inference regarding the statistical significance of the various effects. Because estimation of the model is implemented by using Markov chain Monte Carlo methods, it is possible to evaluate the own- and cross-partial derivative expressions on each pass through the Markov chain Monte Carlo sampler. This allows us to construct a sample of draws from the posterior distribution of the various effects parameters, which can be used for inference. Gelfand *et al.* (1990) showed that one can perform valid inference regarding non-linear parameter relationships by using non-linear combinations of the Markov chain Monte Carlo draws.

A more important set of issues relates to the non-linearity of the probability response in the dependent variable to changes in the explanatory variables of the model. In non-spatial probit regressions a common way to explore the non-linearity in this relationship is to calculate 'marginal effects' estimates by using particular values of the explanatory variables (e.g. mean values or quintile intervals). The motivation for this practice is consideration of how the effect of changing explanatory variable values varies across the range of values that are encompassed by the sample data. Given the non-linear nature of the normal cumulative density function transform on which the (non-spatial) probit model relies, we know that changes in explanatory variable values near the mean may have a very different influence on decision probabilities than changes in very low or high values.

As an example related to our focus on flooding and the probability that a firm will reopen, consider that increasing flood depth from a high value such as 8 ft to a higher value such as 9 ft should have a comparatively small effect on the probability that a firm reopens. This may be quite different for cases where flood depths change from 1 to 2 ft, so variation in the effects that are associated with changes in explanatory variables may differ depending on values taken by these variables.

Unlike the case of non-spatial probit models, our SAR probit model involves a different location for each $x$-variable observation. This is true for spatial regression models where observations are associated with individuals or points or regions, each of which is located in space. An implication of this is that considering different values of the explanatory variables implies a change in location each time that the $x$-variable changes value. Ideally, a spatial regression model would allow us to assess appropriately the effect of changes in location separately from that associated with changes in values of the explanatory variables on the dependent variable.

At one level the two are inextricably linked since values of the explanatory variable change with spatial location. However, the interpretation of effects arising from infinitesimally small changes in an explanatory variable on the dependent variable represents a partial derivative. These measures should hold other factors fixed across the spatial sample of dependent variable values, and one of the other factors that we wish to fix is location. A related consideration is that variable values are not likely to be distributed independently of values taken by the same variable at neighbouring locations, i.e. explanatory variables are likely to exhibit spatial dependence (in addition to the dependent variable, or decision outcomes that have been the focus of our motivation for use of the SAR probit model).

Another contrast with the non-spatial probit model pertains to the degree of non-linearity that is involved. Our spatial probit regression contains an expected value of a latent variable that already involves a non-linear transformation of the dependence parameter $\rho$, specifically $(I_n - \hat{\rho}W)^{-1}X\hat{\beta} = (I_n + \hat{\rho}W + \hat{\rho}^2W^2 + \ldots)X\hat{\beta}$. In the case of the spatial (auto-regressive) probit model, the expected value of the latent variable undergoes further transformation by the non-linear normal cumulative density function. Of course, the contrast with the non-spatial probit situation is that only the single non-linear transformation involving the normal cumulative density function takes place. This suggests that we need to consider changes in the dependent variable probability ('marginal effects') relative to values of $(I_n - \hat{\rho}W)^{-1}x_v$, where $x_v$ denotes the $v$th explanatory variable vector. This should produce a set of outcomes that is much closer to that encountered in non-spatial probit regression models than consideration of the marginal effects relative to values of the explanatory variable $x_v$.

Another important difference that arises in the spatial *versus* non-spatial probit model is the fact that the partial derivatives take the form of a matrix rather than a scalar expression. It is relatively easy to calculate a vector containing the row sums of the matrix of partial derivatives (for each variable $v = 1, \ldots, k$) that reflects the (cumulative) total effects for each observation in our sample. (An average of this vector is used to construct our proposed scalar summary expression for the variable $v$ total effects estimates.) This vector can be used to examine how the total effects vary by location, or values of the variable $x_v$, or expected values of $(I_n - \hat{\rho}W)^{-1}x_v$. This should provide a useful means for investigating how representative inferences based on the global scalar summary total effects estimates are with respect to location, variable levels or expected values levels. We illustrate this in the next section with our application of the SAR probit model to a sample of 673 firms in New Orleans.

## 3. Reopening decisions in the aftermath of Hurricane Katrina

We are interested in exploring factors that influenced the decision of establishments to reopen in the aftermath of Hurricane Katrina. One focus is on the role that is played by spatial dependence. Our SAR probit model produces estimates for the spatial dependence parameter $\rho$, allowing us to draw an inference regarding the presence or absence of spatial interdependence in firms' decisions. A second issue relates to the relative magnitude of direct *versus* indirect effects, i.e.—how important were spatial spillovers in influencing business recovery in New Orleans? The presence of significant spillovers may mean that cost–benefit analyses of the role that is played by disaster assistance systematically underestimates the benefits that are associated with such aid.

As already noted, there is a large literature that emphasizes the importance of interdependence for the case of firm location decisions. Since there are relatively few cases where empirical observations exist regarding business response after a major catastrophe such as Hurricane Katrina, our empirical application allows us to fill this gap in the literature.

### *3.1. Sample data*

The data set that was used for this exploration entails 673 establishments tracked weekly by Richard Campanella during the year following Hurricane Katrina, and then seasonally and annually in subsequent years. Three major New Orleans commercial corridors were selected: St Claude Avenue, from Poland Avenue to Faubourg Tremé, the entire length of Magazine Street, and all of both South and North Carrollton Avenues. These three corridors transect a wide range of socio-economic, historical, topographic and flood damage conditions in the city. St Claude Avenue, which experienced light flooding after Hurricane Katrina, traverses a struggling, working-class downtown neighbourhood; Magazine Street serves middle- to upper-class uptown neighbourhoods and suffered zero flooding; Carrollton Avenue traverses both middle-class and working-class neighbourhoods and suffered extreme flooding in some areas, some flooding in most areas and none in others. The Lower Ninth Ward portion of St Claude Avenue could not be included because it was closed at the time that this study began.

All 16 miles of the three corridors were surveyed by bicycle weekly starting on October 9th, 2005, 6 weeks after Hurricane Katrina and about 2 weeks after unflooded neighbourhoods began to repopulate. Businesses of all types (retail, wholesale, services, etc.) that were visible from the street were recorded by

(a) address,
(b) name,
(c) description and category (food retail, restaurant, spa salon, florist nursery, etc.),
(d) ownership (locally owned independent, regional chain or national chain),
(e) general economic status ('functional', 'mid-range' or 'high end') and
(f) size (1, sole proprietorship with five or fewer employees; 2, about 6–15 employees, such as a typical restaurant; 3, scores of employees, such as a large grocery store; data from the Louisiana Department of Labor were consulted to aid in this judgement).

Finally, and most importantly, the business's status as 'still closed', 'open', 'partially open' (limited hours, by appointment only, etc.), 'new' (a new post-hurricane business) or 'moved' was recorded and re-recorded with each weekly visit. Follow-up phone calls or local inquiries were made when business status proved difficult to discern visually. Data from the Census Bureau, Federal Emergency Management Agency, State of Louisiana and Army Corps of Engineers were then used to record the median household income of the surrounding neighbourhood, the area's topographic elevation and depths of flood after Hurricane Katrina.

The weekly pace of surveys was reduced to biweekly in autumn 2006, because the number of reopenings or new businesses did not warrant a weekly revisit. By 2007–2008, conditions had stabilized to the point that only seasonal or annual visits were made.

Models based on three different dependent variable vectors were estimated, each having the same explanatory variables. One dependent variable vector was constructed to represent the very short period of 0–3 months after Hurricane Katrina, another reflecting the 0–6-month horizon and a final model for the 0–12-month time period. For each period, firms that had reopened were assigned a dependent variable value of 1 and those not reopened a value of 0. During the 0–3-month time horizon we had 300 of the 673 firms opened, during the 0–6-month period 425 firms were open and in the 0–12-month interval 478 firms. The last number reflects a reopening rate of 71% at the 12-month horizon, which matches well to the 66% reopening rate that was derived from larger samples taken of firms that reopened (see Lam *et al.* (2009)).

Although the three different types of dependent variables acknowledge an interaction with time and space, we elected to model this as a spatial problem rather than as a formal spatio-temporal problem. Given the frequency of observations, it would seem ideal to use a spatio-temporal model where each retailer conditioned on nearby observations from previous periods. However, the actual opening is foretold much earlier by rebuilding activity, discussions between store owners and their employees, and other activities that act to smear the temporal nature of the reopening. This introduces a simultaneous nature to the decisions. Spatiotemporal processes with and without a simultaneous spatial component and simultaneous spatial processes have equivalent equilibria (given various assumptions) as discussed in LeSage and Pace (2009) and Pace and LeSage (2010). In particular, LeSage and Pace (2009), chapter 7, considered a spatiotemporal model where the explanatory variables change slowly or deterministically over time as would be the case here. They showed that the steady-state equilibrium for this scenario is a cross-sectional model of the type that is used here.

Explanatory variables used were the flood depth (measured in feet) at the location of the individual establishments, (logged) median income for the census block group in which the store was located, two dummy variables reflecting small and large size firms, with medium size firms representing the omitted class, two dummy variables reflecting low and high socio-economic class of the store *clientèle* (with the middle socio-economic class excluded) and two dummy vari-

**Table 3.** SAR probit model estimates for three time horizons

| *Variable* | *Posterior mean* | *Standard deviation* | *p-level* |
|---|---|---|---|
| *(a) 0–3 months time horizon* | | | |
| constant | −7.6163 | 2.5950 | 0.0020† |
| flood depth | −0.1680 | 0.0443 | 0.0000† |
| log(median income) | 0.7330 | 0.2520 | 0.0015† |
| small size | −0.2765 | 0.1397 | 0.0270† |
| large size | −0.3286 | 0.3212 | 0.1535 |
| low status customers | −0.3293 | 0.1662 | 0.0150† |
| high status customers | 0.0855 | 0.1313 | 0.2690 |
| sole proprietorship | 0.5512 | 0.1957 | 0.0025† |
| national chain | 0.0684 | 0.3785 | 0.4205 |
| $Wy$ | 0.3820 | 0.0938 | 0.0000† |
| *(b) 0–6 months time horizon* | | | |
| constant | −2.9778 | 2.7300 | 0.1425 |
| flood depth | −0.1103 | 0.0349 | 0.0005† |
| log(median income) | 0.3110 | 0.2685 | 0.1270 |
| small size | −0.1087 | 0.1488 | 0.2330 |
| large size | −0.3723 | 0.3316 | 0.1340 |
| low status customers | −0.3423 | 0.1613 | 0.0175† |
| high status customers | 0.0407 | 0.1528 | 0.3920 |
| sole proprietorship | 0.3588 | 0.1813 | 0.0210† |
| national chain | 0.2948 | 0.3806 | 0.2195 |
| $Wy$ | 0.5783 | 0.0842 | 0.0000† |
| *(c) 0–12 months time horizon* | | | |
| constant | −4.3361 | 2.7227 | 0.0600 |
| flood depth | −0.0894 | 0.0343 | 0.0070† |
| log(median income) | 0.4844 | 0.2683 | 0.0395† |
| small size | −0.2142 | 0.1542 | 0.0845 |
| large size | −0.3572 | 0.2981 | 0.1170 |
| low status customers | −0.3209 | 0.1620 | 0.0200† |
| high status customers | −0.1015 | 0.1654 | 0.2665 |
| sole proprietorship | 0.1459 | 0.1886 | 0.2175 |
| national chain | −0.1205 | 0.3893 | 0.3640 |
| $Wy$ | 0.5846 | 0.0929 | 0.0000† |

†Significance at the 95% level.

ables for type of store ownership, one reflecting sole proprietorships and the other representing national chains (with regional chains representing the excluded class).

## 3.2. Model estimates

Given our sample of stores arranged along streets, we wish to determine the appropriate number of neighbours to use when forming the spatial weight matrix $W$ that is used in our model. This was done by using the deviance information criterion (DIC), which is a common approach taken in Bayesian analysis of competing models (Spiegelhalter *et al.*, 2002; Gelman *et al.*, 2004). Models based on varying numbers of nearest neighbours were estimated and the DIC values were calculated, which reflect a measure of fit adjusted for model complexity, similar to other such model comparison measures such as the Akaike information criterion and Bayesian information criterion. Lower values reflect superior models and differences in DIC values of more than 7–10 units between two models are regarded as strong evidence in favour of the model

**Table 4.** SAR probit model effects estimates for the 0–3-month time horizon

| *Variable* | *Lower 0.05* | *Posterior mean* | *Upper 0.95* |
|---|---|---|---|
| *(a) Direct effects* | | | |
| flood depth | −0.0728 | −0.0485† | −0.0255 |
| log(median income) | 0.0677 | 0.2116† | 0.3404 |
| small size | −0.1596 | −0.0802† | −0.0016 |
| large size | −0.2777 | −0.0948 | 0.0854 |
| low status customers | −0.1916 | −0.0950† | −0.0064 |
| high status customers | −0.0492 | 0.0247 | 0.1022 |
| sole proprietorship | 0.0500 | 0.1596† | 0.2651 |
| national chain | −0.1937 | 0.0199 | 0.2371 |
| *(b) Indirect effects* | | | |
| flood depth | −0.0566 | −0.0296† | −0.0110 |
| log(median income) | 0.0318 | 0.1284† | 0.2602 |
| small size | −0.1240 | −0.0499† | −0.0008 |
| large size | −0.2113 | −0.0609 | 0.0473 |
| low status customers | −0.1437 | −0.0583† | −0.0038 |
| high status customers | −0.0334 | 0.0152 | 0.0742 |
| sole proprietorship | 0.0225 | 0.0994† | 0.2188 |
| national chain | −0.1374 | 0.0122 | 0.1595 |
| *(c) Total effects* | | | |
| flood depth | −0.1199 | −0.0782† | −0.0421 |
| log(median income) | 0.1069 | 0.3401† | 0.5569 |
| small size | −0.2685 | −0.1302† | −0.0029 |
| large size | −0.4680 | −0.1558 | 0.1284 |
| low status customers | −0.3210 | −0.1534† | −0.0118 |
| high status customers | −0.0823 | 0.0399 | 0.1647 |
| sole proprietorship | 0.0768 | 0.2591† | 0.4546 |
| national chain | −0.3229 | 0.0321 | 0.3951 |

†Significance at the 95% level.

with the smaller DIC. To calculate the DIC values we used the continuous dependent variable log-likelihood in conjunction with the latent draws for $y^*$. The results point to a model with 11 nearest neighbours at the 3-month horizon and 15 neighbours for the 6- and 12-month horizons. Theory suggests that longer time horizons will via diffusion lead to spatial dependence over a larger area (LeSage and Pace, 2009).

We also examined effects estimates based on models constructed by using 8–14 neighbours for the 3-month horizon model to see how sensitive our estimates and inferences were to the choice of 11 neighbours. A plot with these estimates on the vertical axis *versus* the number of neighbours on the horizontal axis resulted in a nearly horizontal line, and all estimates were within the lower 0.05 and upper 0.95 credible intervals reported for our 11-neighbour model.

The coefficient estimates (posterior means, standard deviations and Bayesian $p$-levels) for the model parameters $\beta$ and $\rho$ are shown in Table 3 for the three different time horizons. As already noted, the coefficient estimates $\beta$ from the SAR probit model cannot be interpreted as representing how changes in the explanatory variables affect the probability that stores will reopen. One point to note is that the coefficient $\rho$ that is associated with the spatial lag of the dependent variable $Wy$ is more than 4 standard deviations away from 0 in all three sets of results. The

**Table 5.** SAR probit model effects estimates for the 0–6-month time horizon

| *Variable* | *Lower 0.05* | *Posterior mean* | *Upper 0.95* |
|---|---|---|---|
| *(a) Direct effects* | | | |
| flood depth | −0.0445 | −0.0276† | −0.0110 |
| log(median income) | −0.0551 | 0.0776 | 0.2104 |
| small size | −0.1092 | −0.0278 | 0.0449 |
| large size | −0.2787 | −0.0943 | 0.0639 |
| low status customers | −0.1679 | −0.0857† | −0.0062 |
| high status customers | −0.0663 | 0.0105 | 0.0857 |
| sole proprietorship | 0.0054 | 0.0906† | 0.1905 |
| national chain | −0.1102 | 0.0745 | 0.2746 |
| *(b) Indirect effects* | | | |
| flood depth | −0.0562 | −0.0343† | −0.0160 |
| log(median income) | −0.0773 | 0.0966 | 0.2836 |
| small size | −0.1428 | −0.0349 | 0.0568 |
| large size | −0.3939 | −0.1209 | 0.0839 |
| low status customers | −0.2319 | −0.1098† | −0.0077 |
| high status customers | −0.0956 | 0.0120 | 0.1201 |
| sole proprietorship | 0.0055 | 0.1183† | 0.2937 |
| national chain | −0.1436 | 0.0999 | 0.4202 |
| *(c) Total effects* | | | |
| flood depth | −0.0913 | −0.0619† | −0.0294 |
| log(median income) | −0.1303 | 0.1743 | 0.4633 |
| small size | −0.2385 | −0.0628 | 0.1018 |
| large size | −0.6412 | −0.2152 | 0.1471 |
| low status customers | −0.3809 | −0.1955† | −0.0134 |
| high status customers | −0.1581 | 0.0226 | 0.2028 |
| sole proprietorship | 0.0108 | 0.2089† | 0.4638 |
| national chain | −0.2474 | 0.1745 | 0.6666 |

†Significance at the 95% level.

magnitudes ranging between 0.38 and 0.58 point to a significant positive spatial dependence in firms' decisions regarding open–closed status so firms nearby exhibit similar decision outcomes regarding reopening. A theoretical upper bound on $\rho$ is strictly less than 1.

The signs (negative or positive) of the estimates reported in Table 3 reflect the signs of the direct effects estimates as well as the indirect and total effects which have the same signs in our SAR probit model. Using the reported signs we see that some variables exhibited consistent signs for all three time horizons. Specifically, flood depth has a negative influence on the probability that firms will reopen, (logged) median household income has a positive influence, small and large firms relative to medium sized firms have a negative influence, low socio-economic status of customers has a negative effect (relative to middle socio-economic status) and sole proprietorships exhibited a positive effect on the probability that stores will reopen.

Two variables exhibited a change in sign between the first two time horizons and the last time horizon. Specifically, high socio-economic status and ownership by national chains, which both had a positive influence on reopening at the first two time horizons, showed a negative effect in the last time horizon.

The effects estimates for models covering the three time horizons are shown in Tables 4–6. These are the basis for inference regarding the effect of changes in the various explanatory

**Table 6.** SAR probit model effects estimates for the 0–12-month time horizon

| *Variable* | *Lower 0.01* | *Posterior mean* | *Upper 0.99* |
|---|---|---|---|
| *(a) Direct effects* | | | |
| flood depth | −0.0353 | −0.0203† | −0.0047 |
| log(median income) | −0.0132 | 0.1113 | 0.2348 |
| small size | −0.1282 | −0.0499 | 0.0205 |
| large size | −0.2336 | −0.0824 | 0.0546 |
| low status customers | −0.1522 | −0.0736† | −0.0024 |
| high status customers | −0.0975 | −0.0232 | 0.0534 |
| sole proprietorship | −0.0560 | 0.0332 | 0.1249 |
| national chain | −0.2135 | −0.0290 | 0.1465 |
| *(b) Indirect effects* | | | |
| flood depth | −0.0472 | −0.0274† | −0.0088 |
| log(median income) | −0.0196 | 0.1541 | 0.3769 |
| small size | −0.2105 | −0.0719 | 0.0276 |
| large size | −0.3497 | −0.1165 | 0.0689 |
| low status customers | −0.2347 | −0.1022† | −0.0040 |
| high status customers | −0.1613 | −0.0337 | 0.0807 |
| sole proprietorship | −0.0757 | 0.0498 | 0.2096 |
| national chain | −0.2995 | −0.0373 | 0.2368 |
| *(c) Total effects* | | | |
| flood depth | −0.0756 | −0.0478† | −0.0140 |
| log(median income) | −0.0310 | 0.2655 | 0.5783 |
| small size | −0.3231 | −0.1219 | 0.0483 |
| large size | −0.5640 | −0.1990 | 0.1192 |
| low status customers | −0.3660 | −0.1758† | −0.0063 |
| high status customers | −0.2598 | −0.0569 | 0.1343 |
| sole proprietorship | −0.1321 | 0.0830 | 0.3243 |
| national chain | −0.4961 | −0.0664 | 0.3790 |

†Significance at the 95% level.

variables on the probability that stores will reopen as well as the (cumulative) spatial spillover effcts on neighbouring stores.

In the 0–3-month horizon, the flood depth variable exerts a negative direct and (cumulative) indirect effect on the probability that stores will reopen, implying a direct effect of a 4.8% decrease in probability of reopening for every 1 ft of flood depth, and an indirect effect around 3%, combining for a total effect around 7.8%. Over time, the direct effects of flooding decrease to 2% by the 0–12-month horizon, whereas the indirect or spillover effects decline by a smaller amount to 2.7% by the 0–12-month horizon. Total effects decline from 7.8% to around 4.8% over time. This seems consistent with the notion that disaster assistance as well as market forces work over time to produce a move back towards equilibrium, ameliorating the effect of flooding with the passage of time.

In the 0–3-month horizon, (logged) median household income of the census block group in which the store was located had a positive and significant direct effect that would raise the probability that stores would reopen by 2.11% for every 10% increase in income and an indirect effect of 1.28% for a total effect around 3.4%. Over the next two time horizons the direct, indirect and total effects of income remained positive but diminished to produce a total effect that was insignificant (using the 0.05 and 0.95 credible intervals).

In the 0–3-month period, small size stores (measured by employment) had a negative direct effect, reducing the probability of reopening. For categorical variables such as the store size, we interpret the magnitude of effects estimates as reflecting how a change in category from the omitted category (in this case medium sized or medium employment stores) would influence the probability of opening. The direct effect of small store size was to decrease the probability of opening by around −8% and the indirect or spillover effect was around −5% for a total effect of −13%. As in the case of the income variable, this variable became insignificant at the 0–6- and 0–12-month horizons.

Low socio-economic status of the store *clientèle* had a negative and significant influence for all three time horizons. During the 0–3-month period, the direct effects were equal to −0.095, with indirect effects of −0.058 for a total effect of −0.15 suggesting a decrease in probability of reopening relative to the omitted class of middle socio-economic status. Over time the direct effects remained about the same during the next two periods, whereas the indirect effects increased to −0.11 and −0.10 in these latter two periods. We note that high socio-economic status *clientèle* had a positive but insignificant effect (relative to the omitted class of middle socio-economic status) for all three time horizons.

The other variable exhibiting significant effects was the ownership dummy variable representing sole proprietorships (relative to the excluded class of regionally owned chains). This variable had a positive direct effect of 0.16 and indirect effects of 0.10, suggesting a positive total effect around 0.26 on the probability of reopening in the 0–3-month horizon relative to the excluded class of regionally owned chains. Sole proprietorships have several features which could account for their faster pace in reopening. First, many of these are businesses that employ multiple members of the same family, which may confer a non-pecuniary benefit to reopening. Second, sole proprietorships may be the main means of support for the owner and family members who have relatively poorer prospects on the open job market due to their firm-specific human capital. Third, sole proprietorships may be able to react faster when making such decisions than larger firms.

Over time, the direct effect declined to 0.09 and 0.03, becoming insignificant during the 0–12-month period. The indirect effect remained high during the 0–6-month period, having a magnitude of 0.12, but it declined to 0.05 and became insignificant in the 0–12-month period. This pattern of relationships led to a total effect in the 0–3-month period of 0.26, declining to 0.21 at the 0–6-month horizon and an insignificant 0.08 in the 0–12-month period. Again, the notion that as time passes the influence of this type of ownership might diminish as recovery leads back to an equilibrium situation seems intuitively plausible.

### *3.3. Investigation of the variability in the effects over space*

As indicated in Section 2.3, we need to consider whether the global scalar summary measures on which we based our inferences in the previous section exhibit large variation over locations. Use of the scalar summary measures might produce inferences that are unrepresentative of the distribution of effects as they vary over our sample space.

As an investigative tool, we produced (cumulative) total effects estimates for each of the 673 store locations in our sample. Since the observations that were used in this application are stores along three different streets, it is relatively easy to examine the total effects estimates visually as we move along each street. In more typical applications where the observations represent regions in space, it should be possible to produce a choropleth map to explore variability over space.

A set of three-dimensional figures was used to plot the total effects estimates for each store location on the vertical axis and the longitude–latitude co-ordinates as the horizontal axes. Fig. 1 shows flood depths for each store on Carrollton and St Claude Avenues, and we note that

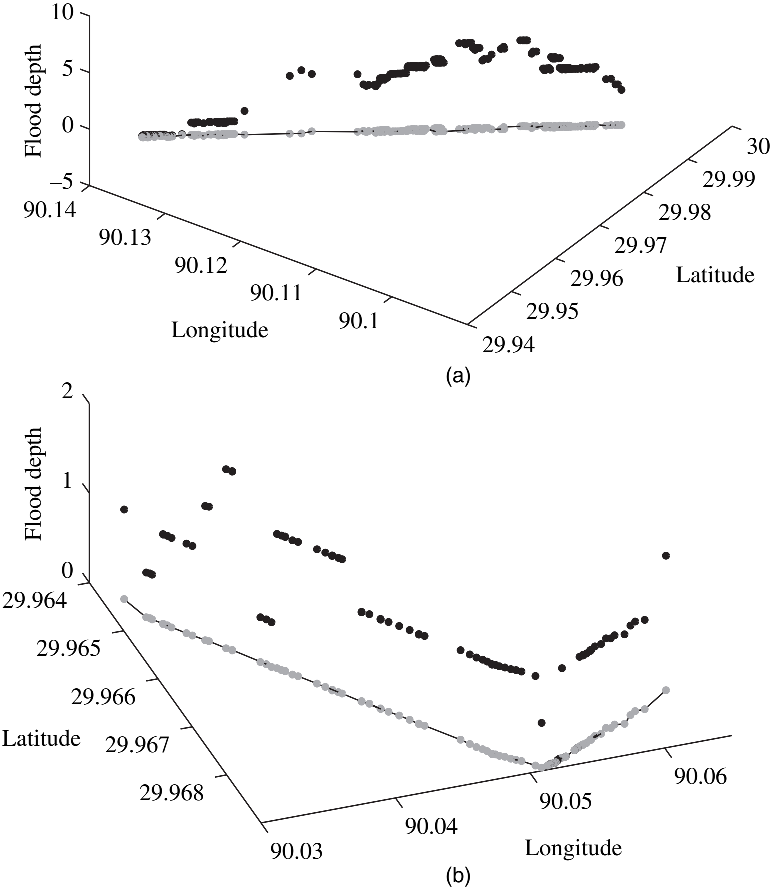


**Fig. 1.** Flood depths (•) and stores (—•—): (a) Carrollton Avenue; (b) St Claude Avenue

Magazine Street experienced no flooding. Intuitively, we might expect the effects from changes in the various explanatory variables on the probability that stores will reopen would vary by flood depth.

We illustrate variability over space for two of our explanatory variables, one being the national chain ownership type where variation does not result in an unrepresentative inference arising from use of the scalar summary measures of the effects estimates. For the other variable, logged median income, we have a situation where there is substantial variation in the effects estimates for stores on Carrollton Avenue.

Both the 0–3-month and 0–12-month estimates are shown in the figures, which allows us to see whether the changes in magnitude of influence for the various variables over time are consistent with our global inferences based on the scalar summary measures.

Fig. 2 shows total effects for the national chain dummy variable (relative to the omitted regional chain ownership dummy variable). Here we see a reversal in sign when moving from the 0–3-month to 0–12-month models, which is consistent with the scalar summary measures of 0.032 and −0.066 for the two periods respectively. In the short term, this type of ownership had a small positive (but not significant) effect on the probability of reopening, whereas the longer-term effect was negative (and not significant). Since the scalar summary measure indicated a lack of significance for the total effects from this variable in both time periods, we should be

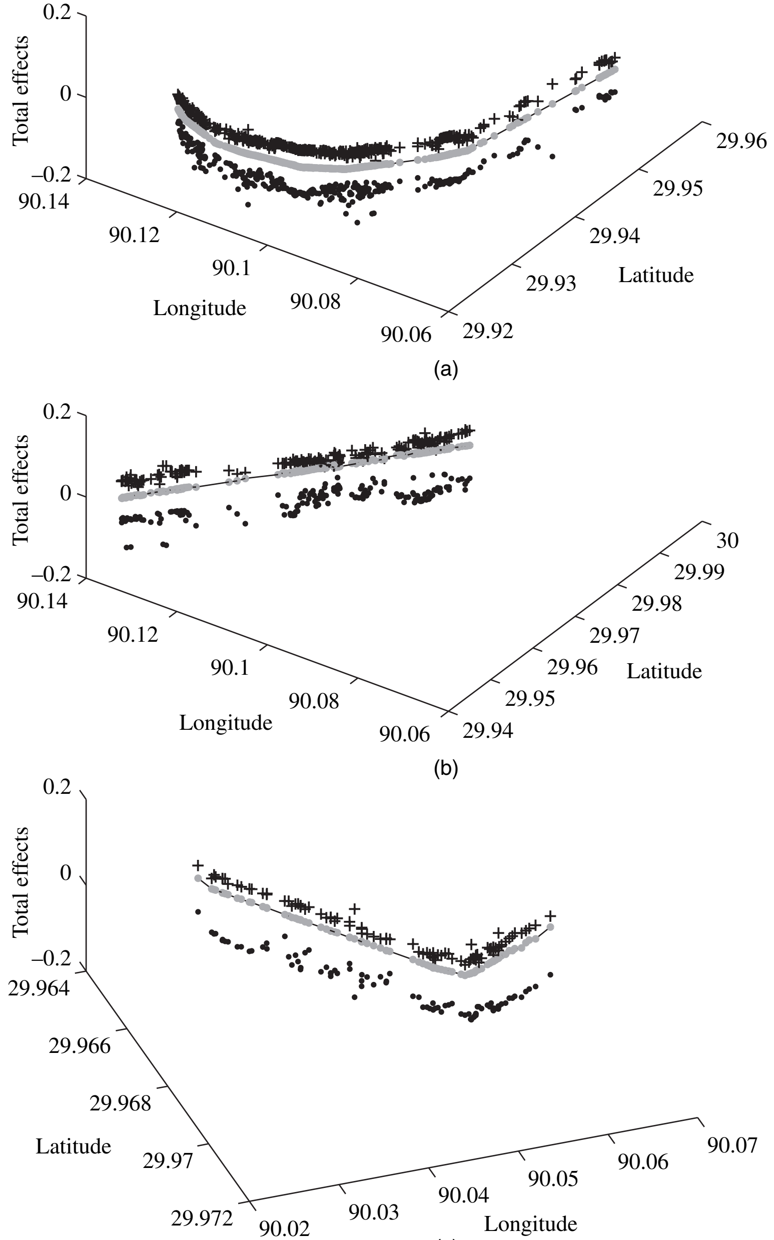


**Fig. 2.** National chains, store level total effects (+, 0–3 months; •, 0–12 months; —•—, stores): (a) Magazine Street; (b) Carrollton Avenue; (c) St Claude Avenue

reluctant to place much emphasis on this set of results. Variation in the magnitude of effects over locations along the three streets is fairly limited, causing no problems for our global scalar summary measures.

Logged median household income is shown in Fig. 3, where we see a positive total effect of this variable on stores reopening. The effects of heavy flooding in the middle portions of Carrollton Avenue are evident as these lead to the 0–3-month estimates falling towards 0. Of course,

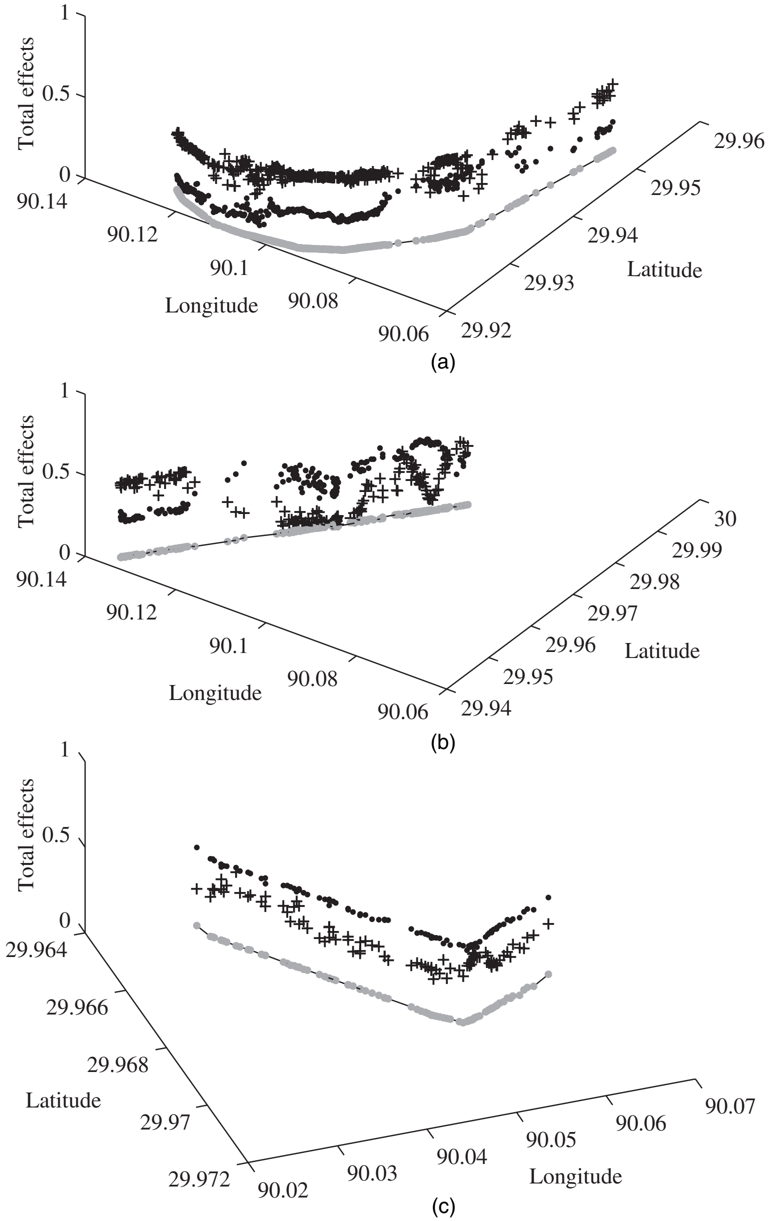


**Fig. 3.** Median income, store level total effects (+, 0–3 months; •, 0–12 months; —•—, stores): (a) Magazine Street; (b) Carrollton Avenue; (c) St Claude Avenue

it seems intuitive that customers' incomes would have less short-term effect in the face of large flood damage to stores. The long-term effect on St Claude Avenue differs from that on the other two streets, becoming larger whereas the longer-term importance of household income diminishes on Magazine Street. For Carrollton Avenue we see patterns for the 0–12-month period that reflect both a decrease and an increase in magnitude relative to the 0–3-month estimates, depending on flood depths associated with store locations. Of course, the mix of stores should have an influence on the role that is played by household income as well as the obvious influence of the flood depth variable. The short-run effect of this variable was positive (0.34) and significant at the 99% level. Long-term total effects were positive (0.265) and different from 0 by using 95% intervals, but not the 99% intervals that are reported in Table 4. The scalar summary measures seem most consistent with the total effects shown for Magazine Street and St Claude Avenue, and least representative for Carrollton Avenue.

By way of summary, it should be possible to use maps to diagnose cases where the scalar summary effects estimates produce a valid global inference regarding the direction and relative magnitude of influence for the model variables on the probability estimates. We found a high degree of variability over the street level locations for the flood depth, median household income and low social status of store customer variables. For all other variables the scalar summary measures of the effects estimates would provide a representative global inference.

An important point is that spatial regression models produce direct and indirect or spatial spillover estimates that vary over space. This seems not to be widely recognized, and in fact there is much confusion regarding the global character of model parameters such as $\beta$, $\rho$ and $\sigma^2$ *versus* the varying nature of the partial derivative effects arising from changes in the explanatory variables of these models. This confusion is often used to justify spatially variable parameter methods such as geographically weighted regression in place of spatial regression models (Fotheringham *et al.*, 2002).

Our analysis shows that the effects (correctly interpreted) arising from changes in explanatory variables vary over spatial locations. This is true of non-probit SAR models as well as the probit variant of this model considered here. Additional non-linearity enters the relationship in the case of a probit model due to the non-linearity of the cumulative density function for the normal distribution. This places some burden on practitioners to examine whether simple scalar summary measures will produce representative global inferences or whether inferences require qualification for some parts of the sample space (locations).

## 4. Conclusions

This study focused on interpretation of estimates arising from use of an SAR probit model, where spatial lags of the dependent variable allow for interdependence in choices. Spatial dependence of household or firm decisions when these economic agents are nearby is likely to be a frequently encountered situation in applied spatial modelling. For example, nearby commuters are likely to make similar commuting choices because of their common location, which leads to similar options regarding access to roadway and public transport systems. Owners of retail establishments near each other share neighbourhood consumers with the same socio-economic and demographic characteristics. Local governments operating in nearby locations confront similar regional economic conditions and common state laws.

Our theoretical development of the way in which changes in the explanatory variables of SAR probit regression models impact choices demonstrated a potential for direct as well as indirect or spatial spillover effects. These are formally represented by an $n \times n$ matrix of potential effects, where $n$ represents the sample size. Constructing useful summary measures of the large number of effects arising from the spatial dependence structure poses a challenge to using these models.

We proposed scalar summary measures of the direct and indirect effects that coincide with the mathematical definition of own- and cross-partial derivatives for the SAR probit model. Successful use of these summary measures for inference requires that practitioners understand the nature of the partial derivative effects that arise in these models. Past focus on the coefficient estimates that are associated with explanatory variables and the scalar parameter measuring the strength of spatial dependence has led to the impression that SAR models produce only a global summary of the regression relationship that is involved. LeSage and Pace (2009) pointed to the numerous influences on neighbouring observations that are associated with a change made in a single observation or location of an explanatory variable. In essence, these models can be viewed as spatially varying effects models, despite the fact that the model relationship is expressed by using a single set of estimated (global) parameters. This aspect of the models becomes more pronounced in the case of the probit variant of SAR models because of the additional (normal) probability transformation.

An examination of scalar summaries for the total effects that are associated with changing the explanatory variables as they vary over spatial locations was illustrated in our application. This is analogous to (but more complicated than) the situation arising in non-spatial probit models where 'marginal effects' are calculated in an effort to consider the non-linearity of response in decision probabilities with respect to changes in the magnitude of the explanatory variables. In the case of SAR probit models, it is more meaningful to consider the non-linearity of response in decision probabilities over spatial locations, since changing location in these models implies a change in variable magnitude because observations and locations are equivalent.

Our findings with regard to factors influencing the return of business in New Orleans after Hurricane Katrina showed that these vary with location as well as passage of time after the disaster. For example, we found that in the short term (0–3 months) the sole proprietorship ownership category had a positive influence on the probability that these firms would reopen, as well as a positive influence on the probability that neighbouring establishments reopened. With the passage of time (0–6 months), the direct effect of this type of ownership (relative to regional chains) diminished whereas the spatial spillover effects on neighbouring firms grew, with both remaining positive and significant. As we might expect, very high levels of flooding at store locations tended to mitigate the positive effects that are associated with this type of ownership. In the longer term (0–12 months) as the effect of disaster aid and other forces bring the economic climate back towards pre-disaster conditions, factors that influenced the probability of reopening in the short term diminished to the point of insignificance in many cases.

A consistent pattern that arises when viewing variation in the effects estimates by location is that decisions to reopen became less dependent on the level of flooding over time, i.e. business responses became more constant (for all factors) over locations that experienced high or low levels of flooding. This might provide a fruitful metric for assessing the status of business recovery from disasters that involve flooding. Formal measures of variation in business response in locations with no flooding *versus* various levels of flooding could be developed in an effort to assess the time horizon when flood depths become unimportant.

## Acknowledgements

The authors acknowledge support for this research provided by the National Science Foundation, grants SES-0554937 and SES-0729264. Additional funding was from the Gulf of Mexico, Texas sea grant NA06OAR41770076 and the Louisiana sea grant GOM/RP-02 programmes. The statements, findings, conclusions and recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation, the National Oceanic and Atmospheric Administration or the US Department of Commerce.

## References

Albert, J. H. and Chib, S. (1993) Bayesian analysis of binary and polychotomous response data. *J. Am. Statist. Ass.*, **88**, 669–679.

Berman, B. and Evans, J. R. (1991) *Retail Management: a Strategic Approach*. New York: Macmillan.
Brown, S. (1989) Retail location theory: the legacy of Harold Hotelling. *J. Retail.*, **65**, 450–470.
Brown, S. (1994) Retail location at the micro-scale: inventory and prospect. *Serv. Indust. J.*, **14**, 542–576.
Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2002) *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. Chichester: Wiley.
Fujita, M. and Smith, T. E. (1990) Additive-interaction models of spatial agglomeration. *J. Regl Sci.*, **30**, 51–74.
Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M. (1990) Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Am. Statist. Ass.*, **85**, 972–985.
Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (1995) *Bayesian Data Analysis*. New York: Chapman and Hall.
Golledge, R. G. and Timmermans, H. (1990) Applications of behavioural research on spatial problems. *Prog. Hum. Geog.*, **14**, 57–99.
Hernandez, T. and Bennison, D. (2000) The art and science of retail location decisions. *Int. J. Retl Distribn Mangmnt*, **28**, 357–367.
Hinloopenaand, J. and van Marrewijk, C. (1999) On the limits and possibilities of the principle of minimum differentiation. *Int. J. Industrl Organizn*, **17**, 735–750.
Hotelling, H. (1929) Stability in competition. *Econ. J.*, **39**, 41–57.
Hunt, M. A. and Crompton, J. L. (2008) Investigating attraction compatibility in an East Texas city. *Progr. Toursm Hosptalty Res.*, **10**, 237–246.
Lam, N., Pace, K., Campanella, R., LeSage, J. and Arenas, H. (2009) Business return in New Orleans: decision making amid post-Katrina uncertainty. *PLOS ONE*, **4**, no. 8, article e6765.
Lee, M. L. and Pace, R. K. (2005) Spatial distribution of retail sales. *J. Real Est. Finan. Econ.*, **31**, 53–69.
LeSage, J. P. and Pace, R. K. (2009) *Introduction to Spatial Econometrics*. New York: Taylor and Francis–CRC Press.
McGoldrick, P. J. and Thompson, M. G. (1992) *Regional Shopping Centres: In-town versus Out-of-town*. Aldershot: Avebury Press.
Nelson, R. L. (1958) *The Selection of Retail Locations*. New York: Dodge.
Pace, R. K. and LeSage, J. P. (2010) Spatial econometrics. In *Handbook of Spatial Statistics* (eds A. E. Gelfand, M. Fuentes, P. Guttorp and P. Diggle), pp. 245–260. Boca Raton: Chapman and Hall–CRC.
Reilly, W. J. (1931) *The Law of Retail Gravitation*. New York: Knickerbocker.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *J. R. Statist. Soc.* B, **64**, 583–639.