Newton’s laws of motion provide the analogical basis for most spatial interaction modelling in geographical systems.
Spatial interaction model

Scientists and philosophers posited that the forces that occurred in the social world could be modelled the same way as in the physical world.

H.C. Carey (1793-1879):

*Man tends of necessity to gravitate towards his fellow man. Of all animals he is the most gregarious, and the greater the number collected in a given space the greater is the attractive force there exerted. ... Gravitation is here, as everywhere else in the material world, in the direct ratio of mass, and in the inverse one of the distance.*
Ravenstein’s laws of migration

• The most explicit application was made by Ravenstein in 1888, who defined laws of migration.

Migrants enumerated in certain centre of absorption will consequently grow less with distance proportionately to the native population which furnishes them. (Ravenstein, 1885)
Gravity model

\[ F_{ij} = kP_i P_j / d_{ij}^2 \]

Where \( F_{ij} \) is the number of migrants recorded between place \( i \) and place \( j \)
\( P_i \) is population of \( i \)
\( P_j \) is population of \( j \)
\( D_{ij} \) is distance between \( i \) and \( j \)
\( K \) is a coefficient
Simple form:

$$F_{ij} = kP_i P_j / d_{ij}^2$$

Generalised to:

$$F_{ij} = kP_i^{b_1} P_j^{b_2} / d_{ij}^{b_3}$$
Gravity model

\[ F_{ij} = k P_i^{b_1} P_j^{b_2} / d_{ij}^{b_3} \]

\[ \ln F_{ij} = \ln(k P_i^{b_1} P_j^{b_2} / d_{ij}^{b_3}) \]

\[ \ln F_{ij} = \ln k + b_1 \ln P_i + b_2 \ln P_j - b_3 \ln d_{ij} \]

\[ \ln F_{ij} = b_0 + b_1 \ln P_i + b_2 \ln P_j - b_3 \ln d_{ij} \]
Gravity model

\[ F_{ij} = k P_i^{b_1} P_j^{b_2} / d_{ij}^{b_3} \]

\[ \ln F_{ij} = \ln (k P_i^{b_1} P_j^{b_2} / d_{ij}^{b_3}) \]

\[ \ln F_{ij} = \ln k + b_1 \ln P_i + b_2 \ln P_j + b_3 \ln d_{ij} \]

\[ \ln F_{ij} = b_0 + b_1 \ln P_i + b_2 \ln P_j + b_3 \ln d_{ij} \]
OLS (Ordinary Least Squares)

\[
\ln F_{ij} = b_0 + b_1 \ln P_i + b_2 \ln P_j + b_3 \ln d_{ij} + e_{ij}
\]

\[
e_{ij} \sim N(0, \sigma^2)
\]
Poisson model

\[ F_{ij} = \exp(b_0 + b_1 \ln P_i + b_2 \ln P_j + b_3 \ln d_{ij}) + e_{ij} \]

\[ e_{ij} = \text{Poisson}(u_{ij}) \]
Measuring distance

Distance measurements:
What type of centroid to use?
geometric
population weighted
How to measure distance?
Euclidean
Network
Mixed: Euclidean and network
road network
Estuary problem
Island effect

Assume Euclidean distance results in over-estimates of flows between, into and out of islands.

In fact, the model for all Scottish wards shows these flows are under-estimated.
Comparison between migration model results with different distance measures

<table>
<thead>
<tr>
<th></th>
<th>Euclidean distance</th>
<th>Mixed distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviance</td>
<td>921422</td>
<td>922090</td>
</tr>
<tr>
<td>Degree of freedom</td>
<td>1002998</td>
<td>1002998</td>
</tr>
<tr>
<td>Proportion explained</td>
<td>0.5811</td>
<td>0.5801</td>
</tr>
<tr>
<td>Constant</td>
<td>1.3390</td>
<td>1.1811</td>
</tr>
<tr>
<td>Logged distance</td>
<td>-1.3888</td>
<td>-1.3644</td>
</tr>
<tr>
<td>Logged origin population</td>
<td>0.7144</td>
<td>0.7124</td>
</tr>
<tr>
<td>Logged destination population</td>
<td>0.6714</td>
<td>0.6697</td>
</tr>
</tbody>
</table>

Data source: 1991 SMS Set 1, Scotland
Intra-zone flows

• Possible to include intra-zone flows in the model
• Need to define how to measure distance for intra-zone flows
• Radius of a circle with the same area as the zone
  – Radius
  – ½ radius
  – ¼ radius
Choice of explanatory variables

– Deviance
– Z test
– theoretical
Factors affecting migration

• Push factors
  – Decline in a national resource or the prices it commands; decreased demand for a particular product or service
  – Loss of employment
  – Discriminatory treatment
  – Cultural alienation from a community
  – Poor marital opportunities
  – Retreat due to catastrophe
Factors affecting migration

• Pull factors
  – Improved employment opportunities
  – Superior income-earning opportunities
  – Opportunities for training/education
  – Preferable environment or general living conditions
  – Movement as a result of dependency on someone else who has moved (tied)
  – Novel, rich or varied cultural, intellectual or recreational environment
    • Bogue, 1969, Lewis 1982
MIGMOD

• Led by Office of the Deputy Prime Minister
• Experts from University of Newcastle, University of Leeds, and Greater London Authority
• The aim was to investigate the impacts of alternative economic and policy scenarios on population flows between the English regions, and to provide some advice about urban–rural shift.
Variable categories in MIGMOD

• Spatial structure
• Demography
• Economic
• Employment
• Housing
• Social
• Environment
• Access to service, amenities
Inter-council migration flow data, 2001

Observation=992
Mean=114.7
Min=0
Max=3207
Std dev=273.0
Inter-council migration flow data, 2001

Observation=992
Mean=3.57
Min=0
Max=8.1
Std dev=1.53
Fit OLS regression

Stata:
model 1: $R^2=0.6262$

\[ \text{reg lng1 lopop ldpop ldist} \]

model 2: $R^2=0.6769$

\[ \text{reg lng1 lopop ldpop ldist contig} \]
Results from the OLS regression, model 1

|          | Coef.  | Std. Err. | t      | P>|t|  | 95% CI    |
|----------|--------|-----------|--------|------|-----------|
| lopop    | 0.757  | 0.042     | 18.130 | 0.000| 0.675-0.839|
| lrpop    | 0.862  | 0.042     | 20.660 | 0.000| 0.780-0.944|
| ldist    | -0.613 | 0.038     | -16.250| 0.000| -0.687-0.539|
| _cons    | -12.539| 0.836     | -14.990| 0.000| -14.180-10.897|

\[
LnF_{ij} = -12.539 + 0.757lnP_i + 0.862lnP_j - 0.613lnd_{ij}
\]
Results from the OLS regression, model 2

|       | Coef. | Std. Err. | t     | P>|t| | 95% CI     |
|-------|-------|-----------|-------|-----|-------------|
| lopop | 0.783 | 0.039     | 20.140| 0.000 | 0.707 - 0.859 |
| lrpop | 0.888 | 0.039     | 22.840| 0.000 | 0.812 - 0.964 |
| ldist | -0.387| 0.040     | -9.800| 0.000 | -0.465 - -0.309|
| contig| 1.313 | 0.106     | 12.440| 0.000 | 1.106 - 1.520 |
| _cons | -14.326 | 0.791 | -18.110| 0.000 | -15.879 - -12.773|

$$LnF_{ij} = -14.326 + 0.783lnP_i + 0.888lnP_j - 0.387lnd_{ij} + 1.313contig$$
Fit the Poisson model

• Softwares:
  – SPSS: GENLIN command
  – STATA: Poisson or GLM command
  – SAS: GENMOD
  – R: glim()
Goodness of fit for Poisson model

• Null Model
  – Deviance = 267542.82

• Model 1
  – Deviance = 86187.036
  – $G^2 = \frac{(267542.82 - 86187.036)}{267542.82} = 67.8\%$

• Model 2
  – Deviance = 73010.64
  – $G^2 = \frac{(267542.82 - 73010.64)}{267542.82} = 72.7\%$
Goodness of fit for Poisson model

- Model 1
  - Deviance = 86187.036

- Model 2
  - Deviance = 73010.64

- Deviance difference = 86187.036 - 73010.64 = 13176.396 > 3.84
### Results from the Poisson regression, model 1

|       | Coef. | Std. Err. | t     | P>|t| | 95% CI |
|-------|-------|-----------|-------|------|--------|
| lopop | 0.817 | 0.004     | 184.51| 0.000| 0.808  | 0.826  |
| lrpop | 0.845 | 0.004     | 190.86| 0.000| 0.836  | 0.854  |
| ldist | -0.882| 0.003     | -253.15| 0.000| -0.889 | -0.875 |
| _cons | -11.513| 0.085   | -134.07| 0.000| -11.682| -11.345|

\[ LnF_{ij} = -11.513 + 0.817 \ln P_i + 0.845 \ln P_j - 0.882 \ln d_{ij} \]
Results from the Poisson regression, model 2

|          | Coef. | Std. Err. | z    | P>|z| | 95% CI      |
|----------|-------|-----------|------|------|-------------|
| lopop    | 0.811 | 0.004     | 181.220 | 0.000 | 0.803       | 0.820       |
| lrpop    | 0.840 | 0.004     | 187.530 | 0.000 | 0.831       | 0.849       |
| ldist    | -0.511 | 0.005    | -109.070 | 0.000 | -0.520      | -0.502      |
| contig   | 0.980 | 0.009     | 115.230 | 0.000 | 0.963       | 0.997       |
| _cons    | -13.140 | 0.089   | -147.550 | 0.000 | -13.314     | -12.965     |

\[ LnF_{ij} = -13.140 + 0.811lnP_i + 0.840lnP_j - 0.511lnd_{ij} + 0.980contig \]
## Largest positive residuals from the Poisson gravity model

<table>
<thead>
<tr>
<th>origin</th>
<th>destination</th>
<th>Observed flow</th>
<th>Predicted flow</th>
<th>deviance residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aberdeen City</td>
<td>Aberdeenshire</td>
<td>3207</td>
<td>727.0</td>
<td>4759.9</td>
</tr>
<tr>
<td>Aberdeenshire</td>
<td>Aberdeen City</td>
<td>2997</td>
<td>725.6</td>
<td>4251.0</td>
</tr>
<tr>
<td>Dundee City</td>
<td>Angus</td>
<td>1124</td>
<td>291.3</td>
<td>1517.6</td>
</tr>
<tr>
<td>Edinburgh, City of</td>
<td>West Lothian</td>
<td>1989</td>
<td>954.0</td>
<td>1461.3</td>
</tr>
<tr>
<td>Aberdeen City</td>
<td>Edinburgh, City of</td>
<td>883</td>
<td>177.7</td>
<td>1415.7</td>
</tr>
<tr>
<td>Edinburgh, City of</td>
<td>East Lothian</td>
<td>1530</td>
<td>610.6</td>
<td>1405.5</td>
</tr>
<tr>
<td>Edinburgh, City of</td>
<td>Fife</td>
<td>1590</td>
<td>666.2</td>
<td>1383.2</td>
</tr>
<tr>
<td>Highland</td>
<td>Edinburgh, City of</td>
<td>758</td>
<td>152.6</td>
<td>1214.8</td>
</tr>
<tr>
<td>Fife</td>
<td>Edinburgh, City of</td>
<td>1436</td>
<td>671.0</td>
<td>1092.6</td>
</tr>
<tr>
<td>Edinburgh, City of</td>
<td>Midlothian</td>
<td>1416</td>
<td>795.9</td>
<td>815.9</td>
</tr>
</tbody>
</table>
Largest negative residuals from the Poisson gravity model

<table>
<thead>
<tr>
<th>origin</th>
<th>destination</th>
<th>Observed flow</th>
<th>Predicted flow</th>
<th>deviance residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Lanarkshire</td>
<td>Glasgow City</td>
<td>1163.0</td>
<td>2469.2</td>
<td>-875.6</td>
</tr>
<tr>
<td>Glasgow City</td>
<td>North Lanarkshire</td>
<td>1555.0</td>
<td>2427.8</td>
<td>-692.8</td>
</tr>
<tr>
<td>Renfrewshire</td>
<td>Glasgow City</td>
<td>962.0</td>
<td>1783.6</td>
<td>-593.9</td>
</tr>
<tr>
<td>South Lanarkshire</td>
<td>Glasgow City</td>
<td>1635.0</td>
<td>2347.3</td>
<td>-591.3</td>
</tr>
<tr>
<td>South Lanarkshire</td>
<td>North Lanarkshire</td>
<td>1225.0</td>
<td>1723.1</td>
<td>-418.0</td>
</tr>
<tr>
<td>Glasgow City</td>
<td>Renfrewshire</td>
<td>1284.0</td>
<td>1722.8</td>
<td>-377.5</td>
</tr>
<tr>
<td>North Lanarkshire</td>
<td>Falkirk</td>
<td>303.0</td>
<td>699.7</td>
<td>-253.6</td>
</tr>
<tr>
<td>North Lanarkshire</td>
<td>West Lothian</td>
<td>250.0</td>
<td>675.4</td>
<td>-248.5</td>
</tr>
<tr>
<td>Falkirk</td>
<td>North Lanarkshire</td>
<td>172.0</td>
<td>715.8</td>
<td>-245.3</td>
</tr>
<tr>
<td>West Lothian</td>
<td>North Lanarkshire</td>
<td>164.0</td>
<td>689.2</td>
<td>-235.5</td>
</tr>
</tbody>
</table>
Poission vs OLS

1. Theoretically Poisson is better as the response variable is discrete count variable, not continuous

2. Consistency, OLS under-predicts while Poisson keeps the total sum as observed

<table>
<thead>
<tr>
<th></th>
<th>observed</th>
<th>OLS</th>
<th>Poisson</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>113824</td>
<td>82260</td>
<td>113824</td>
</tr>
</tbody>
</table>

3. Stability: Log-transformation of OLS may introduce further distortion, especially for sparse flow matrix
Problems in fitting Poisson models

- Multicollinearity
- Goodness of fit does not work well for sparse matrices
- Spatial autocorrelation
- MAUP:
  - Scale
  - Zonation
More on visualisation of flow data

• It is difficult to map flow data using the commercial GIS packages

• Common to draw desire lines where origins and destinations are connected by straightlines

• Special packages:
  – Flowmap [http://flowmap.geo.uu.nl/](http://flowmap.geo.uu.nl/)
  – ONS: CommuterView
Fig. 4. UK migration 2000–2001 as a line density raster.
Fig. 11. Intra-city migration: Glasgow.
Internal migration in England and Wales

http://www.neighbourhood.statistics.gov.uk/HTMLDocs/dvc25/
TECHNOLOGIES FOR MIGRATION AND COMMUTING ANALYSIS

Spatial Interaction Data Applications

JOHN STILLWELL, OLIVER DUKE-WILLIAMS, & ADAM DENNETT
Reading list

- ODPM 2002 Development of a migration model, November, Office of the Deputy Prime Minister
- Senior 1979 From gravity modelling to entropy maximizing. Progress in Human Geography 3 175-210


Acknowledgements

• Thanks go to ESRC AQMeN Collaborative Small Grants Scheme, which provides support to this training.
• We also thank Dr Oliver Duke-Williams of UCL who provides insightful advice and comments on the census flow data.
• Census data are crown copyright
• Boundary data are from Edina
• Thanks also go to Alan Fleming of NRS who provided 2001 travel to study data
• Thanks also go to NRS for their support.