The Gravity Model: Derivation and Calibration

Philip A. Viton

October 19, 2012

Introduction

- We turn now to the Gravity Model of trip distribution.
- As previously noted, this is most widely used model in current practice.
- From our perspective, it is based on two important ideas, one old and one new.
  1. It is based on a *parametrized theoretical model*, meaning that the theory depends (unlike the Average Growth Factor model) on constants that need to be estimated.
  2. It utilizes the same idea of improvement via iteration as the Average Growth Factor model, though, as we will see, in a slightly different context.

Derivation of the Gravity Model — Physics

- The gravity model is based on Newton's gravitational theory from physics, interpreted in a transportation context.
- In physics, the attractive force between two bodies is directly proportional to their masses, and inversely proportional to the square of the distance between them:

$$h_{ij} = g\frac{m_i m_j}{d_{ij}^2}$$

where:

  - h_{ij} is the attractive force between bodies i and j.
  - m_i and m_j are the bodies' masses.
  - d_{ij} is the distance between them.
  - g is Newton's gravitational constant, approximately 6.7×10^{-8} in cgs (centimeters/grammes/seconds) units.

Derivation of the Gravity Model — Transportation

The "derivation" of the transportation version of the gravity model is a simple argument from analogy:

- The Newtonian attractive force term (h_{ij}) is analogized to interzonal trip-making: $h_{ij} \mapsto T_{ij}$.
- The masses are analogized to total trips in and out of zones, so that $m_i \mapsto O_i$ and $m_j \mapsto A_j$.
- The distance term is retained, but in an attempt to be more general (whether there is a serious theoretical basis for this is a question) the power-2 term is permitted to be arbitrary.
- There is no reason to believe that the Newtonian proportionality term applies here, so we replace it by a different one.
- The result is the first form of the (transportation) gravity model:

$$T_{ij} = \theta \frac{O_i A_j}{d_{ij}^{\beta}}$$

Gravity Model with Conservation of Origins I

We now show that by requiring the gravity model to satisfy conservation of origins, we can eliminate the proportionality constant θ.

- Sum both sides of the model over destinations j:

$$\sum_j T_{ij} = \sum_j \theta \frac{O_i A_j}{d_{ij}^{\beta}}$$

- On the left, we have O_i. On the right, we can take terms not involving j outside the summation:

$$O_i = \theta O_i \sum_j \frac{A_j}{d_{ij}^{\beta}}$$

Gravity Model with Conservation of Origins II

- Now divide both sides by O_i

$$1 = \theta \sum_j \frac{A_j}{d_{ij}^{\beta}}$$

and solve for θ :

$$\theta = \frac{1}{\sum_j A_j / d_{ij}^{\beta}}$$

Gravity Model with Conservation of Origins III

- Now insert this into the gravity model. Since we are interested in T_{ij} (which already uses j) we re-label the summation index in the expression for θ. The result is:

$$T_{ij} = \frac{1}{\sum_m A_m / d_{im}^{\beta}} \frac{O_i A_j}{d_{ij}^{\beta}}$$

- In order to make this a bit more readable, let's define the *travel-time factor* F_{ij} as

$$F_{ij} = 1 / d_{ij}^{\beta}$$

- Then we have our second form of the transportation gravity model:

$$T_{ij} = \frac{O_i A_j F_{ij}}{\sum_m A_m F_{im}}$$

Gravity Model with Conservation of Origins IV

There are two important things to note about this form of the model:

1. It is guaranteed to satisfy conservation of origins exactly. Therefore we will never need to check whether this condition is satisfied or not. Of course, the same does *not* apply to conservation of attractions: there is no reason to suppose that this condition will be satisfied; and in fact it usually will not be.
2. The travel-time factors F_{ij} all depend on an unknown parameter (the parameter β). This means that the interpretation of these F's is not simply as a growth rate, which is something we could calculate given data on the present and some predictions about the future. Unlike the growth rates, the travel-time factors require us to estimate something.

Gravity Model with Zonal Adjustment Factors

- As an empirical matter, early users of the gravity model observed that it did not appear to yield very accurate predictions.
- In order to improve the predictive accuracy, they introduced a set of *zonal adjustment factors*, to be denoted K_{ij}. When these were incorporated, the transportation gravity model took the form

$$T_{ij} = \frac{O_i A_j F_{ij} K_{ij}}{\sum_m A_m F_{im} K_{im}} \tag{1}$$

- As to the zonal adjustment factors, note that:
  1. They have no basis whatever in theory, unlike the travel-time factors. They are introduced simply to improve predictions.
  2. The matrix $K \equiv \{K_{ij}\}$ is another $Z \times Z = Z^2$ parameters which need to be estimated. It might reasonably be asked whether this will be asking too much of the available data.

Calibration I

- We turn now to the task of calibrating the gravity model of trip distribution in its form (1).
- The standard method for doing this is an iterative method developed by the US Bureau of Public Roads (now the Federal Highway Administration), known as the BPR method.
- As we have seen, the transportation form of the gravity model requires us to estimate two sets of unknown parameters: the travel time factors ($F \equiv \{F_{ij}\}$) and the zonal adjustment factors ($K \equiv \{K_{ij}\}$).

Calibration II

- The underlying idea of the BPR method is to choose these unknowns in order to make the gravity model "work" as well as possible *for our present-day data*. That is, unlike the growth factor model calibrations, the BPR calibration method makes no use whatever of future data.
- Once we have calibrated the model (ie found estimates of the travel-time and zonal adjustment factors, which we will denote by $\hat{F} \equiv \{\hat{F}_{ij}\}$) and $\hat{K} \equiv \{\hat{K}_{ij}\}$ respectively), we will assume that these are stable (ie unchanged) as between our current period and the future period for which we wish to distributed the trips.
- This makes prediction simply a matter of plugging in the data and utilizing our estimated $\hat{F}$ and $\hat{K}$.

BPR Method : Summary I

The BPR method involves the following steps:

0. (optional) reduce the dimensionality of the problem by imposing similarity assumptions on the different F-factors. In our case this will be based on similar interzonal travel times, and we will also require additional data: the current interzonal travel times. These similar sets of F-factors define what we will call superzones..
1. Iterative step: estimate the F's based on superzonal total trips. That is, we use the gravity model formulation to reproduce as closely as we wish, the observed *superzonal* total trips. The result of this step (Step 1) is estimates of the travel-time factors, to be denoted $\hat{F}$.

BPR Method : Summary II

2. At the end of step 1 we have a candidate trip matrix (an "estimate" of our target T^0). However it will usually not satisfy conservation of attractions. So we perform another iterative step, called row-and column factoring, to approximately balance the trip matrix. This is Step 2.
3. At the end of step 2 we have an approximately balanced trip matrix. But it still does not completely reproduce the original T^0. We reproduce T^0 by using the one set of parameters we have so far ignored, namely the zonal adjustment factors (the K matrix). This is Step 3.

BPR Method : Example

- In my view, the simplest way to understand the fairly complicated BPR calibration method is to develop the formal expressions for the calibration steps alongside an actual example.
- So that is what we will do here.
- **Note** : I will post to the course website an abbreviated version of the example, containing just the computations. You may find it helpful to refer to this as you review the details.

Example Data

Trip-interchange matrix:

$$T^0 = \begin{array}{ccc|c} 100 & 350 & 100 & 550 \\ 240 & 150 & 210 & 600 \\ 60 & 120 & 200 & 380 \\ \hline 400 & 620 & 510 & 1530 \end{array}$$

Interzonal travel-times matrix:

$$t^0 = \begin{bmatrix} 1 & 6 & 11 \\ 7 & 3 & 12 \\ 15 & 13 & 4 \end{bmatrix}$$

We shall use a convergence criterion of 5% in all iterative steps ($\alpha_L = 0.95$, $\alpha_H = 1.05$).

Step 0

We will assume that pairs of zones with similar interzonal travel times involve the same travel-time factor.

We operationalize this by assuming that all pairs of zones i and j satisfying

$$0 \leq t_{ij}^{0} < \delta$$

for a given δ in travel time, utilize the same travel-time factor, F_1. And we take all zonal pairs satisfying

$$\delta \leq t_{ij}^{0} < 2\delta$$

to utilize a second travel time factor, F_2. And then

$$2\delta \leq t_{ij}^{0} < 3\delta$$

are assumed to involve a third travel-time factor F_3. And so on.

Step 0

In our example we shall take $\delta = 5$ minutes.

This defines the following superzones:

Superzone	Zonal Pairs
1	$I_1 = \{1, 1\}, \{2, 2\}, \{3, 3\}$
2	$I_2 = \{1, 2\}, \{2, 1\}$
3	$I_3 = \{1, 3\}, \{2, 3\}, \{3, 1\}, \{3, 2\}$

with superzonal trip totals:

$$O_S = [450, 590, 490]$$

Step 0

In our example, we will be calculating 3 distinct travel-time factors. It is a simple matter to parlay these three into a full set F of factors. Formally, we define a mapping (rule) ϕ that takes our 3 estimates of the travel time factors (F_1, F_2, F_3) and produce an estimate of the full F_{ij} matrix. In our case this rule is:

$$\phi(F_1, F_2, F_3) = \begin{bmatrix} F_1 & F_2 & F_3 \\ F_2 & F_1 & F_3 \\ F_3 & F_3 & F_1 \end{bmatrix}.$$

Finally we set

$$\begin{aligned} K &\equiv 1 \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{aligned}$$

(This amounts to ignoring the K's for now, since they enter the gravity model multiplicatively).

Step 1

- In Step 1 we will estimate the distinct travel-time factors F_i, as determined by step 0.
- In this step we work exclusively with superzonal aggregated travel.
- But within that context, the method is very much like the way we calibrated the growth-factor models: we generate an estimate of T, and we compute error factors by comparing actual superzonal total (ie those implied by T^0) with those implied by the current T.
- However, the updating rule, as we will see, is slightly different.

Step 1

An iteration in Step 1 consists of:

1. Use previously computed (or assumed) F's and the gravity model to generate a new T matrix.
2. Compare the superzonal total travel implied by T with the totals implied by T^0
3. If convergence is not satisfied, update the F's and try again.

Step 1, Iteration 1

Initial estimate: $F_i^0 = (1, 1, 1)$. Apply ϕ and get

$$F^0 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Step 1, Iteration 1

Apply the gravity model and find:

$$T^1 = T_{ij}^1 = \frac{O_i^0 A_j^0 F_{ij}^0}{\sum_m A_m^0 F_{im}^0}$$

$$= \begin{array}{ccc|c} 143.791 & 222.876 & 183.333 & 550 \\ 156.863 & 243.137 & 200 & 600 \\ 99.3464 & 153.987 & 126.667 & 380 \\ \hline 400 & 620 & 510 & 1530 \end{array}$$

Step 1, Iteration 1

Convergence check; against the *superzone* totals

Target	450	590	490
Actual	513.595	379.739	636.667
Error ratio E_i^1	0.876177	1.5537	0.769634

Note that the "targets" are the superzonal trip totals. We see that we have not converged.

Step 1, Iteration 2

New superzonal factors F_i^1 are the *previous* F's "corrected" by the error ratios:

$$\begin{aligned} F_i^1 &= F_i^0 \times E_i^1 \\ &= [1.0, 1.0, 1.0] \times [0.876177, 1.5537, 0.769634] \\ &= [0.876177, 1.5537, 0.769634] \end{aligned}$$

Apply ϕ to derive the full set of travel-time factors:

$$F^1 = \begin{bmatrix} 0.876177 & 1.5537 & 0.769634 \\ 1.5537 & 0.876177 & 0.769634 \\ 0.769634 & 0.769634 & 0.876177 \end{bmatrix}$$

Step 1, Iteration 2

$$T^2 = T_{ij}^2 = \frac{O_i^0 A_j^0 F_{ij}^1}{\sum_m A_m^0 F_{im}^1}$$

$$= \begin{array}{ccc|c} 112.97 & 310.507 & 126.522 & 550 \\ 239.457 & 209.307 & 151.236 & 600 \\ 94.9643 & 147.195 & 137.841 & 380 \\ \hline 447.392 & 667.009 & 415.599 & 1530 \end{array}$$

Step 1, Iteration 2

Convergence test:

Target	450	590	490
Actual	460.119	549.964	519.917
Error ratio E_i^2	0.978009	1.0728	0.942458

Still no convergence

Step 1, Iteration 3

New interzonal F_i :

$$\begin{aligned} F_i^2 &= F_i^1 \times E_i^2 \\ &= [0.876177, 1.5537, 0.769634] \\ &\quad \times [0.978009, 1.0728, 0.942458] \\ &= [0.856\,909,\ 1.\,666\,8, 0.725\,347] \end{aligned}$$

Apply ϕ :

$$F^2 = \begin{bmatrix} 0.856909 & 1.6668 & 0.725347 \\ 1.6668 & 0.856909 & 0.725347 \\ 0.725347 & 0.725347 & 0.856909 \end{bmatrix}$$

Step 1, Iteration 3

Apply the gravity model:

$$T^3 = \begin{array}{ccc|c} 107.966 & 325.512 & 116.522 & 550 \\ 255.134 & 203.306 & 141.56 & 600 \\ 93.6824 & 145.208 & 141.11 & 380 \\ \hline 456.782 & 674.026 & 399.191 & 1530 \end{array}$$

Step 1, Iteration 3

Convergence check:

Target	450	590	490
Actual	452.381	580.647	496.972
Error ratio	0.994736	1.01611	0.985971

These have all converged.

Step 1, Final Result

At the end of Step 1 we have our final estimate of the travel-time factors $\hat{F} \equiv \hat{F}_{ij}$, which is the F matrix we obtained at the convergent iteration of our step-1 computations, ie

$$\hat{F} = \begin{bmatrix} 0.856909 & 1.6668 & 0.725347 \\ 1.6668 & 0.856909 & 0.725347 \\ 0.725347 & 0.725347 & 0.856909 \end{bmatrix}$$

Step 2

- At the end of step 1 we have estimated the F's.
- We also have, associated with those F's (and under the assumption that the K's are all 1) an estimate of the trip matrix.
- As compared with T^0 (the trip matrix we are trying to reproduce) the current estimate T satisfies conservation of origins (automatically) but:
  - does not satisfy conservation of attractions
  - does not reproduce the individual elements of T^0.
- Step 2 is an attempt to remedy the first of the two problems (conservation of attractions).

Step 2

If we need to do step 2 (ie if we don't satisfy conservation of origins even approximately at the end of Step 1) then an iteration of Step 2 consists of:

1. A column factoring : this adjusts the individual elements to ensure conservation of attractions. But in doing this, we destroy conservation of origins; so
2. We perform a row factoring. This restores conservation of origins, but destroys conservation of attractions.

At the end of a Step-2 iteration we check conservation of attractions. If it is not approximately satisfied, then we perform another iteration; and we continue until conservation of attractions is approximately satisfied.

Step 2, Setup I

Continuing our example, based on $\hat{F}$ fron Step 1 and $K \equiv 1$ we have an "estimate" of the travel-time matrix we're trying to reproduce, namely the final T matrix from the previous step:

$$T^3 = \begin{array}{ccc|c} 107.966 & 325.512 & 116.522 & 550 \\ 255.134 & 203.306 & 141.56 & 600 \\ 93.6824 & 145.208 & 141.11 & 380 \\ \hline 456.782 & 674.026 & 399.191 & 1530 \end{array}$$

Step 2, Setup II

Is this good enough? We need only check the columns (conservation of attractions) for convergence:

Target	400	620	510
Actual	456.782	674.026	399.191
Error ratio	0.875691	0.919845	1.27758

Note that we check against the actual A^0 — we have no more use for the superzones.

Clearly not good enough. So we need to do the balancing step.

Step 2, Setup III

To simplify the notation, let's rename our starting point (T^3) to be T^1.

At each iteration of step 2 (given that we need to do it at all, as we do in this example) we must do:

1. a column factoring
2. a row factoring
3. a convergence check on the columns.

Step 2, Iteration 1, Column Factoring

The column factors are the previous error ratios, so

$$0.875691, 0.919845, 1.27758$$

We multiply the elements in each column by its column factor giving

$$T^{1a} = \begin{array}{ccc|c} 94.5446 & 299.421 & 148.866 & 542.832 \\ 223.419 & 187.01 & 180.854 & 591.283 \\ 82.0368 & 133.569 & 180.279 & 395.885 \\ \hline 400 & 620 & 510 & 1530 \end{array}$$

Example: the $(1,2)$ element of T^1 $(= 325.512)$ is in column 2, so we use the second column factor, giving $325.512 \times 0.919845 = 299.421 = T^{1a}_{12}$.

Note that this satisfies conservation of attractions, but no longer satisfies conservation of origins.

Step 2, Iteration 1, Row Factoring

The row factors are the row-wise “error factors” of T^{1a} :

Target	550	600	380
Actual	542.832	591.283	395.885
Error ratio	1.0132	1.01474	0.959875

and multiplying each row by its row factor gives

$$T^{1b} = \begin{array}{ccc|c} 95.793 & 303.375 & 150.832 & 550 \\ 226.712 & 189.767 & 183.521 & 600 \\ 78.7451 & 128.209 & 173.046 & 380 \\ \hline 401.25 & 621.351 & 507.398 & 1530 \end{array}$$

Example: the $(1,2)$ element of T^{1a}, $(= 299.421)$ is in row 1, so we use the first row factor, giving $299.421 \times 1.0132 = 303.375 = T^{1b}_{12}$.

Step 2, Iteration 1, Convergence Check

We need to check the columns only, since the row factoring assures conservation of origins.

Target	400	620	510
Actual	401.25	621.351	507.398
Error ratio	0.996884	0.997825	1.00513

We have converged.

Step 3

We now have an approximately balanced T matrix (ie T^{1b}) derived on the hypothesis that $K \equiv 1$, namely

$$T^{1b} = \begin{array}{ccc|c} 95.793 & 303.375 & 150.832 & 550 \\ 226.712 & 189.767 & 183.521 & 600 \\ 78.7451 & 128.209 & 173.046 & 380 \\ \hline 401.25 & 621.351 & 507.398 & 1530 \end{array}$$

This "estimate" satisfies conservation of origins, satisfies conservation of attractions to within our convergence criterion, but still does not reproduce the individual elements of T^0.

Step 3

In order to reproduce the individual elements of T^0 we use the zonal adjustment factors, which we have ignored up to now. The zonal adjustment (K) factors are now computed as

$$\hat{K}_{ij} = T^0_{ij} \div T^{1b}_{ij}$$

(ie on the last (convergent) T matrix from step 2) and where $\div$ means element-by-element division . Then

$$\begin{aligned}
\hat{K} &= \begin{bmatrix} 100 & 350 & 100 \\ 240 & 150 & 210 \\ 60 & 120 & 200 \end{bmatrix} \div \begin{bmatrix} 95.793 & 303.375 & 150.832 \\ 226.712 & 189.767 & 183.521 \\ 78.7451 & 128.209 & 173.046 \end{bmatrix} \\
&= \begin{bmatrix} 1.04392 & 1.15369 & 0.662989 \\ 1.05861 & 0.790443 & 1.14429 \\ 0.761952 & 0.93597 & 1.15576 \end{bmatrix}
\end{aligned}$$

BPR Method: Summary of Example Results

For our calibration of the gravity model using the BPR method we have found:

$$\hat{F} = \begin{bmatrix} 0.856909 & 1.6668 & 0.725347 \\ 1.6668 & 0.856909 & 0.725347 \\ 0.725347 & 0.725347 & 0.856909 \end{bmatrix}$$

(from the final result of Step 1); and

$$\hat{K} = \begin{bmatrix} 1.04392 & 1.15369 & 0.662989 \\ 1.05861 & 0.790443 & 1.14429 \\ 0.761952 & 0.93597 & 1.15576 \end{bmatrix}$$

(from Step 3).

Discussion : Step 2

- Remember that the motivation for the BPR method is to construct the F and K factors to allow us to exactly reproduce our observed trip matrix T^0.
- However, a quick calculation shows that we have not succeeded in this: we find, plugging in our calibrated results:

$$T_{ij}^* = \frac{O_i^* A_j^* \hat{F}_{ij} \hat{K}_{ij}}{\sum_m A_m^* \hat{F}_{im} \hat{K}_{im}} = \begin{array}{ccc|c} 109.62 & 365.25 & 75.14 & 550 \\ 273.38 & 162.66 & 163.96 & 600 \\ 73.24 & 139.44 & 167.33 & 380 \\ \hline 456.23 & 667.35 & 406.42 & 1530 \end{array}$$

which does not reproduce our original observed trip matrix.

Discussion : Step 2

- If you think about this for a minute you realize that the problem seems to lie in our Step 2, the row-and-column factoring.
- For one thing, this step has nothing to do with the gravity model at all: it's just an ad-hoc way of arriving at a balanced interim trip matrix.
- For another, you might reasonably say, why do it at all? If we can exactly reproduce T^0 by doing Step 1 and then adjusting the results via the zonal adjustment factors in Step 3, won't we thereby automatically have a balanced trip matrix?
- It seems to me that the answer to this is Yes; suggesting that step 2 is dispensable.

Calibration Without Step 2

At the end of Step 1 we had an interim estimate of T^0 as

$$T^3 = \begin{array}{ccc|c} 107.966 & 325.512 & 116.522 & 550 \\ 255.134 & 203.306 & 141.56 & 600 \\ 93.6824 & 145.208 & 141.11 & 380 \\ \hline 456.782 & 674.026 & 399.191 & 1530 \end{array}$$

implying that we could take

$$\begin{aligned} \hat{K} &= \begin{bmatrix} 100 & 350 & 100 \\ 240 & 150 & 210 \\ 60 & 120 & 200 \end{bmatrix} \div \begin{bmatrix} 107.966 & 325.512 & 116.522 \\ 255.134 & 203.306 & 141.56 \\ 93.6824 & 145.208 & 141.11 \\ \end{bmatrix} \\ &= \begin{bmatrix} 0.92622 & 1.07523 & 0.858208 \\ 0.940682 & 0.737804 & 1.48347 \\ 0.640461 & 0.826402 & 1.41734 \end{bmatrix} \end{aligned}$$

Calibration Without Step 2

And a simple calculation shows that if we plug *these* estimates into the gravity model:

$$
\begin{aligned}
\hat{F} &= \begin{bmatrix} 0.856909 & 1.6668 & 0.725347 \\ 1.6668 & 0.856909 & 0.725347 \\ 0.725347 & 0.725347 & 0.856909 \end{bmatrix} \\
\hat{K} &= \begin{bmatrix} 0.92622 & 1.07523 & 0.858208 \\ 0.940682 & 0.737804 & 1.48347 \\ 0.640461 & 0.826402 & 1.41734 \end{bmatrix}
\end{aligned}
$$

we do indeed exactly reproduce our original trip matrix T^0.

Projection I

- Projection in the case of a calibrated gravity model is simply a matter of using predicted future data (here, the origins O^* and the attractions A^* and possibly future interzonal travel times if we believe that these are changing) and plugging them into the gravity model, since we are assuming that the $\hat{F}$ and $\hat{k}$ matrices are stable over time. That is,we take

$$T^*_{ij} = \frac{O^*_i A^*_j \hat{F}_{ij} \hat{K}_{ij}}{\sum_m A^*_m \hat{F}_{im} \hat{K}_{im}}$$

Projection II

- There *is* a case for row-and-column factoring (Step 2) when we are doing a projection, since we will usually get an unbalanced result, ie a result that is in a sense inconsistent with the predicted attractions (A^*). (Remember that conservation of origins is guaranteed).
- For consistency, then, it makes sense to perform an artificial balancing step on the results even though we know that this is not strictly mandated by the gravity model itself. That is, we first compute T^* according to the above formula, and then we go through Step 2 (row-and-column factoring) to arrive at an approximately balanced T^{**}. We report T^{**} as our final prediction.