# The Average Growth Factor Model

Philip A. Viton

October 12, 2012

# Introduction

- This note discuses the Average Growth Factor model of trip distribution. The interest of this model lies primarily in two ideas that will figure again in our discussion of the Gravity Model:
    1. The idea of checking how well we've done in our task of generating a predicted $T$ matrix.
    2. The idea of improving our results through iteration
- Data required: $T^0$ (observed present-day trip matrix), plus estimates $O^* = (O_1^*, O_2^*, \ldots, O_Z^*)$ of all future originations by zone. Note that this is more data than was needed for the uniform growth model.
- The Average Growth Factor model is based on idea that we expect that the growth in interzonal travel between any two zones is related to the growth in travel in each zone individually.

# The Model

- Specifically, define the individual zonal growth rates for any two zones $i$ and $j$ as:

$$
\begin{aligned}
F_i^0 &= O_i^* / O_i^0 \\
F_j^0 &= O_j^* / O_j^0
\end{aligned}
$$

- Then we assume that the growth rate in travel *between* the two zones is equal to the average of the two zonal growth rates:

$$F_{ij}^0 = \frac{F_i^0 + F_j^0}{2}$$

- And that an estimate of the future interzonal travel between zones $i$ and $j$ is given by

$$T_{ij}^1 = T_{ij}^0 \times F_{ij}^0$$

# Conservation of Origins

- Suppose we generate a new trip matrix $T^1$ according to this scheme. How can we assess our answer?
- Since we don't know anything about the "true" future zone-to-zone $T_{ij}$'s, the only way to do this would seem to be via the originations: we compare the originations implied by $T^1$ with the *known* future originations $O^*$.
- In the jargon, we are asking if our estimate $T^1$ satisfies *conservation of origins.* That is, do the originations implied by $T^1$ add up (to those we "know" ie to the vector $O^*$).

# Error Ratios I

- Since we are interested in checking $T^1$ for conservation of origins, this suggests that we look at the "error ratios"

$$E_i^1 = \frac{O_i^*}{O_i^1} \qquad i = 1, 2, \ldots, Z$$

- But how do we assess these? Clearly, the best possible answer would be to find $E_i^1 = 1$ for all $i$; but this is probably too much to hope for in practice. So we should probably attempt to say when a particular set of originations (here, the vector $O_i^1$) is good enough.

# Error Ratios II

- Let's try to formalize the idea of being "good enough".
- We will do so by defining two *convergence criteria* $\alpha_L, \alpha_H$. We will say that a particular $E_i^1$ is "good enough" if

$$\alpha_L \leq E_i^1 \leq \alpha_H$$

  and that our $T^1$ is an acceptable answer if

$$\alpha_L \leq E_i^1 \leq \alpha_H \qquad i = 1, 2, \ldots, Z$$

- Note that this is a form of *approximate* satisfaction of conservation of origins.

# Error Ratios III

- But where do these convergence criteria come from?
- The short answer is that that's up to you: you need to decide just how accurate you want your answers to be.
- It is important to understand that "accurate" refers only to whether the originations implied by our candidate trip matrix ($T^1$) agree with the known/assumed future originations: there is no question of comparing the individual elements of the candidate trip matrix with anything, since there's nothing to compare them to.
- In practice, many analysts use one of two sets of convergence criteria:

$$
\begin{aligned}
(\alpha_L, \alpha_H) &= (0.95, 1.05) \quad \text{called a 5\% criterion} \\
(\alpha_L, \alpha_H) &= (0.99, 1.01) \quad \text{called a 1\% criterion}
\end{aligned}
$$

# Iteration I

- Suppose we've done all this: computed the average growth factors and generated $T^1$.
- And suppose we find that the condition

$$\alpha_L \leq E_i^1 \leq \alpha_H \qquad \text{for all } i = 1, 2, \ldots, Z$$

is *not* satisfied for all zones $i$ (it may be satisfied for some of them).
- What can we do about it? The key here is to recognize that having computed $T^1$, and found that at least one of the error factors $E_i^1$ is outside our convergence bounds, we have not exhausted our information.
- Note that this was not true of the Uniform model: there we were guaranteed that (in our new notation) $S(T^*) = S(T^1)$, so in this case our calibration attempt *did* exhaust the information.
  This would still be true if we also had zone-by-zone predictions of the originations, since the Uniform Growth Model would ignore that detail, and just add up the origins to predict total trip making.

# Iteration II

- So at the insight is to observe that having generated $T^1$ and observing that we do not satisfy conservation of origins, we recognize that $T^1$ can be considered as a new starting point on our way to a prediction, just as $T^0$ was.
- And the suggestion is that we simply try again, using $T^1$ as our new starting point.

# Iteration III

So what we are proposing goes like this:

1. Based on $T^0$ and $O^*$, compute the zonal growth factors $F_i^0 = O_i^* / O_i^0$ (for each $i$), and then the average growth factors matrix $F_{ij}^0 = (F_i^0 + F_j^0)/2$.
2. Generate a new candidate future trip matrix $T_{ij}^1 = T_{ij}^0 \times F_{ij}^0$.
3. Compute the error ratios $E_i^1 = O_i^* / O_i^1$. Check to see whether they satisfy $\alpha_L \leq E_i^1 \leq \alpha_H, \quad$ for all $i = 1, 2, \ldots, Z$.
4. If they do, then we are done: we take $T^* = T^1$.
5. If not, then start again with $T^1$ as our base.
6. Continue until we have converged.

# Iteration Summary

We write the model, at the stage (iteration) where we are going to be computing the $k+1$-th trip matrix $T^{k+1}$ using as our base trip matrix the previously computed $T^k$, as

$$
\begin{aligned}
F_i^k &= E_i^k = O_i^* / O_i^k && \text{(the zonal growth factors)} \\
F_{ij}^k &= \frac{F_i^k + F_j^k}{2} && \text{(the average growth rates matrix)} \\
T_{ij}^{k+1} &= T_{ij}^k \times F_{ij}^k
\end{aligned}
$$

and we compute $F_{ij}^k$ and the next iteration $T_{ij}^{k+1}$ just in case the condition $\alpha_L \le E_i^k \le \alpha_H$ does *not* hold for at least one $i$.

(We sometimes write the iteration step as $T^{k+1} = T^k \otimes F^k$, where $\otimes$ is term-by-term matrix multiplication, not the matrix product, also known as the Hadamard product).

# Illustrative Data Example

We turn now to a step-by-step illustration of the calibration procedure. The data needed is:

- A base (observed) trip matrix (just as for the uniform growth factor model). For illustration we assume:

$$T_{ij}^0 = \begin{array}{cccc|c} 25 & 12 & 10 & 18 & 65 \\ 10 & 30 & 14 & 6 & 60 \\ 8 & 12 & 27 & 14 & 61 \\ 6 & 13 & 17 & 32 & 68 \\ \hline 49 & 67 & 68 & 70 & 254 \end{array}$$

- Predicted originations by zone: we assume

$$O_i^* = [75, 45, 80, 95]$$

- Convergence: we shall assume $(\alpha_L, \alpha_H) = (0.95, 1.05)$.

# Iteration 1 — Growth Factors

Zonal growth factors:

$$\begin{aligned} F_i^0 &= O_i^* / O_i^0 \\ &= [75, 45, 80, 95] \div [65, 60, 61, 68] \\ &= [1.15385, 0.75000, 1.31148, 1.39706] \end{aligned}$$

Average growth factors matrix:

$$F_{ij}^0 = \begin{bmatrix} 1.15385 & 0.95192 & 1.23266 & 1.27545 \\ 0.95192 & 0.75000 & 1.03074 & 1.07353 \\ 1.23266 & 1.03074 & 1.31148 & 1.35427 \\ 1.27545 & 1.07353 & 1.35427 & 1.39706 \end{bmatrix}$$

Example: $F_2^0 = 0.75000$, $F_3^0 = 1.31148$, so $F_{23}^0 = (0.75000 + 1.31148)/2 = 1.030\,74$.

# Iteration 1 — New Trip Matrix

$$T_{ij}^1 = T_{ij}^0 \times F_{ij}^0$$

$=$

| | | | | |
|---|---|---|---|---|
| 28.8462 | 11.4231 | 12.3266 | 22.9581 | 75.5540 |
| 9.5192 | 22.5000 | 14.4303 | 6.4412 | 52.8907 |
| 9.86129 | 12.3689 | 35.4098 | 18.9597 | 76.5997 |
| 7.65271 | 13.9559 | 23.0225 | 44.7059 | 89.3370 |
| 55.8794 | 60.2478 | 85.1893 | 93.0649 | 294.3810 |

Example: $T_{23}^1 = F_{23}^0 \times T_{23}^0 = 1.03074 \times 14 = 14.4303.$

# Iteration 1 —- Convergence Check

| | | | | |
|---|---|---|---|---|
| Target, $O_i^*$ | 75 | 45 | 80 | 95 |
| Actual, $O_i^1$ | 75.554 | 52.8907 | 76.5997 | 89.337 |
| Error ratio $E_i^1$ | 0.99267 | 0.85081 | 1.04439 | 1.06339 |

Since these are not all within our convergence bounds, we have not converged.

# Iteration 2 — Growth Rates

The new zonal growth factors are the error ratios from the convergence check:

$$F_i^1 = E_i^1 = [0.992668, 0.850811, 1.04439, 1.06339]$$

New average growth factor matrix:

$$F_{ij}^1 = \begin{bmatrix} 0.99267 & 0.92174 & 1.01853 & 1.02803 \\ 0.92174 & 0.85081 & 0.94760 & 0.95710 \\ 1.01853 & 0.94760 & 1.04439 & 1.05389 \\ 1.02803 & 0.95710 & 1.05389 & 1.06339 \end{bmatrix}$$

# Iteration 2 — New Trip Matrix

$$T_{ij}^2 = T_{ij}^1 \times F_{ij}^1$$

| | | | | |
|---|---|---|---|---|
| 28.6346 | 10.5291 | 12.5550 | 23.6016 | 75.3204 |
| 8.7743 | 19.1432 | 13.6742 | 6.16485 | 47.7565 |
| 10.0440 | 11.7207 | 36.9817 | 19.9815 | 78.7279 |
| 7.86721 | 13.3572 | 24.2632 | 47.5397 | 93.0273 |
| 55.3201 | 54.7502 | 87.4741 | 97.2877 | 294.8320 |

# Iteration 2 — Convergence Check

| | | | | |
|---|---|---|---|---|
| Target, $O_i^*$ | 75 | 45 | 80 | 95 |
| Actual, $O_i^2$ | 75.3204 | 47.7565 | 78.7279 | 93.0273 |
| Error ratio, $E_i^2$ | 0.99575 | 0.94228 | 1.01616 | 1.02121 |

Since $E_2^2$ is (just) outside our bound (0.95) we have not yet converged.

# Iteration 3 — Growth Factors

Zonal growth factors:

$$F_i^2 = E_i^2 = [0.99575, 0.94228, 1.01616, 1.02121]$$

Average growth factors matrix:

$$F_{ij}^2 = \begin{bmatrix} 0.99575 & 0.96901 & 1.00595 & 1.00848 \\ 0.96901 & 0.94228 & 0.97922 & 0.98174 \\ 1.00595 & 0.97922 & 1.01616 & 1.01868 \\ 1.00848 & 0.98174 & 1.01868 & 1.02121 \end{bmatrix}$$

# Iteration 3 — New Trip Matrix

$$T_{ij}^3 = T_{ij}^2 \times F_{ij}^2$$

| | | | | |
|---|---|---|---|---|
| 28.5128 | 10.2028 | 12.6297 | 23.8017 | 75.1471 |
| 8.50236 | 18.0383 | 13.3900 | 6.0523 | 45.9830 |
| 10.1038 | 11.4772 | 37.5792 | 20.3548 | 79.5150 |
| 7.9339 | 13.1133 | 24.7165 | 48.5478 | 94.3115 |
| 55.0529 | 52.8316 | 88.3155 | 98.7566 | 294.9570 |

# Iteration 3 — Convergence Check

| | | | | |
|---|---|---|---|---|
| Target, $O_i^*$ | 75 | 45 | 80 | 95 |
| Actual, $O_i^3$ | 75.1471 | 45.983 | 79.515 | 94.3115 |
| Error ratio, $E_i^3$ | 0.998043 | 0.978623 | 1.00610 | 1.00730 |

Since for all of these we have $\alpha_L \leq E_i^3 \leq \alpha_H$, we conclude that we have converged.

# Final Result

Our final prediction for the distributed trips is therefore:

$$T^*_{ij} = T^3_{ij} = \begin{array}{cccc|c} 28.5128 & 10.2028 & 12.6297 & 23.8017 & 75.1471 \\ 8.50236 & 18.0383 & 13.3900 & 6.05229 & 45.9830 \\ 10.1038 & 11.4772 & 37.5792 & 20.3548 & 79.5150 \\ 7.9339 & 13.1133 & 24.7165 & 48.5478 & 94.3115 \\ \hline 55.0529 & 52.8316 & 88.3155 & 98.7566 & 294.9570 \end{array}$$

Note that this satisfies conservation of total trips, but it satisfies conservation of origins only approximately. We could of course continue to iterate, and this would improve convergence (provide a better approximation to conservation of origins).