

## **P/A and O-D Matrices**

### **Integrating Surveyed and Synthesised Data for Demand Modelling**

Stephen Marsh  
Jacobs Consultancy  
9<sup>th</sup> February 2007

IC Seminar: The Challenges for Today's Transport Modellers

### **Background**

- **Jacobs Consultancy commissioned by the Tyne and Wear Authorities (T&W) to develop a new generation transport planning model**
- **Later commissioned by T&W and separately by Durham County Council to support their TIF applications**
- **This work is being done in association with John Bates and Miles Logie**

## Structure of Talk

- Introduction and Objectives
- The problems of good P/A Base Matrices
- Our solution
- Creation of Synthetic Matrices
- Incorporating local survey data
- Success of the T&W Base Matrix development
- Questions?

**JACOBS**  
CONSULTANCY

## Key Transport Modelling/Appraisal Ingredients

- 1. Reliable Base Matrices**
2. Appropriate Demand Segmentation
3. Validated Highway and PT Networks
4. Incremental Demand Model
5. Allowing for changes over time

**JACOBS**  
CONSULTANCY

## Objectives

- **Build and test transport modelling tools consistent with TIF guidance**
- **Develop innovative techniques that can be transferred to other studies**
- **Building base matrices using a combination of synthetic and observed data**

**JACOBS**  
CONSULTANCY

## Importance of Building Good P/A Base Matrices

- **Base year matrices are produced as part of “four-stage” model**
- **Incremental demand models pivot off base matrices**
- **P/A Base Matrices are the only rational form for demand modelling**
- **Poor base demand will lead to poor model application in forecast years**
- **Recommended in DfT guidance**

**JACOBS**  
CONSULTANCY

### Issues relating to P/A Base Matrices

- How to construct validated base matrices using a variety of data sources
- The widespread lack of emphasis on demand modelling
- Straightforward estimation of 'matrices from counts'
- But the information gained in calibrating this O-D matrix is not put back to the P/A
- Some models only use the P/A demand model to adjust a validated O-D matrix

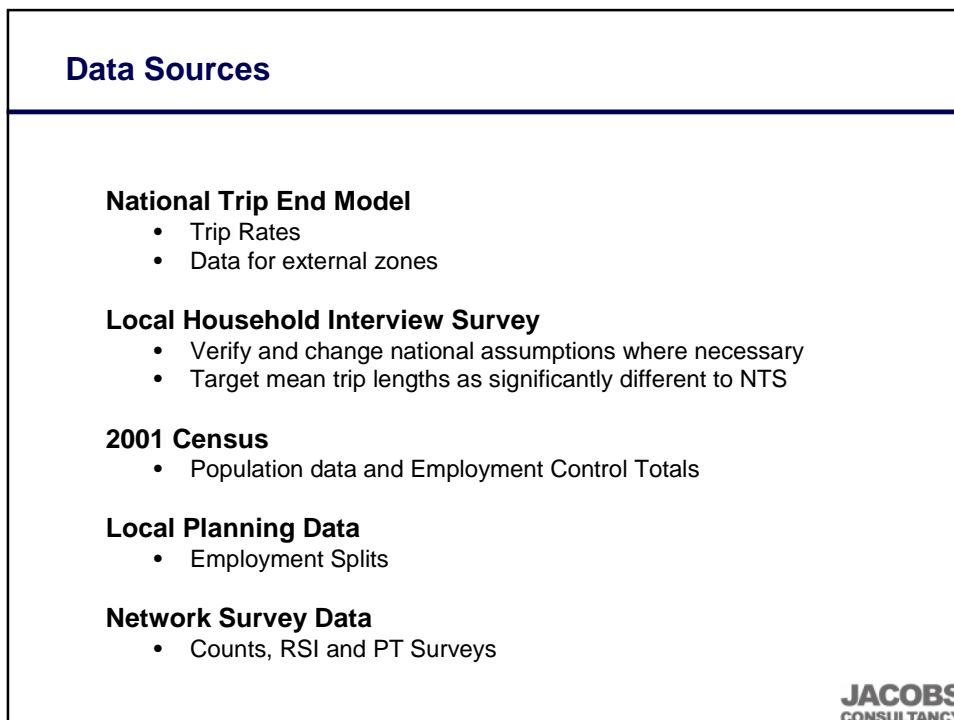
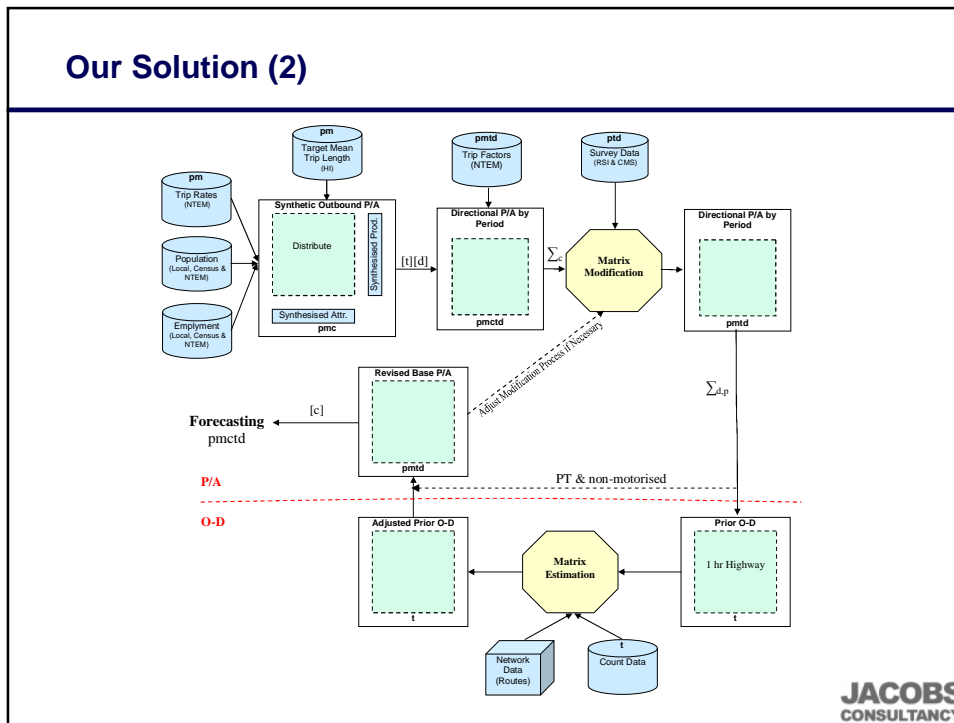
**JACOBS**  
CONSULTANCY

### Our Solution (1)

Our challenge was to develop a practical methodology that:

- Produces consistent P/A and O-D base matrices
- Exploits the strengths and minimises the weaknesses of synthetic and observed data
- Uses statistical confidence estimates that reflect survey data errors
- Any adjustment gained from O-D 'matrix estimation' can be passed on to the P/A matrices

**JACOBS**  
CONSULTANCY



## Synthetic Matrix Creation (1)

### 1. Production Trip Ends:

- Get local estimates of Populations for each zone in study area
- Use National Trip End Model (NTEM) to populate external zones
- Generate detailed car availability estimates using NATCOP
- Generate trip ends using NTS trip rates (NTEM) and modal factors

### 2. Attraction Trip Ends:

- Get local estimates of Employment for each zone in study area
- Use National Trip End Model (NTEM) to populate external zones
- Generate trip ends using NTS trip rates (NTEM) and modal factors

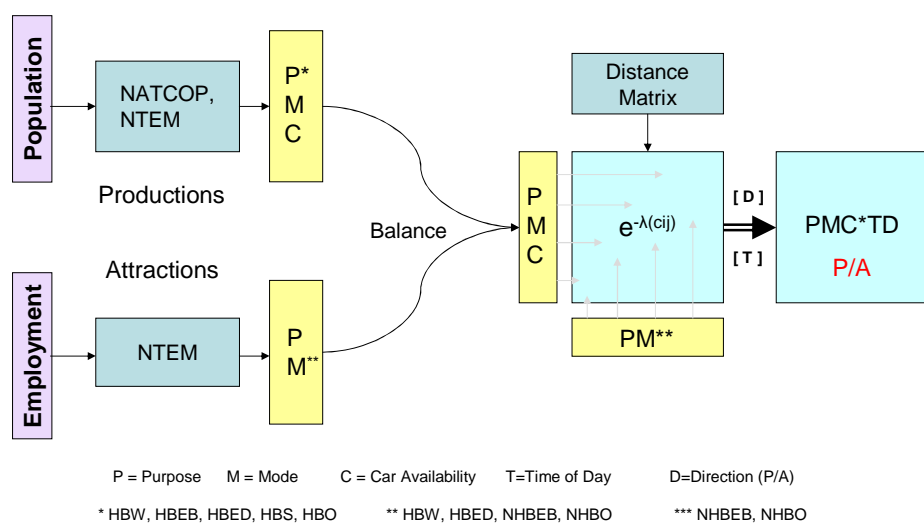
### 3. Distribute using negative exponential:

- Use modelled network distances to distribute over
- Use local Household Interview data for target mean trip lengths

➤ **P/A Synthetic Matrices by Purpose (9), Mode (3), Time of Day (4), Car Availability (4) and Direction (2)**

**JACOBS**  
CONSULTANCY

## Synthetic Matrix Creation (2)



**JACOBS**  
CONSULTANCY

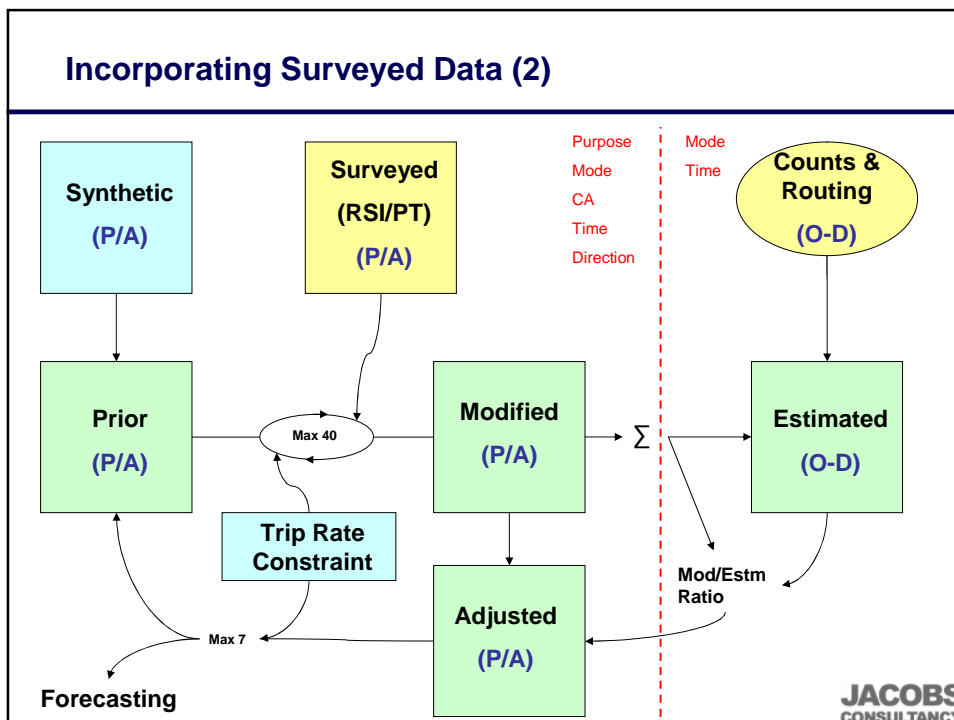
## Incorporating Surveyed Data (1)

**Distribution models are generally poor so Local Survey data is used to improve the distribution modelling**

**There are two types of local survey data used:**

- 1. P/A Intercept Data**
  - Matrix Modification
  - Confronts Synthetic Data with Cell Observation
  - Operates on a cell basis
  - Uses statistical confidence bounds to combine data sets
  - Separately for each Purpose, Mode, Time and Direction
- 2. O-D Count Data & Routing Info**
  - Matrix Estimation
  - by time period (Highway only for T&W)
  - Changes transferred back to P/A matrix

**JACOBS**  
CONSULTANCY



### Success of the T&W Base Matrix Development (1)

- **Practical application of methodology has proved a success**
- **Creation of validated Base Matrices completed**
- **Matrices are being used in supply model calibration**
- **Results of supply model validation look promising**

**JACOBS**  
CONSULTANCY

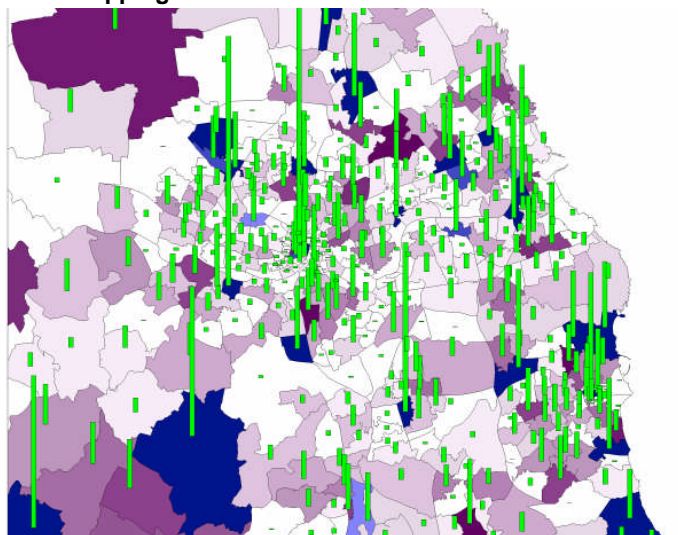
### Matrix Validation/Sense Checking

- **Input Data**
  - Counts & Intercept Surveys
  - Population data & Employment data
- **Procedural Checks**
  - Trip Ends vs TEMPRO and 2001 Census
  - Mode, Time of Day and Car Availability
  - Base Matrices vs Synthetic Matrices and Trip Ends
  - Productions & Attractions vs Populations and Employment
- **Reality Checks**
  - Desire Lines between Sectors by Purpose and Mode
  - Sector Matrices & Cross-Tyne Movements by Mode
  - Sector Matrices for Mode, TOD and Car Availability Splits
  - Select Link Flows across Tyne and to City Centres
  - Trip Length Distributions by Mode and Purpose
  - Screenline Comparisons

**JACOBS**  
CONSULTANCY

### Success of the T&W Base Matrix Development (2)

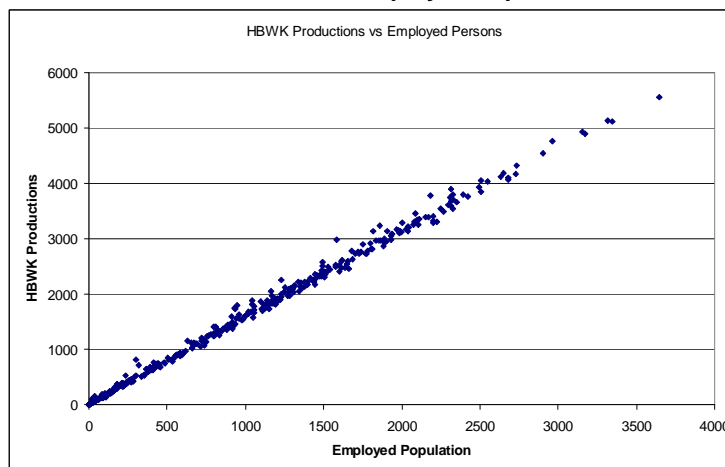
Home Based Shopping Attractions vs Retail Jobs



JACOBS  
CONSULTANCY

### Success of the T&W Base Matrix Development (3)

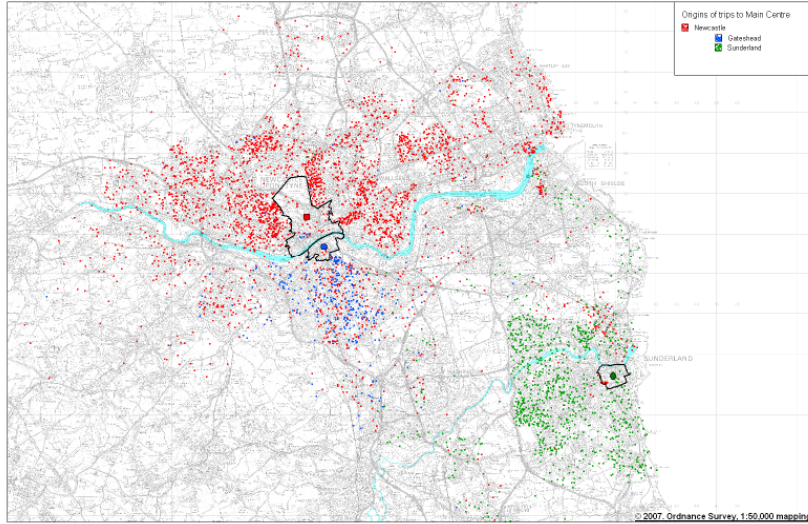
Home Based Work Productions vs Employed Population



JACOBS  
CONSULTANCY

### Success of the T&W Base Matrix Development (4)

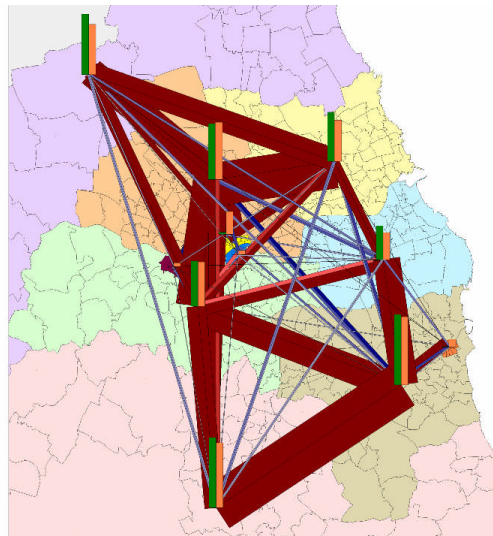
Origins of trips to Main Centres - Home Based Work (CAR) 8-9 am



**JACOBS**  
CONSULTANCY

### Success of the T&W Base Matrix Development (5)

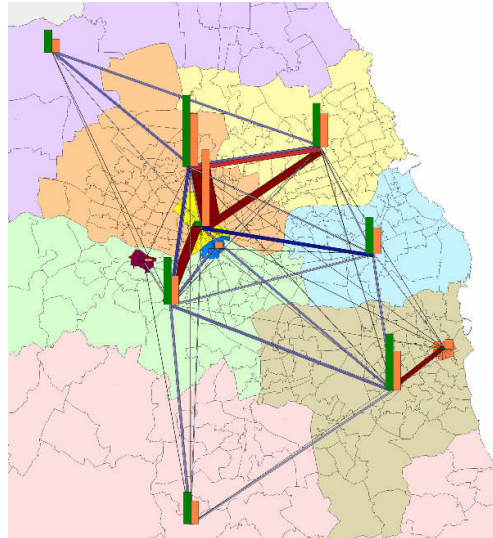
Home Based Work  
24hr Car Trips



**JACOBS**  
CONSULTANCY

### Success of the T&W Base Matrix Development (6)

Home Based Work  
24hr Public  
Transport Trips

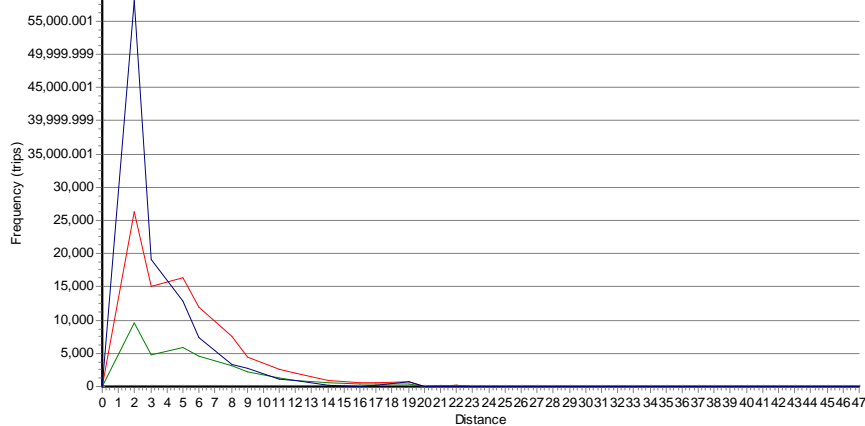


JACOBS  
CONSULTANCY

### Success of the T&W Base Matrix Development (7)

home\_education trips for MP Average Trip Lengths (km):

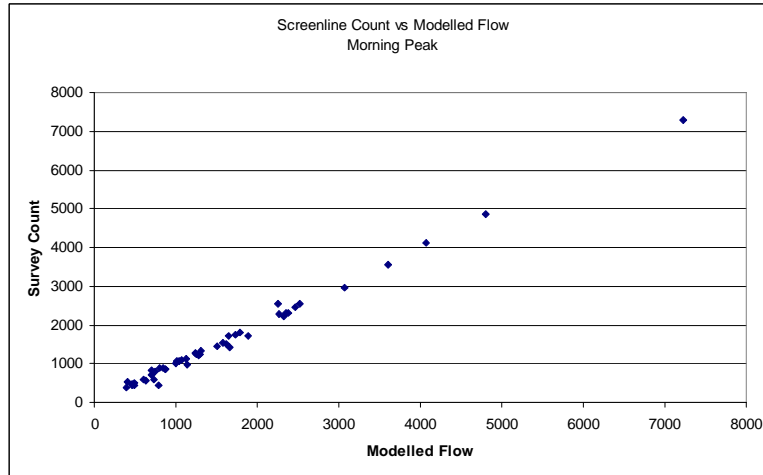
Vehicle\_Person (Ave Trip Length = 4.21)    Public\_Transport (Ave Trip Length = 4.47)  
Slow (Ave Trip Length = 2.18)



JACOBS  
CONSULTANCY

## Success of the T&W Base Matrix Development (8)

### Screenline Comparison – Morning Peak



**JACOBS**  
CONSULTANCY

Thank You

**JACOBS**  
CONSULTANCY

**P/A and O-D Matrices**  
**Integrating Surveyed and Synthesised Data for Demand Modelling**

ICE Seminar: The Challenges for Today's Transport Modellers

Stephen Marsh  
Jacobs Consultancy  
9<sup>th</sup> February 2007-01-31

**Slide** **Script**

1. My name is Stephen Marsh and I am a Senior Consultant with Jacobs Consultancy in London. We have been working on a method to, as the title suggests, integrate local survey data with Synthetic Data, to create validated Production/Attraction base matrices.
2. The background to this work is that in 2004 Jacobs Consultancy were commissioned by the Tyne and Wear Authorities to develop a state of the art transport planning model. The model specification was amended in 2005/2006 to take account of the introduction of TIF funding and guidance, for which this model will form a key component of the Tyne and Wear and separately the Durham County Council TIF Bids. We have been supported in this work by John Bates and Miles Logie using OmniTRANS as the platform for most of the components.
3. So, today's talk will be structured as follows: firstly, I'll give a broader introduction to the topic and the objectives of the work; second, I'll talk about the importance of creating P/A matrices for use in supply modelling, but also the issues in doing so; next, I'll outline the solution that we have developed to create validated P/A base matrices and then give some details on the procedures developed to create a set of synthetic matrices and how we combine these with local surveyed data; and, finally, I'll give some details of the success we have had in creating base matrices for the Tyne and Wear model.
4. This slide shows the main ingredients needed in transport modelling and appraisal: Base Matrices; Appropriate Segmentation; Validated Networks; a Demand Model; and, allowing for changes over time. I won't go through 2 to 5 in any detail, but it highlights the importance of creating reliable base matrices for the success of any transport modelling and appraisal work. The remainder of the talk concentrates on the creation of base matrices.
5. This slide shows some of the objectives of the work. We were originally commissioned by the Tyne and Wear Authorities (T&W) to build a model to assist their LTP submissions. We then developed this model to assist in TIF applications.
6. So why is it so important to create good P/A base matrices as opposed to the more common O-D form?  
Base Matrices are the foundation of all incremental demand models. Incremental models pivot off the base matrices to create forecast demand which is then used in policy testing. So, the base matrices have to accurately represent the demand in the base, not just provide good observed count validation. As John has talked about, the matrices have to be P/A format, rather than origin – destination as this is the only rational basis to pivot from. Incremental models are widely used and are generally recommended by the DfT WebTag guidance and this means creating P/A matrices

for the base year. If the base year demand is poor or inconsistent with forecasting methods then any model application in forecast years will also be poor. As it is important to have the base matrices in P/A format it is important to combine synthetic and observed data, as is regularly done, but, crucially, maintain the P/A relationship. So why has this not been done in every study?

7. Firstly, combining data of different types from different sources is a significant problem. And then being able to store, process and analyse all the data that is needed to construct P/A Base Matrices is another problem – in the creation of the Tyne and Wear Base Matrices we have generated over 30,000 individual matrices as well as many other datasets of different types.  
Secondly, there has been a general lack of emphasis on demand modelling in recent decades. Instead the emphasis has been on Supply side models, especially highway models, and often for the morning peak only.  
Third, there exist many software packages that provide matrix estimation from counts.  
While there is no fundamental problem with using matrix estimation software they are generally only applied to O-D type matrices (often just created from intercept survey data) without putting the information back into input P/A matrices, and so are not suitable for demand modelling purposes.  
All this has meant that P/A demand models are used to predict changes in demand and this change is then used to adjust a validated O-D matrix.  
So, what could be done to get around these issues?
8. Faced with this challenge we set about defining a practical means by which we could:  
First, combine synthetic matrices with available survey data whilst maintaining the P/A relationship in the base.  
Secondly, use the inherent strengths and weaknesses of the data types to combine the data in the most effective way. The main strength of synthetic data is the trip ends, whereas the distribution is poor. Intercept data have very poor trip end information but much better distributional data. Count data have no purpose or P/A information but are good for ensuring overall consistency with network flows.  
Third, quantify the inherent measurement, random and systematic errors in the survey data to produce confidence estimates for individual data items.  
Finally, and most importantly, allows the use of O-D matrix estimation tools to directly modify the P/A matrices thus maintaining consistency between the P/A and O-D base matrices.
9. This slide gives an overview of the process. I won't spend much time on this slide as a larger version of the diagram is included with the notes that you can pick up later if you wish. I'll describe in some detail the process in the next slides.  
In summary, we create a set of national synthetic P/A matrices as the starting point in the top left corner.  
We then combine this data with intercept survey data in the upper octagon named matrix estimation.  
The matrices are then summed to create assignment matrices, we move below the red line from P/A matrices into O-D matrices, so we can use OmniTRANS matrix estimation software.  
The adjustments from the matrix estimation are then transferred back to the P/A matrices above the red line.  
The process is then repeated.
10. Before talking about the procedures I'll outline the wide variety of data sources we have used. The NTEM model is used for trip rate information and as input data in the external zones.

The Tyne and Wear household interview survey is used to verify the national assumptions. The mean trip lengths available from the NTS data were significantly different to local data so local data was used for target mean trip lengths in distribution.

The census data was used to create population and employment estimates and also in sense checking.

Local planning data was used to create employment data by type

Finally, Count Data, RSI Data and PT Intercept Survey data were included.

11. The synthetic matrix is a major component of the base matrix building procedures. I won't spend a lot of time on this slide as the next slide shows the process as a diagram.

In summary:

We create national synthetic production and attraction trip ends from a variety of data sources. These are then distributed with a negative exponential gravity model before applying time of day and directional factors.

The synthetic matrices that come out of this process are segmented by 9 purposes, 3 modes, 4 times of day, 4 car availability groups and 2 directions.

12. The first step is to create synthetic trip productions.

These are based on population and household numbers and person type splits in each zone.

The data is sourced from a combination of local data (local planning departments), 2001 Census data and NTEM data in the external zones.

We then use the National Car Ownership program to create estimates of car availability in each zone.

We then use trip rates from NTEM to create trip productions in each zone by Purpose, Mode and Car Availability

The next step is to create synthetic attractions.

This uses employment data for each zone, which again is sourced from a combination of local data, 2001 Census data, ABI data and NTEM data for the external zones.

There is no car availability in the attractions model so we apply NTEM attraction rates directly to this data to create trip attractions by mode and purpose

We then distribute the trips, calibrating the deterrence function to a mean trip length target from the T&W household interview survey.

This provides our synthetic matrix, which is the basis of the base matrices.

13. To account for the poor distribution in the synthetic matrices we statistically combine the synthetic matrices with observed data from surveys.

Again, the next slide contains a diagram of the process. We use two types of survey data in this process.

The first is intercept survey data from RSI surveys and a continuous monitoring survey on the public transport network. This process, called matrix modification, combines the information in the P/A observed matrices with the synthetic matrices on a cell basis.

The second type of data is O-D count information which is used in matrix estimation. In the Tyne and Wear model we only have highway counts whereas in the Durham model we are able to run matrix estimation for both highway and PT.

14. So, the process works like this.

We start with the synthetic matrices in the top left corner. These become the prior matrices for the matrix modification. We then combine the observed and synthetic matrices.

To do this we calculate confidence bounds on the observation based on the inherent systematic, measurement and random errors in the datasets.

We then evaluate each cell in the synthetic matrices against the observed data and alter the cell value accordingly.

Next we apply the synthetic trip rate constraint by purpose. This ensures that the synthetic trip ends (which we know are very stable) are maintained and any undue changes are minimised. This process iterates a maximum of 40 times.

After 40 iterations we have the modified matrix. We then sum over purpose, Car Availability and Direction and run matrix estimation. The matrices are now in O-D format.

The ratio between the matrix estimation inputs and the outputs is calculated and applied back to the P/A components of the aggregate O-D matrices.

This creates an adjusted matrix. We then re-apply the trip rate constraint to stabilise the system and if necessary this becomes the new prior matrix for the process to run iteratively.

This process creates validated Base Matrices in a P/A format that are directly linked to the O-D assignment matrices making use of many data sources.

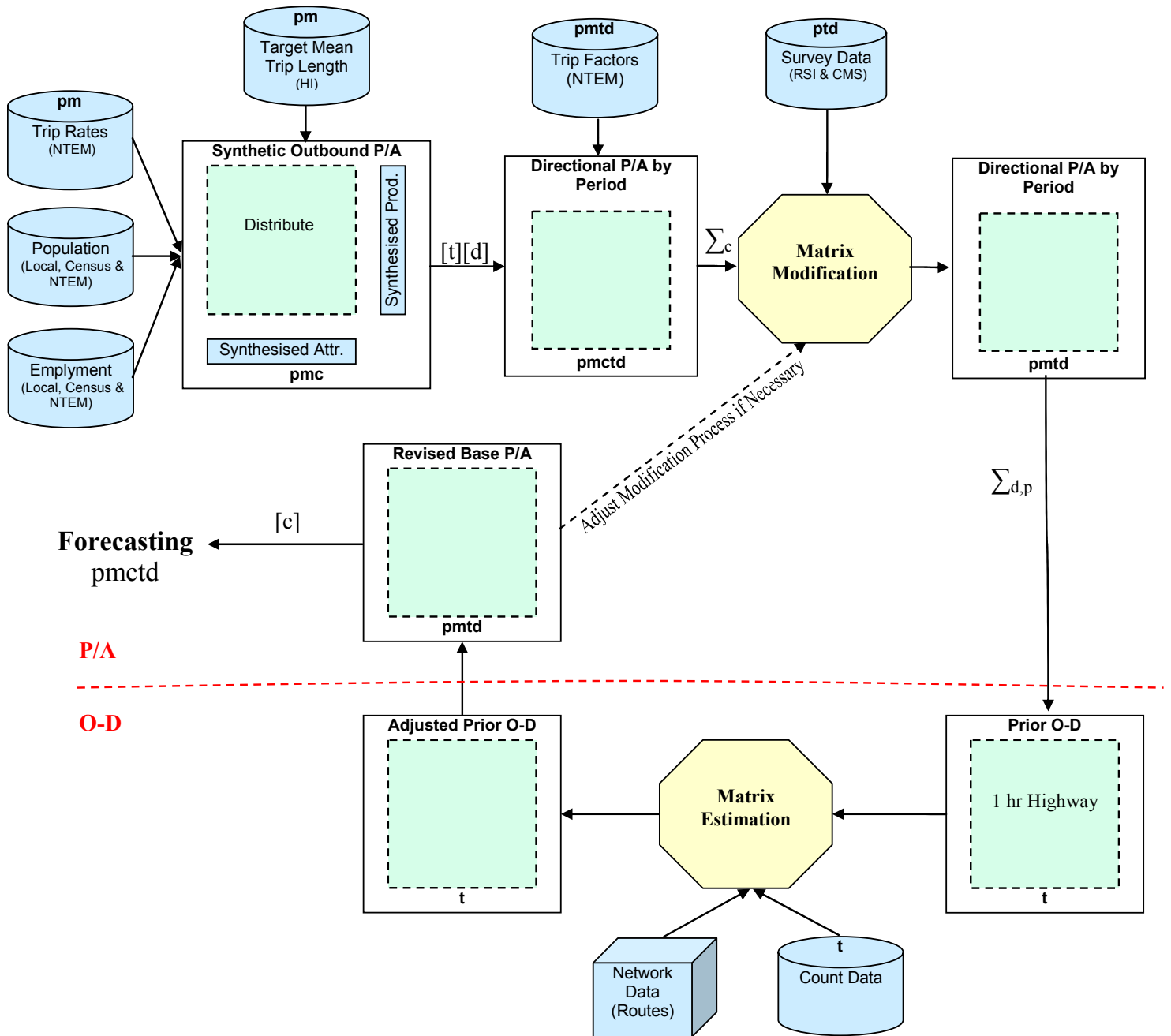
It uses stable trip end modelling but allows the best use of costly survey data.

Therefore allowing modal and TOD shares to be adjusted by local data.

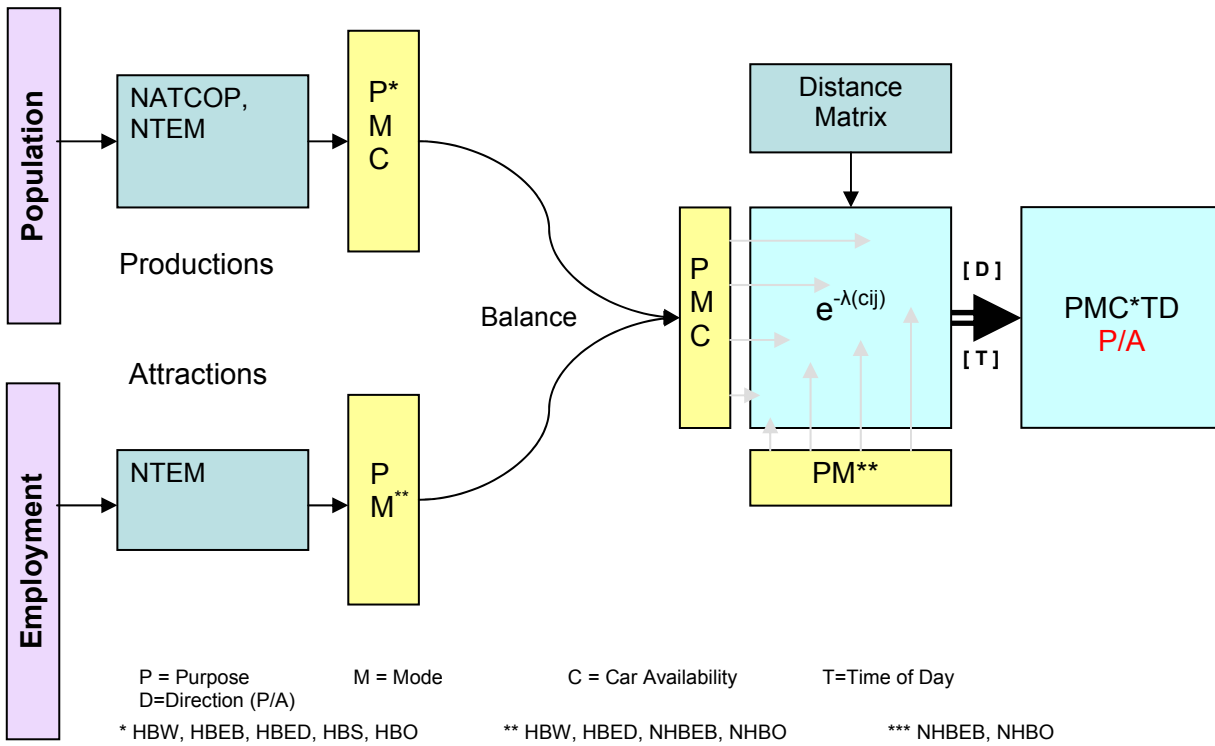
And, the base demand is consistent with forecasting methodologies.

15. We have spent some time getting this procedure right and have produced a set of validated base matrices for the T&W model.  
The initial results of the supply model calibration using these matrices look promising, so we can consider the process we have developed to be a success.  
I'll now quickly go through the sense checking and a couple of outputs from the matrix validation, if you'd like any more information feel free to speak to me after the seminar.
16. To validate the base matrices we developed a systematic procedure that covers the following items:
  - Input Data  
Counts and Intercept as well as Population and Employment Data
  - Procedural Checks  
Does the model produce outputs in line with the inputs and with TEMPRO and Census data
  - And Reality Checks  
Things like desire lines and screenline flows.
17. It would be worth thinking how many studies carry out all of these checks during model development. I don't have any time to go into this so please come to talk to us later if you have any questions, there is little guidance on this area so I imagine other studies might be able to use this process of give us some feedback from their studies. We certainly found a number of interesting details about our model that we wouldn't have found out without carrying out these checks.
18. The following slides show some of the outputs from the Base Matrix validation. I'll go through these quite quickly.
19. Thank you for listening and I hope that you've enjoyed the presentation.

## Base Matrix Creation Overview



### Synthetic Matrix Creation



### Matrix Modification/Estimation

