DATA COLLECTION NEEDS AND TECHNIQUES

Outline

1. Problems with Transit Data
2. Framework for Data Collection
3. Data Needs
   - Precision
4. Manual Data Collection Techniques

Problems in current data collection programs

Great variation in data collection resources

Variation in techniques used:
automated, manual, mixed

No statistical approach incorporating
- **Required accuracy**
- **Data variability**

Little trust in data

Inefficient use of data

Major limitation on use of other analytic methods

Size of Checker Force by System Size

	< 50 buses	51-200 buses	201-500 buses	501-1,000 buses	> 1,000 buses	Total
None	32/27*	18/11	4/7	1/6	2/7	58/58
1-5	6/12	16/22	8/5	1/0	2/0	33/39
6-11	1/1	0/1	3/2	3/1	1/1	8/6
12-20	0	0	0/1	3/1	1/0	4/2
over 20	0	0	0	1/1	4/2	5/3
Total	40	34	15	9	10	108

*** Full Time/Part Time**

Source: TCRP Synthesis Report "*Bus Route Evaluation Standards*", Synthesis of Transit Practice 10, Transit Cooperative Research Program, 1995.

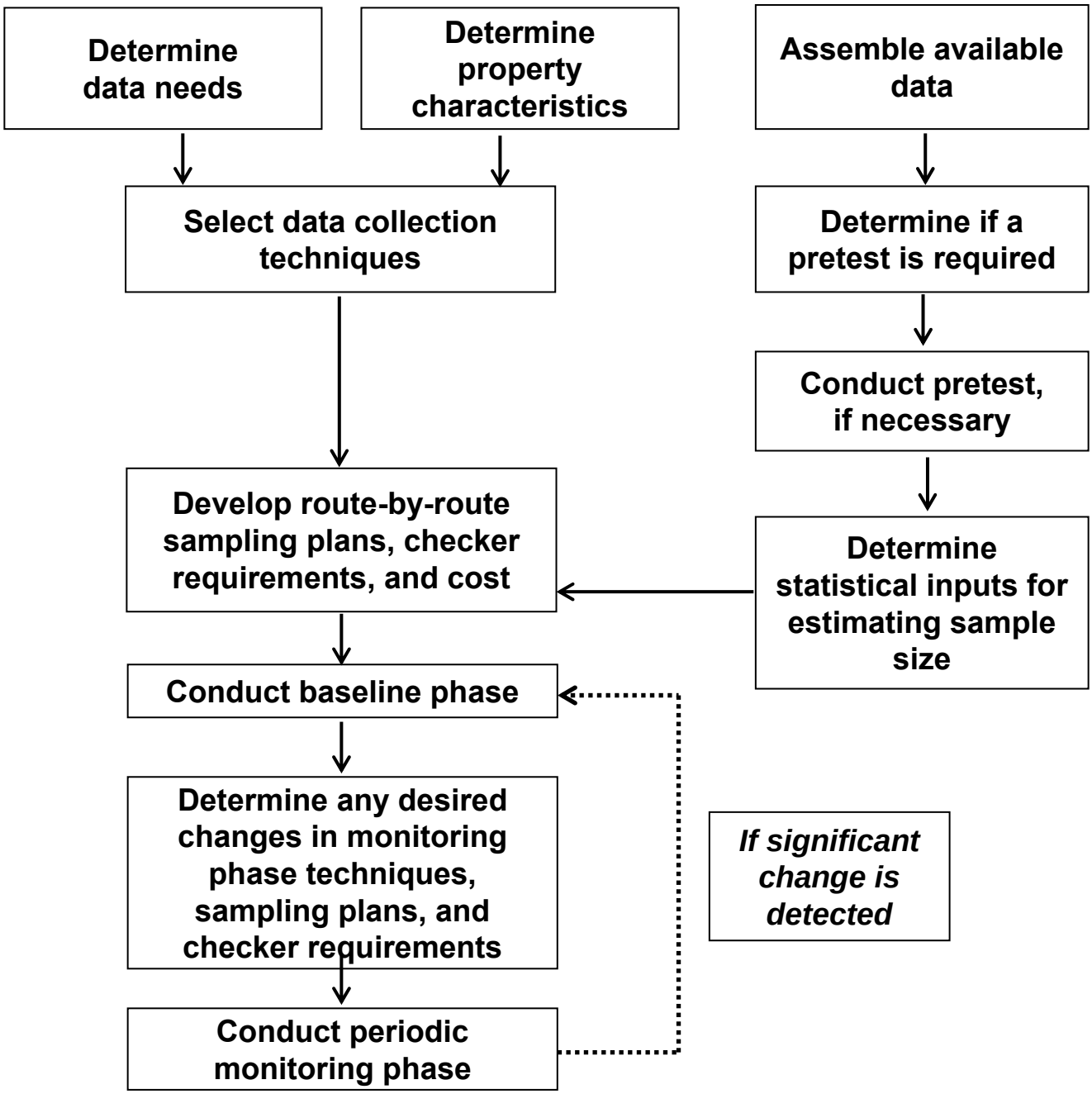
Weekday Check Frequency by System Size

	< 50 buses	51-200 buses	201-500 buses	501-1,000 buses	> 1,000 buses	Total
1 set per pick per year	9/9*	7/10	2/1	4/1	2/0	24/21
2 sets per pick per year	5/6	3/3	1/0	2/1	5/1	16/11
1 set per alternate pick	1/0	1/2	2/1	0	0	4/3
Once per year	11/8	10/5	1/5	2/4	0/4	24/26
Every other year or less frequently	7/8	9/13	9/8	1/3	3/5	29/37

*** Point Check/Ride Check**

Source: TCRP Synthesis Report "*Bus Route Evaluation Standards*", Synthesis of Transit Practice 10, Transit Cooperative Research Program, 1995.

Summary of Data Collection Program Design and Implementation



Data Needs in Baseline Phase

A. Route (or Stop) Specific

Load (at peak point -- other key points)*

Running time

Schedule adherence

Total boardings (i.e., passenger-trips)

Revenue

Boardings (or revenue) by fare category

Passenger boarding and alighting by stop

Transfer rates between routes

Passenger characteristics and attitudes

Passenger travel patterns

B. System Wide

Unlinked Passenger Trips

Passenger-miles

Linked Passenger trips

* at specified points; not averaged throughout a trip

Precision and Confidence Level

Accuracy of an estimate has two dimensions.
"Mean boardings per trip is 33.1."

Exactly 33.1???
"Mean boardings per trip is 33.1, plus or minus 10%"
– precision
"Mean boardings per trip is 33.1, plus or minus 3.3"
– tolerance

Are you sure?
"I'm 95% confident that mean boardings per trip is 33.1, plus or minus 10"
– precision and confidence level

To simplify matters:

- **hold confidence level fixed (90%)**
- **vary precision to reflect different levels of accuracy**

National Transit Database specification for annual boardings, pass-miles: ±10% precision at 95% confidence level

Desired Accuracy

System boardings for management purposes:

____±3% quarterly – equivalent to ± 1.5% for annual estimate

On-time performance, systemwide:

Suppose percent on time is 80%. Choose tolerance:

80% ± %? (I'll choose 4%)

Convert to "absolute equivalent tolerance"

AET = 0.5 tol / sqrt[p*(1-p)], where p = expected proportion

AET = 0.5 (4%) / sqrt[0.8*(0.2)] = 5%

Tolerance naturally improves as the proportion moves toward the extremes (0%, 100%). AET is the tolerance you'd get if the proportion were 50%.

Expected proportion	Tol corresponding to ±5% AET
50%	5%
60% or 40%	4.9%
70% or 30%	4.6%
80% or 20%	4%
90% or 10%	3%
95% or 5%	2.2%

Recommended Tolerances

Peak Load (also boardings):

Routes with 1-3 buses	±30%
Routes with 4-7 buses	±20%
Routes with 8-15 buses	±10%
Routes with >15 buses	± 5%

Vehicle trip time:

Routes with trip time ≤ 20 mins	±10%
Routes with trip time ≥20 mins	± 5%

On-time performance ±0.1
AET

Passenger Counting Techniques

Operator (trip cards)

Traffic Checker (with handheld device)

- ride check (on/off and running time)
- point check (load and headway)

Fare System

- passenger counts
- revenue counts only

Automatic Passenger Counters

Passenger Surveys

Types of Counts and Readings

Type of Count/ Reading	Description	Corresponding Deployment Options
On/off count	Ons and offs by stop; also time at time points. In rare cases, ons may be by fare category	Ride check/APC
Boarding counts	Boardings by trip, by fare category, may be also by stop	Ride check Driver count Fare systems
Load counts	Load on bus as it passes a point; also time at that point	Point check
Revenue count	Revenue by fare type	Fare systems
Transfer counts	Count of transfer tickets sorted by original and final route	Fare systems
Route origin/ Destination count	Count of passengers by O/D stop pair	Special
Survey	Passengers respond to questions, either written or verbal	Special

Method of Checking by System Size

	< 50 buses	51-200 buses	201-500 buses	501-1,000 buses	> 1,000 buses	Total
Manual	25	16	8	5	7	61
Manual with Hand Helds	5	9	4	3	2	23
Automated Entirely	1	0	0	0	0	1
Automated with Manual Supplements	0	5	2	1	1	9
Fare system only	9	4	1	0	0	14
Total	40	34	15	9	10	108

Source: TCRP Synthesis Report "*Bus Route Evaluation Standards*", Synthesis of Transit Practice 10, Transit Cooperative Research Program, 1995.

Use of Passenger Counting Technologies and Procedures

Number of systems	Total	> 1000 buses	250-1000 buses	< 250 buses
Technology/Procedure				
Checkers, Pencil and Paper	23	8	11	4
Electronic Registering Fareboxes	23	6	11	6
On-Board Surveys	15	2	11	2
Vehicle Operator Trip Cards	14	3	9	2
Estimate from Passenger Revenue	13	5	8	0
Checkers and Hand-Held Units	13	3	7	3
APC in use	8	2	5	1
APC Testing	5	1	3	1
Smart Cards	2	0	1	1
Number of Systems	33	9	17	7

Source: TCRP Synthesis Report "*Passenger Counting Technologies and Procedures*", Synthesis of Transit Practice 29, Transit Cooperative Research Program, 1998.

Mapping Data Needs to Data Collection Techniques

		Load		Unlinked Pass.		Revenue	
Collection Technique	Counts and Readings	Peak Point	Other	Total	By Fare Category	By Trip	By Segment
Ride Check	On/off counts	=	=	=			
Point Check							
-peak point	Load count		=				
-end point	Load count						
	Farebox reading	=		=	=		
-multiple point	Load count						
	Farebox reading	(1)	=			(2)	=
Driver Check	Boarding Count				=	=	=
Electronic Farebox	Boarding Count			=	=	=	=
APCs	On/off count	=	=	=			
Special	O-D count	=	=	=			
	Survey						

KEY: = if applicable, blank if not applicable

NOTES: (1) If peak point is included
(2) If end points are included

Mapping Data Needs to Data Collection Techniques (cont'd)

				Running Time			
Collection Technique	Counts and Readings	Ons, Offs by Stop	Pass. Miles	By Trip	By Segment	Schedule Adherence	Travel Patterns, Transfer Rates
Ride Check	On/off counts	=	=	=	=	(1)	
Point Check							
-peak point	Load count						
-end point	Load count					=	
	Farebox reading						
			(2)		=		
-multiple point	Load count						
	Farebox reading						
		(3)	(4)	=	=		
Driver Check	Boarding Count						
Electronic Farebox	Boarding Count						
APCs	On/off count	=	=	=	=	=	
Special	O-D count Survey	=	=	=	=	=	=

KEY: = if applicable, blank if not applicable

NOTES:
(1) Cannot be used to check headways unless consecutive trips are checked
(2) Round trip only if one end point is checked
(3) Can provide a rough estimate
(4) If end points are included

Improving Traffic Checker Data

Point check observed 70 passengers on a trip.

- **Uncontrolled: load could be anywhere between 41 and 81.**
- **Large random variation compounded by systematic overcount**
- **Controlling error: verification counts, immediate feedback, retraining**
- **Have the checker board the bus to count**

Improving Traffic Checker Data (cont'd)

Preprinted forms:

- scheduled trips, stop lists

Handheld devices

- reduce real-time coding errors
- error detection
- load checks

Have checkers code their own data

- immediate graphical feedback

Watch for fabricated data