Air-travel Itinerary Shares: Relevant Factors & Competition Dynamics

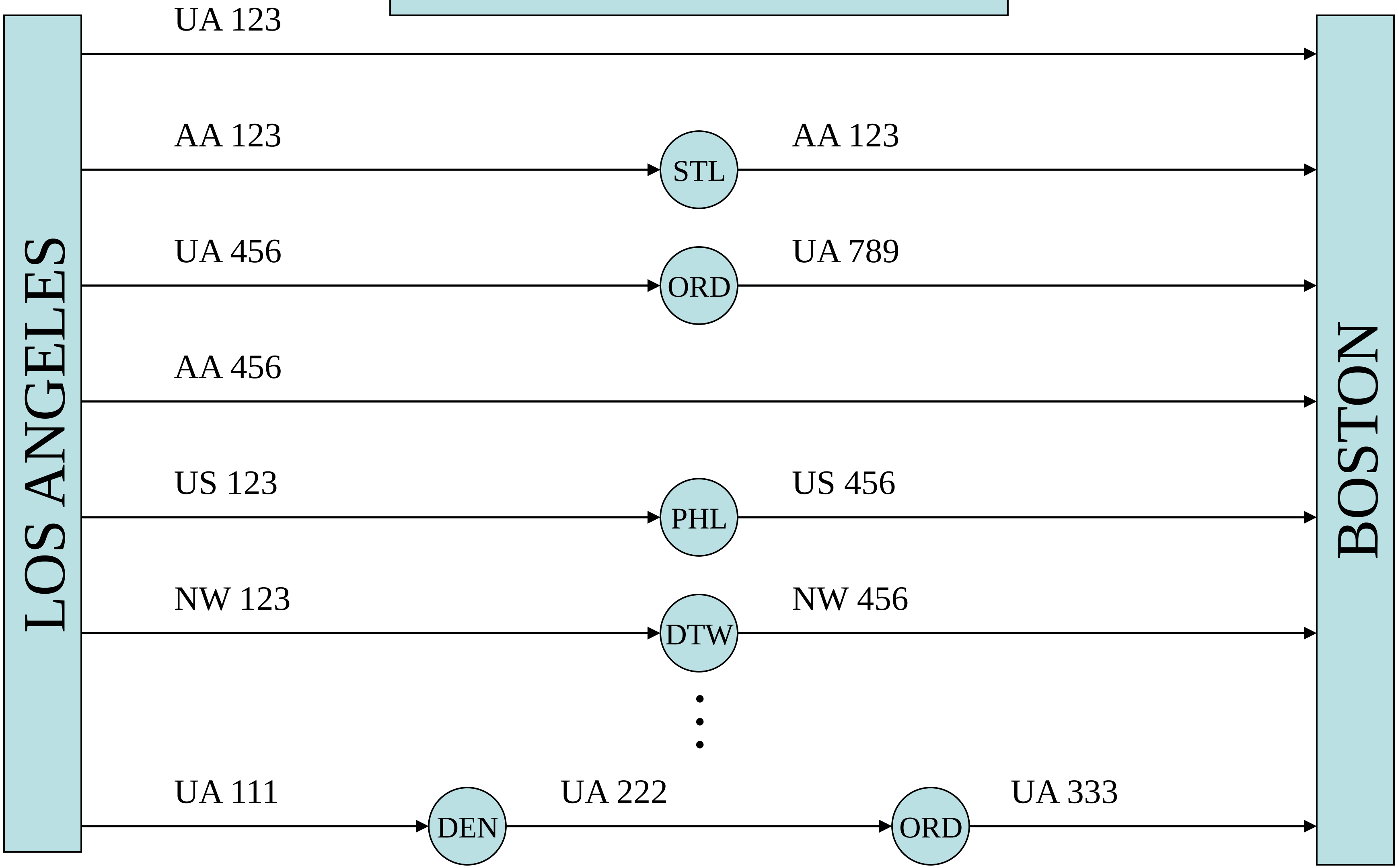
Gregory M. Coldren, Ph.D.
Assistant Professor of Mathematics
Frederick Community College

MMATYC Spring 2015 Conference
Frederick Community College
May 29, 2015

Air-travel Itinerary Share Models

- Forecast the share of passengers expected to travel on each itinerary between any airport-pair
- Support air-carrier strategic & tactical decisions, such as:
  - M&A / Codeshare decisions (Who to fly with)
  - Network decisions (Where/How/How often to fly)
  - Fleet decisions (What to fly)
  - Time-of-day decisions (When to fly)

~1,000 Daily Passengers



Data Assembly

- Passenger bookings data obtained from CRS (MIDT) sources
- A major carrier’s itinerary building engine was used to generate the itineraries between all North American airport-pairs
- Generated itineraries were merged with booked itineraries to assemble the estimation datasets

DCA Conceptual Framework

- Discrete choice analysis scenarios are described by four elements:
  - A decision-maker
  - The discrete alternatives available to the decision-maker
  - Attributes of these alternatives
  - A decision rule

DCA Conceptual Framework (Continued)

- Following convention:

$$U_i = V_i + \varepsilon_i$$

$$V_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_n X_{ni}$$

$$f(x) = \mu e^{-\mu(x-\eta)} e^{-e^{-\mu(x-\eta)}}$$

MNL Model Formulation

$$P(i:C) = P(U_i \geq U_j; \forall j) = P(V_i + \varepsilon_i \geq V_j + \varepsilon_j; \forall j) =$$

$$P(\varepsilon_j \leq V_i - V_j + \varepsilon_i; \forall j \neq i) =$$

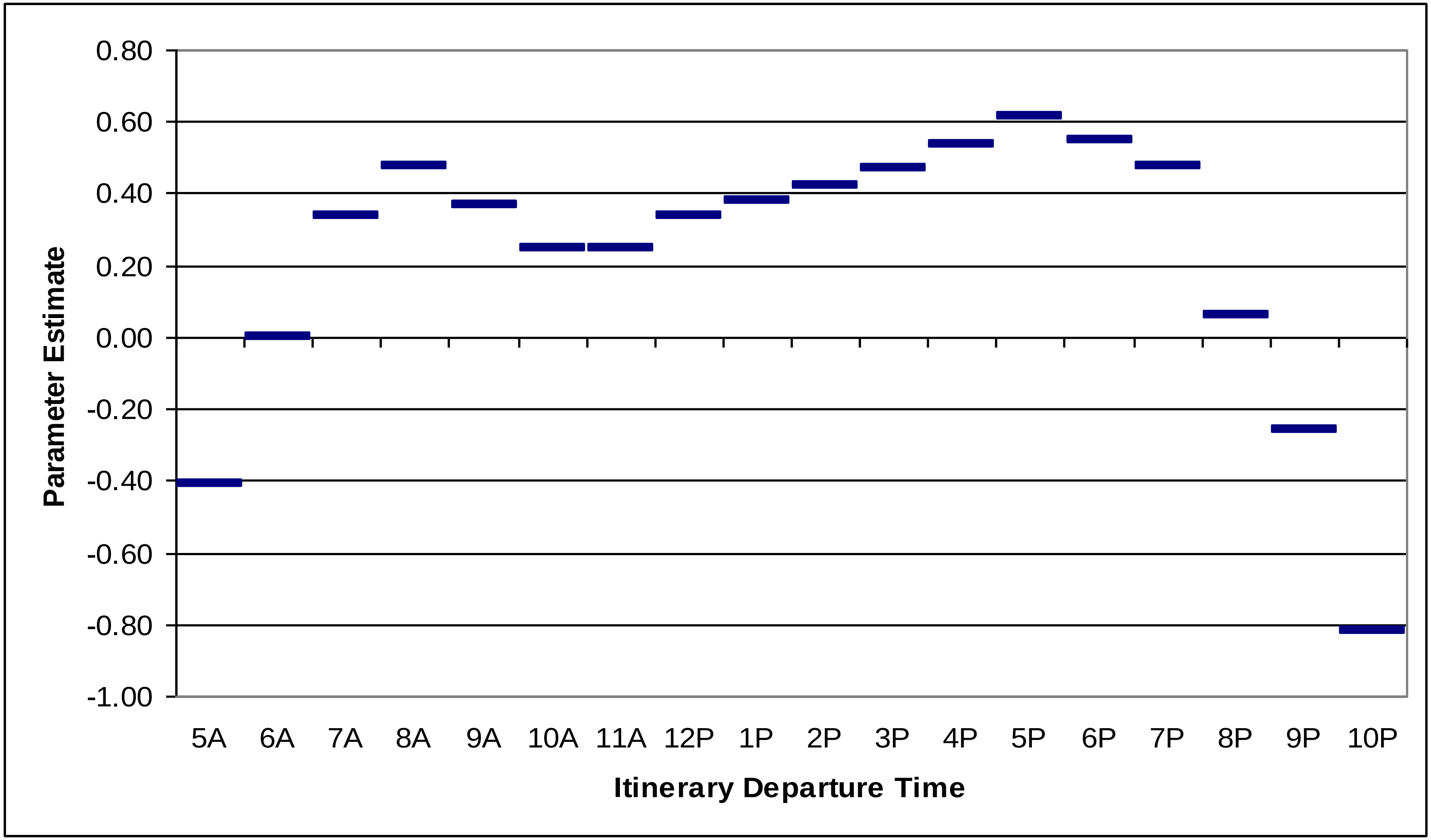
$$\int_{\varepsilon_i=-\infty}^{\infty} \int_{\varepsilon_j=-\infty}^{V_i - V_j + \varepsilon_i} f(\varepsilon)\, d\varepsilon_j \cdots d\varepsilon_1 d\varepsilon_i = \cdots$$

$$= \frac{e^{\mu V_i}}{\sum_j e^{\mu V_j}}$$

Impact of Service Factors on Itinerary Share

- Stops / Connections
- Connection quality (for connecting itineraries)
- Carrier, Airport presence, Fares, Codeshare factors
- Equipment
- Departure time / Day of Week

Passenger Departure Time Preferences from Time Period Model

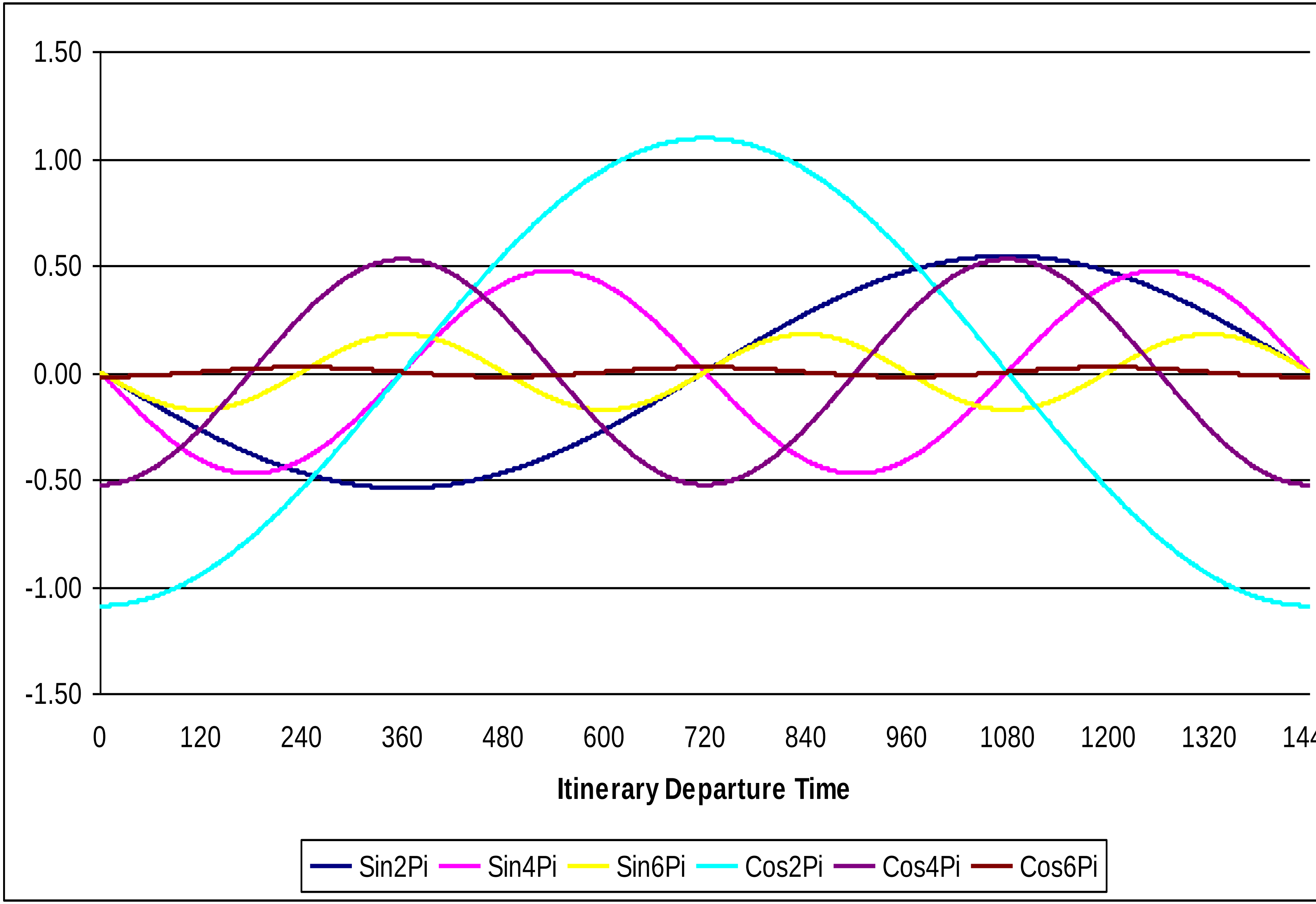


Employing Sin-Cos Curves to Model Departure Time Preferences

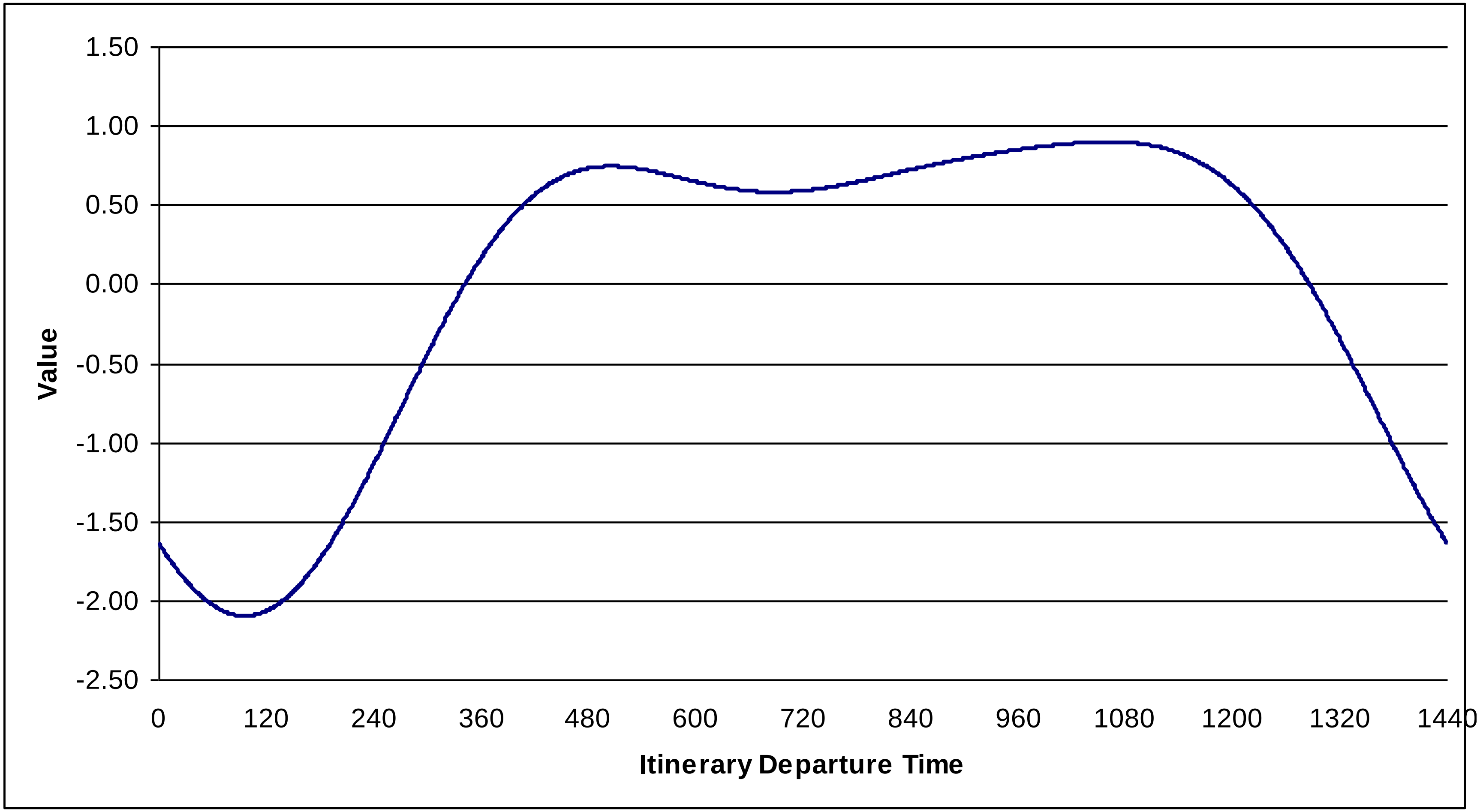
- Following Zeid *et al.* (2005), Sin & Cos curves are included in the value functions
- Sin-Cos model is specified as:

$$V_i = \cdots + \beta_1 Sin\left(\frac{2\pi t_i}{1440}\right) + \beta_2 Sin\left(\frac{4\pi t_i}{1440}\right) + \beta_3 Sin\left(\frac{6\pi t_i}{1440}\right) +$$

$$\beta_4 Cos\left(\frac{2\pi t_i}{1440}\right) + \beta_5 Cos\left(\frac{4\pi t_i}{1440}\right) + \beta_6 Cos\left(\frac{6\pi t_i}{1440}\right) + \cdots$$


1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
0
120
240
360
480
600
720
840
960
1080
1200
1320
1440
Itinerary Departure Time
Sin2Pi
Sin4Pi
Sin6Pi
Cos2Pi
Cos4Pi
Cos6Pi

Passenger Departure Time Preferences from Sin-Cos Model



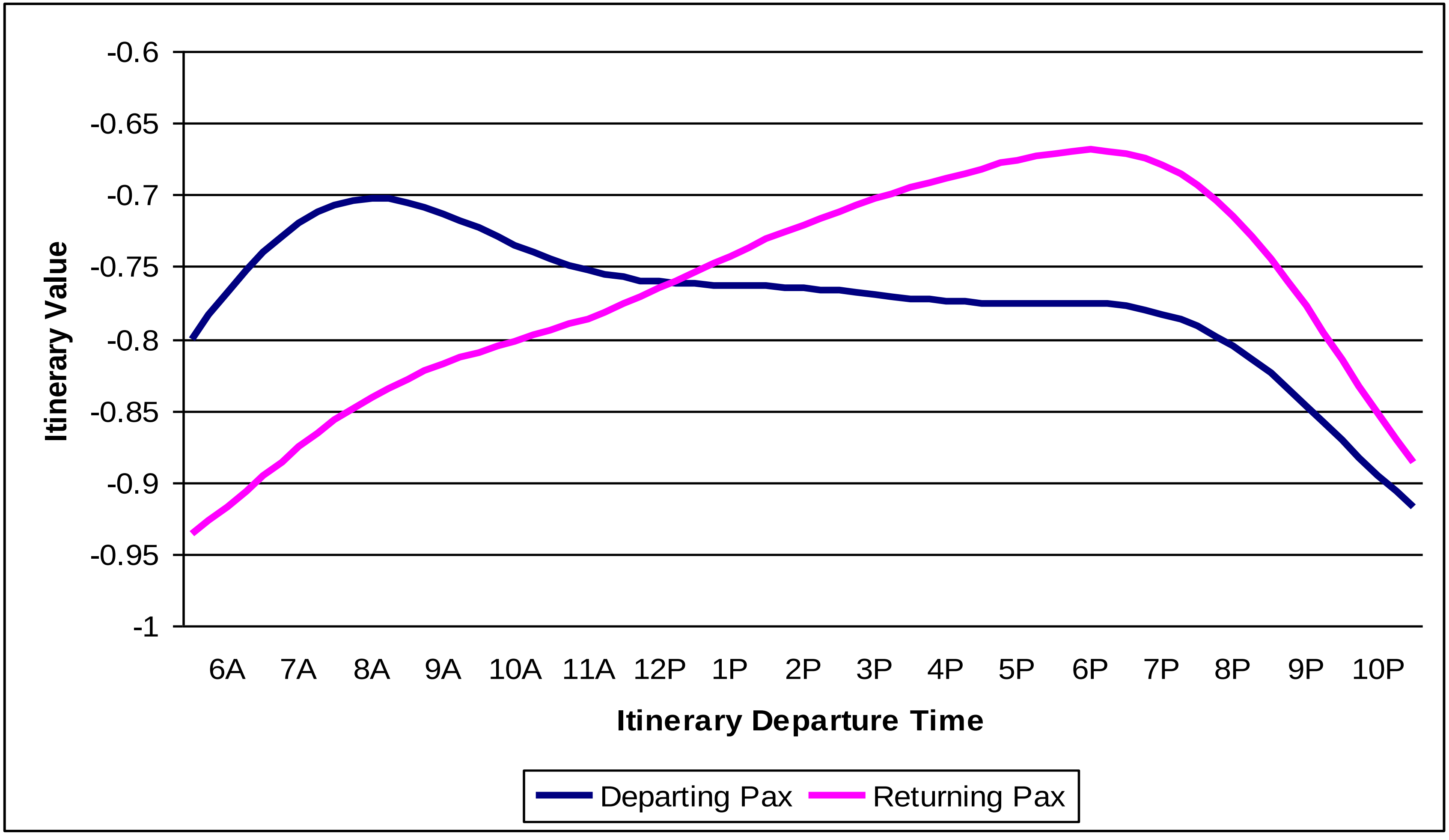
Results: Base Models

	Model		
	No Time Variables	Time Periods	Sin-Cos
Adjusted Log Likelihood at Convergence	-104,000	-103,169	-103,164

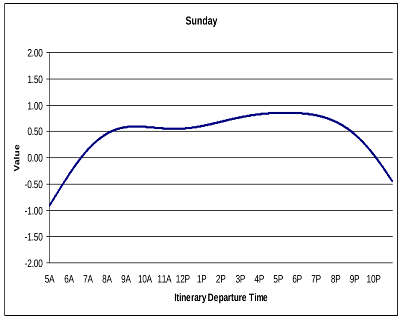
Results: All Pax vs. Differentiating between Departing and Returning Pax

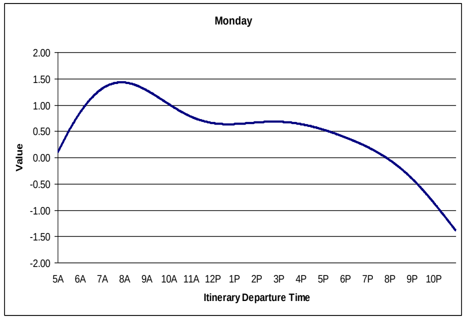
	Model	
	All Passengers	**Segmented**
Adjusted Log Likelihood at Convergence	-103,164	-100,346

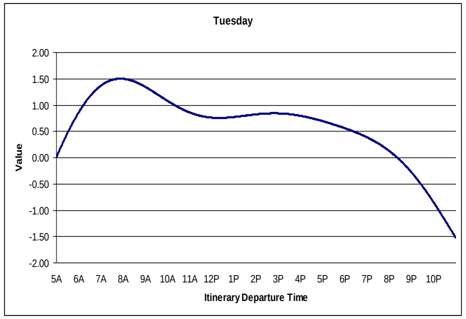
Departure Time Preferences for Departing and Returning Passengers

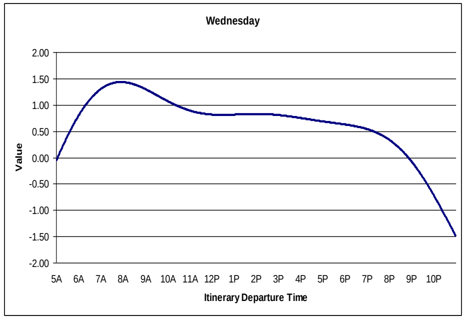


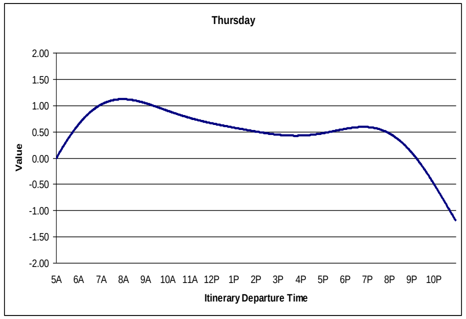
Departing Passengers:
Day-of-week Preferences

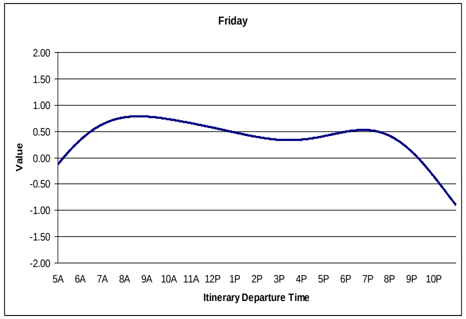


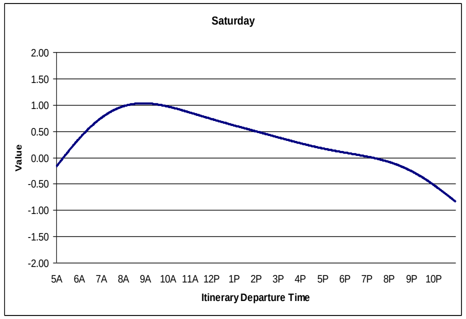




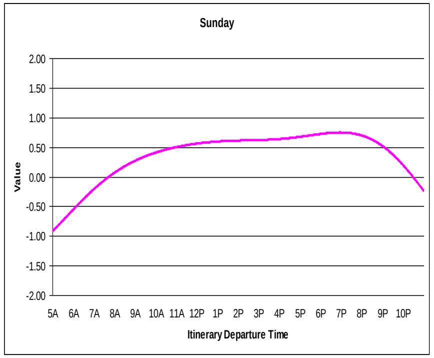


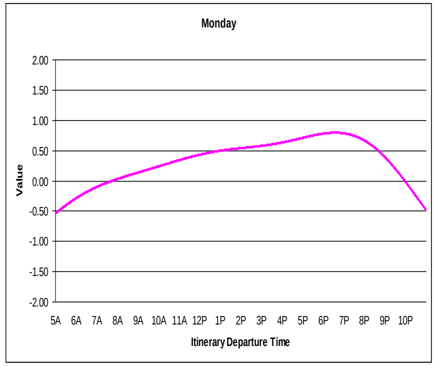


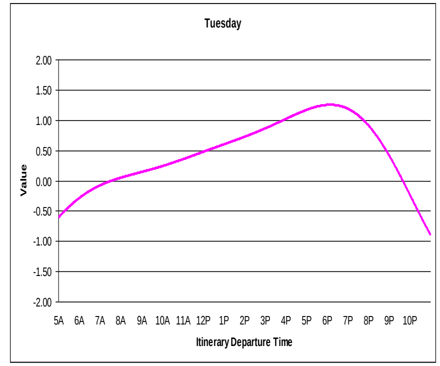


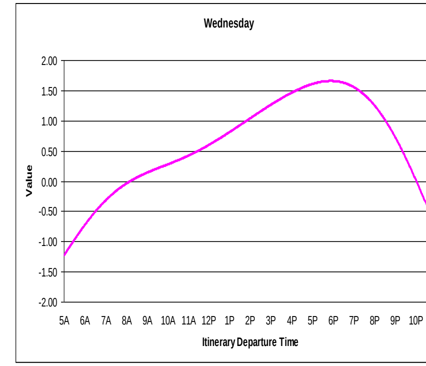


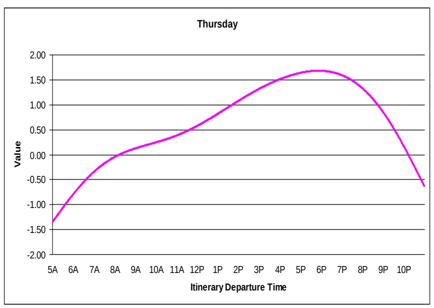
Returning Passengers: Day-of-week Preferences

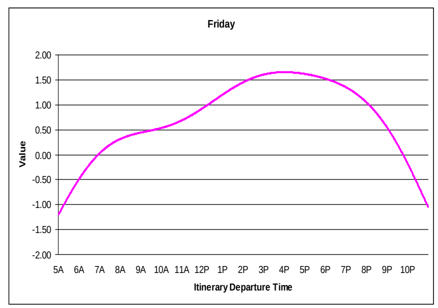


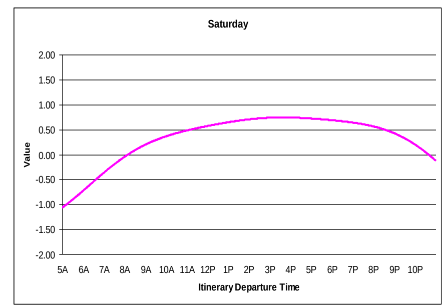












Results: Segmented vs. Segmented w/ Day-of-week Differentiation

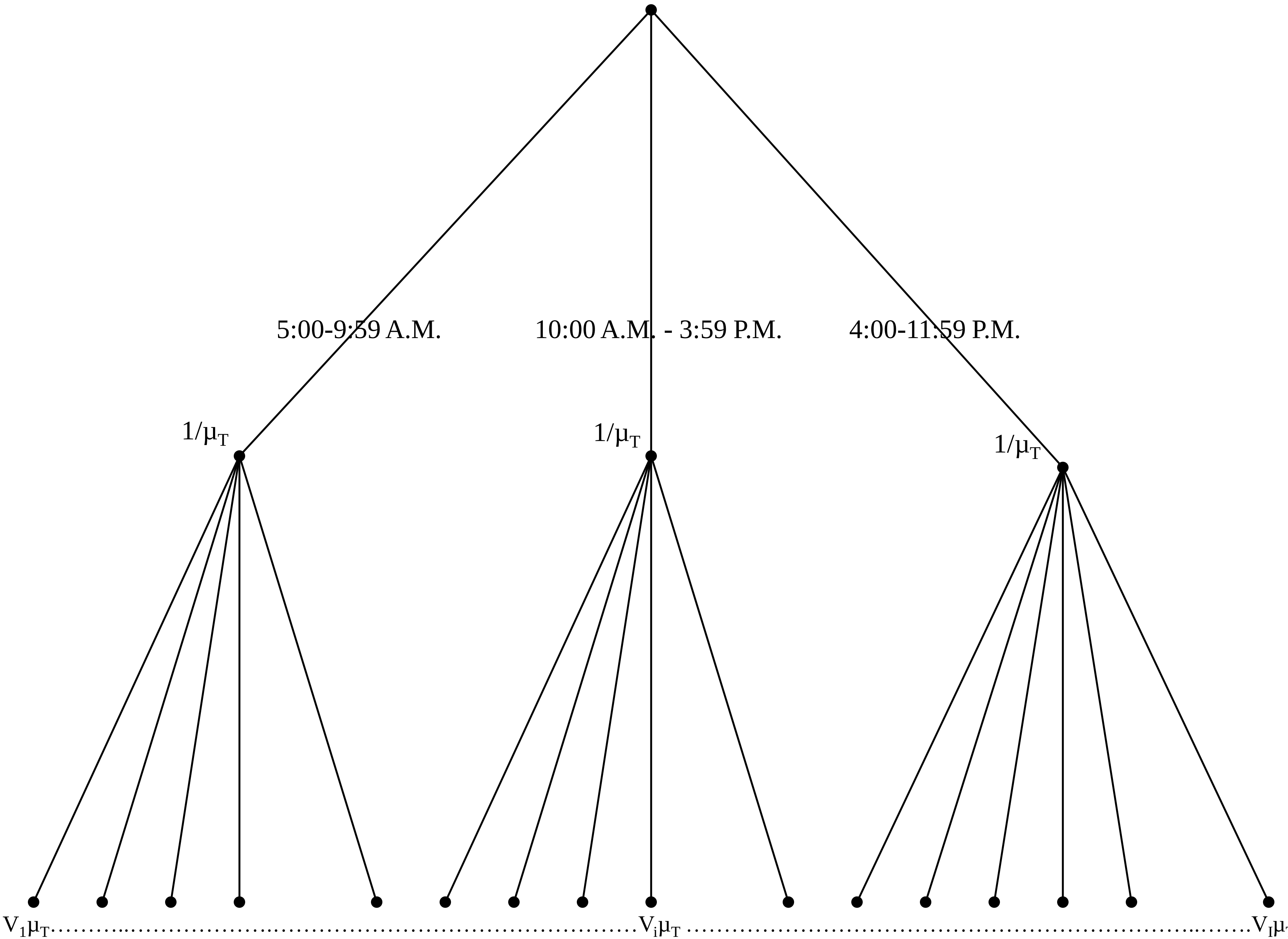
	Model	
	Segmented	Segmented w/ DOW Differentiation
Adjusted Log Likelihood at Convergence	-100,346	-99,747

Modeling the Competitive Dynamic among Itineraries with NL Models

- MNL models are unrealistic:

$$\eta^{P_j}_{X_{ik}} = \frac{\partial P_j}{\partial X_{ik}} \frac{X_{ik}}{P_j} = -P_i \beta_k X_{ik}$$

- Inter-itinerary competition shown to be differentiated by proximity in departure time and/or carrier (Coldren & Koppelman 2005)
- NL itinerary share models allow a carrier to better model the true competitive structure of the network


5:00-9:59 A.M.
10:00 A.M. - 3:59 P.M.
4:00-11:59 P.M.
$1/\mu_T$
$1/\mu_T$
$1/\mu_T$
$V_1\mu_T$
$V_i\mu_T$
$V_I\mu_T$

NL Market Shares

$$S_i = S_n \times S_{i|n} = \frac{\exp\left(\frac{1}{\mu}\Gamma_n\right)}{\sum_{n' \in N} \exp\left(\frac{1}{\mu}\Gamma_{n'}\right)} \times \frac{\exp(\mu V_i)}{\sum_{i' \in n} \exp(\mu V_{i'})}$$

$$S_i = S_m \times S_{n|m} \times S_{i|n}$$

$$= \frac{\exp\left(\frac{1}{\mu_m}\Gamma_m\right)}{\sum_{m' \in M} \exp\left(\frac{1}{\mu_m}\Gamma_{m'}\right)} \times \frac{\exp\left(\frac{\mu_m}{\mu_n}\Gamma_n\right)}{\sum_{n' \in N} \exp\left(\frac{\mu_m}{\mu_n}\Gamma_{n'}\right)} \times \frac{\exp(\mu_n V_i)}{\sum_{i' \in n} \exp(\mu_n V_{i'})}$$

2-level NL Model
Within-nest Elasticity

$$\eta_{X_{ik}}^{P_j} = \frac{\partial P_j}{\partial X_{ik}} \frac{X_{ik}}{P_j} = -P(i \mid n')\beta_k X_{ik}\left[P(n') + (\mu - 1)\right]$$

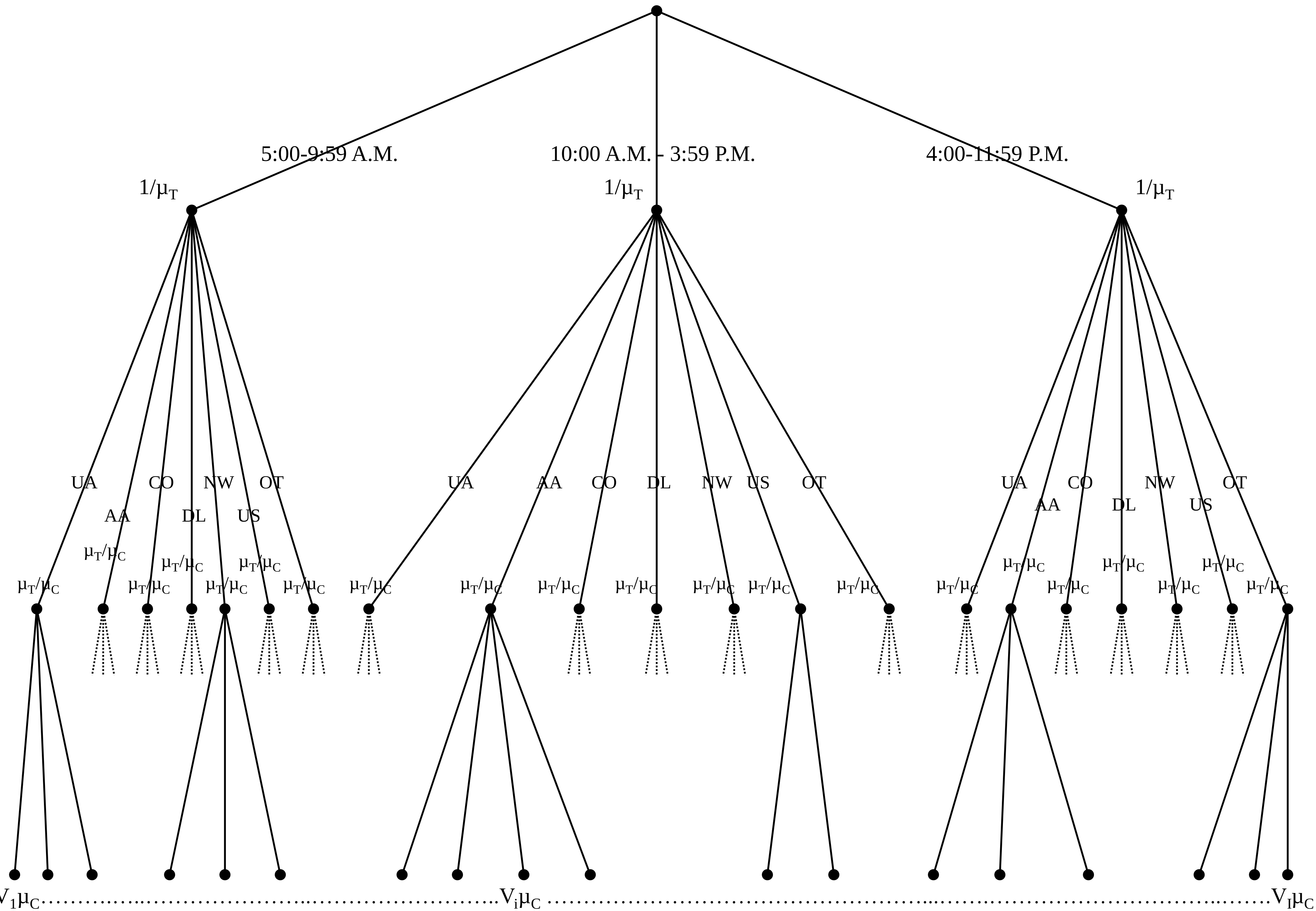
$$\left|-P_i \beta_k X_{ik}\right| \le \left|-P(i \mid n')\beta_k X_{ik}\left[P(n') + (\mu - 1)\right]\right|$$

$$P_i \le P(i \mid n')\left[P(n') + (\mu - 1)\right]$$

$$P_i \le P_i + P(i \mid n')(\mu - 1)$$

Results: MNL and 2-Level NL

	Model	
	MNL Segmented w/ DOW Differentiation	NL Segmented w/ DOW Differentiation
Inverse Logsums (5A-9A)	--------	1.44 (Dep) 1.18 (Ret)
Inverse Logsums (10A-3P)	--------	1.42 (Dep) 1.22 (Ret)
Inverse Logsums (4P-11P)	--------	1.41 (Dep) 1.20 (Ret)
Adjusted Log Likelihood at Convergence	-99,747	-99,613

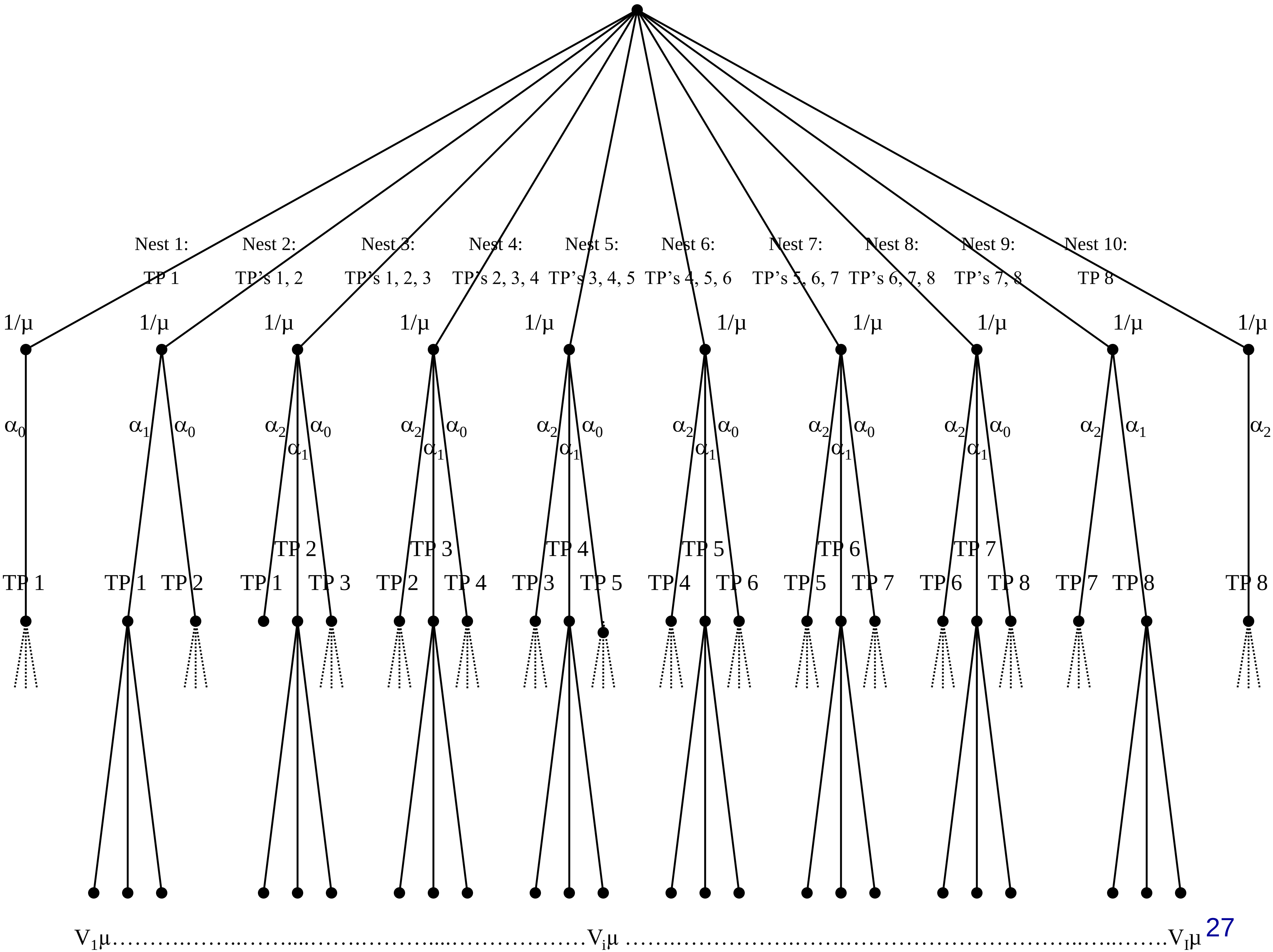

5:00-9:59 A.M.
10:00 A.M. - 3:59 P.M.
4:00-11:59 P.M.
1/μT
1/μT
1/μT
UA
AA
CO
DL
NW
US
OT
UA
AA
CO
DL
NW
US
OT
UA
AA
CO
DL
NW
US
OT
μT/μC
V1μC
ViμC
VIμC

Results: 2 and 3-Level NL's

	Model	
	2-level NL Segmented w/ DOW Differentiation	**3-level NL Segmented w/ DOW Differentiation**
Upper-level Inverse Logsums (5A-9A)	1.44 (Dep) 1.18 (Ret)	1.31 (Dep) 1.00 (Ret)
Upper-level Inverse Logsums (10A-3P)	1.42 (Dep) 1.22 (Ret)	1.23 (Dep) 1.07 (Ret)
Upper-level Inverse Logsums (4P-11P)	1.41 (Dep) 1.20 (Ret)	1.27 (Dep) 1.11 (Ret)
Lower-level Inverse Logsums (Carrier)	--------	1.51 (Dep) 1.30 (Ret)
Adjusted Log Likelihood at Convergence	-99,613	-99,472

Modeling the Competitive Dynamic among Itineraries w/ OGEV Models

- NL models impose unrealistic constraints on the time-of-day competition dynamic
- Itineraries exhibit “proximate covariance” (Small 1987)
- OGEV models (Small (1987), Coldren & Koppelman (2005)) capture this property
- These models consist of overlapping time periods where itineraries are allocated to contiguous nests according to allocation parameters


Nest 1:
TP 1
Nest 2:
TP's 1, 2
Nest 3:
TP's 1, 2, 3
Nest 4:
TP's 2, 3, 4
Nest 5:
TP's 3, 4, 5
Nest 6:
TP's 4, 5, 6
Nest 7:
TP's 5, 6, 7
Nest 8:
TP's 6, 7, 8
Nest 9:
TP's 7, 8
Nest 10:
TP 8
1/μ
α0
α1
α2
TP 1
TP 2
TP 3
TP 4
TP 5
TP 6
TP 7
TP 8
V1μ……Viμ……VIμ

Summary

- Many factors influence itinerary share
- Sin-Cos curves are behaviorally superior to (and provide better goodness-of-fit (GofF)) than 18 time period dummy variables
- Differentiating passengers by departing and returning dramatically improves model GofF
- Model GofF significantly improved by differentiating between the different days of the week
- NL models are shown to be highly significant and are better able to model (vs. MNL) the competitive structure of the network