# Continuous Departure Time of Day Preferences for Continental U.S. Airline Markets Segmented by Distance, Direction of Travel, Number of Time Zones, Day of Week and Itinerary Type


Virginie Lurkin, University of Liege, HEC-Management School, 4000 Liege, Belgium, Email: v.lurkin@ulg.ac.be

Laurie A. Garrow, Georgia Institute of Technology, School of Civil and Environmental Engineering, 790 Atlantic Drive, Atlanta, GA 30332-0355, Ph: (404) 385-6634, Email: laurie.garrow@ce.gatech.edu

Matthew J. Higgins, Georgia Institute of Technology, Ernest Scheller Jr. College of Business, 800 West Peachtree NW, Atlanta GA 30308 & National Bureau of Economic Research, 1050 Massachusetts Ave., Cambridge, MA 02138, Email: matt.higgins@scheller.gatech.edu

Jeffrey P. Newman, Georgia Institute of Technology, School of Civil and Environmental Engineering, 790 Atlantic Drive, Atlanta, GA 30332-0355, Ph: (404) 385-6634, Email: jpn@gatech.edu

Michael Schyns, University of Liege, HEC-Management School, 4000 Liege, Belgium, Email: M.Schyns@ulg.ac.be



*Last update: September 23, 2016*

**Abstract**

Airlines use itinerary choice models to allocate the total number of passengers in a city pair to specific itineraries. In a related paper, we estimated a multinomial logit (MNL) itinerary choice model using database of more than 3 million tickets for Continental U.S. markets provided by the Airlines Reporting Corporation that accounted for price endogeneity. The size and comprehensiveness of our database allowed us to estimate highly refined continuous departure time of day preference curves that account for distance, direction of travel, the number of time zones traversed, departure day of week and itinerary type (outbound, inbound or one-way). This paper and accompanying Excel spreadsheet located at http://garrowlab.ce.gatech.edu contain the results of all model coefficients (including the 1260 time of day parameters) and summarize results in a series of ten figures. These highly-refined time of day preference curves can be used by airlines, researchers, and government organizations in the evaluation of demand-management and other policies.

## Introduction

In a related paper (Lurkin, et al. 2016), we use a ticketing database provided by the Airlines Reporting Corporation to estimate an itinerary choice model similar to those currently used in practice. Specifically, we estimate a multinomial logit (MNL) model that uses a control function to account for price endogeneity. The independent variables included in our final specification are described in Table 1 and model results (excluding carrier constants that are suppressed for confidentiality reasons and departure time of day variables that are suppressed for presentation purposes) are shown in Table 2.

Due to the size of our analysis database, we were able to estimate detailed departure time of day preference curves that are segmented by distance, direction of travel, number of time zones traveled, day of week, and itinerary type (outbound, inbound or one-way). To the best of our knowledge, these curves represent the most refined publicly-available estimates of airline passengers' departure time of day preferences. This paper and accompanying Excel spreadsheet contain the results of all model coefficients, which can be used by other researchers for applications that require an understanding of how airline customers make tradeoffs among itinerary attributes including departure times, price, elapsed time, equipment types, carriers, level of service (nonstop vs. connecting), etc.

## Formulation of departure time of day preferences

As described in Lurkin et al. (2016), there are multiple approaches that can be used to model departure time preferences. The first is a discrete formulation that uses a series of indicator variables, one for each departure hour. The second is a continuous formulation that combines sine and cosine functions. We model time of day preferences using a continuous time of day formulation and follow the approach originally proposed by Zeid, et al. (2006) for intracity travel and adapted by Koppelman, et al. (2008) for itinerary choice models[1] by including three sine and three cosine functions representing frequencies of $2\pi, 4\pi,$ and $6\pi$. For example, the $\sin 2\pi$ term is given as:

$$\sin 2\pi = sin\{(2\pi \times \text{departure time})/1440\}$$

where departure time is expressed as minutes past midnight and 1440 is the number of minutes in the day. Similar logic applies to the $\sin 4\pi, \sin 6\pi, \cos 2\pi, \cos 4\pi,$ and $\cos 6\pi$ terms.

[1] Carrier (2008) uses four sine and four cosine functions to model departure time preferences for European markets.

We allow departure time preferences to vary according to several dimensions. Specifically, we create ten segments that capture how time of day preferences vary as a function of distance, direction of travel, and the number of time zones traveled in addition to the itinerary type (outbound, inbound, or one-way) and the departure day of week. Descriptive statistics for our ten segments are shown in Table 3.

[ Insert Table 3 about here ]

Figures 1 – 10 show the results of the departure time of day preferences for each segments. The accompanying Excel spreadsheet contains the parameter estimates and t-statistics for the 1260 time of day coefficients (10 segments x 6 sine/cosine terms x 3 itinerary types x 7 departure days = 1260 parameters).

Results are intuitive and show that departure time of day preferences vary across many dimensions. For example, Figure 1 shows the results of the departure time of day preferences for those itineraries in the same time zone with distances ≤ 600 miles. The curves show distinct departure time preferences by day of week and itinerary type. There is a strong preference for morning flights for outbound itineraries departing Monday, Tuesday and Wednesday and a strong preference for evening flights for inbound itineraries returning Wednesday or Thursday and to a lesser extent on Friday. These patterns likely correspond to people traveling for business who depart early in the week on morning flights and return later in the week on evening flights. Departure time preferences are not as pronounced on Saturday, although Saturday does exhibit a preference for morning departures for outbound itineraries, likely corresponding to the start of leisure trips. Departure time preferences are least pronounced for Sunday. The time of day preferences for one-way itineraries are not as strong as those for outbound and inbound itineraries (and typically fall between the two curves). This is expected, as the one-way itineraries may represent either the outbound or inbound portion of a trip (but is unknown to the researcher).

Similar patterns are observed for different segments, although the exact interpretation and peak periods differ depending on the segment. In general, the shorter the itinerary and the fewer time zones crossed, the stronger the departure time of day preferences are (note the scale of the y-axis is the same for all figures). For example, departure time of day preferences are not as strong for itineraries traveling two or three time zones. This is likely due to increased travel times (combined with the loss of hours due to time zone changes traveling eastbound) that make it difficult to depart in the morning and arrive for an early morning meeting (outbound) or return at the end of the day and arrive home at a reasonable hour (inbound). A similar pattern is observed for itineraries that are in the same time zone or one time zone apart in the sense that those itineraries that correspond to shorter distances ($\leq$ 600 miles) have stronger departure time of day preferences than longer itineraries.

**Acknowledgements**

This research was supported in part by "Fonds National de la Recherche Scientifique" (FNRS - Belgium).

**References**

Carrier, E. (2008) *Modeling the Choice of an Airline Itinerary and Fare Product Using Booking and Seat Availability Data*, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Koppelman, F.S., Coldren, G.C. and Parker, R.A. (2008) Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B*, 42(3), 263-73.

Lurkin, V., Garrow, L.A., Higgins, M. and Schyns, M. (2016) Accounting for price endogeneity in airline itinerary choice models: An application to Continental U.S. markets. *Working paper*, Georgia Institute of Technology, Atlanta, GA.

Zeid, M.A., Rossi, T.F. and Gardner, B. (2006). Modeling time of day choice in the context of tour and activity based models. *Paper presented at the 85th annual meeting of the Transportation Research Board*, Washington, DC.

**List of Tables and Figures**

Table 1: Independent Variables and Definitions

Table 2: Descriptive Statistics by Segment

Table 3: Model Results (Excluding Carrier and Time of Day Coefficients)

Figure 1: Departure Time of Day Preferences: Same TZ, distance ≤ 600 miles

Figure 2: Departure Time of Day Preferences: Same TZ, distance > 600 miles

Figure 3: Departure Time of Day Preferences: One TZ Westbound, distance ≤ 600 miles

Figure 4: Departure Time of Day Preferences: One TZ Westbound, distance > 600 miles

Figure 5: Departure Time of Day Preferences: One TZ Eastbound, distance ≤ 600 miles

Figure 6: Departure Time of Day Preferences: One TZ Eastbound, distance > 600 miles

Figure 7: Departure Time of Day Preferences: Two TZ Westbound

Figure 8: Departure Time of Day Preferences: Two TZ Eastbound

Figure 9: Departure Time of Day Preferences: Three TZ Westbound

Figure 10: Departure Time of Day Preferences: Three TZ Eastbound

Table 1: Independent Variables and Definitions

| Variable | Definition |
|---|---|
| ***Travel Time, Number of Connections, Connection, and Equipment Attributes*** | |
| Elapsed time | Elapsed time is defined as the difference between the arrival time at the itinerary destination and the departure time at the itinerary origin. All arrival and departure times are reported in Coordinated Universal Time (UTC), which accounts for time zone differences. |
| Number of connections | Number of itinerary connections. A value of zero indicates a nonstop itinerary and a value of one (two) indicates a single (double) connection. |
| Direct flight | A "direct flight" is one that has two flight legs. The operating carrier and operating flight number of the two flight legs are the same. A direct flight is defined to have zero connections. |
| Wide- or narrow-body<br><br>Regional jet or propeller | Equipment types include two categories. The first includes wide-body and narrow-body aircraft (no regional jets). The second includes narrow-body regional jets and propellers. For itineraries with more than one leg, the smallest equipment type is used. |
| ***Departure Time of Day*** | |
| $\text{Sin2pi\_DOW}_i\text{\_TripType}_j$<br>…<br>$\text{Cos6pi\_DOW}_I\text{\_TripType}_J$ | Departure time preferences are modeled using 126 terms. Three sin (sin2pi, sin4pi, sin6pi) and three cosin functions (cos2pi, cos4pi, cos6pi) apply to each departure day of week *i*=1,2,…,7 and three trip types *j*=outbound, inbound, one-way. |
| ***Price*** | |
| Average high yield fare<br><br>Average low yield fare | We calculate separate prices for high yield and low yield fare products. We include First, Business, and Unrestricted Coach products as high yield fares and the Restricted Coach and Other/Unknown products as low yield fares. We calculate average high yield and average low yield fares for each itinerary origin, destination, carrier, and level of service (nonstop/direct, single connection, and double connection). |
| ***Marketing Relationships*** | |
| Online | An online itinerary is one that has the same marketing and same operating carrier for all legs. |
| Codeshare | A codeshare itinerary is one that has the same marketing carrier for all legs, but different operating carriers. |
| Interline | An interline itinerary is one that has different marketing carriers. Only itineraries with two or more legs can be interline itineraries. |
| ***Carrier Preference*** | |
| Carrier_1<br>Carrier_2<br>…<br>Carrier_9 | For *k*=1,2,…,9 the indicator variable Carrier_k =1 if the itinerary operating carrier associated with an itinerary is carrier *k* and 0 otherwise. The itinerary operating carrier is defined as the carrier that operates the longest flight leg. The first eight terms represent carriers that each have more than 1% market share in the estimation data. All other carriers are combined into the Carrier_9 term. |

Table 2: Descriptive Statistics by Segment

| Segment | City pairs | Choice sets | Itins | Pax | Distance | | | Choice sets | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Mean | Max | Min Alts | Mean Alts | Max Alts |
| Same TZ, distance ≤ 600 mi. | 4,703 | 30,943 | 711,282 | 2,219,511 | 31 | 415.0 | 600 | 2 | 10.8 | 95 |
| Same TZ, distance > 600 mi. | 3,524 | 22,861 | 520,481 | 1,848,742 | 601 | 839.3 | 1,534 | 2 | 14.3 | 105 |
| One TZ Westbound, distance ≤ 600 mi. | 859 | 5,617 | 112,615 | 306,119 | 84 | 463.6 | 600 | 2 | 10.6 | 64 |
| One TZ Westbound, distance > 600 mi. | 3,864 | 24,820 | 498,999 | 1,466,815 | 601 | 993.9 | 1,925 | 2 | 15.1 | 127 |
| One TZ Eastbound, distance ≤ 600 mi. | 863 | 5,630 | 115,187 | 312,265 | 84 | 462.0 | 600 | 2 | 10.3 | 63 |
| One TZ Eastbound, distance > 600 mi. | 3,898 | 25,062 | 501,345 | 1,446,807 | 601 | 993.7 | 1,925 | 2 | 14.5 | 137 |
| Two TZ Westbound | 1,860 | 11,505 | 239,936 | 681,666 | 643 | 1,576.4 | 2,451 | 2 | 17.1 | 133 |
| Two TZ Eastbound | 1,823 | 11,267 | 233,113 | 684,627 | 643 | 1,571.4 | 2,451 | 2 | 15.3 | 93 |
| Three TZ Westbound | 1,121 | 6,732 | 165,428 | 509,346 | 1,578 | 2,203,3 | 2,774 | 2 | 21.3 | 156 |
| Three TZ Eastbound | 1,091 | 6,619 | 167,159 | 559,037 | 1,578 | 2,210.9 | 2,774 | 2 | 19.2 | 138 |
| TOTAL | **23,606** | **151,056** | **3,265,545** | **10,034,935** | | | | | | |

Table 3: Model Results (Excluding Carrier and Time of Day Coefficients)

| | **Uncorrected Model** | | **Corrected Model** | |
|---|---|---|---|---|
| **Variable** | **Parameter** | **t-statistic** | **Parameter** | **t-statistic** |
| Delta Air Lines (DL) (ref.) | 0 | **-** | **0** | **-** |
| United Airlines (UA) | -0.0411 | -30.8100 | -0.0076 | -5.5300 |
| American Airlines (AA) | 0.3802 | 284.5700 | 0.4308 | 305.9900 |
| US Airways (US) | 0.1917 | 142.5000 | 0.2002 | 148.4700 |
| Alaska Airlines (AS) | -0.0189 | -4.8900 | -0.0814 | -20.9400 |
| Jetblue Airways (B6) | -0.3048 | -104.2000 | -0.3788 | -126.6800 |
| Frontier Airlines (F9) | -0.2950 | -88.7600 | -0.4279 | -122.0600 |
| AirTran Airways (FL) | -0.9756 | -278.0200 | -1.0632 | -296.1600 |
| Other airlines | -0.3656 | -92.1300 | -0.4033 | -101.3000 |
| Average high yield fare ($) | -0.0025 | -257.8400 | -0.0036 | -265.4800 |
| Average low yield fare ($) | -0.0049 | -382.6400 | -0.0069 | -327.3900 |
| Elapsed time (min) | -0.0053 | -503.0300 | -0.0050 | -455.1000 |
| Number of connections | -2.4892 | -1,193.5500 | -2.5582 | -1,179.0400 |
| Direct flight | -2.2624 | -374.9800 | -2.3311 | -384.4800 |
| Regional jet or propeller (ref.) | 0 | - | **0** | - |
| Wide- or narrow-body | 0.4150 | 384.1400 | 0.3889 | 352.6900 |
| Online (ref.) | 0 | - | **0** | - |
| Codeshare | 0.2742 | 207.8700 | 0.2825 | 213.8000 |
| Interline | -0.2342 | -35.3800 | -0.1297 | -19.4200 |
| $\hat{\boldsymbol{\delta}}$ (residuals) | - | **-** | 0.0020 | **118.6000** |
| $\boldsymbol{LL}(\mathbf{0})$ | -32,652,846.05 | | -32,652,846.05 | |
| Final $\boldsymbol{LL}$ | **-26,239,664.32** | | **-26,232,323.64** | |
| Adj. $\boldsymbol{\rho}^2$ | 0.1964 | | 0.1966 | |

LL= log likelihood, Adj. $\boldsymbol{\rho}^2$ = 1 – (Final LL - #Attributes) / $\boldsymbol{LL}(\mathbf{0})$

Figure 1: Departure Time of Day Preferences: Same TZ, distance ≤ 600 miles

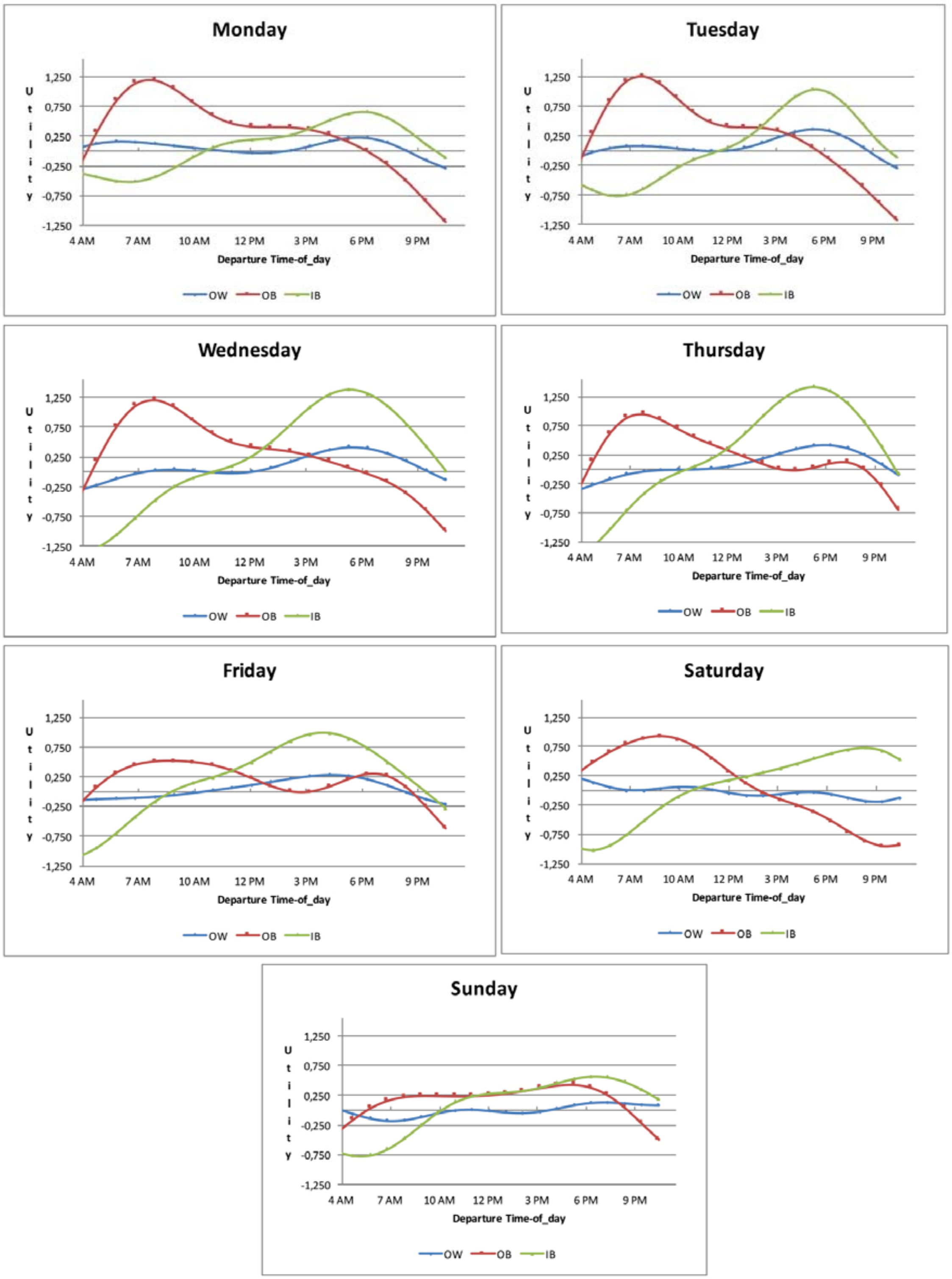

Figure 2: Departure Time of Day Preferences: Same TZ, distance > 600 miles

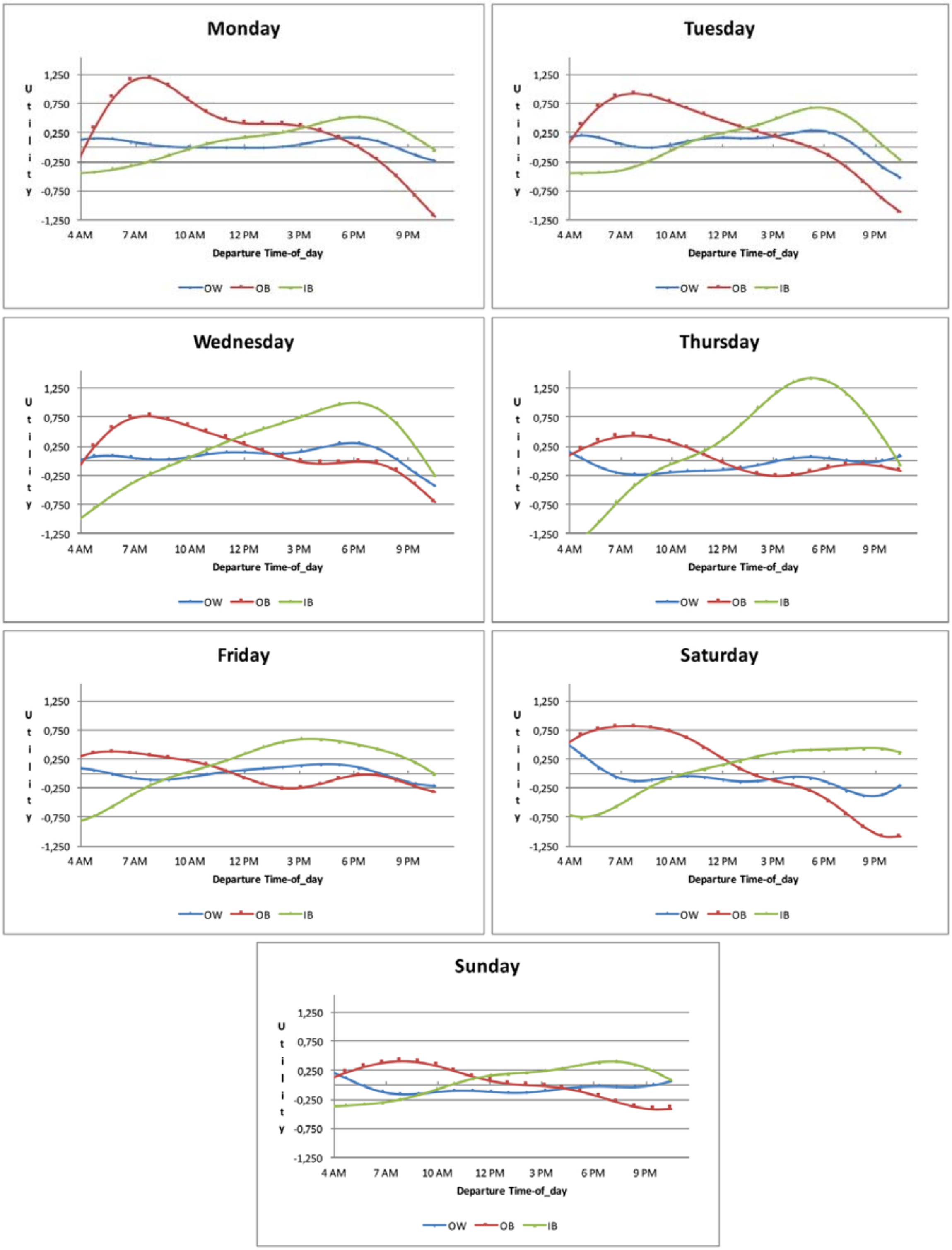

Figure 3: Departure Time of Day Preferences: One TZ Westbound, distance ≤ 600 miles

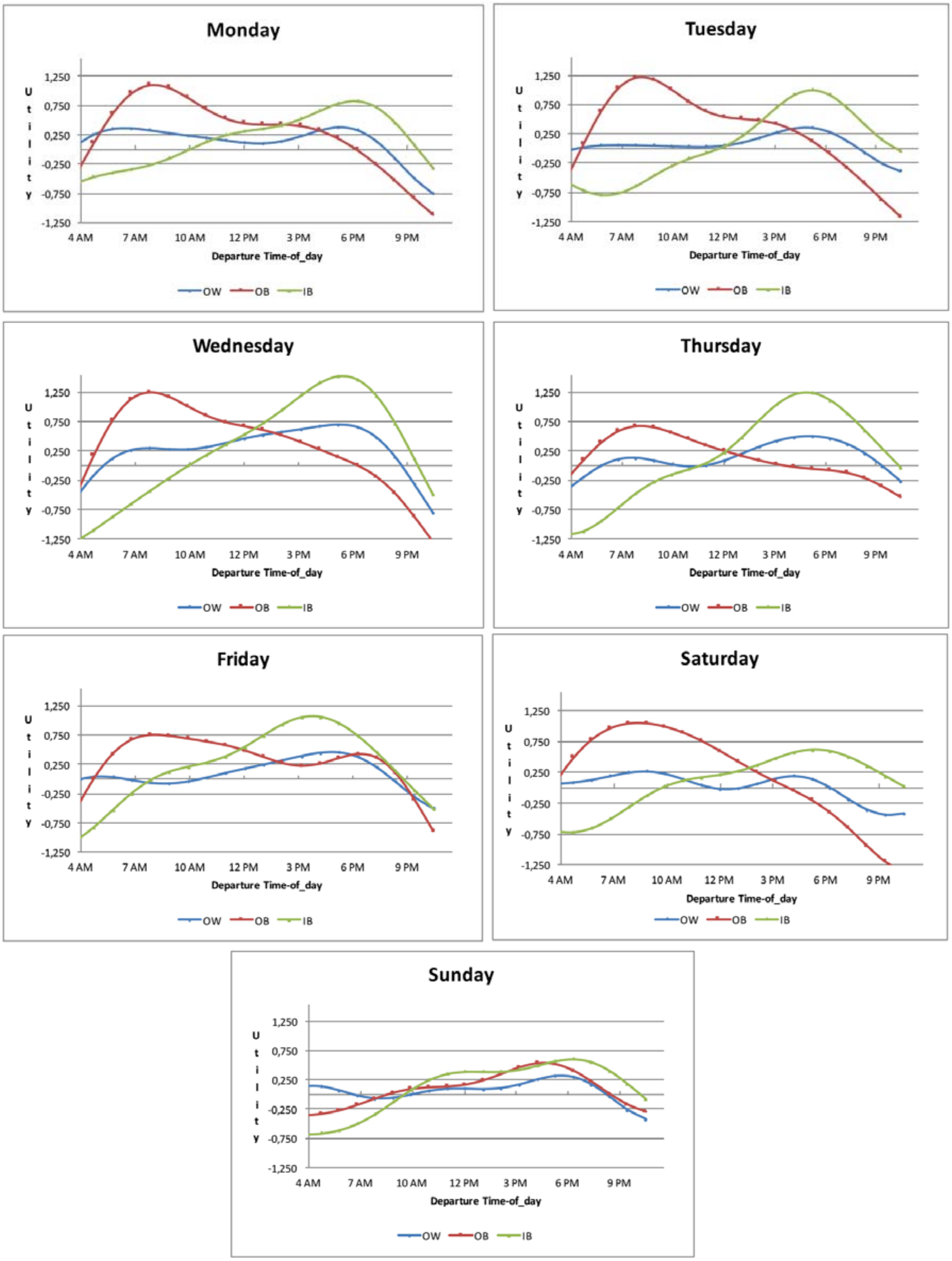

Figure 4: Departure Time of Day Preferences: One TZ Westbound, distance > 600 miles

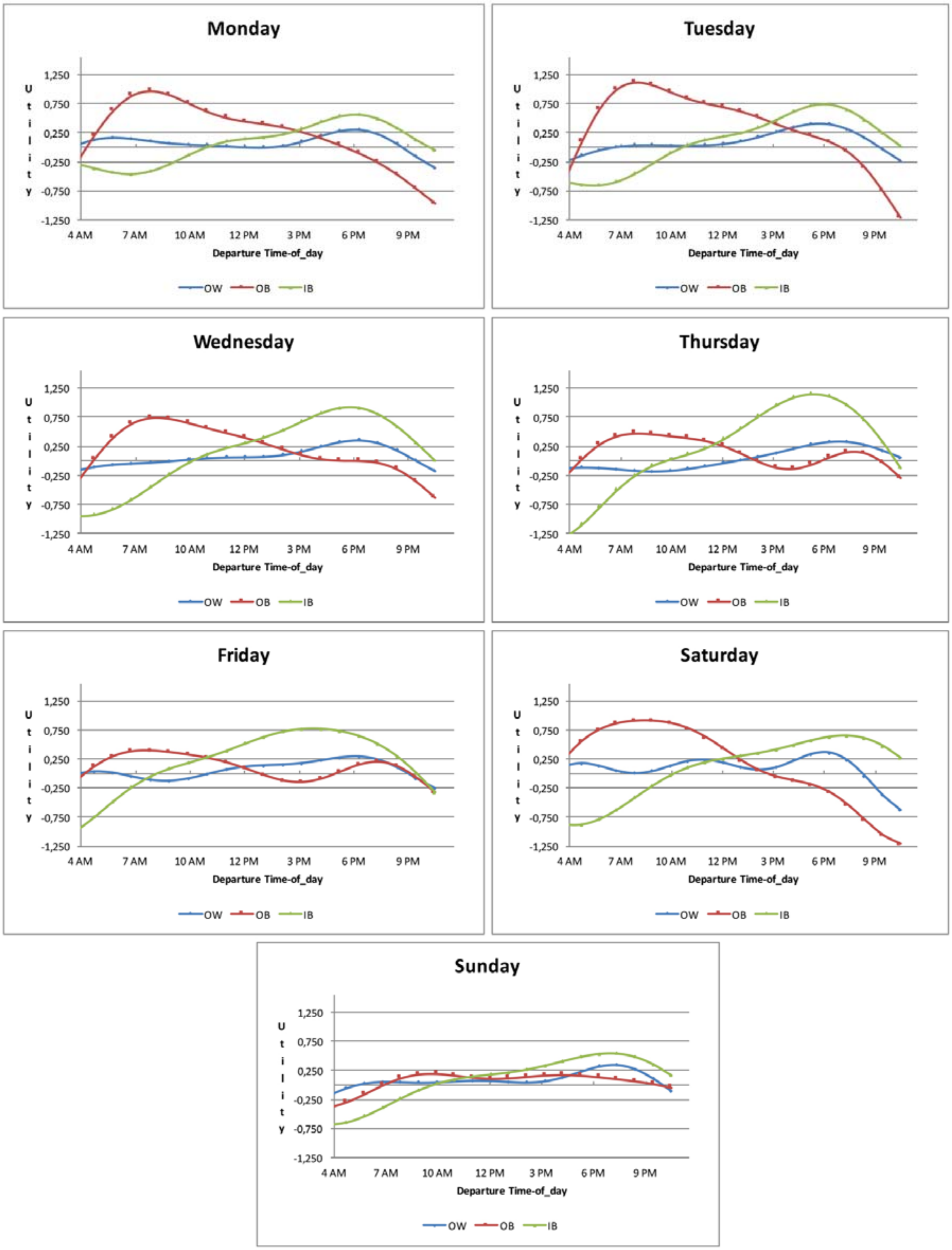

Figure 5: Departure Time of Day Preferences: One TZ Eastbound, distance ≤ 600 miles

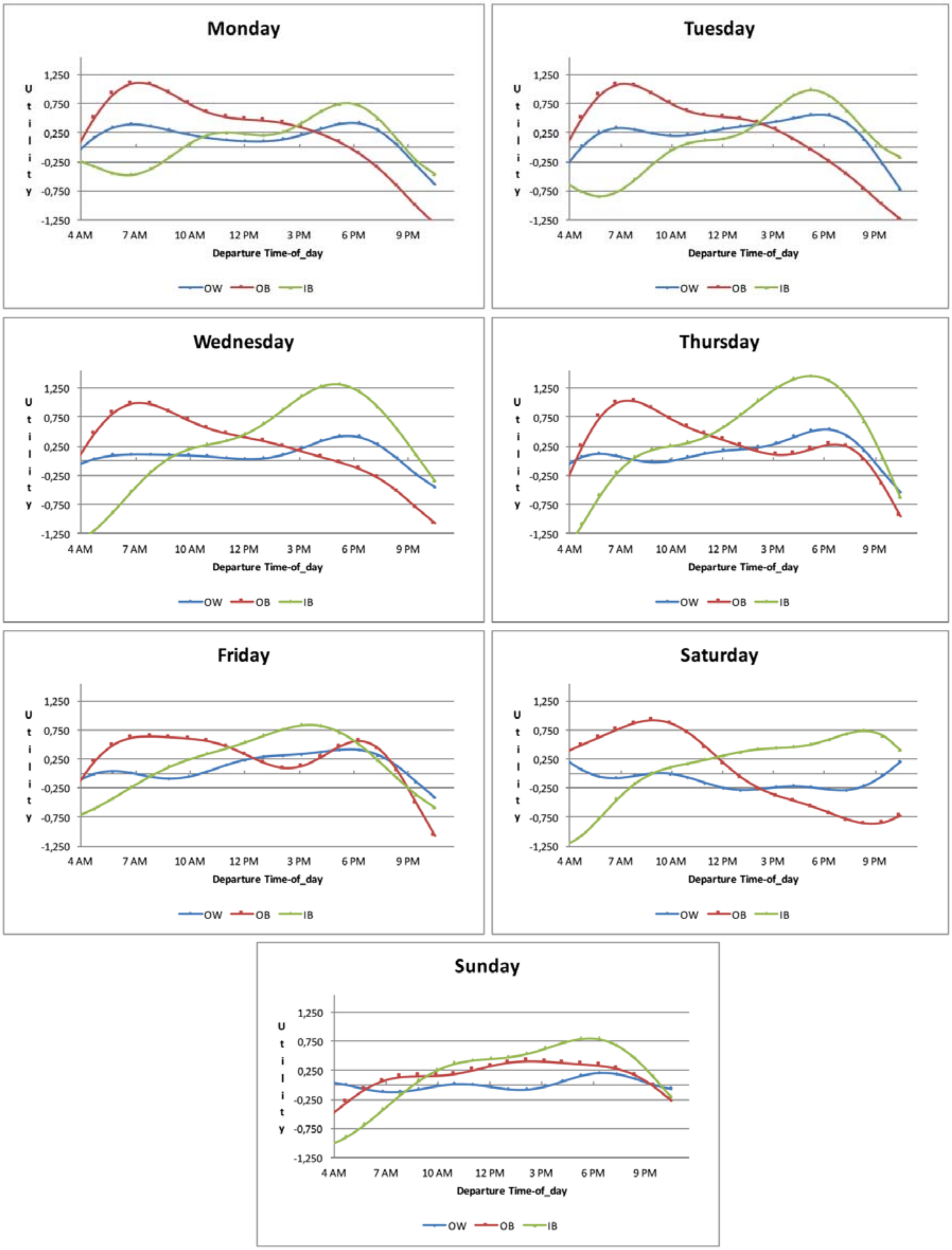

Figure 6: Departure Time of Day Preferences: One TZ Eastbound, distance > 600 miles

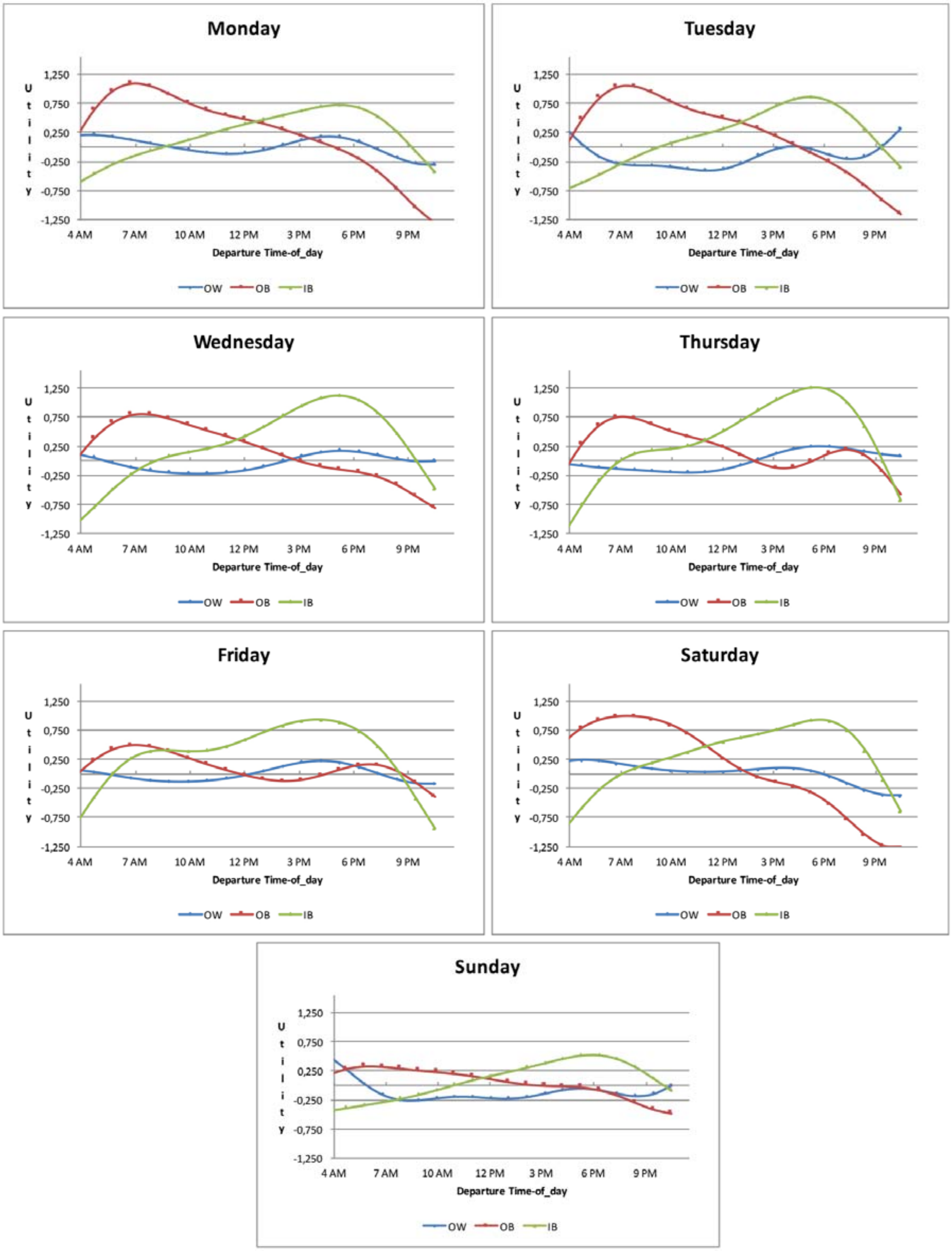

Figure 7: Departure Time of Day Preferences: Two TZ Westbound

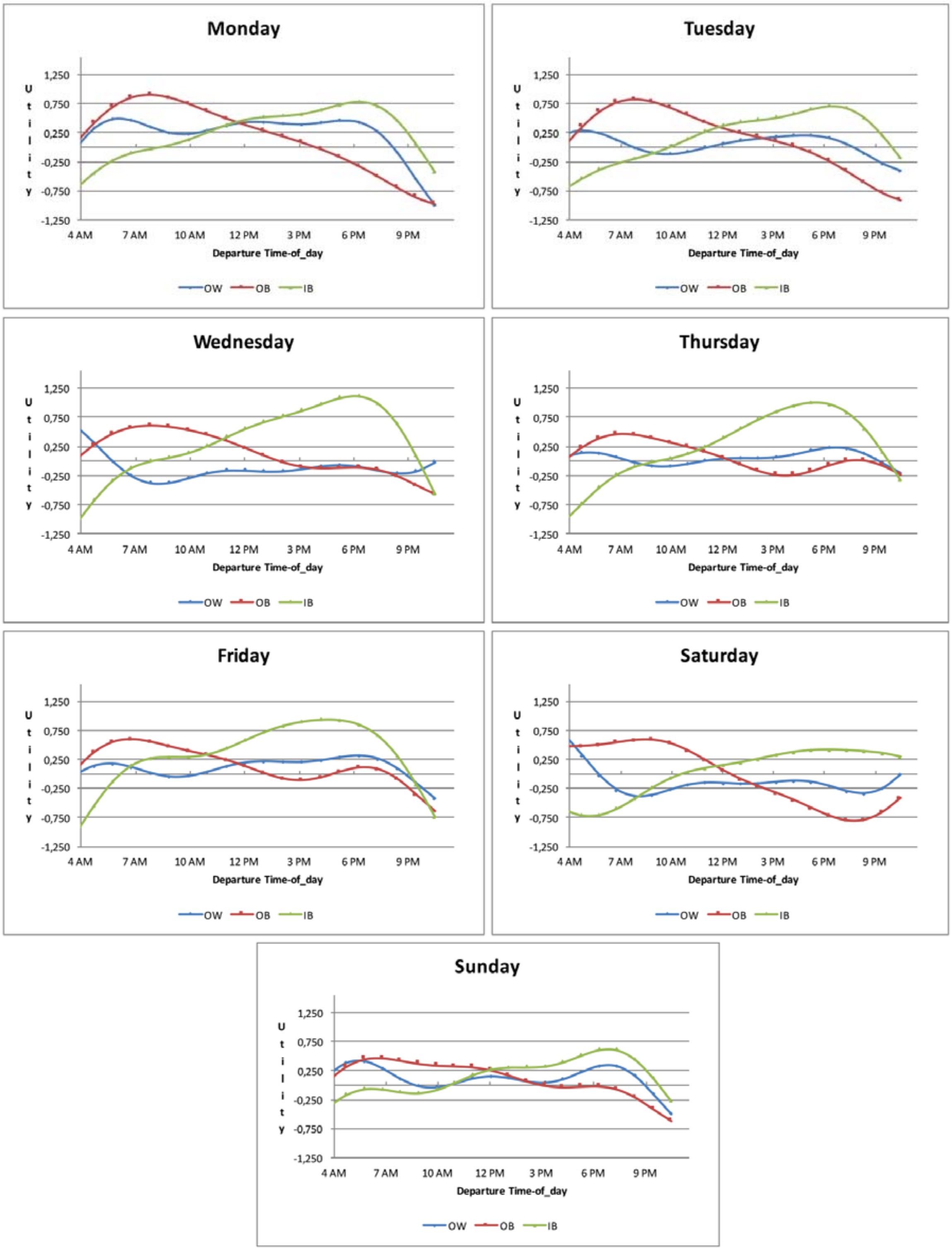

Figure 8: Departure Time of Day Preferences: Two TZ Eastbound

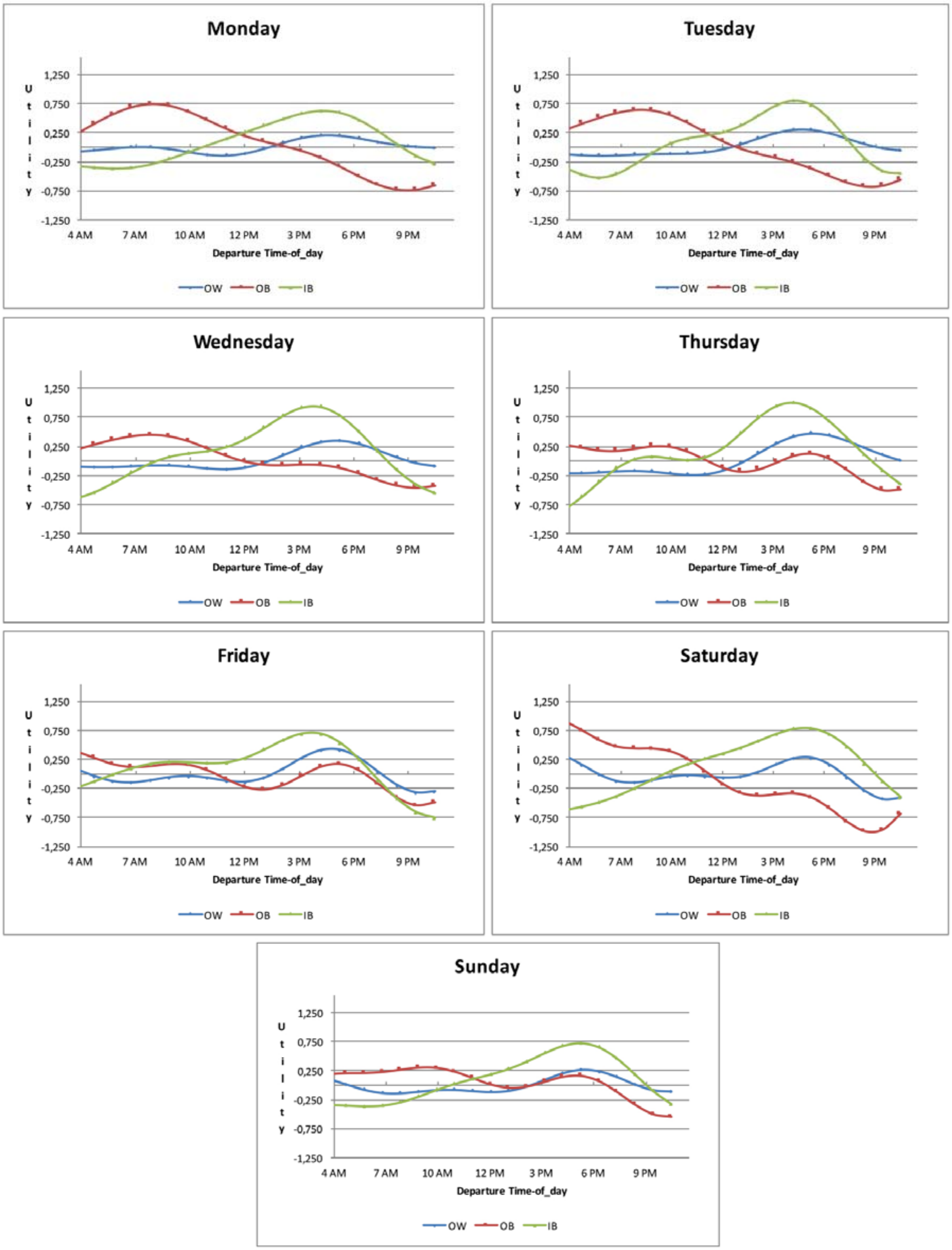

Figure 9: Departure Time of Day Preferences: Three TZ Westbound

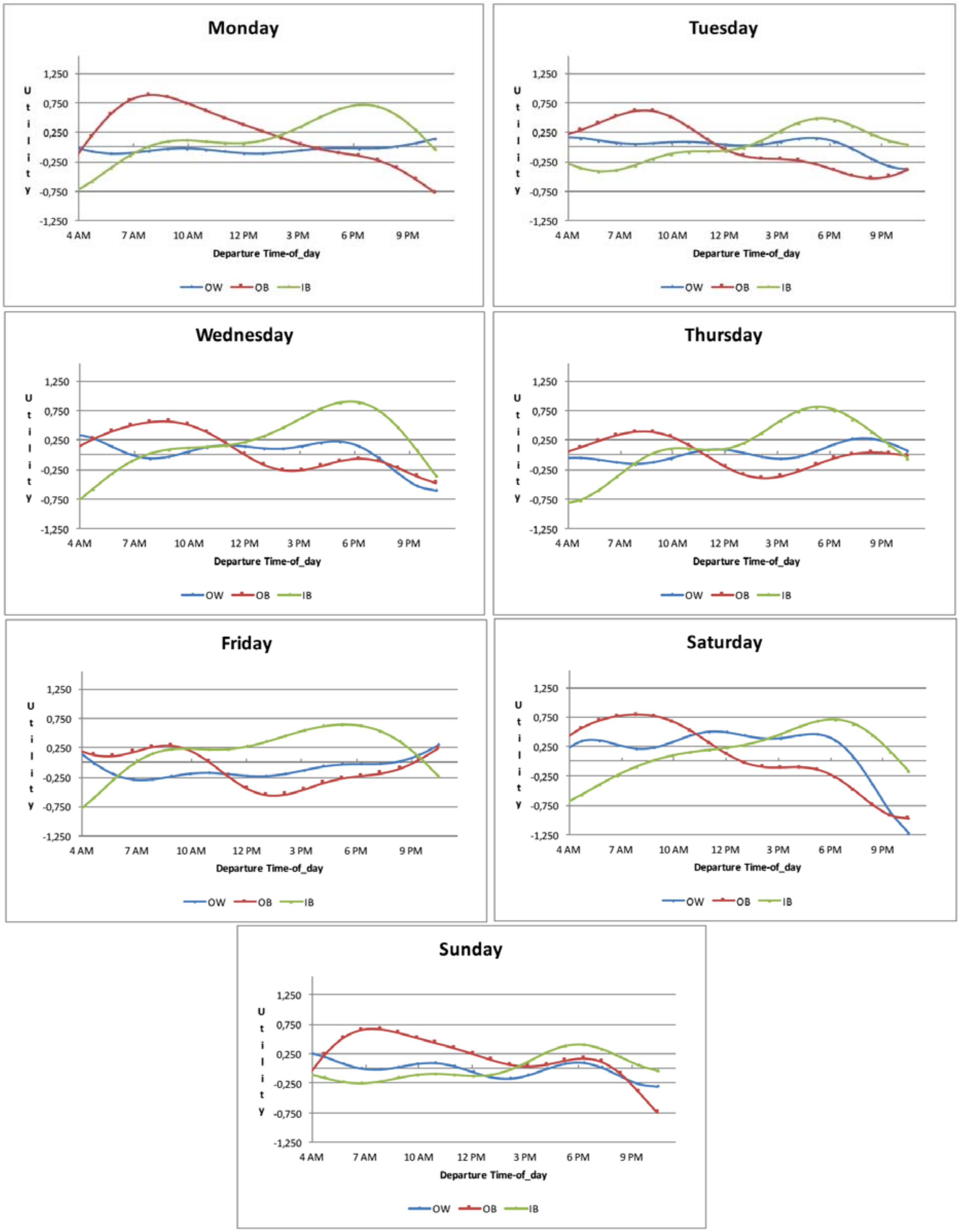

Figure 10: Departure Time of Day Preferences: Three TZ Eastbound