Are Transit Trips Symmetrical in Time and Space?

Evidence from the Twin Cities

Sang Gu Lee and Mark Hickman

This study exploits electronic fare collection data to examine the symmetry of boardings and alightings along a transit route. The symmetry of boardings and alightings is arguably the most important concept in the estimation of travel distances such as average trip lengths and passenger miles from data from entry-only fare collection systems. The paper shows the ways such data can be used to examine the symmetry of boardings and alightings through travel patterns in spatial and temporal dimensions. A novel method for aggregating stops, especially for the nearest stops in the opposite direction, is used to compare boardings in one direction with alightings in the opposite direction. Spatially, the method allows examination of the characteristics of boardings and alightings in a spatial dimension. Temporally, the method examines whether a specific and symmetric passenger flow is observed between specific periods (e.g., between morning and afternoon peaks). A case study of the Minneapolis–Saint Paul, Minnesota, region is performed by using automatically collected data from Metro Transit. Automatic fare collection data reveal considerable variation in passenger flow between specific periods. The use of automated passenger-counting data shows this variation to be statistically significant when both temporal and spatial symmetry are examined on an individual day.

Transit passengers' origins and destinations, sometimes represented by their boardings and alightings, provide indispensable information for most transportation applications, from strategic planning to traffic control and management. This information is often used to estimate passengers' travel distance, such as average trip lengths and passenger miles, that is a mandatory component of reporting in the National Transit Database. In most cases, the symmetry of boardings and alightings (SBAs) is arguably the most important concept, and this concept always assumes, on the basis of the general daily activity pattern (*1*), that the movements of transit passengers are reversible over the course of the day (*2–4*). This assumption, however, may be examined in a more disaggregate manner through (*a*) each individual's behavior or activity location and (*b*) time-varying (within-day) effects. Specifically, this paper investigates whether such symmetry occurs between specific periods within the day (e.g., within midday periods or between the morning and afternoon peak periods).

Previously, a disaggregate approach of SBA was complicated by the fact that bus stops (differing from rail stations) are usually represented as separate locations that serve different directions. In the stop-level model, bus stops often have a directional component, which can be treated as symmetric and nonsymmetric in relation to whether stops are considered the same in both directions (*5*). In parcel-level modeling, Furth et al. emphasize the need to determine separate service areas for a stop's boardings and alightings, as well as for each direction of travel (*6*). Because the two opposite stops often are not directly across from each other, it is not easy to assess the one-to-one matching between stops in opposing directions.

With data from Google's general transit feed specification (GTFS), a novel method for aggregating stops is proposed. More specifically, one-to-one matching between stops in opposite directions can be developed to capture the characteristics of boardings and alightings in space. This model is applied to investigate an effect of SBA over time with a much larger data set that has been automatically collected from various electronic technologies: automated fare collection (AFC) systems and automated passenger-counting (APC) systems. This approach can demonstrate an effect of SBA between specific periods over the course of the day. The goal of this study was to explore and implement a potential method of examining the symmetry of boardings and alightings by using automatically collected and other openly shared data.

LITERATURE REVIEW

Assumption of SBA

Boarding is commonly related to trip production (e.g., a residential area in the morning for an initial trip), while alighting is related to trip attraction (e.g., a workplace). A number of authors have presented the concept of symmetry in boarding and alighting for which the number of boardings in one direction at one location is equal to the number of alightings in the opposite direction at that same location, commonly over the course of a 24-h day.

Because of the lack of alighting information with electronic fare card data, Navick and Furth propose the symmetry assumption: that the boarding pattern for a route in one direction is equivalent to the alighting pattern in the opposite direction over the course of an entire day (*2*). They mention that the symmetry of boarding and alighting can be a valuable tool for estimating the alighting pattern on many routes, while some routes do not exhibit such symmetry. In any case, additional checking is required. Navick and Furth present a Kolmogorov–Smirnov (K-S) test to determine whether two boarding and alighting data sets show symmetry, through differences between the cumulative distribution of boardings in one direction and of alightings in the opposite direction.

Civil Engineering and Engineering Mechanics, University of Arizona, 1209 East 2nd Street, P.O. Box 210072, Tucson, AZ 85721-0072. Corresponding author: M. Hickman, mhickman@email.arizona.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2382, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 173–180.
DOI: 10.3141/2382-19

Johnson assumes that people start their return trip from the same place they ended the beginning trip, so that only boarding counts are needed to estimate average distance traveled (*3*). Notably, this assumption is affirmed via visual confirmation of boarding and alighting maps and tables. Chu emphasizes that alighting may not be always similar to boarding in the opposite direction for a given location (*7*). These quantities, however, are likely to be of a similar magnitude. Richardson also adapts the symmetry assumption that passenger flows on most public transport routes are reversible over the course of the day (*4*). Lu and Reddy assume the symmetric daily activity pattern ("conservation of passengers" and "equal and opposite passenger activities in opposing directions") (*1*) for the same reason as Navick and Furth (*2*): their data set has no record of alighting locations.

Transit Data

A huge amount of data is readily available from automated data collection systems. These systems include AFC and APC systems that describe spatial and temporal patterns of passengers' behavior. Manipulation and synthesis of these records can make more meaningful data sets that provide insight into passenger boarding, alighting, origins and destinations, and transfer behavior. These data sets can be easily integrated with other openly shared data such as GTFS and parcel-level land use data.

A large body of literature exists on the application of automatically collected and openly shared data for public transportation planning. That work, focusing on AFC and APC data, can be grouped into two categories: customer behavior analysis or travel pattern analysis (*8–13*) and demand forecasting and origin–destination estimation (*14–21*).

The idea of using GTFS data for various modeling purposes is becoming more common. Tomer et al. conduct a nationwide study to examine transit accessibility to jobs by using 168 GTFS data sets (*22*). Puchalsky et al. develop the Delaware Valley regional transit forecasting model by using GTFS data (*23*). The University of Arizona Transit Research Unit extensively uses GTFS data for detecting alighting locations, transfer stop locations, or both with fare card transaction data (*24*) and for aggregating transit stops (*25*). That unit has also developed a transit trip–based shortest-path algorithm that exploits GTFS data (*26*).

Stop Aggregation Model

Although data collection technology in public transit often resolves spatially to the level of the individual stop, this approach may have limitations. A single stop is often associated with a single direction and perhaps a single route. From the point of view of access to land and activities, other stops in the immediate vicinity, serving other directions and other routes, may also provide access to the same land uses and activities (*27*). The goal of the proposed stop-aggregation model (SAM), therefore, is to define a generalized definition of a "stop" that more closely matches the nature of locations serving as passenger origins and destinations. This definition could include landmarks (such as large trip generators or major intersections) or some combination of stops that collectively serve a passenger's activity location.

An aggregate area around a transit stop or station is defined by three parameters: (*a*) distance or proximity, measured by using euclidean and network distances in geographic information systems; (*b*) text in the description of the stop, queried using database tools in SQL; and (*c*) the catchment area. More details are described in Lee et al. (*25*). SAM provides considerable potential in various research areas, as shown in Figure 1. The use of an "aggregate" stop reduces some of the complexity of the analysis and also provides a more behavioral interpretation of stop locations and passenger behavior. Most importantly for this study, SAM provides a way of aggregating stops in opposite directions so that a given aggregate stop (or "stop group") can represent a given set of transit trip generators. When symmetry is examined, a boarding from a given stop group can be compared with an alighting within the same stop group.

DATA

Google's GTFS

GTFS is an open data format for transit schedules that was first released in 2005. GTFS has been used by Google since 2006 to incorporate transit information (e.g., stops, routes, and schedules) into the Google Maps application and is typically presented as a series of text files with comma-separated values. Currently, many major transit agencies, including Metro Transit in the Minneapolis–Saint Paul, Minnesota,

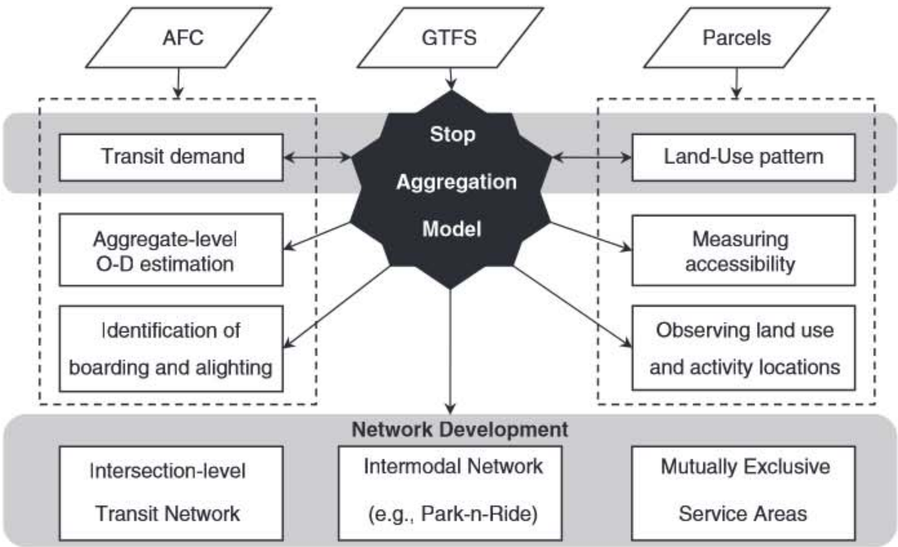

FIGURE 1 Various applications of SAM (O-D = origin–destination).

metropolitan area, make their data available through GTFS (*28*). These data help researchers, developers, and transit agencies efficiently share and retrieve transit schedule and service data.

AFC and APC Systems

The data for this study were obtained from Metro Transit and covered 1 month (November 2008) of data. In the AFC system, 2.17 million records were made by 79,775 fare cards (each with a unique identifier). At the time, Metro Transit operated a fleet of 1,010 buses over 186 routes (3.4 million records in APC system).

AFC data are the detailed fare card transaction data exhibiting peoples' transport habits over time. The AFC data include fare card identifiers; transaction dates and times; the type of fare paid; route and vehicle identifiers; and boarding locations, alighting locations, or both. The 24-h span from 3 a.m. on 1 day to 3 a.m. on the next day is considered a "day" because many buses end their trips after midnight and very few night owl trips are operated. Accordingly, weekday GTFS schedule data, with 488,105 records connecting bus stops with specific times of a bus arrival, are associated with location information of 14,601 stops served by Metro Transit. That agency's Route 6, for example, has 155 stop groups (from SAM).

APC data include the number of boardings and alightings at a stop and are recorded for about 30% of the operated bus trips by Metro Transit. Chu emphasizes that APC data do not distinguish between initial and transfer boardings (*7*). Despite this disadvantage, APC data can capture the travel characteristics of cash users who are not recorded in the AFC system but who still are strongly associated with transit trip generators. As a case study, Route 6 of Metro Transit in the Minneapolis–Saint Paul region is selected to implement the proposed methodology through use of both AFC and APC data. The route itself runs north–south along Hennepin Avenue from downtown Minneapolis to the southern suburb of Edina.

Estimation of Stop-Level Origins–Destinations

In an earlier study, the authors found that, by looking at successive fare card transactions of each passenger, his or her origins and destinations could be inferred at the stop level (*24*). In addition, some transactions that are indicated as a transfer within the AFC system can actually be regarded as new trips because of the duration between alighting on the initial trip and boarding on the connecting trip. The results of these transit origin–destination estimates were 28,260 inferred linked trips from 33,514 transactions on November 10, 2008. For this study, a subset of these linked trips (17,642 of 28,260), for which the passengers made exactly two linked trips during the day, are examined. This subset provides 8,821 unique fare card identifiers within the AFC data (Table 1).

TABLE 1 Fare Card Types by Number of Linked Trips

	Frequency of Linked Trips					
Farecard Type	1	2	3	4	5	6
Metro pass	2,512	3,036	72	10	0	0
Stored value FF	1,120	1,277	70	13	0	0
U-pass	1,994	3,104	385	180	15	6
C-pass	607	534	99	28	2	1
31D FF $85	350	398	44	12	3	1
31D ADA	181	162	19	12	3	1
31D FF $59	184	146	31	6	1	0
Other	213	164	22	2	1	0
Total	7,161	8,821	742	263	25	9

NOTE: FF = full fare; ADA = Americans with Disability Act.

METHODOLOGY

Spatial Pattern Analysis

To capture the actual location of a transit user for describing his or her activities and transit trips, SAM is used to provide one-to-one matching between stops in opposite directions. Figure 2*a* shows one possible case that can occur when a transit user makes two trips during the day, neither of which involves a transfer.

Figure 2*a* shows that the first AFC transaction is made during the early morning at a specific location on a specific route [outbound boarding (OB)]. The second AFC transaction (the assumed return trip) is made during the late afternoon at another location on the same route [inbound boarding (IB)]. Accordingly, the outbound alighting (OA) stop for the first trip and the inbound alighting (IA) stop for the second trip can be inferred. Two pairs, OB-IA and OA-IB, are thus established, and symmetry in space is assured. SAM captures the stop pairs of boarding and alighting. Some cases are fully captured (for Type 1 in Figure 2*b*, both OB-IA and OA-IB are in symmetry in space) or partially captured (for Types 2 and 3 in Figure 2, *c* and *d*, OB-IA and OA-IB, respectively, are in symmetry in space).

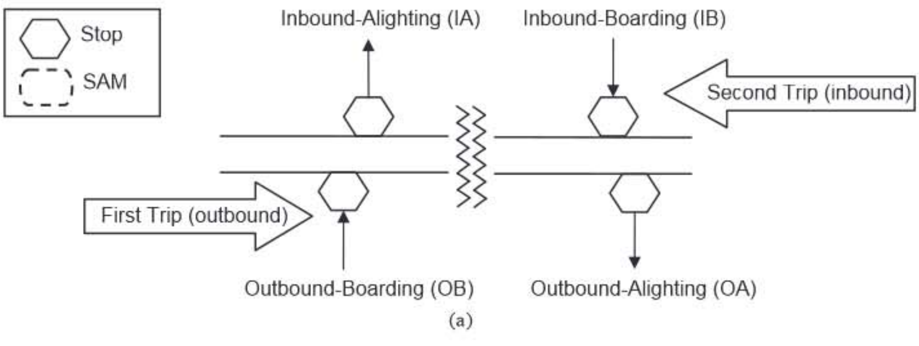

FIGURE 2 Approach of one-to-one matching between stops in opposite directions: (a) two pairs of boarding and alighting.

(continued on next page)

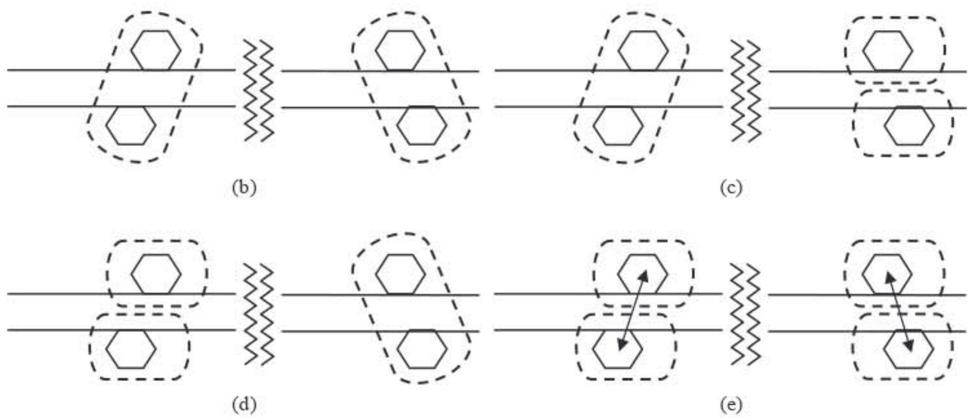

FIGURE 2 *(continued)* **Approach of one-to-one matching between stops in opposite directions: (b) Type 1 (OB-IA and OA-IB), (c) Type 2 (OB-IA only), (d) Type 3 (OA-IB only), and (e) Type 4 (none).**

However, other stop locations may not be successfully aggregated proximally by using SAM if there is a complicated configuration of a specific route for which SAM rules are not applied (e.g., a large separation of stops in opposing directions). In this case, the catchment-based SAM (for Type 4 in Figure 2*e*, none of the pairs is in symmetry in space, so temporal analysis is applicable) may still find the nearest stop in the opposite direction (*25*).

Temporal Pattern Analysis

To examine SBA across periods, specific periods are assumed. The periods used in this study are based on the Metro Transit fare policy (in which peak hours are between 6 and 9 a.m. and between 3 and 6:30 p.m.), as follows:

- Morning prepeak, before 6:00 a.m.;
- Morning peak, 6:00 to 9:00 a.m.;
- Midday, 9:00 a.m. to 3:00 p.m.;
- Afternoon peak, 3:00 to 6:30 p.m.; and
- Afternoon postpeak, after 6:30 p.m.

On the basis of the fare card transaction times, all trips (OB, OA, IB, and IA) are assigned to each period. APC data are also applied to these same periods. As a result, the symmetry associated with the combination of temporal and spatial patterns (e.g., OB at the morning peak, IA at the afternoon peak) can be examined.

K-S Test

The K-S test quantifies a distance between the empirical distribution functions of two samples. To accommodate Type 4, catchment-based SAM was applied to a specific trip on Route 6. In the test sets, which are boardings in one direction with alighting in the other direction (87 pairs from 99 northbound stops and 95 southbound stops), the test statistic can be compared with the critical value at a 95% level of confidence (a significance level of .05):

$$\text{K-S}_{0.05} = 1.36\sqrt{\frac{(n_1 + n_2)}{n_1 n_2}}$$

where n_1 and n_2 are the number of passengers counted in the two directions, respectively, and K-S_{max} is the greatest absolute deviation between the cumulative distribution of boardings in the subject direction and alightings in the opposite direction (*2*). The approach here to performing the K-S test focuses mainly on a specific passenger flow between the specific period pairs. This approach differs from the authors' previous work, which focused on symmetry across the entire day.

RESULTS

Analysis of Spatial and Temporal Patterns

The authors applied SAM to the full transit network in the Minneapolis–Saint Paul region by using the stop list from the Metro Transit GTFS. The result of SAM shows the reduction in the network complexity: 14,601 individual stops in the network are reduced to 7,951 aggregated stop groups (*25*). This study's AFC-based stop-level linked trips (17,642 trips by 8,821 unique identifiers) are applied to the aggregated stop groups for spatial (Types 1 through 4) and temporal (five periods) analyses.

Table 2 presents the results of spatial and temporal travel pattern analyses that used the AFC data. Spatially, while 2,300 of 8,821 identifiers (26%) fall into Type 4, 2,715 of 8,821 identifiers (31%) are fully addressed (Type 1). The spatial travel patterns for Types 2 and 3 show a significant difference in percentage of identifiers; this result suggests that home-based trips are more likely to capture an OB-IA pair. Temporally, morning and afternoon peak period trips encompass about 50% of the 17,642 trips.

Lee and Hickman found that activity and travel patterns in spatial and temporal dimensions differ significantly across fare card types, at least comparing Metro Pass (only available only to participating employers) and Stored Value cardholders with U-Pass (available only to University of Minnesota students) and C-Pass (offered to students at participating colleges) cardholders (*29*). Different patterns by fare card type seem to be consistent with the patterns observed in this SBA analysis.

After application of OB-IA pairs, OA-IB pairs, or both to a spatial model, considerable passenger flows between specific period pairs for Type 1 by fare card type is also investigated. Table 3 illustrates the four highest passenger flows of Type 1 for the period pairs of morning to midday, morning to afternoon, midday to midday, and midday to afternoon, by fare card type. This allows observation of the use of a specific type of fare card during specific period pairs. For Metro Pass holders, more than 90% of travel is concentrated in the morning-to-afternoon pair, which seems to be consistent with a

TABLE 2 Results of Spatial and Temporal Travel Pattern Analyses

Time Period		Type 1 2,715[a]		Type 2 2,564[a]		Type 3 1,242[a]		Type 4 2,300[a]	
From OB-OA	To IB-IA	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)
Morning prepeak	Morning peak	0	0	0	0	0	0	0	0
	Midday	1	1	1	1	1	1	1	0
	Afternoon peak	2	1	2	1	3	1	1	1
	Afternoon postpeak	0	0	0	0	0	0	0	0
Morning peak	Morning peak	0	0	0	0	1	1	1	1
	Midday	8	8	7	7	10	9	9	8
	Afternoon peak	49	48	58	58	33	34	47	46
	Afternoon postpeak	4	3	7	5	4	4	6	5
Midday	Midday	8	9	5	6	10	12	7	8
	Afternoon peak	15	16	9	11	16	16	16	17
	Afternoon postpeak	7	6	6	6	8	8	6	6
Afternoon peak	Afternoon peak	2	2	1	1	5	5	3	3
	Afternoon postpeak	4	4	3	3	8	8	4	3
Afternoon postpeak	Afternoon postpeak	0	1	0	0	1	1	0	0

NOTE: 8,821 fare cards.
[a]Identifiers.

conventional commuter travel pattern. For U-Pass holders, travel during the morning-to-afternoon pair is much less concentrated.

K-S Test Result

Finally, tests are performed for Route 6 during normal weekdays to compare the cumulative distribution of northbound (southbound) boardings with southbound (northbound) alightings, by using only the APC data. Because 1 day may have unique characteristics, the K-S test is performed by using data of the "specific time period pairs" versus "across the entire day." Comparison between K-S$_{max}$ and K-S$_{.05}$ (Table 4) shows different results in SBA. Most cases of symmetry across the entire day fail the K-S test, while many of the tests between specific period pairs indicate symmetry at a 95% confidence interval with the K-S test. As an example, the test for Route 6 in both directions on November 17, 2008, is shown in Figure 3.

The K-S test results seem to suggest symmetry, at least between the travel during the morning and afternoon peak periods, between the midday and the afternoon postpeak periods, and during midday. This symmetry seems to be confirmed at least on specific days. However, if one examines the total for all weekdays (the Total rows in Table 4 for each period pair), in general, these fail the K-S test, indicating that the hypothesis of symmetry cannot be accepted with at least 95% confidence. One might then conclude that, on any given day, symmetry by period may be observed, while on an average day, symmetry by time period is not observed.

Discussion of Results

Although the authors have had considerable success at inferring transit origins and destinations from AFC data, not all fare cards illustrate full tours in the course of a given day (*24*). These data can lead to some confusion about the true origins and destinations for some of the linked trips. For example, the 7,161 linked trips (one linked trip per day reported in the AFC data) in Table 1 seem to indicate that passengers made only a single one-way trip. However, the majority of these fare cards indicate a second or subsequent transaction. Although interpretation of these data is somewhat ambiguous,

TABLE 3 Passenger Flows in Type 1 by Fare Card Type over Selected Periods

	Morning Peak to Midday		Morning Peak to Afternoon Peak		Midday to Midday		Midday to Afternoon Peak		Total	
Farecard Type	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)	OB-IA (%)	OA-IB (%)
Metro pass	8	10	56	57	4	4	7	10	37	37
Stored value FF	7	9	17	17	11	11	12	12	14	15
U-pass	68	65	18	17	67	65	70	68	38	37
C-pass	7	5	2	2	5	8	5	4	3	4
31D FF $85	3	3	5	5	1	2	2	3	4	4
31D ADA	6	6	0	0	4	4	1	1	1	1
31D FF $59	2	2	0	0	4	4	1	0	1	1
Other	0	0	1	1	4	3	2	2	1	1
Total	100	100	100	100	100	100	100	100	100	100

TABLE 4 Test Results for Symmetry Hypothesis for Route 6 by Periods

Date	Boarding Direction	Number of Trips	Total Boardings (n_1)	Alighting Direction	Total Alightings (n_2)	K-S_{max}	K-$S_{0.05}$	Accept Symmetry Hypothesis?
Morning Peak to Afternoon Peak								
11/17/08	NB	2	141	SB	142	0.1392	0.1617	Yes
	SB	2	145	NB	162	0.1411	0.1555	Yes
11/18/08	NB	2	146	SB	144	0.1867	0.1597	No
	SB	2	142	NB	149	0.1488	0.1595	Yes
11/19/08	NB	2	139	SB	132	0.1002	0.1653	Yes
	SB	2	127	NB	151	0.1468	0.1637	Yes
11/20/08	NB	2	156	SB	152	0.0693	0.1550	Yes
	SB	2	145	NB	172	0.0562	0.1533	Yes
11/21/08	NB	1	66	SB	79	0.1623	0.2268	Yes
	SB	1	78	NB	70	0.0945	0.2239	Yes
Total	NB	9	648	SB	649	0.0543	0.0755	Yes
	SB	9	637	NB	704	0.0806	0.0744	No
Midday to Afternoon Postpeak								
11/17/08	NB	1	41	SB	44	0.1585	0.2952	Yes
	SB	1	41	NB	47	0.3939	0.2906	No
11/18/08	NB	1	60	SB	67	0.1953	0.2417	Yes
	SB	1	64	NB	58	0.1633	0.2466	Yes
11/19/08	NB	1	54	SB	71	0.1886	0.2456	Yes
	SB	1	66	NB	51	0.2326	0.2536	Yes
11/20/08	NB	1	62	SB	84	0.2281	0.2277	No
	SB	1	80	NB	53	0.2608	0.2409	No
11/21/08	NB	1	48	SB	67	0.1595	0.2572	Yes
	SB	1	65	NB	52	0.2346	0.2530	Yes
Total	NB	5	265	SB	333	0.1510	0.1120	No
	SB	5	316	NB	261	0.1967	0.1138	No
Midday to Midday								
11/17/08	NB	5	274	SB	342	0.1144	0.1103	No
	SB	5	324	NB	288	0.1065	0.1101	Yes
11/18/08	NB	3	182	SB	199	0.1270	0.1395	Yes
	SB	3	192	NB	197	0.0953	0.1379	Yes
11/19/08	NB	5	335	SB	328	0.1198	0.1056	No
	SB	5	316	NB	360	0.0878	0.1048	Yes
11/20/08	NB	5	333	SB	308	0.0823	0.1075	Yes
	SB	5	300	NB	330	0.0700	0.1085	Yes
11/21/08	NB	4	216	SB	286	0.1242	0.1226	No
	SB	4	257	NB	229	0.0920	0.1236	Yes
Total	NB	22	1,340	SB	1,463	0.0936	0.0514	No
	SB	22	1,389	NB	1,404	0.0668	0.0515	No

NOTE: NB = northbound; SB = southbound.

the fare card transactions with two linked trips per passenger per day are used as a starting point for this study and represent more than 60% of the entire data set (17,642 of 28,260 linked trips). Of course, some bias exists in this data set because of the assumed spatial symmetry of two linked trips per day. Additional analysis of the AFC data to infer individuals' movements is a promising line of future study.

The APC data show a possible imbalance problem (i.e., in one vehicle trip, the total number of boardings may not equal the total number of alightings). This discrepancy results either from APC device errors (intrinsic to APC technology) or from some passengers possibly remaining on board after a given vehicle trip to travel in the reverse direction. Despite the limitations and caveats that accompany the APC data, they provide important information for understanding boarding and alighting activities at the stop level.

CONCLUSIONS

This study demonstrates boarding and alighting patterns in a more disaggregate manner and examines individuals' movements in spatial and temporal dimensions. In contrast to previous studies, this one provides a systematic matching technique for aggregating stops (SAM) that allows for a more structured analysis of the symmetry of boardings and alightings. A much larger data set that has been automatically collected from various electronic technologies (AFC and APC systems) is applied to examine the symmetry of boardings and alightings spatially and temporally.

The results of the analysis of spatial travel patterns support a part of Lu and Reddy's assumption of "equal and opposite passenger activities in opposing directions" (*1*). Analysis of temporal travel patterns at a disaggregate level can reduce aggregation errors (e.g., data from the

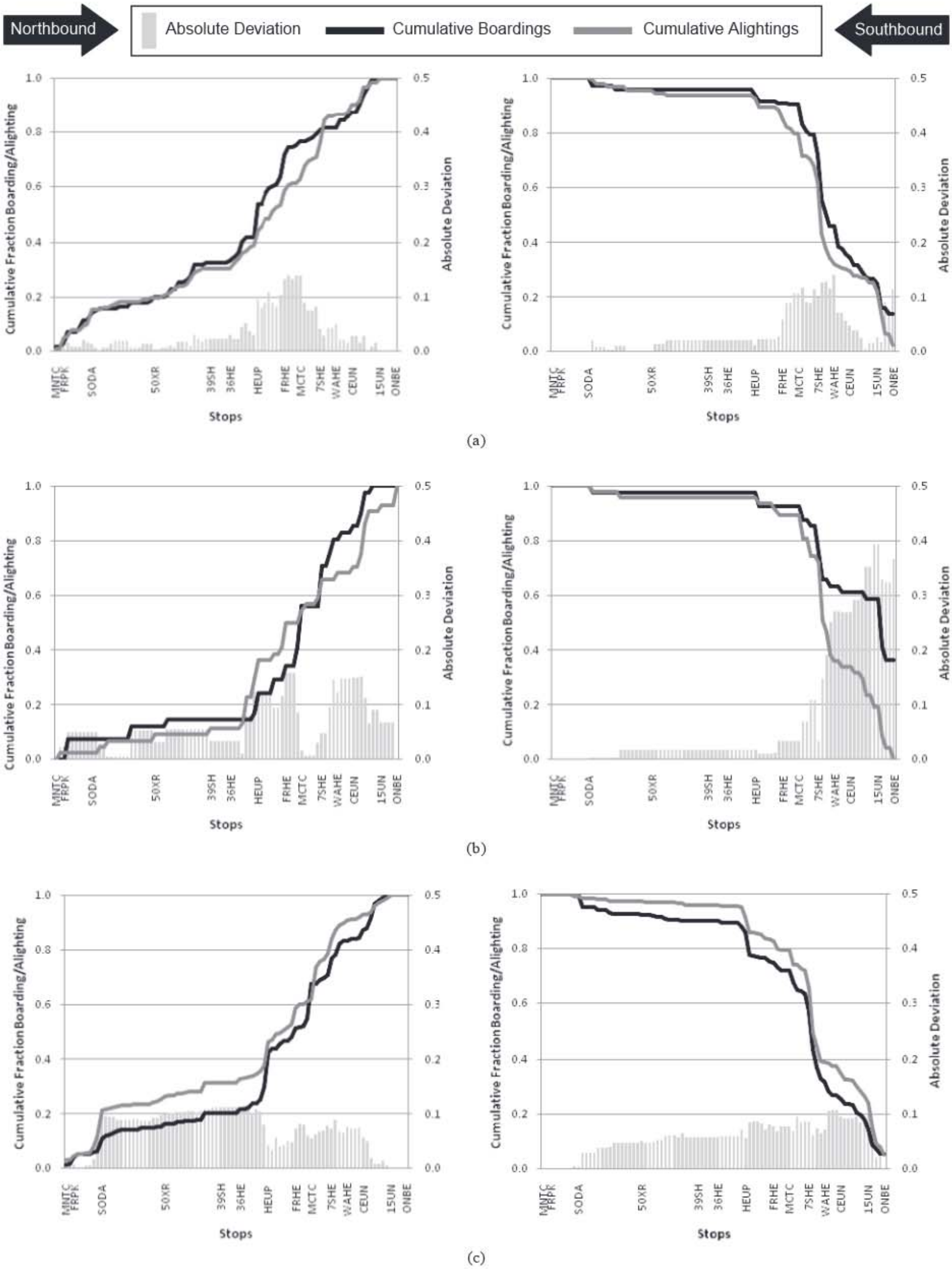

FIGURE 3 Selected K-S tests for Route 6 (November 17, 2008): (a) morning peak to afternoon peak, (b) midday to afternoon postpeak, and (c) midday to midday.

course of the entire day), and the authors observe a considerable passenger flow that depends on fare card type between specific periods in the AFC data. Finally, by using APC data to examine both temporal and spatial symmetry, the authors observe symmetry by period pairs on an individual day. However, the authors find that symmetry is not generally found between these same period pairs when total boardings and alightings for a full set of weekdays are compared.

ACKNOWLEDGMENTS

The authors thank Metro Transit and MetroGIS for providing the data analyzed in the study. Specific recognition goes to Paul Krech and John Levin of the scheduling department at Metro Transit, who generated the GTFS data set from the November 2008 schedules.

REFERENCES

1. Lu, A., and A. V. Reddy. Algorithm to Measure Daily Bus Passenger Miles Using Electronic Farebox Data for National Transit Database Section 15 Reporting. In Transportation Research Record: Journal of the Transportation Research Board, No. 2216, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 19–32.
2. Navick, D. S., and P. G. Furth. Estimating Passenger Miles, Origin–Destination Patterns, and Loads with Location-Stamped Farebox Data. In Transportation Research Record: Journal of the Transportation Research Board, No. 1799, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 107–113.
3. Johnson, A. Bus Transit and Land Use: Illuminating the Interaction. Journal of Public Transportation, Vol. 6, No. 4, 2003, pp. 21–39.
4. Richardson, T. Estimating Average Distance Travelled from Bus Boarding Counts. Road and Transport Research Journal, Vol. 14, No. 4, 2005, pp. 77–89.
5. Peng, Z., and K. J. Dueker. Spatial Data Integration in Route-Level Transit Demand Modeling. Journal of Urban and Regional Information Systems Association, Vol. 7, No. 1, 1995, pp. 26–37.
6. Furth, P. G., M. C. Mekuria, and J. L. SanClemente. Parcel-Level Modeling to Analyze Transit Stop Location Changes. Journal of Public Transportation, Vol. 10, No. 2, 2007, pp. 73–91.
7. Chu, X. Ridership Models at the Stop Level. Report BC137-31. National Center for Transit Research, Florida Department of Transportation, Tallahassee, 2004.
8. Agard, B., C. Morency, and M. Trépanier. Mining Public Transport User Behavior from Smartcard Data. Presented at 12th International Federation of Automatic Control Symposium on Information Control Problems in Manufacturing: INCOM 2006, Saint-Etienne, France, May 17–19, 2006.
9. Morency, C., M. Trépanier, and B. Agard. Analysing the Variability of Transit Users Behaviour with Smart Card Data. Presented at 9th International Intelligent Transportation Systems Conference, Toronto, Ontario, Canada, Sept. 17–20, 2006.
10. Utsunomiya, M., J. Attanucci, and N. H. M. Wilson. Potential Uses of Transit Smart Card Registration and Transaction Data to Improve Transit Planning. In Transportation Research Record: Journal of the Transportation Research Board, No. 1971, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 119–126.
11. Morency, C., M. Trépanier, and B. Agard. Measuring Transit Use Variability with Smart-Card Data. Transport Policy, Vol. 14, No. 3, 2007, pp. 193–203.
12. Agard, B., C. Morency, and M. Trépanier. Mining Smart Card Data from an Urban Transit Network. In Encyclopedia of Data Warehouse and Mining, 2nd ed. (J. Wang, ed.), Information Science Reference, Hershey, Pa., 2008, pp. 1292–1302.
13. Seaborn, C., J. P. Attanucci, and N. H. M. Wilson. Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. In Transportation Research Record: Journal of the Transportation Research Board, No. 2121, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 55–62.
14. Barry, J. J., R. Newhouser, A. Rahbee, and S. Sayeda. Origin and Destination Estimation in New York City with Automated Fare System Data. In Transportation Research Record: Journal of the Transportation Research Board, No. 1817, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 183–187.
15. Zhao, J., A. Rahbee, and N. H. M. Wilson. Estimating a Rail Passenger Trip Origin–Destination Matrix Using Automatic Data Collection Systems. Computer-Aided Civil and Infrastructure Engineering, Vol. 22, No. 5, 2007. pp. 376–387.
16. Trépanier, M., N. Tranchant, and R. Chapleau. Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, Vol. 11, No. 1, 2007, pp. 1–14.
17. Chu, K. K. A., and R. Chapleau. Enriching Archived Smart Card Transaction Data for Transit Demand Modeling. In Transportation Research Record: Journal of the Transportation Research Board, No. 2063, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 63–72.
18. Farzin, J. M. Constructing an Automated Bus Origin–Destination Matrix Using Farecard and Global Positioning System Data in São Paulo, Brazil. In Transportation Research Record: Journal of the Transportation Research Board, No. 2072, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 30–37.
19. Barry, J. J., R. Freimer, and H. Slavin. Use of Entry-Only Automatic Fare Collection Data to Estimate Linked Transit Trips in New York City. In Transportation Research Record: Journal of the Transportation Research Board, No. 2112, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 53–61.
20. Reddy, A., A. Lu, S. Kumar, V. Bashmakov, and S. Rudenko. Entry-Only Automated Fare-Collection System Data Used to Infer Ridership, Rider Destinations, Unlinked Trips, and Passenger Miles. In Transportation Research Record: Journal of the Transportation Research Board, No. 2110, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 128–136.
21. Wang, W., J. P. Attanucci, and N. H. M. Wilson. Bus Passenger Origin–Destination Estimation and Related Analyses Using Automated Data Collection Systems. Journal of Public Transportation, Vol. 14, No. 4, 2011, pp. 131–150.
22. Tomer, A., E. Kneebone, R. Puentes, and A. Berube. Missed Opportunity: Transit and Jobs in Metropolitan America Cities. Presented at Brookings Institution's Metropolitan Opportunity Series, Washington, D.C., May 12, 2011.
23. Puchalsky, C. M., D. Joshi, and W. Scherr. Development of a Regional Forecasting Model Based on Google Transit Feed. Presented at 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012.
24. Nassir, N., A. Khani, S. G. Lee, H. Noh, and M. Hickman. Transit Stop-Level Origin–Destination Estimation Through Use of Transit Schedule and Automated Data Collection System. In Transportation Research Record: Journal of the Transportation Research Board, No. 2263, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 140–150.
25. Lee, S. G., M. Hickman, and D. Tong. Stop Aggregation Model: Development and Application. In Transportation Research Record: Journal of the Transportation Research Board, No. 2276, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 38–47.
26. Khani, A., S. Lee, M. Hickman, H. Noh, and N. Nassir. Intermodal Path Algorithm for Time-Dependent Auto Network and Scheduled Transit Service. In Transportation Research Record: Journal of the Transportation Research Board, No. 2284, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 40–46.
27. Lee, S., M. Hickman, and D. Tong. Development of a Temporal and Spatial Linkage Between Transit Demand and Land Use Patterns. Journal of Transport and Land Use, forthcoming.
28. General Transit Feed Specification Reference. Google. http://code.google.com/transit/spec/transit_feed_specification.html. Accessed March 2011.
29. Lee, S., and M. Hickman. Travel Pattern Analysis Using Smart Card Data of Regular Users. Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.

All errors remain with the authors.

The Traveler Behavior and Values Committee peer-reviewed this paper.