Extraboard team sizing: An analysis of short unscheduled absences among regular transit drivers

Ehab I. Diab
Ph.D. candidate
School of Urban Planning
McGill University
Suite 400, 815 Sherbrooke St. W.
Montréal, Québec, H3A 2K6
Canada
Tel.: 514-549-0093
Fax: 514-398-8376
E-mail: ehab.diab@mail.mcgill.ca

Rania A. Wasfi
Ph.D. candidate
Department of Geography
McGill University
805 Sherbrooke St. West
Montreal, QC
H3A 2K6
Tel.: 514-398-8741
Fax: 514-398-8376
Email: rania.wasfi@mail.mcgill.ca

Ahmed M. El-Geneidy
Associate Professor
School of Urban Planning
McGill University
Suite 400, 815 Sherbrooke St. W.
Montréal, Québec, H3A 2K6
Canada
Tel.: 514-398-8741
Fax: 514-398-8376
E-mail: ahmed.elgeneidy@mcgill.ca

April 2014

Word count 6163, 3 tables and 1 figure

Paper published in Transportation Research Part A: Policy and Practice

For citation please use: Diab, E., Wasfi, R., & El-Geneidy, A. (2014). Extraboard team sizing: An analysis of short unscheduled absences among regular transit operators. *Transportation Research Part A: Policy and Practice*, 66, 27-38.

Research Highlights

1- This article focuses on regular bus drivers' short-duration unscheduled absences.
2- It analyzes absenteeism data at the aggregate level of garage-day-period.
3- A multilevel regression model is generated to investigate regular drivers' absence
4- Sensitivity analyses are conducted to evaluate the benefits of using modeling.
5- It provides transit planners with a methodology to support extraboard planning practice.

ABSTRACT

Several factors contribute to short-duration unscheduled absences of bus transit drivers. This article aims to understand these factors at the aggregate level and to anticipate future total absence that will need to be filled for a large-size transit operator. The aggregate level is defined as the total number of regular driver absences per garage, day of week and time period that need to be covered by the extraboards. This study analyzes absenteeism data obtained from OC Transpo, the transit provider of the city of Ottawa, Canada. A multilevel regression model is generated to investigate regular drivers' absences. The short-unscheduled absence is estimated in relation to temporal factors, drivers' personal characteristics, aspects of assigned work, and service delivery characteristics. Furthermore, using the model's coefficients, sensitivity analyses are conducted to demonstrate the advantages of this technique over traditional ones adopted by various transit agencies. This study provides transit planners and policy makers with a practical methodology that can be used to support extraboard planning practice and help reduce the incidence of missed trips due to absences while having the appropriate size of extraboard drivers.

Keywords: Absence, short-duration absence, unscheduled absence, workforce planning, transit drivers, extraboard planning, spare-operators planning.

INTRODUCTION

Transit providers maintain a reserve pool of back-up drivers, known as the spare-board or extraboard, to fill open work assignments when regular-duty drivers are absent, non-scheduled service occurs, or work left open after the booking process has been completed. Non-scheduled service, work left open after the booking process, and employees' planned scheduled absences are known far enough in advance for workforce planning, which are not in this paper's scope. In contrast, unscheduled absences among bus drivers vary considerably on a daily basis, requiring transit agencies to anticipate for them in extraboard planning. Thus, extraboard planning predicts the amount of open work per day that will need to be filled during a 2 to 4 months booking period, which lead in many cases to acquiring more extraboard drivers to cover all short unscheduled absences as well as other needs. Short unscheduled absences are defined as unscheduled absences extending from one to three days, which account for most of the day-to-day variation in the amount of open work that must be filled by extraboard drivers. Short unscheduled absences occur for many reasons, such as short-term illness, family commitments, late arrival to work, social reasons and many other unreported reasons.

The extraboard planning process, which is known also as 'sizing,' is an essential exercise related to predicting unscheduled absences among regular-duty drivers. Transit agencies might have to cancel trips or decrease frequency along certain routes if the size of the extraboard is insufficient to deal with the variation in these absences. Trip cancelation has a direct impact on passengers' perception of service quality, since waiting time is doubled and buses will be crowded. If this gap between extraboard and absences increases, the number of canceled trips increases, leading to a decrease in passenger trust and loyalty, which consequently decreases transit use and agencies' revenues (Furth & Muller, 2007; Perk, Flynn, & Volinski, 2008; Vuchic, 2005). Some agencies try to cover this increase by paying some regular drivers overtime to cover for the differences between the absences and the extraboard size. Overtime is generally paid at a higher rate and is bounded by union regulations due to the burden falling on the regular drivers making it even harder in some cases for an agency to cover all the absences. On the other hand, the presence of too many extraboard drivers compared to the amount of open work (e.g. unscheduled absences) presents a burden on the transit agency through increasing the general operating costs. Accordingly there is a need to generate a methodology that can accurately predict the amount of extraboard drivers needed per day, time period and garage for every transit agency to efficiently ensure service delivery in an efficient way.

This paper predicts the short term absences among regular-duty drivers to help in defining a more accurate extraboard size that can meet the day to day variations. The study will help in reducing the need for overtime drivers and/or excess number of extraboard drivers where no work is open for them. It analyzes short-duration absences among bus drivers at the OC Transpo, the transit operator (and provider) for the city of Ottawa, Ontario, Canada, using data drawn from the agency's human resource and scheduling databases. The paper starts with a literature review of drivers' absenteeism and strategies in sizing the extraboard drivers. This is followed by a description of the case study and data used. The next section pertains to the methodology used to prepare and analyze the data for predicting drivers' absence and anticipating the required size. It is then followed by a discussion of those results and finally a conclusion.

LITERATURE REVIEW

Extraboard drivers can be defined as the 'backup drivers', which is the pool of reserved drivers who fill open work resulting from unscheduled absences and other causes (DeAnnuntis & Morris, 2008). The strategies and processes of covering regular driver absence (or estimating the extraboard) consist of three integrated stages according to practice (Kaysi & Wilson, 1990; Koutsopoulos, 1990). The first is the strategic stage, which is generally related to estimating or hiring the optimum size of extraboard workforce. The second is the tactical stage, which is related to answering the questions of where and when these extraboard drivers should be (garage and day), according to the expected open work that needs to be filled. These two stages are usually done every 2 to 3 months according to the booking time frame. The last stage is the operational stage, which is related to daily unexpected open work based on specific times of day. This sometimes requires extraboard and regular drivers to work unscheduled overtime, which affected by many complicated factors, including work rules and a transit agency's unscheduled overtime work policy (Strathman, Kwon, & Callas, 2012). Generally, an optimum balance between unscheduled overtime assignments and extraboards are required in order to increase transit agencies' efficiency without service interruptions (Koutsopoulos & Wilson, 1987).

Some researchers theoretically interpret absence as a response to work dissatisfaction (Muchinski, 1977; Porter & Steers, 1973), others link it to unpleasant or hazardous job conditions (Allen, 1981). In other cases absences can be related to social (family) emergencies. Social absences are generally more tolerated by colleagues and management (Chadwick-Jones, Nicholson, & Brown, 1982). In this theoretical framework, the greatest potential to reduce absenteeism lies in recognizing the implicit social relationship between both drivers and management. Stress, traffic congestion, delays and dealing with difficult passengers (Gardell, Aaronson, & Barkloff, 1982; Greiner, Krause, Ragland, & Fisher, 1998; Long & Perry, 1985; Volinski, 1999) are correlated with increasing short-term absence frequency. These previous theoretical factors of absence, to an extent, affect all drivers in a transit agency, since workloads are shifted from one driver to another during their span of service.

Empirical studies have shown that drivers' personal situation factors have an effect on their absences. Among these factors are gender, experience, age and whether or not they are on probation (for drivers during the first six-months of employment) (Allen, 1981; Drago & Wooden, 1992; Keller, 1983; Leigh, 1986). Temporal factors such as seasonality (winter to summer), months, days of the week, and holidays influence drivers' absence, which consequently affects the number of extraboards needed during these periods (Kenyon & Dawkins, 1989; Shiftan & Wilson, 2001; Strathman, Broach, & Callas, 2009a; Strathman et al., 2012). Assigned work and service delivery characteristics (e.g. assignment type, garage, driving time, assignment time of day, and rotating shift work between time periods) have an effect on drivers' absence behavior (Fitzgibbons & Moch, 1980; Strathman et al., 2009a). However, these previous studies focused on individuals at the disaggregate level (Shiftan & Wilson, 2001; Strathman et al., 2009a), which is more interesting from a human resources point of view. For this disaggregate approach to be useful to transit drivers assigning bookings, predictions have to be made using the same unit of analysis, which is the individual driver. In other words, a prediction matrix needs to be generated for each driver during every booking while accounting for all the previously mentioned characteristics, which can be time consuming and adds complexity to the process. Also the prediction will lead to a probability value, which can impose further challenges on transit planners. Since the goal of this process is to reach an estimate of the extraboard size in a particular day in a booking, such estimations can be generated directly at the aggregate level to reduce the complexity in the prediction process and to generate a practical tool that is easy to use by transit operators. Therefore, the main purpose of the present study is to model the factors that impact regular driver's absence at the aggregate level of garage, day, and time period to anticipate the total absence that will need to be filled by extraboard drivers. In addition, the study demonstrates the advantage of this technique over traditional ones adopted by various transit agencies, which to our knowledge, is rarely addressed in the literature.

With the acknowledgment of the highly fluctuating factors affecting bus driver absences, the question of how transit agencies are sizing their extraboard becomes a valid and a critical one. According to a survey conducted by Long and Perry (1984) from twenty-one transit agencies within California, most agencies were using subjective and judgmental methods based on individual experience in determining the size of the extraboard. Therefore, the results were sometimes questionable, since they presented too large or too small a size of the extraboard staff. DeAnnuntis and Morris (2008) confirmed these results later using 35 transit agencies responding to a survey in the US. These agencies ranged from large (more than 250 buses) to medium-sized (more than 50 and fewer than 250 buses) and small systems (fewer than 50 buses). The majority of these agencies were using historical data and experience to estimate the size of the extraboard. In addition, most of them indicated that they were utilizing automated scheduling software without a component for extraboard management. The present study provides transit planners with a practical and systematic methodology that can be used directly to support extraboard planning practices, reducing the incidence of missed trips. This was done by using the aggregate level of analysis, which is useful from an agency's managerial and policy makers' standpoint. It provides a total number of drivers that can be acquired according to the transit agency's current levels of absenteeism.

DATA AND METHODOLOGY

The data used in this study comes from OC Transpo's archived human resource, scheduling, and absence databases. This data was mainly derived from Hastus' work-scheduling software package. OC Transpo has four bus garages, and it operates 195 bus routes serving 5,800 stops over a total distance of 5,584 kilometers, covering 446 km^2 in area (OC Transpo, 2013b). The largest bus depot, with a total capacity 482 buses, is the St. Laurent garage, where OC Transpo's headquarters is located. The other garages are the Industrial, Merivale and Pinecrest garages. The latter two garages have a total capacity of 275 and 193 buses, respectively, and are primarily used by employees working during the rush hour, with no use over the weekends. Industrial garage opened recently in 2010 with a total capacity of 167 buses, and it is used mainly for articulated buses (OC Transpo, 2013a). The analyzed data was collected between April 21st, 2008 and July 31st, 2012. After removing drivers' strike periods, around 4 years of data were kept for analysis.

The unit of analysis in this paper is the total driver absence per garage, day of week, and time period. Days of the week have been distinguished according to the booking week number. In OC Transpo, two week numbers are used, namely, *week one* to identify the odd week numbers, and *week two* to identify the even week numbers, allowing drivers to change their assignment between week one and week two during a booking. All variables were summarized according to the previous criteria, for example the average drivers' age during the early morning hours (before 6 A.M.) for the second bus garage during week one's Fridays of a booking. The use of week number allowed us to accurately account for any changes between weeks in drivers' schedules, and therefore absence, during a booking. In addition, to ensure robustness of the generated data, a group of 10 drivers per garage, day and time period was used as a threshold for a group. Accordingly, the garage-day-period records that have 10 drivers or fewer were deleted from the data set. After this process and after removing holidays and days with missing data, 14,305 groups of garage-day-period observations were generated with an average group size of 69.1 drivers per observation and standard deviation of 48.9 drivers. In addition, around 1400 randomly selected garage-day-period observations (around 10% of the sample) have been removed from the data and saved to test the quality of the generated models in estimating the extraboard size at the end of the study.

Five different time periods were identified in order to determine the amount of daily open work that might occur. The first is the early morning period which starts at 2:30 A.M. and ends at 6:00 A.M. The second is the morning period which starts at 6:00 A.M. and ends at 9:30 A.M. The third is the late morning period which starts at 9:30 A.M. and ends at 12:30 P.M. The fourth is the afternoon period which starts at 12:30 P.M. and end at 4:30 P.M. The fifth, and last period, is the evening period that starts 4:30 P.M. These periods were identified according to the mainstreams of drivers' starting work time. The average number of garage-day-period observations per period was 2861 observations with a minimum of 1347 and maximum of 4333 observations. Furthermore, five garages were identified in order to determine the amount of daily open work that might occur. Four of them represent the four main bus depots (St. Laurent, Merivale, Pinecrest, and Industrial garages), while the fifth represents the work that starts from any location along the network (rather than a specific garage).

This study employs a multilevel regression model to capture, isolate, and estimate the total absences at the aggregate of the garage-day-period level. The multilevel model is particularly appropriate for research designs where the data for participants is organized at more than one level or structure (Bickel, 2007; Gelman & Hill, 2007). Such description apparently fits to this case, since the obtained data is organized at different structures represented by the different garages. A likelihood ratio test (LR test) is used to compare the multilevel regression model to the linear regression model; if it is significant, it means that the multilevel model is more appropriate for the analysis. The multilevel model allows us to accurately control and isolate the average variation between entities (or garages in this case) and to provide a better fit for the analyzed data than a regular linear regression model. In other words, the multilevel modelling allows us to differentiate between the variation that is caused within the garage from the variation between the garages. Table 1 includes the list of variables incorporated in the statistical analysis. Other variables were tested but they were eliminated from the study due to their insignificance and/or correlation (with a Pearson coefficient of greater than 0.65) to other used variables, such as *working 7 days a week*, *working 8 consecutive days*, *weather conditions, paid time* and *drivers' experience*, or due to missing data, such as *driving distance*.

According to previous studies, there are various factors affecting drivers' absences that can be included in the analysis. Temporal factors such as day of the week, month and year, and whether the workday occurred after or before holidays, or on the day before or after the drivers' regular day off, have an impact on absence. Drivers' personal characteristics, such as age, gender, having a spouse and/or child and probationary status, are expected to have an impact on short-duration unscheduled absence among drivers. Assigned work characteristics, such as the total number of working days per week, whether drivers shift between time periods and garages during the week, whether drivers working as spare drivers during a week, and the day location during a booking, may affect their absences. Service delivery characteristics include assignment spread time and assignment starting time during a day (Strathman et al., 2009a). Furthermore, the multilevel regression model contains various interaction variables that represent the total number of drivers per garage, which may have an effect on absence. The model specification is:

1. *Absence = f(a. temporal factors, b. drivers personal characteristics, c. assigned work characteristics, d. service delivery characteristics, e. garage interactions)*
   Where*:*
   a. *Temporal factors* = Day $_{i}$, Year $_{t}$, Month $_{j}$, Day before a Holiday, Day after a Holiday, Regular work day, Day after regular day off (%), Day before regular day off (%)
   b. *Personal characteristics* = Drivers with spouse (%), Drivers with child (%), Female (%), Average Age, Average age squared, Probationary status (%)
   c. *Assigned work characteristics* = Total number of assignments, Drivers working more than 5 days a week (%), Drivers with spare assignments (%), Time shifting (%), Garage shifting (%), Booking second half, Booking first half.
   d. *Service delivery characteristics* = Assignments average time spread (min.), Assignments average time spread square (min.), Assignment from 2.30 A.M. to 6:00 A.M., Assignment from 6.00 A.M. to 9:30 A.M., Assignment from 9.30 A.M. to 12:30 P.M., Assignment from 12.30 P.M. to 4:30 P.M., Assignment after 4.30 P.M.
   e. *Garage Interactions*= Garage 9002, Garage 9003, Garage 9004, Garage 9005, and Garage 9006

The second part of this analysis uses the coefficients from the above model to conduct a sensitivity analysis to estimate the total number of drivers' absences during the random 1400 garage-day-period observations that were excluded from the previous model, while keeping all variables constant at their mean values. In addition, to understand the quality of the models, a difference in means t-test was used to compare the actual absences recorded by the agency (during the excluded 1400 garage-day-period observations) versus the estimated total absences resulting from the model. Afterwards, different scenarios are presented in order to provide the agency with a threshold number of required extraboard drivers according to the level of service.

Table 1: Variables used in the model

Variable Name	Description
Absence	The dependent variable, which is the total drivers' short unscheduled absence per period, garage and day.
Temporal Factors	
Day $_i$	A dummy variable that equals 1 when the drivers' assigned workday occurred on day *i*, where *i* represents a day ranging from Sunday through Saturday; it equals zero otherwise.
Year $_t$	A dummy variable that equals 1 when the drivers' assigned workday occurred in year *t*, where *t* represents a year ranging from 2008 through 2012; it equals zero otherwise.
Month $_j$	A dummy variable that equals 1 when the drivers' assigned workday occurred in month *j*, where *j* represents a month ranging from January through December; it equals zero otherwise.
Day before a Holiday	A dummy variable that equals 1 when the drivers' assigned workday occurred before a holiday and zero otherwise.
Day after a Holiday	A dummy variable that equals 1 when the drivers' assigned workday occurred after a holiday and zero otherwise.
Day after regular day off (%)	The percentage of drivers' assigned workday occurred the day after their regular day off per garage-day-period.
Day before regular day off (%)	The percentage of drivers' assigned workday occurred the day before their regular day off per garage-day-period.
Drivers Personal Characteristics	
Drivers with spouse (%)	The percentage of drivers' that have a spouse per garage-day-period.
Drivers with child (%)	The percentage of drivers' that have a child per garage-day-period.
Female (%)	The percentage of female drivers per garage-day-period.
Average age	Average drivers' age per garage-day-period.
Average age square	Average drivers' age squared per garage-day-period.
Probationary Status (%)	The percentage of the new drivers during the first six-months of employment) per garage-day-period.
Assigned Work Characteristics	
Total assignments	The total number of assignments that should be delivered per garage-day-period.
Working more than 5days a week (%)	The percentage of drivers working more than 5 days a week per garage-day-period.
With spare assignments (%)	The percentage of drivers that working as a spare for one day or more per garage-day-period.
Time shifting (%)	The percentage of drivers shifting between time periods during a week per garage-day-period.
Garage shifting (%)	The percentage of drivers shifting between garages during a week per garage-day-period.
Booking second half	A dummy variable that equals 1 when the drivers' assigned workday occurred in the second half of a booking and zero otherwise.

Booking first half	A dummy variable that equals 1 when the drivers' assigned workday occurred in the first half of a booking and zero otherwise.
Service delivery characteristics	
Assignment time spread (minutes)	The average assignment time spread per minutes per garage-day-period.
Assignment time spread square (minutes)	The average assignment time spread per minutes squared per garage-day-period.
Assignment from 2.30 A.M. to 6:00 A.M.	A dummy variable that equals 1 when the drivers' assigned workday started at 2:30 A.M. and before 6:00 A.M. and zero otherwise.
Assignment from 6.00 A.M. to 9:30 A.M.	A dummy variable that equals 1 when the drivers' assigned workday started at 6:00 A.M. and before 9:30 A.M. and zero otherwise.
Assignment from 9.30 AM to 12:30 P.M.	A dummy variable that equals 1 when the drivers' assigned workday started at 9:30 A.M. and before 12:30 P.M. and zero otherwise.
Assignment from 12.30 P.M. to 4:30 P.M.	A dummy variable that equals 1 when the drivers' assigned workday started at 12:30 P.M. and before 4:30 P.M. and zero otherwise.
Assignment after 4:30 P.M.	A dummy variable that equals 1 when the drivers' assigned workday started at or after 4:30 P.M. and zero otherwise.
Garage Interactions	
Garage 9002	The total number of drivers' per period and day that are starting from the St. Laurent garage.
Garage 9003	The total number of drivers' per period and day that are starting from the Industrial garage.
Garage 9004	The total number of drivers' per period and day that are starting from the Pinecrest garage.
Garage 9005	The total number of drivers' per period and day that are starting from the Merivale garage.
Garage 9006	The total number of drivers' per period and day that are starting from any location along the network (rather than a specific garage).

ANALYSIS

Table 2 presents summary statistics for the continuous variables used in the analysis. The table compares the data used in the model with the random sample that we excluded from the model to use later on for a sensitivity analysis exercise. The values presented here are obtained after the data cleaning process that was mentioned earlier. Generally, it can be noticed that the random sample data averages and standard deviation are closely related to the used data in the model. Independent sample t-tests are used to compare each variable in the dataset utilized in generating the models to the random sample left out to measure the accuracy of our predictions and conduct the sensitivity analysis with. The t-tests results show that there are no statistically significant differences between the means for any of the variables used in the model and the random sample at 95% Confidence Interval.

Regarding the data used in the model, for the dependent variable, the average absence among the regular-duty drivers is 9.9 drivers with standard deviation of 10.4 drivers. This high standard deviation value indicates high level of variation is present compared to the mean when it comes to absences, making it harder to predict. Around 80% of drivers per garage-day-period have a spouse with a standard deviation of 8.5% drivers. Females represent 10.4% of the total number of drivers per garage-day-period with standard deviation of 5.2%. The average percentage of drivers that work more than 5 days a week is around 39% per garage-day-period with standard deviation of 18.0%. On OC Transpo the average assignment time spread per garage-day-period is 502 minutes (8.3 hours), with standard deviation of around 128 minutes (2.1 hours).

Table 2: Descriptive statistics

	Data used in the model		Random sample data	
Variables*	**Mean**	**Std. Deviation**	**Mean**	**Std. Deviation**
Absence (number)	9.90	10.37	9.84	10.34
Day after regular day off (%)	20.96	28.92	21.80	29.51
Day before regular day off (%)	22.30	28.09	21.61	27.55
Drivers with spouse (%)	80.36	8.51	80.48	8.56
Drivers with child (%)	72.53	9.51	72.60	9.57
Female (%)	10.37	5.19	10.26	5.20
Average age (years)	47.74	3.54	47.81	3.58
Average age square (years square)	2292.0	342.5	2298.2	346.9
Probationary status (%)	1.68	4.65	1.82	5.07
Total assignments (number)	69.02	48.84	69.43	48.99
Working more than 5 days a week (%)	38.80	17.87	38.67	18.30
With spare assignments (%)	4.35	4.73	4.40	4.93
Time shifting (%)	38.32	22.09	38.72	22.54
Garage shifting (%)	36.69	20.77	36.86	21.27
Assignment time spread (min.)	502.84	128.23	500.46	129.49
Assignment time spread square (min.)	269286	120191	267212	120164
Garage 9002** (number)	85.42 (3488)	50.38	88.05 (367)	51.37
Garage 9003** (number)	33.26 (588)	11.18	33.88 (67)	12.05
Garage 9004** (number)	41.31 (2094)	25.89	42.54 (235)	28.08
Garage 9005** (number)	50.15 (2177)	26.75	49.42(238)	27.55
Garage 9006** (number)	82.81 (4558)	55.35	82.87 (495)	54.32
Number of cases	12,905		1,400	

* Reported values are per garage-day-period.
** Not considering the zero values of the interactions. The number of cases is reported between brackets

MULTILEVEL REGRESSION MODEL

In order to present the best fit to the generated data, two types of statistical models were tested using the total number of regular driver absences per garage-day-period as the dependent variable. The first was a linear regression model and the second was a multilevel regression model that uses garages as group indicators. The linear regression model had an r-squared value of 0.81 suggesting a significant high explanatory power by the model compared to other related models in the literature (Strathman et al., 2009a). In addition, the F-Test results showed that the p-value is almost equal to zero. Therefore, we rejected the null hypothesis with confidence above 99.99%, and we concluded that the independent variables as a set have a significant relationship with the dependent variable. However, the LR test that compares the multilevel regression model to the linear regression model was statistically significant, which validated that it is important to take into consideration that drivers absences vary across different garages. Therefore, a multilevel regression model was employed in this study. Table 2 presents the results of this model.

As seen in Table 2, regarding the temporal variables, days of the week have a significant impact on driver absence. At the weekday level, all days have a statistically significant coefficient value compared to Wednesday. Friday has the highest coefficient value, which indicates that drivers are more likely to be absent on this day than on any other day of the week, increasing the total absence per garage-day-period by 1.9 drivers. Monday and Thursday also have a positive value and increase the total driver absences per garage-day-period by 0.9 and 0.4 drivers, respectively. In contrast, Tuesday has a negative value, decreasing the total driver absences by 0.5 drivers per garage-day-period. These findings are consistent with previous studies (Strathman et al., 2009a). At weekend day level, interestingly, Sunday has no significant impact on driver absence compared to Wednesday, while Saturday has a significant negative coefficient value, decreasing the absence by 2.2 drivers. In other words the drivers who signed up to work on weekends are less likely to be absent compared to the ones who sign up on a weekday.

Months also have an effect on regular drivers' absence totals, with the peak level of absences occurring in October and November. They increase the number of drivers absent per by garage-day-period by 0.8 and 1.0 drivers compared to December, respectively. On the other hand, all of the other months have a significant negative coefficient value compared to December. January, February and March have the highest negative value decreasing the absence by 2.7, 2.0, and 1.5 drivers per garage-day-period.

At the annual scale, the total number of driver absences was highest in 2012. While in 2008, 2009, 2010 and 2011 variables have significant negative coefficient values compared to 2012, decreasing the driver absence by 4.2, 3.7 and 1.3 and 2.0 drivers per garage-day-period, respectively. It is important to note that drivers' strike periods from December 2008 to February 2009 have been removed from the dataset. This may account for the high negative values of the total absence during these years, indicating that drivers tend to decrease their short-duration unscheduled absence before and after their strikes.

Days following holidays and days before holidays show statistically significant negative values, which decreased the total driver absences by 0.68 and 0.69 drivers per garage-day-period, respectively, compared to days that follow or precede regular days. Therefore, the model suggests that policies which allow scheduled interruptions (such as training activities) on the days after and before holidays, are accepted by OC Transpo. Furthermore, the percentage of drivers' assigned workdays the day before and after their regular day off had a significant effect, decreasing the total absence by 0.02 and 0.03 drivers, respectively, for every additional one percent per garage-day-period. These results suggest that drivers are not likely to be absent the day before or after their day off at OC Transpo. This may reflect a management practice in OC Transpo that stresses the drivers' attendance before and after their regular day offs.

Regarding the drivers' personal characteristics, the one percent of increase in the drivers with children has a significant positive value, increasing the total driver absence by 0.03 drivers per garage-day-period. Therefore, If we consider that the average percentage of drivers with children per garage-day-period is 72.5% (as indicated by table 2), then the mean absences is expected to be 2.2 drivers.. The drivers' average age per garage-day-period decrease the absence by 2.8 drivers for every additional year. However, the average age square term, which is used to understand the non-linear relationship of the variables, indicates a statistically significant positive value, increasing the absence by 0.03 drivers for each additional year, which is consistent with previous research (Strathman et al., 2009a; Strathman, Broach, & Callas, 2009b). Several other personal characteristics—the percentages of female drivers, drivers in the first six-months of employment, and drivers with spouse per garage-day-period— seem to increase the total absence, however they were not significant.

In terms of the assigned work characteristics, the total number of assignments has a significant positive coefficient value, increasing the total regular driver absences by 0.13 drivers for every additional assignment. The percentage of drivers that work more than 5 days a week increases the total absence by 0.03 drivers for every additional one percent, compared to the percentage of drivers that work 5 days a week or less. Therefore, policies that minimize the number of drivers that work more than 5 days a week are recommended. The percentage of drivers with spare assignments increases the total absence by 0.09 drivers for every additional one percent per garage-day-period. The percentage of drivers that shifts between time periods decreases the absence by 0.04 drivers for every additional one percent per garage-day-period. This reflects that drivers enjoy the schedule time flexibility. The percentage of drivers that shifts between garages increases the absence by 0.01 drivers for every additional one percent per garage-day-period. This may be due to the high stress level that drivers face in maneuvering between different garages. The day location within the second half of a booking increases the total absence by 0.5 drivers per garage-day-period compared to the first half of booking, while keeping all other variables at their mean values.

Concerning the service delivery characteristics, while the assignment time spread per minutes decrease the total driver absence per garage-day-period by 0.06 drivers for every additional minute, the spread square term increase the drivers' absence by 0.01 drivers. This indicates a non-linear relationship between the assignment time spread and the total absence. Therefore, OC Transpo may increase the assignment time spread while using caution. Compared to the early morning assignments that start before 6 AM, the morning assignments (from 6:00 AM to 9:30AM), the afternoon assignments (from 12:30 PM to 4:30 PM) and the evening assignments (after 4:30 PM) decrease the total driver absence by 1.4, 3.3, and 1.7 drivers per garage-day-period, respectively. In contrast, the late morning period increases the total absence per garage-day-period by 1.1 drivers, which may indicate a trend by drivers.

Garages have an impact on absence. As expected, the total number of drivers working from any location along the network (garage 9006) has the highest positive impact value, increasing the total absence by 0.1 drivers for every additional drivers compared to the main garage. Garage 9003 (Industrial) also increases the absence by 0.04 drivers per period and day compared to the main garage. This may be because of this garage is used mainly for articulated buses. These buses, due to their size, may have an impact on drivers' stress level compared to other regular buses, leading to increases in their absence rate. Therefore, a detailed study concerning driver's behavior and perception is recommended to understand the impacts of operating articulated buses on driver's fear of collision, stress level and/or comfort. In contrast, the Garage 9005 (Merivale) has a significant negative impact value, decreasing the total absence by 0.02 drivers for every additional drivers compared to the main garage, while 9004 (Pinecrest) did not show a statistically significant impact on total driver absences in our model.

The random part of the multilevel regression model shows the standard deviations of the intercept and residuals (error term). In general, the idea of the random coefficient demonstrates that the overall error variance consists of two parts: the first results from the random variation of the intercept (standard deviation of the constant) and the second results from the variance of the error (standard deviation of the residual). The Intraclass correlation coefficient is a statistic that measures the degree of dependence among observations nested within garages. The interclass correlation coefficient explains the proportion of variability of drivers' absence days that occurs between garages rather than within garages. The model suggests that 27.6% of the variability in drivers' absence is due to differences between garages. In addition, it was estimated that 95% of the random coefficient of the total absence per garage-day-period intercept varied between 0.8 days and 3.4 drivers per garage-day-period, suggesting significant variability in drivers' absence between garages.

Table 3: The multilevel regression model results

Variable	Coefficients	t	[95% Conf.	Interval]
Constant	77.00 ***	9.56	61.22	92.78
Temporal Factors				
Sunday	0.63	1.52	-0.18	1.44
Monday	0.87 ***	2.03	0.03	1.71
Tuesday	-0.48 ***	-3.45	-0.75	-0.21
Wednesday (base reference)	---	---		
Thursday	0.42 ***	3.09	0.15	0.69
Friday	1.85 ***	5.81	1.23	2.47
Saturday	-2.20 ***	-4.72	-3.11	-1.29
Y2008	-4.20 ***	-20.07	-4.61	-3.79
Y2009	-3.76 ***	-18.57	-4.15	-3.36
Y2010	-1.36 ***	-9.84	-1.63	-1.09
Y2011	-2.01 ***	-13.87	-2.30	-1.73
Y2012 (base reference)	---	---		
January	-2.67 ***	-9.54	-3.21	-2.12
February	-1.98 ***	-7.52	-2.50	-1.46
March	-1.50 ***	-5.65	-2.02	-0.98
April	-1.28 ***	-5.60	-1.73	-0.83
May	-0.58 ***	-2.77	-0.98	-0.17
June	-0.41 ***	-2.12	-0.80	-0.03
July	-1.48 ***	-7.38	-1.88	-1.09
August	-0.86 ***	-4.29	-1.25	-0.46
September	-0.50 ***	-2.37	-0.91	-0.09
October	0.76 ***	3.69	0.36	1.16
November	1.03 ***	5.19	0.64	1.42
December (base reference)	---	---		

Variable	Coefficients	t	[95% Conf. Interval]	
Day before a Holiday	-0.68 ***	-2.60	-1.19	-0.17
Day after a Holiday	-0.69 ***	-2.98	-1.14	-0.23
Regular days (base reference)	---	---		
Day after regular day off (%)	-0.02 ***	-4.47	-0.04	-0.01
Day before regular day off (%)	-0.03 ***	-5.68	-0.04	-0.02
Drivers Personal Characteristics				
Drivers with spouse (%)	0.01	0.98	-0.01	0.02
Drivers with child (%)	0.03 ***	5.25	0.02	0.04
Female (%)	0.01	0.70	-0.01	0.03
Average age	-2.80 ***	-8.52	-3.45	-2.16
Average age square	0.03 ***	9.14	0.02	0.04
Probationary status (%)	0.01	0.97	-0.01	0.04
Assigned Work Characteristics				
Total assignments	0.13 ***	60.35	0.12	0.13
Working more than 5 days a week (%)	0.03 ***	6.43	0.02	0.04
With spare assignments (%)	0.09 ***	6.15	0.06	0.12
Time shifting (%)	-0.04 ***	-8.84	-0.05	-0.03
Garage shifting (%)	0.01 ***	3.33	0.01	0.02
Booking second half	0.46 ***	4.18	0.25	0.68
Booking first half (base reference)	---	---		
Service delivery characteristics				
Assignment time spread (min.)	-0.06 ***	-14.86	-0.07	-0.05
Assignment time spread square (min.)	0.01 ***	14.41	0.00	0.00
Assignment from 2.30 A.M. to 6:00 A.M. (base reference)	---	---		
Assignment from 6.00 A.M. to 9:30 A.M.	-1.38 ***	-9.54	-1.67	-1.10
Assignment from 9.30 A.M. to 12:30 P.M.	1.14 ***	5.53	0.73	1.54
Assignment from 12.30 P.M. to 4:30 P.M.	-3.31 ***	-20.47	-3.62	-2.99
Assignment after 4.30 P.M.	-1.70 ***	-7.81	-2.12	-1.27
Garage Interactions				
Garage 9002 (base reference)	---	---		
Garage 9003	0.04 ***	2.28	0.01	0.07
Garage 9004	0.01	1.58	0.00	0.02
Garage 9005	-0.02 ***	-4.50	-0.03	-0.01
Garage 9006	0.09 ***	37.44	0.08	0.09
Number of records	12,905			
St. dev. of constant	1.72		0.85	3.48
St. dev. of residuals	4.48		4.43	4.54
Intraclass correlation	27.6%			

*** Significant at 99% ** Significant at 95% * Significant at 90%

SENSITIVITY ANALYSIS

Using the coefficients derived from the previous model, it is possible to estimate the total number of regular driver absences by conducting a sensitivity analysis that predicts the daily changes in absence for each garage-day-period, while keeping all variables at their mean values. This sensitivity analysis, in part, is used to validate the model results and to demonstrate the advantages of the modeling technique over traditional ones utilized by various transit agencies. It is done for the random 1400 garage-day-period records that were not included in generating the previous model. Next, a paired difference in mean t-test and Pearson correlation test were used to compare the *actual* recorded total absence by the agency to the *estimated* number of absences resulting from the sensitivity analysis during these 1400 periods.

For the random sample, the average for the actual total driver absences was 9.8 drivers, while the average for estimated number of absences was 10.0 drivers per garage-day-period, with a mean difference of 0.19 drivers. However, these averages do not reflect the reality due to the high standard deviation. The standard deviation of the actual absences from the average was 10.3 drivers, while it was 9.6 drivers for the estimated absences. This indicates greater dispersion in actual absences than the estimated absences by 0.7 drivers. The Pearson correlation test shows a statistically significant positive correlation of 0.9 between the actual and estimated absence. This implies a strong relationship between them. Thus, estimated absence increases as actual absence increases per garage-day-period. In addition, a two-tailed paired samples *t* test revealed that there is no statistically significant difference between the actual and the estimated absence per garage-day-period, t (1339) =1.6, p>.05. Figure 1 shows the average total number of drivers, the actual used extraboard number according to the current policy adopted by the agency, and the estimated and actual absence per garage and per time period.

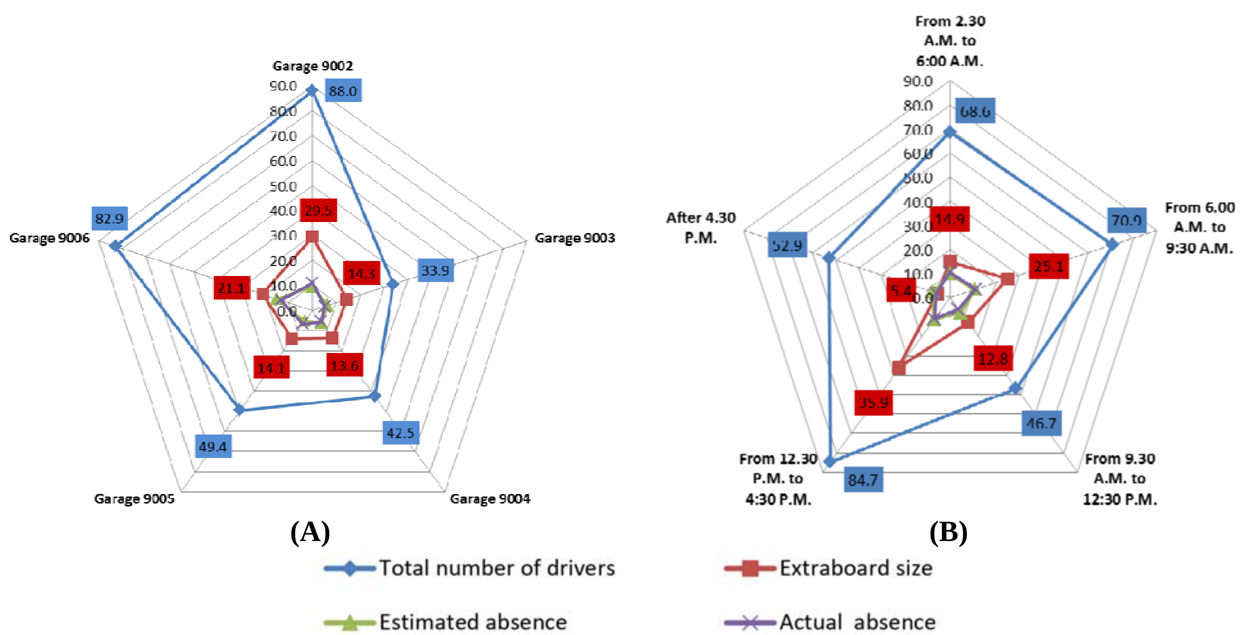

Figure 1: (A) Drivers per garage and (B) Drivers per time period

The complicated nature of studies on absence levels and the imperfect data collection apparently does not resolve all of the potential problems related to estimation of absences

(Shiftan & Wilson, 1994). Therefore, seven scenarios are represented to size the extraboard teams in order to cover the high level of standard deviation differences between the estimated and the actual average absences per garage-day-period. These scenarios were developed using the coefficients and values derived from the previous model in this study, and provide a range of possibilities for the agency to meet the target coverage of the total driver absence. The first scenario keeps the model outputs as is without any changes (± 0%). The second scenario adds 20% to the model estimation. The third scenario added 40%, while the fourth, fifth, sixth and seventh scenarios added 60%, 80%, 100%, and 120%, respectively, to the models estimated output. It should be noted that the estimated total absences that have a value of less than or equal one have been normalized to a value of 2 absences in order to make the scenarios counting more effective. This helped to overcome the problem of covering 2.5% of the absence by the analysis scenarios. Table 3 shows these scenarios along with the agency current extraboard sizing practice.

As seen in table 3, the last gray column shows the current transit agency practice, which is included for benchmarking. It should be noted that due to data limitations the available actual used extraboard data covers 912 garage-day-periods out of the 1400 records. For the current practice, there was a significant difference between the provided extraboard drivers and the regular drivers absence by 11.3 extraboard drivers on average per garage-day-period, $t(911) = 28.4$, $p \leq .05$. This is justifiable since these extraboards should be used to cover all types of open work assignments, including the short-term unscheduled absences. However, this extraboard team size did not cover all the garage-day-periods' unscheduled absences, instead it only covered 90.6% of the sample, leaving 9.4% of these garage-day-periods assignments for the overtime drivers to cover or were considered missed trips.

The ratio between extraboards and regular drivers' absence was 3.6 extraboards per one absence. The extraboards percentage in relation to the total number of regular drivers was 33.2% per garage-day-period. These numbers coincide with the percentages reported in the literature. According to a survey recently done by five agencies, large agencies can have an extraboard percentage as big as 34% (with an average of 23%) and a ratio of 4 extraboard drivers per piece of open work (with an average of 2.9 extraboard drivers) (DeAnnuntis & Morris, 2008). Although the survey used a different level of analysis (the day level rather than garage-day-period level used in this paper), it is worth reporting to provide a general context. The maximum number of overtime drivers per garage-day-period needed was 34 drivers with a frequency of two incidents. This number represents 25% of the total regular drivers' number per related garage-day-period. This indicates that while OC Transpo has a large set of extraboards, big gaps still occur between the planned extraboard and the driver's unscheduled absence.

Regarding the different scenarios, as seen in table 3, the first scenario will cover 70.6% of the 1400 garage-day-period absences, leaving 29.4% of them with a total number of extraboards less than the number of actual absences. The average difference between the estimated and the actual absences recorded by the agency equals 1.1 drivers. The ratio between the estimated and actual absences is 1.5 drivers per piece of open work. The estimated percentage of extraboards in relation to the total number of regular drivers is 15.5% per garage-day-period. In this scenario, the maximum number of regular drivers needed to be overtime is 27 drivers with a frequency of one garage-day-period. This number represents 11% of the total number of drivers per related garage-day-period. Comparing this scenario's results to the total actual extraboard team size, indicates that the transit agency will be able to keep around 10.2 drivers (11.3 drivers provided by the agency minus 1.1 drivers estimated by the model) to cover other proposes, while achieving this level of service.

In the second scenario, the expected number of garage-day-periods that is underestimated by the model predictions is 18.1%. This means that the agency will have enough extraboard drivers to cover 82% of all days by providing an extra 2.1 drivers compared to the previous scenario. Yet this scenario is not ideal since the high number of days with fewer extraboards than the number needed may lead to a substantial amount of canceled trips during different days in the year.

Table 4: Model scenarios

	SENSITIVITY ANALYSIS							Current transit agency practice
	Scenario 1 (± 0%)	**Scenario 2** (+ 20%)	**Scenario 3** (+ 40%)	**Scenario 4** (+ 60%)	**Scenario 5** (+ 80%)	**Scenario 6** (+100%)	**Scenario 7** (+120%)	
Totally covered periods (%)	**70.6%**	**81.9%**	**92.5%**	**95.8%**	**99.0%**	**99.7%**	**100%**	**90.6%**
Average difference between estimated and actual absence*	1.1	3.2	5.4	7.6	9.8	11.9	14.1	**11.3**
Ratio of total extraboard to actual absence	1.5	1.8	2.2	2.4	2.8	3.1	3.3	**3.6**
The extraboard percentage**	15.5%	18.3%	21.9%	24.7%	28.3%	31.0%	33.8%	**33.2%**
Maximum number of drivers needed to be Overtime	27.0	16.0	7.0	5.0	3.0	1.0	0.0	**34.0**
Frequency*** and (%)	1 (11.1%)	1 (6.6%)	1 (7.3%)	1 (5.2%)	1 (3.1%)	4 (0.5%)	0.0	**2 (25.2%)**

* The average of the estimated absence minus the actual absence recorded by the agency. This calculation is done for every day-period during the random 1400 garage-day-periods, and then the average outcome has calculated. For the *transit current practice* column it was calculated as the provided extraboard drivers minus the actual regular drivers absence

** The estimated extraboard size divided by the total number of regular drivers.

*** Frequency is the number of incident that the maximum number of drivers needed to be overtime occurs. The percentage of the needed number of drivers to be overtime compared with the total number of regular drivers per related garage-day-period.

The third and fourth scenario will cover 92.5% and 95.8% of the 1400 garage-day-period absences, respectively. Compared to the current transit agency practice, these scenarios will be covering an additional 2% and 5% of the absence, respectively. This enhancement in coverage is combined with a significant additional saving of 5.9 and 3.7 extraboard drivers of the current team sizes for other purposes. Furthermore, in these scenarios, the ratio between the extraboard size and the actual absence are 2.2 and 2.4 drivers, which is less than the transit agency ratio by 1.4 and 1.2 drivers per absence, respectively. The estimated extraboards percentage in relation to the total number of regular drivers is 22% and 25%, which indicate a decrease of 11% and 8.5%

of the current transit agency percentages, respectively. Finally, the maximum needed number of regular drivers needed to be overtime in these two scenarios is dropped significantly to reach 7 and 5 drivers per garage-day-period with a frequency of one incident, respectively, which is much less than the transit agency current practice maximum needed number of extraboard (which was 34 drivers with a frequency of two incidents). The maximum number of overtime drivers needed is 7%, and 5% of the total number of regular drivers for the related garage-day-period, which is much less than the current transit agency practice (of 25%).Therefore, the transit agency can implement one of these scenarios while being confident that it will have a better fit of the extraboards in relation to absences, covering only a relatively small percentage of absence by the overtime drivers. In the last three scenarios (the fifth, sixth, and seventh scenarios) the percentage of days covered by the estimated extraboard will increase tremendously to reach more than 99% of the total days by providing an extra 10 drivers compared to the first scenario, while saving around 2 drivers from the current extraboard team size for other purposes. Finally, it should be noted that, by using the sixth scenario, the transit agency could adjust and relocate the same amount of extraboard drivers that it uses now to cover more than 99% of the regular driver absence, instead of the 90.5% that is covered by the agency's current extraboard sizing practice, without any additional extraboards. Due to the high standard deviation in the absences covering 100% of the absences by extraboard drivers might not be the best solution for a transit agency. A combination between extraboard drivers and overtime ones is recommended.

To help in highlighting the best scenario from a financial stand point the following section explores a rough financial example that can be used to show the power of our model and estimation technique in finding the best scenario. These numbers can vary based on extraboard driver's cost as well as overtime hourly compensations from one agency to another. To cover all days during the year with an optimum cost, the total cost of each scenario is calculated based on a. the total cost of unscheduled overtime and b. the total cost of hiring extraboards. Figure 2 shows different cost estimations scenarios. First, the total unscheduled overtime per sample is calculated based on the total number of underestimated garage-day-period multiplied by the average needed drivers per period and by the average number of hours per assignment (8.3 hours, as indicated in table 2). In addition, if we consider that the random sample is roughly 0.4 of a year (it is a tenth of 4 years), the annual unscheduled overtime will be the total overtime per sample multiplied by a factor of 2.5. Second, the total extraboards hiring cost is the average number of needed extraboard drivers (based on the average difference between estimated and actual absence (Table 4 - second row) and the actual absence which equals 9.8 operators) multiplied by the average driver's annual income. In OC Transpo, the average wage for drivers rests at $27 per hour and $ 40.5 per overtime hour (Lanktree, 2013), with an average annual income of $70,000 (including the benefits) (Reevely, 2012).

As seen in figure 2 the third scenario shows the lowest total cost option compared to other scenarios and compared to the current practice used by the transit agency. In the third scenario, 7.5% of garage-day-periods (105 out of 1400 garage-day-periods) needed to be covered by overtime drivers, with a total of 1826 hours (the average needed drivers per garage-day-period is 2.1 drivers). Therefore, the total overtime cost will be $74,000 based on the 1400 garage-day-periods random sample, and $185,000 per year. The total cost of hiring about 15.3 extraboard drivers (9.8 plus 5.4 extraboards) is $1,070,000. Therefore, the total cost in this scenario is around $1,255,000. The total cost of the second, fourth, fifth and sixth scenario is around $1,462,000, $1,293,000, $1,393,000 and $1,528,000, respectively. Accordingly, it is recommended that the transit agency should prioritize the size of the extraboard based on the third scenario to ensure all days during the year are covered with an optimum cost. Meanwhile the remaining tasks will be covered by overtime drivers at a cost that is less than hiring an additional extraboard driver. Comparing the third scenario to the current practice shows a total savings of around savings $ 900,000, in favor of the third scenario, while covering 100% of the absences.

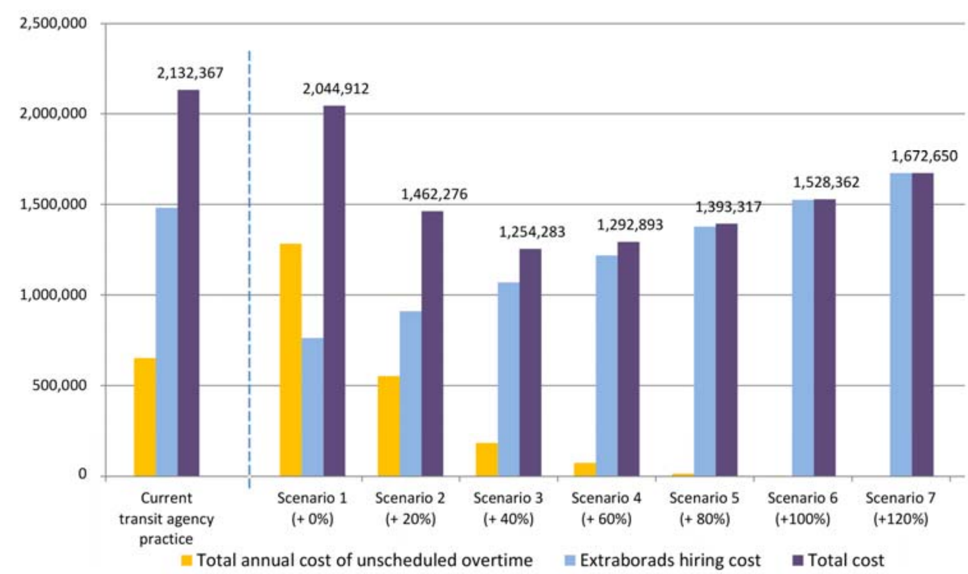

Figure 2: Cost estimations

CONCLUSION

This article aims to understand the factors that contribute to the short-duration unscheduled absences of bus transit drivers at the aggregate level, and to anticipate the future total absences that will need to be filled by an extraboard team. Five different time periods were identified in order to determine the amount of daily open work that might occur. Then, a multilevel regression model was generated to investigate regular drivers' short-unscheduled absences while accounting for the differences between the garages. The short-unscheduled absence was estimated in relation to temporal factors, drivers' personal characteristics, aspects of assigned work, and service delivery characteristics. Lastly, using these models coefficients, sensitivity analyses were conducted, offering the agency several counts of extraboard drivers per day to cover a wide range of assignment thresholds.

On the tactical planning level at OC Transpo, at the weekday level, the model results suggest that it would be wise to avoid Thursday, Friday, and Monday scheduled interruptions by the agency, such as training activities or business meetings, while it is encouraged to allow these interruptions during Tuesdays and Wednesdays. In addition, it is suggested to minimize scheduled interruptions for regular drivers as much as possible during the months of December, November, and October. Furthermore, the results indicate that it would be wise to decrease the number of drivers that work spares for one day or more during the week. It is also recommended to decrease the number of drivers that shifts between garages, while accepting the time shifting between the time periods. In addition, it is also suggested that more than five days a week of scheduled work should be reduced, since it increases the drivers' probability of being absent. In addition, the model results suggest that it would be effective to avoid scheduled interruptions during the second half a booking. Furthermore, providing incentives to reduce drivers' absence to those who start their work during the early morning and late morning period assignments and who work from any location along the network is recommended.

Additionally, the research suggests various levels of service for the extraboard teams that can be applied by the agency in order to cover their regular drivers' absences. For instance, using the third scenario, around 93% of the total absences per garage-day-periods will be covered while saving the agency 6 extraboards and achieving an additional 2% coverage of the total absence per garage-day-periods compared to the current policy adopted by the agency. This scenario will require the transit agency to cover only 7% of the total absence per garage-day-periods by the overtime drivers, with a maximum of 7 drivers per garage-day-period (around 7% of the total number of regular drivers), which is much less than the agency's current practice. Finally, hypothetically, the transit agency could adjust and relocate the same amount of extraboard drivers using the sixth scenario to cover more than 99% of the regular driver absence without any additional extraboards.

To conclude, this study provides transit planners with a practical methodology that can be used directly to support extraboard planning practices, thus reducing the incidence of missed trips and ensuring the number of extraboard drivers is enough to cover the missed trips during the entire year. By examining trends at the aggregate level, this study overcomes some of the shortcomings of previous research that focused only on individuals at the disaggregate level. This aggregate level analysis is useful from an agency's managerial standpoint. The recommendations from this study are not limited to OC Transpo, as our findings were similar when studying extraboard sizing for another transit agency in Canada using a similar approach although differences in scale and context was present. Transit planners and operators can benefit from the case study described in this paper by adopting a similar method to manage their extraboard team more effectively, since the method introduce depends mainly on readily available data (work scheduling data) and is flexible enough to include different setups, institutional framework, and regulations.

ACKNOWLEDGMENT

We would like to acknowledge *OC Transpo,* for partially funding this research, and for providing the data used in the project. Furthermore, special thanks to Michel Courval, Hugh Muller and Sophie Frigon from CSched for their feedback and collaboration during this project. Thanks to Professor James Strathman of Portland State University for his advices and guidance at the beginning of this study. The authors would like to thank Colin Stewart and Ana Tepavac for their help in proof reading the article. Last but not least we would like to thank the anonymous reviewers for their valuable feedback on the earlier version of the manuscript.

REFERENCES

Allen, S. (1981). An empirical model of work attendance. *The Review of Economics and Statistics, 63*(1), 77-87.

Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* . New York: The Guilford press.

Chadwick-Jones, J., Nicholson, N., & Brown, C. (1982). *Social psychology of absenteeism*. New York: Praeger.

DeAnnuntis, C., & Morris, W. (2008). Transit Extraboard Management: Optimum Sizing and Strategies. *Transportation Research Record*(2072), 110-124. doi: 10.3141/2072-12

Drago, R., & Wooden, M. (1992). The determinants of labor absence: Economic factors and workgroup norms across countries. *Industrial and Labor Relations Review, 45*(4), 764-778.

Fitzgibbons, D., & Moch, M. (1980). Employee absenteeism: A multivariate analysis with replication. *Organizational Behavior and Human Performance, 26*(3), 349-372.

Furth, P., & Muller, T. (2007). Service reliability and optimal running time schedules. *Transportation Research Record*(2034), 55-61. doi: 10.3141/2034-07

Gardell, B., Aaronson, G., & Barkloff, K. (1982). The working environment for local public transport personnel. Stockholm: Swedish Work Environment Fund.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, United Kingdom: Cambridge University Press.

Greiner, B., Krause, N., Ragland, D., & Fisher, J. (1998). Objective stress factors, accidents, and absenteeism in transit operators: A theoretical framework and empirical evidence. *Journal of Occupational Health Psychology, 3*(2), 130–146.

Kaysi, I., & Wilson, N. (1990). Scheduling transit extraboard personnel. *Transportation Research Record*(1266), 31-43.

Keller, R. (1983). Predicting absenteeism from prior absenteeism, attitudinal factors, and nonattitudinal factors. *Journal of Applied Psychology, 68*(3), 536-540.

Kenyon, P., & Dawkins, P. (1989). A time series analysis of labour absence in Australia. *The Review of Economics and Statistics, 71*(2), 232-239.

Koutsopoulos, H. (1990). Scheduling of extraboard operators in transit systems. *Transportation Science, 24*(2), 87-104.

Koutsopoulos, H., & Wilson, N. (1987). Operator workforce planning in the transit industry. *Transportation Research Part A, 21*(2), 127-138.

Lanktree, G. (2013). Hard work no breaks pushes OC Transpo drivers pay past 100k, says union boss. *Metro* Retrieved April 12, 2013 from http://metronews.ca/news/ottawa/618512/hard-work-no-breaks-pushes-oc-transpo-drivers-pay-past-100k-says-union-boss/

Leigh, J. (1986). The effects of unemployment and the business cycles on absenteeism. *Journal of Economics and Business, 37*(2), 159-170.

Long, L., & Perry, J. (1984). Extraboard scheduling, workers' compensation, and operator stress in public transit: Research results and managerial implications. Irvine, California Institute of Transportation Studies,University of California.

Long, L., & Perry, J. (1985). Economic and occupational causes of transit operator absenteeism: A review of research. *Transport Reviews, 5*, 247-267.

Muchinski, P. (1977). Employee absenteeism: A review of the literature. *Journal of Vocational Behavior, 10*(3), 316–340.

OC Transpo. (2013a). About OC Transpo: Garages, from www.octranspo1.com/about-octranspo

OC Transpo. (2013b). About octranspo: bus fleet, from www.octranspo1.com/about-octranspo

Perk, V., Flynn, J., & Volinski, J. (2008). Transit Ridership, Reliability, and Retention. Florida: National Center For Transit Research (NCTR), University of South Florida.

Porter, L., & Steers, R. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. *Psychological Bulletin, 80*(2), 151–176.

Reevely, D. (2012). So what's an OC Transpo driver really make? , 13 April, 2013, from http://blogs.ottawacitizen.com/2012/11/29/so-whats-an-oc-transpo-driver-really-make/

Shiftan, Y., & Wilson, N. (1994). Absence, overtime, and reliability relationships in transit workforce planning. *Transportation Research Part A: Policy and Practice, 28*(3), 245-258.

Shiftan, Y., & Wilson, N. (2001). The absence consequences of overtime in the transit industry. *Journal of Public Transportation, 3*(4), 25–40.

Strathman, J., Broach, J., & Callas, S. (2009a). Analysis of short-duration unscheduled absences of transit operators: TriMet case study *Transportation Research Record*(2111), 185-194. doi: 10.3141/2111-21

Strathman, J., Broach, J., & Callas, S. (2009b). Evaluation of short-duration, unscheduled absences among transit operators: TriMet case study. Portland, Oregon: Oregon Transportation Research and Education Consortium (OTREC).

Strathman, J., Kwon, S., & Callas, S. (2012). Extraboard performance: TriMet case study. Portland, Oregon: Oregon Transportation Research and Education Consortium (OTREC).

Volinski, J. (1999). Synthesis of transit practice 33: Practices in assuring employee availability. Washington, D.C.: RB, National Research Council.

Vuchic, V. (2005). *Urban transit: Operations, planning and economics*. New York: John Wiley & Sons.