**Comments from the editors and reviewers:**

1. **Reviewer 1**

This paper addresses a relevant topic and offers some insightful results. It does need however in my view to undergo substantial revisions before becoming suitable for publication. In the following, I note 9 comments which I consider major and thereafter list minor comments.

Major comments:

* 1. The paper is currently unnecessarily lengthy. I think the essence of the paper and its substance are such that do not require 16 figures and 29 pages. The paper requires some considerable trimming.
  2. The authors overstate their results. In essence, the obviously naive approach of planning to arrive the same second as the bus is expected to arrive can be expected to often (and even a majority of cases if there is even a slight systematic bias towards over-predicting the remaining time for arrival) result with missing the bus. I find the so-called the insurance buffer to be the more innovative element in this study. Note its affinity with the notion of 'hidden waiting time' which refers to a high (e.g. 95th) percentile of the expected distribution.

This is a fair comment. We cut down the size of the analysis part and only keep the essence of the results.

Per reviewer 1’s question of the reason why we dedicated many paragraphs to greedy tactic (the tactic that reviewer 1 addressed in the comment body): greedy tactic is in fact the default tactic that most transit planning apps suppose their users to use. The default scenario of using a transit planning apps is: the app will provide a home departure time / leaving time calculated from real-time data

* 1. Section 2.1 is redundant. I suggest the authors instead review the methods of quantifying the impacts of RTI on waiting times and thereafter report their findings. Also the relevance of the paragraph on surveys and their costs in 2.2 can be removed without any loss of information, unless the authors intend to specifically comment on those conducted in the context of RTI and waiting times.
  2. The literature review I miss analytical approaches such as the studies performed by a study published in this journal ("Evaluating the added-value of online bus arrival prediction schemes") and a related study published in the Journal of ITS ("Real-time bus arrival information system: An empirical evaluation") by the same authors. These studies also considered the impact of RTI on waiting times along the line, as well as function of the prediction horizon which is equivalent to the analysis in relation to walking time performed in this paper. Moreover, the comparison with a waiting time resulting from following the schedule has also been performed. The authors should better acknowledge how their work relates to previous work.
  3. The presentation of the method can be dramatically shortened. GTFS and APC are by now very standard data sources in transit research. The notion of "reclaimed delay" is also not new. The impact of initial delays on running times further downstream has been extensively studied with conflicting evidence, see for example:

El-Geneidy, A. M., J. G. Strathman, T. J. Kimpel, and D. T. Crout. 2006. “Effects of bus stop consolidation on passenger activity and transit operations.” Transportation Research Record

El-Geneidy, A. M., J. Horning, and K. Krizek. 2011. “Analyzing transit service reliability using detailed data from automatic vehicular locator systems.” Journal of Advanced Transportation.

Cats, O. 2018. "Determinants of bus riding time deviations: Relations between driving patterns and transit performance". Journal of Transportation Engineering.

Similarly, departure strategies that have been reported elsewhere like AT can be described briefly.

* 1. The authors refer to measurement error and schedule recovery efforts as an explanation for an over-estimation of bus arrival time. Even in the absence of those, an over-estimation could simply result from traffic conditions including short signals and skipping stops (no boarding and alighting passengers).
  2. A critical point is that IB is introduced only in PT but none of the other TPSs includes an element that is conscious of risk-taking. ST can also include an IB term, i.e. avoiding just missing the bus. This applies also to ET.
  3. Notations throughout the manuscript are sloppy. For example, introducing if definitions without indicating the value taken otherwise. There are also other matters, needs to be revisited carefully.
  4. Some of the conclusions may not be transferable, can the authors please reflect on that? For example, the performance of ST obviously depends on the on-time performance of the service under consideration (in particular, the share of early arrivals). The spatial pattern discussed in 4.3.3. is also clearly caused by the deterioration of the on-time performance further downstream.

Minor comments:

* 1. Suggest to shorten the title. After the question mark can simply have only "An empirical analysis"
  2. TPS is mentioned in the last paragraph of Section 1 but has not been introduced yet
  3. Add axes titles in figures 1 and 2
  4. Broken references in section 3.3
  5. Frumin and Zhao (2012) is not the original source for Eq. 3. Please refer to the original contributor.
  6. The description accompanying Figure 3 is not sufficiently clear, please revisit.
  7. Please revisit also the last paragraph in the conclusions, it is not clear to me what is meant by this.

1. **Reviewer 2**

This paper presents an analysis of a single bus route in Columbus, Ohio to explore the impacts of different passenger trip planning strategies, including those using real-time information, on passenger wait times.  Overall, I found the manuscript to contain numerous noteworthy flaws. Specifically, the analysis relies on some unusual assumptions that may be driving the results; furthermore, the scope of the analysis is limited to a single bus route in a single city, limiting the generalizability of the findings.  Moreover, the authors have not validated their theory and findings with real world behavioral data, such as from surveys or focus groups. In light of these weaknesses, I recommend significant revisions to the paper.  My specific comments are detailed below corresponding to the page number (since there were no line numbers in the manuscript).

* 1. Page 2, Paragraph 3

-How did the authors come up with the idea that RTI apps can diminish waiting times to zero? Have the authors conducted a survey/focus groups/interviews of riders to demonstrate that this “greedy” strategy is something riders actually do?  Numerous prior studies of waiting times have been cited in the literature review section of this paper, and they all included reality high average wait times (e.g., Watkins et al. found 9.23 minutes for RTI users compared to 11.21 minutes for non-users).  It seems unrealistic to expect riders to minimize their wait times to zero.

* 1. Page 3, Section 2.1

-The authors do not differentiate between perceived versus actual wait time differences in the literature review, which is an important distinction in prior research on the impacts of real-time information.  Please add a brief discussion.

* 1. Page 6, Paragraph 1

-The authors claim that APC data is more accurate in terms of arrival/departure time at each stop compared to GTFS-realtime. I found this surprising.  How did the authors come to this finding? What analysis did you conduct to demonstrate the accuracy?  Additionally, did you verify a sample of the data with real world observations (e.g., ride the bus and manually record the stop times, then compare them to APC and GTFS)? Please justify.

* 1. Page 6, Paragraph 2

-COTA’s data is updated once per minute, which seems quite long compared to many other agencies. For example, the MBTA in Boston updates their bus location data every 5 seconds (see [https://medium.com/@sjbarbeau/introducing-the-gtfs-realtime-validator-e1aae3185439](https://urldefense.com/v3/__https:/medium.com/@sjbarbeau/introducing-the-gtfs-realtime-validator-e1aae3185439__;!!KGKeukY!mfZjvfj7c1QUhjnD-mhYCT1dZD5VB15xS-B7ucYmAzfx91W6mSkP4e1DvL7DxyWCSg$)). This is likely an important data limitation from COTA that is driving some of your results (discussed more later).  Please add discussion of typical update/refresh times from other transit agencies.

* 1. Page 6, Last Paragraph

-The authors make the assumption that the walking process “is linear with respect to distance.” They later explain that the only way to change the walking time is to depart the home at a different time. However, the assumption of constant walking speed seems highly unrealistic. If a rider sees a bus approaching and they think they might miss it, they are very likely to speed up and potentially even run to meet the bus. Indeed, a study by Dziekan and Kottenhoff (2007) of the subway in Stockholm, Sweden observed passengers entering subway stations and counted the number of passengers running and walking when RTI signage outside the subway station was on, and the results reveal that significantly more people run when the RTI signage was on rather than when the signage was off. In light of this real world evidence, the assumption of constant walking speed seems very unrealistic, and the authors should test the sensitivity of their results to altering this assumption.

- Similar to the previous comment, another recent study by Ferris et al. (2010) found that RTI provided on mobile devices may impact a passenger’s decision of where to board the transit vehicle, which would impact the passenger’s walking distance to access transit. On a survey of RTI users conducted in Seattle, Washington, 78% of respondents reported they were more likely to walk to a different stop based on RTI (Ferris et al., 2010).  Can the authors test this (e.g., changing where to board) in their modelling framework?

* 1. Page 9, Table 1

- How did the authors arrive at these 5 trip planning strategies?  Similar to my previous comment, have the authors conducted a survey/focus groups/interviews of riders to demonstrate that these are strategies riders actually use?

* 1. Page 10, Equation 3

- It should be noted that the average waiting time formula for random arrivals is generally only applied to high frequency transit routes (e.g., headways less than 10-15 minutes).

* 1. Page 3, Equation 5

- The authors state that the “bus will rarely if ever leave a stop earlier than the scheduled time.” Did the authors verify this statement empirically, such as comparing the GTFS schedule to GTFS-realtime? Drivers occasionally do run “hot.” Please justify the assumption that they don’t.

* 1. Page 11, Figure 3

- What data was used to create the visualization shown in Figure 3?

* 1. Page 13, Equation 9

- The authors state that “for PT family, insurance buffer should be at least equal to the expected waiting time.”  Please explain why this is the case. It wasn’t clear to me.

* 1. Page 14, Last 2 Paragraphs

- The authors state that there is a large computational burden to conduct the analysis, so they only selected one bus route for the analysis. This greatly limits the generalizability of the research.  Instead of using 1 year of data, why not use 1 week of data and run the analysis for multiple bus routes? I strongly encourage the authors to consider a larger geographic sample.

* 1. Page 17, Figure 7

- Figure 7 shows high sensitivity to the 60 second update of real-time data, which, as previously noted, seems to be a reality high value (e.g., the MBTA updates every 5 seconds).  Is there a way you can test the sensitivity of this in your model? At a minimum, it should be discussed as a drawback of the case study of COTA, as it appears to be driving the results shown in Figure 7 and may not apply to other agencies with better real-time data.

* 1. Page 21, Figures 11 and 12

- I found the GT results counterintuitive, and I would have expected them to be similar to the PT results. I suspect this is likely due to the assumptions built into your models (e.g., constant walking speed so riders won’t run for an approaching bus and the long 60-sec threshold of real time updates).

* 1. Page 22, Figure 13

- For the route shown in Figure 13, how many timepoints are there along the route?  Where are the timepoints located? How do the timepoints relate to your findings (in this figure and the subsequent maps of the route)?  Please add a discussion of timepoints.

- For Figure 13, please make sure the colors in the legends correspond to the same numerical values. It is difficult to compare GT with ST, AT and ET since the ranges for the colors are different. This comment also applies to subsequent figures.

Minor Comments

* 1. On page 9, please fix the reference errors to the figures/tables.
  2. The authors introduce many new acronyms throughout the paper, which can be confusing for readers. Please try to limit the use of acronyms that are not commonly found in the prior literature (e.g., consider removing IB, TPS, HDT, etc.).