Exhibit D – Research Project Requirement Template

An Application-Agnostic Investigation of Location-Based Services Data Quality: Using Synthetic Data and Empirical Benchmarking

Recipient/Grant (Contract) Number: The University of Texas at Austin; University of Michigan/Grant # 69A3552344815 and 69A3552348320

Center Name: National Center for Understanding Future Travel Behavior and Demand (TBD)

Research Priority: Improving Mobility of People and Goods

Principal Investigator(s): Atiyya Shaw

Project Partners: N/A

Research Project Funding: $146,250

Project Start and End Date: 06/2025-05/2027

Project Description: Over the past decade, the widespread adoption and use of mobile services and devices have enabled the large-scale, passive collection of human location data. The most widely available of these, termed location-based services (LBS) data, are largely GPS enabled location "pings" collected from third party smartphone applications. These data have already proven to be useful for a wide range of mobility applications – from the joint estimation of aggregate transportation network flows (i.e., origin destination data) to examining special event and disaster evacuation movements. However, because LBS data are not representative at multiple levels (e.g., spatial/geographic, sociodemographic, behavioral), their use for individual-level (i.e., disaggregate) analyses has been limited. Nonetheless, the potential of these data for disaggregate behavior and demand modeling is a well-recognized priority area, particularly in light of growing limitations to traditionally relied upon survey data. Mobility analysis with LBS data requires preprocessing, but prior work often arbitrarily selects thresholds (for preprocessing) based on their specific end applications and unique characteristics of their vendor data. In this study, we take a methodological approach to documenting the sources of error in LBS data and developing iterative methods for evaluating preprocessing thresholds applied to this data. We propose to evaluate these decisions according to both sociodemographic and behavioral representativeness. This investigation will facilitate follow up work that entails the application of algorithms for addressing data sparsity and representativeness, and thereafter for fusing these data with travel survey data. This body of work is intended to complement existing efforts by a limited set of researchers working to advance the use of LBS data for disaggregate behavioral analysis.

US DOT Priorities: The methodological and applied aims of this research align well with the USDOT priority of transformative research – the primary aim of this project is to develop new, robust data sources for transportation policy development and decision-making. In addition, this research also contributes to technology transfer/deployment as it will be conducted with input from MPO and state transportation officials and outputs of this effort are intended to be implemented within the transportation planning and modeling process.

Outputs: This work is expected to generate multiple conference papers and presentations and at least two research journal papers. We will also interface with state and regional transportation planning and modeling staff to assess the best path forward for sharing relevant findings and approaches.

Outcomes/Impacts: The anticipated outcomes of this work entail: (1) a framework for transportation researchers to use to evaluate and describe LBS (and more generally passive) data; and (2) recommendations for filtering and preprocessing standards to prioritize the reduction of specific types of representation errors and biases.

Final Research Report: A URL link to the final report will be provided upon completion of the project.