PRSE
Potsdam Research Seminar in Economics

ABSTRACT:

MACHINE LEARNING FROM BIG GPS DATA ABOUT THE COST OF CONGESTION

DR. NICOLAS KOCH

MCC BERLIN

15 December 2020 from 02.15-03.30 p.m.
via Zoom

We estimate the welfare loss of road congestion using a representative machine-learned panel data set of GPS-coded trips in Berlin. To overcome the accuracy, scope, and granularity limitations of traditional travel surveys, we rely on precise information on 32 million trips gathered from connected cars and cellular networks over the full year 2017. We leverage unsupervised machine learning to assign groups of trips to individual drivers. Using density-based spatial clustering, we find about 37,000 individual clusters, which we use to account for unobserved heterogeneity in individual travel behavior. Identification of the causal effect of traffic density on the time cost of travel relies on a new instrumental variable strategy exploiting intra-weekday traffic patterns observed throughout the year. We find large temporal heterogeneity in marginal external costs of congestion, ranging during daytime between 0.10 and 1.77€ per vehicle km. Optimal congestion charges range from 1 to 41€ cents per km depending on time. A time-specific congestion charge could increase welfare by at least 777 million € per year. Only about 12% of welfare gain is achievable by a second-best, uniform congestion charge of 11€ cents per km.