

Time Trends in the Use of Administrative Data for Empirical Research

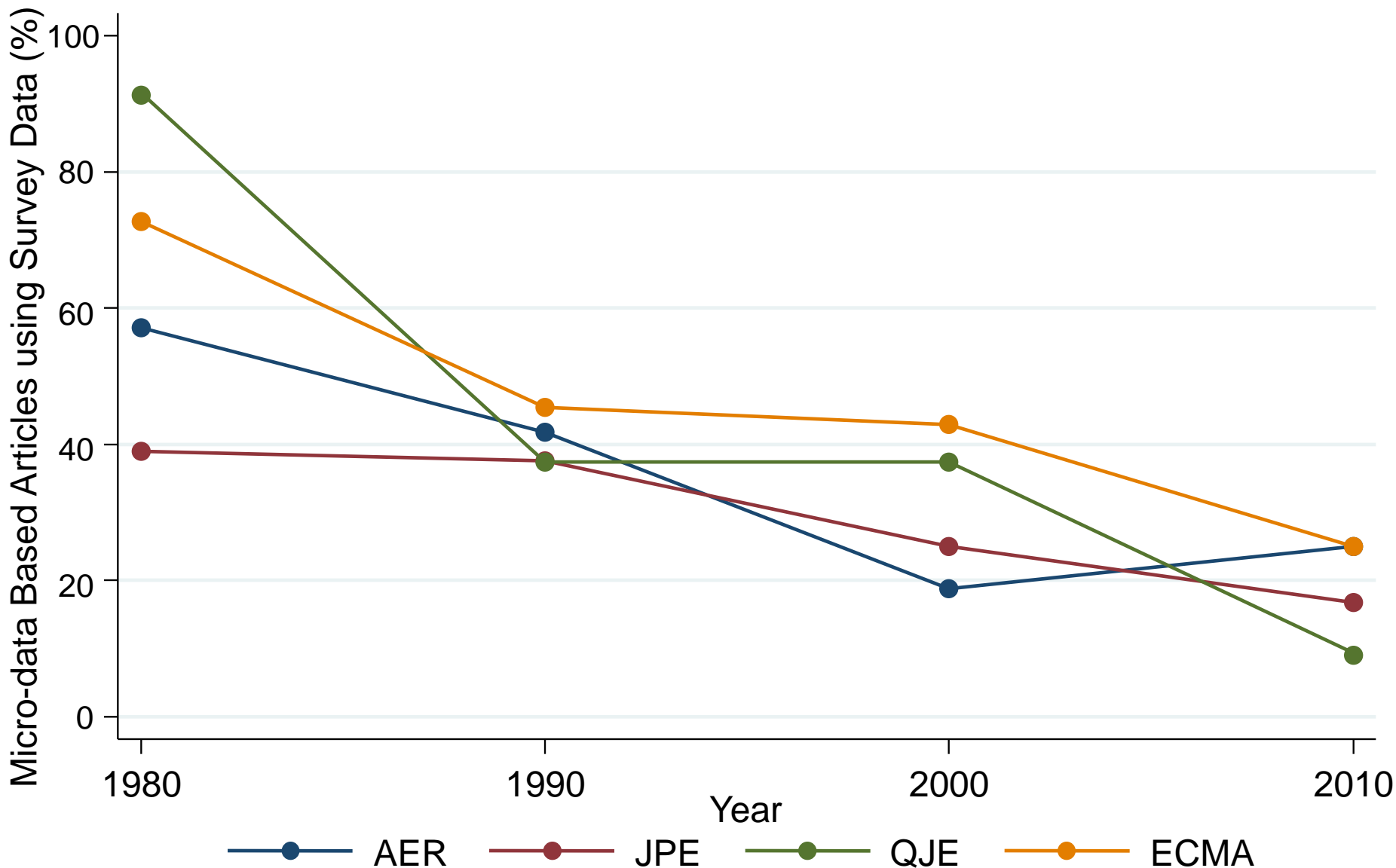
Raj Chetty, Harvard and NBER

July 2012

What are the Benefits of Administrative Data?

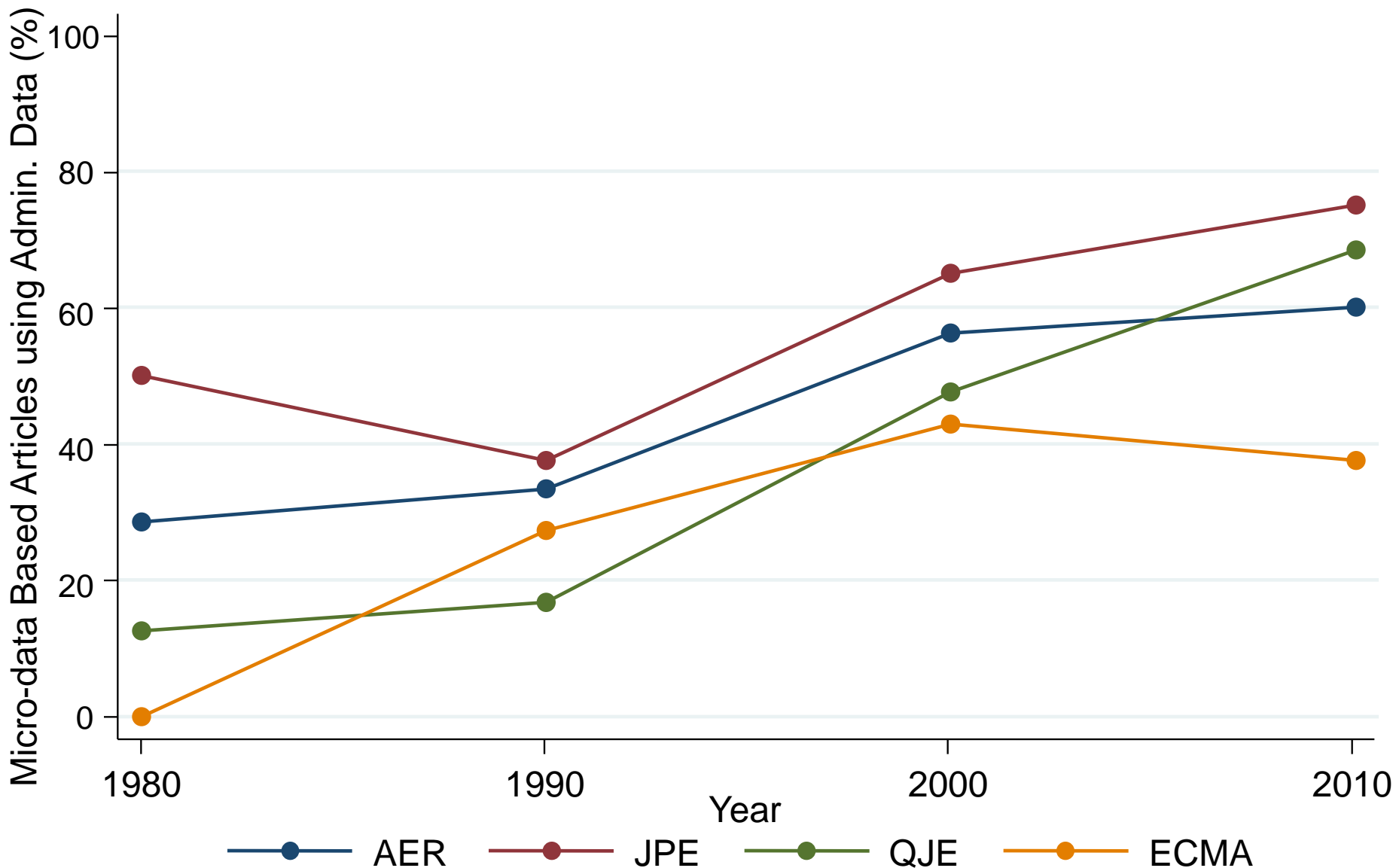
1. Higher quality information: virtually no missing data or attrition
 - CPS non-response rate now 31% for income
 - SIPP attrition rate exceeds 30% within three years
2. Longitudinal tracking over long periods
 - Match rates of 95% over 20+ years in studies of long-term impacts of early childhood education [Chetty et al. 2011, Chetty, Friedman, Rockoff 2012]
3. Very large sample sizes: 2,000 times the size of the CPS
 - Can develop new non-parametric, quasi-experimental research designs
 - The following slides document the rising use of administrative data in leading economics journals

Use of Pre-Existing Survey Data in Publications in Leading Journals, 1980-2010



Note: "Pre-existing survey" datasets refer to micro surveys such as the CPS or SIPP and do not include surveys designed by researchers for their study. Sample excludes studies whose primary data source is from developing countries.

Use of Administrative Data in Publications in Leading Journals, 1980-2010



Note: "Administrative" datasets refer to any dataset that was collected without directly surveying individuals (e.g., scanner data, stock prices, school district records, social security records). Sample excludes studies whose primary data source is from developing countries.