

The South African Revenue Service and National Treasury Firm-level Panel: Firm-level data compiled from tax and customs administrative records

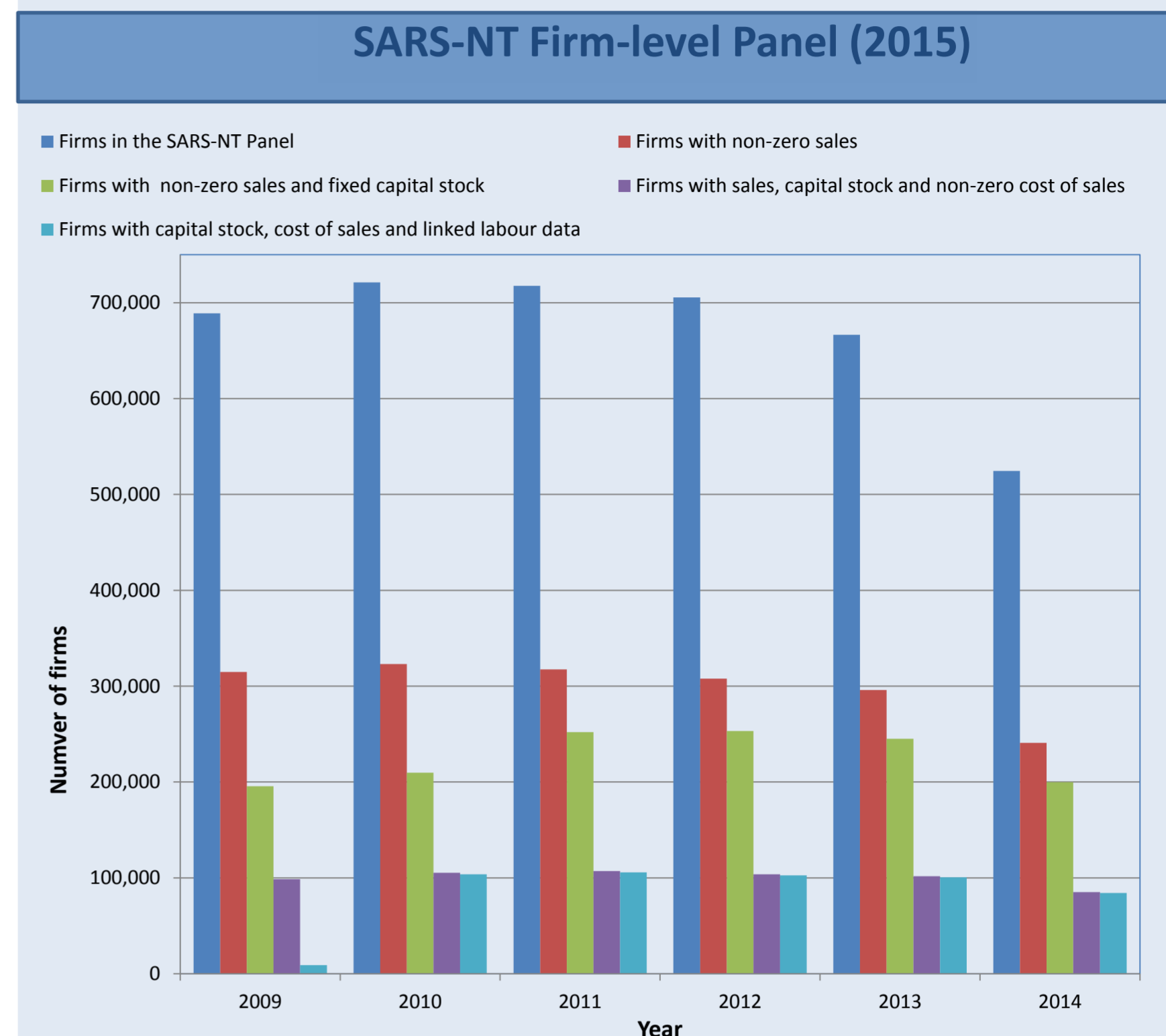
Duncan Pieterse, C. Friedrich Kreuser and Elizabeth Gavin

How is the panel constructed?

- Replace names and administrative identifiers in tax and trader records with non-intelligent identifiers.
- Identify each entity that provides a Corporate Income Tax (CIT) return as a **firm**.
- Link CIT return data provided per financial (tax) year with associated Value-added Tax (VAT) and Pay-as-you-earn (PAYE) records, as well as commodity-level import and export records.
- Derive the number of jobs and full-year-equivalent jobs associated with a firm for each of its financial years from employers' PAYE records for all individuals who receive payment from them.

Points to note:

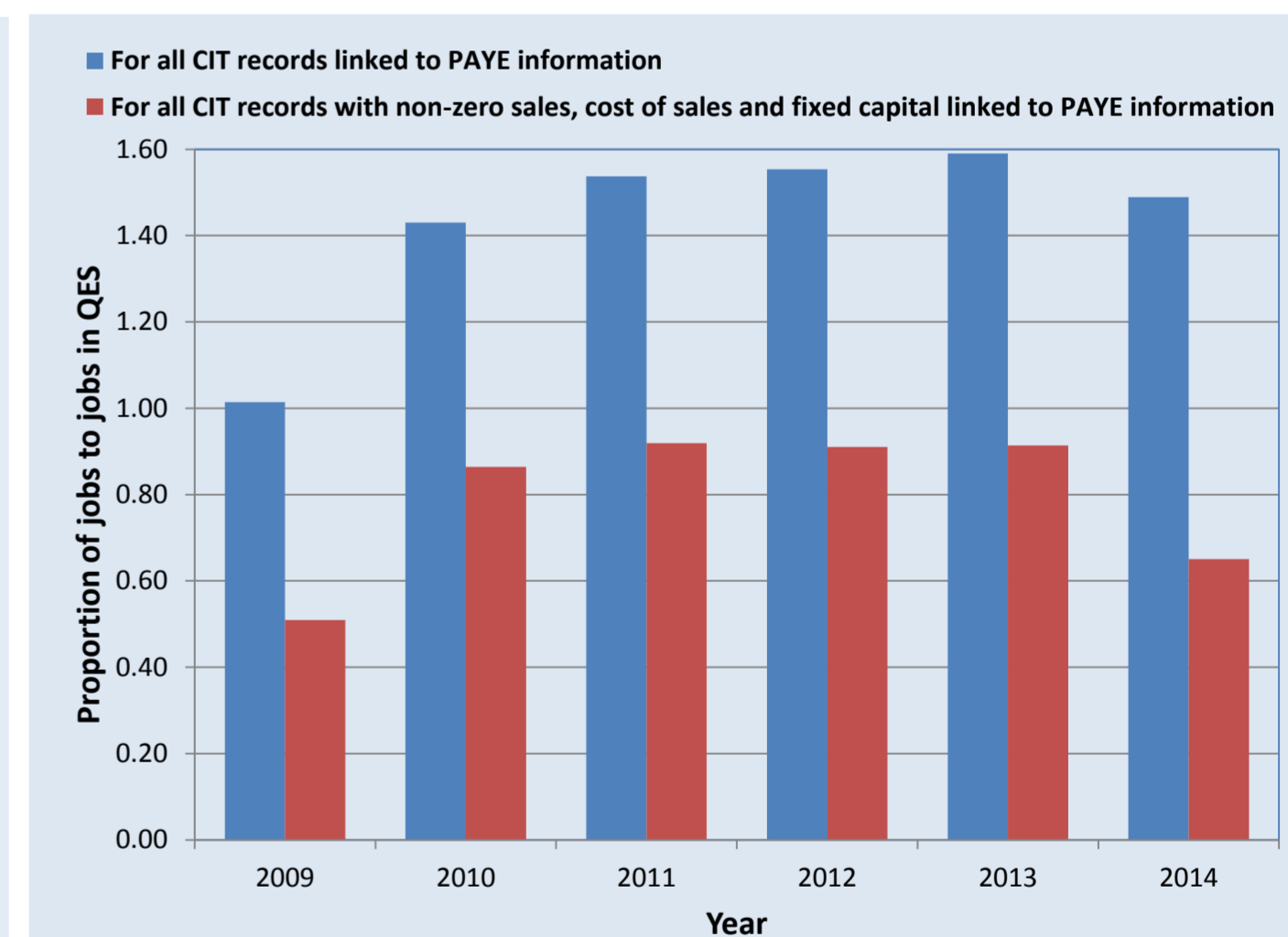
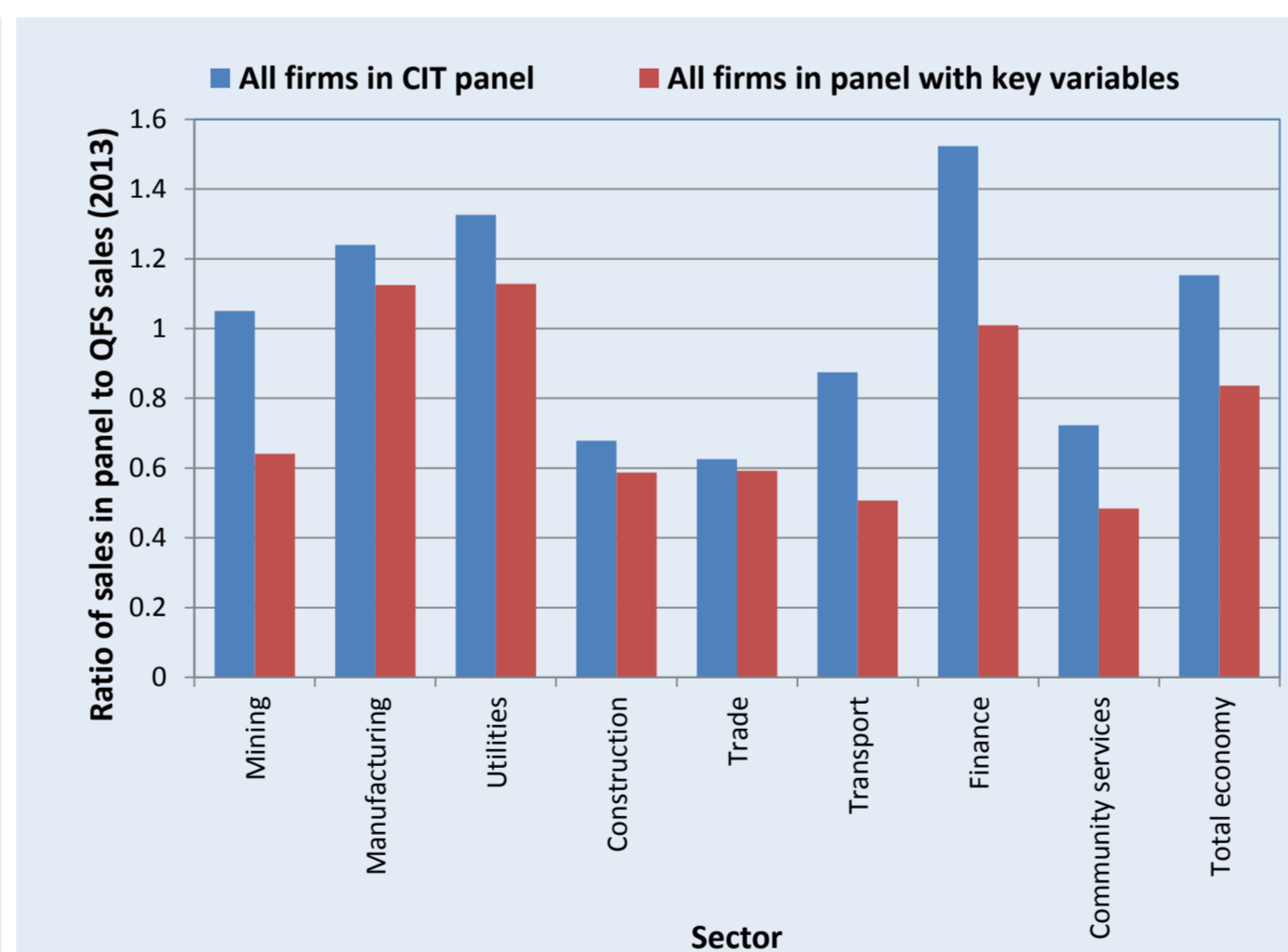
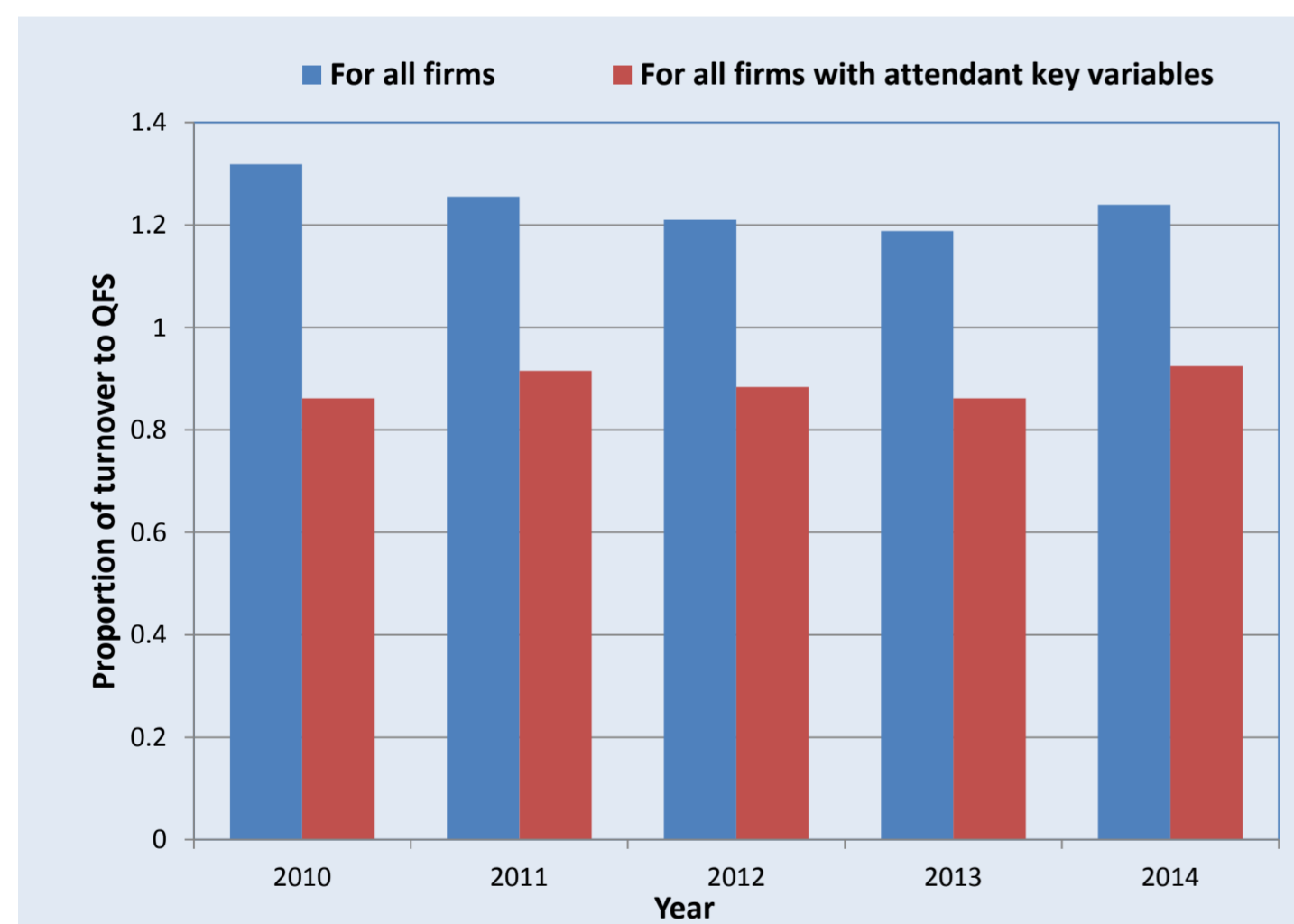
- The tax and trade data in the current panel was extracted in 2014 and covers late 2007 to 2014. The data for later years is less complete.
- Not all CIT returns will be associated with PAYE, VAT or trader records. On the other hand, more than one PAYE or VAT registration may be associated with a single CIT-registered entity.
- Firms that have not been active may submit CIT returns, and which thus show zero turnover and costs.
- The firm's financial year may not coincide with the tax year for Personal Income Tax (March – February), and this needs to be considered in deriving the number of jobs associated with the firm during its financial year.
- Tax forms change over time, as do the variables available and definitions thereof. Variable naming conventions were developed to reflect these distinctions in the raw fields and variables derived to facilitate time-series analysis.
- The sector classification from CIT, VAT and PAYE records may differ.



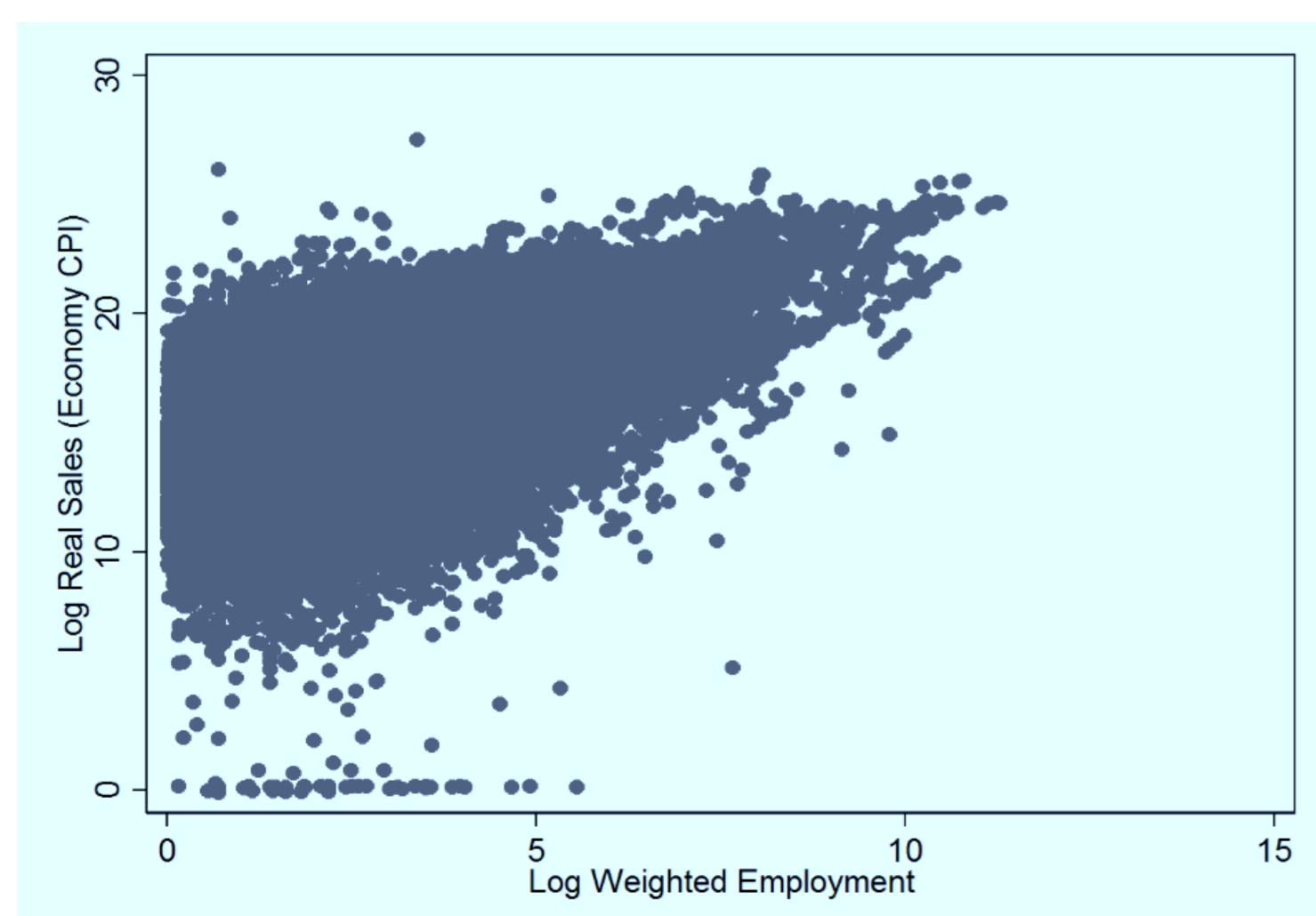
How does the firm-level panel compare with other firm data?

- The sales (turnover) for firms in the panel were aggregated to compare with the aggregates of the Quarterly Financial Survey (QFS).
- This was done for all firms, as well as for those firms with key variables required for productivity analysis (sales, cost of sales, capital, labour information).
- Although many records do not contain these variables, the aggregate sales for those that do are comparable to the QFS. However, large differences emerge when turnover is examined by sector.
- The number of jobs per firm in the panel were aggregated to compare with the job levels from the Quarterly Employment Survey (QES).
- This was done for all firms, as well as for those firms that had records of the key variables for productivity estimates (sales, cost of sales, capital).

- Points to note:
- The comparisons exclude agriculture and community and social services.
 - The QES excludes employers who are not VAT active.
 - Employees cannot be distinguished from recipients of pensions in PAYE data from 2010 to 2012 on the basis of income source codes alone, so figures for these years may be overstated.



What advantages does the firm panel hold for research?



- The combining of different administrative records provides a richer set of variables on a firm than is typically available from a sample survey, e.g. both the turnover and employee numbers of a firm may be used in examining firm behaviour in relation to firm size.
- Tax and trade administrative records naturally yield longitudinal firm data enabling investigation into firm dynamics. The period covered by the panel will grow with time.
- Because the administrative records cover the full population of tax active firms, rare 'events' or behaviours are represented in the panel data, e.g. the uptake of an infrequently used tax incentive, such as learnerships.