

# Production of business statistics using administrative data

Administrative Data Seminar

22<sup>nd</sup> February 2010

Caitríona O'Brien

# Why use administrative data for business statistics?

- Reduce response burden on small enterprises
- Reduce the size and cost of surveys
- Data available for a larger number of small businesses
- More detail at regional level
- Successfully done in other EU member states

# Structural Business Surveys

- Services
  - All enterprises with 20 or more persons
  - Sample of smaller enterprises
  - 20,000 surveyed from total population of 100,000
- Industry
  - All enterprises with 3 or more persons
  - Approx 5,500 enterprises surveyed
  - Approx 5,500 enterprises with less than 3 not surveyed
- Building & Construction
  - All enterprises with 20 or more persons
  - Approx 1,300 enterprises surveyed
  - Approx 48,000 enterprises with less than 20 not surveyed

# Revenue data Sources

Revenue File	Variables Available
Corporation Tax (CT)	Turnover, Gross profit Turnover – Gross profit = Cost of sales
Income Tax (IT)	Turnover, Gross profit Turnover – Gross profit = Cost of sales
P35	Employment, earnings of employees

# Linking Revenue and CSO files

- CSO and Revenue have different unit identifiers
- Central Business Register (CBR) number links to more than one record on tax files
  - Multiple customer numbers
  - Multiple time periods
- Accounting period information
  - Different reference years
  - Accounting periods longer or shorter than 1 year

# Linking Revenue and CSO files – Contd.

- CBR numbers with no match on tax files
  - No tax return made?
  - Late tax return?
  - Tax return available but no turnover information on it
  - Link between identifiers not available

# Plan for Services Sectors

- Do not survey single person enterprises
- Obtain **turnover, gross profit, employment and wages and salaries** from administrative sources for single person enterprises
- Use regression to explain correlations between detailed survey variables and variables available from administrative sources

# Plan for Services Sectors contd.

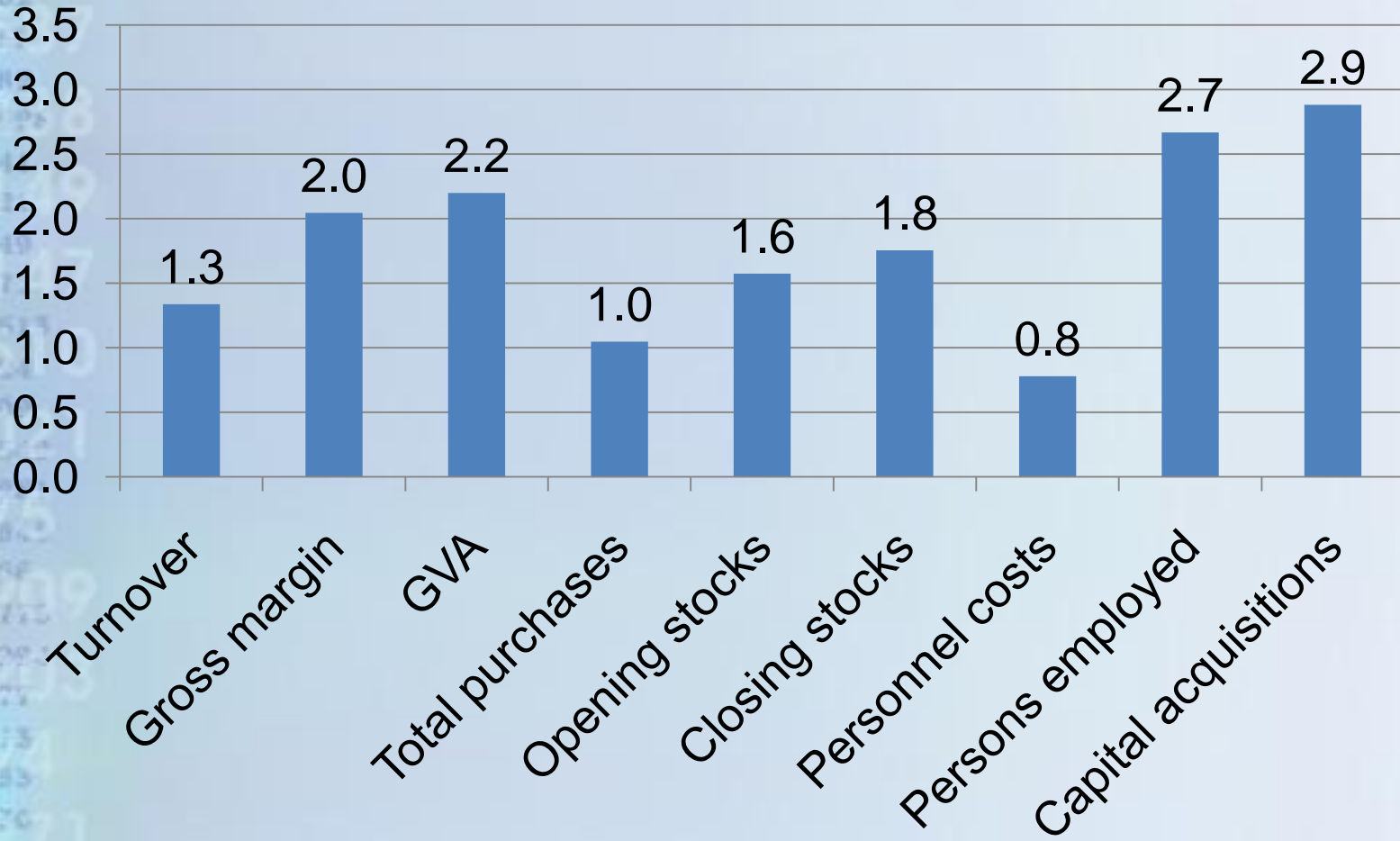
- Use survey data for 2 and 3 person enterprises to calculate model parameters
- Use model to calculate variables not available from administrative files for single person enterprises
- Gross up results to ASI register population



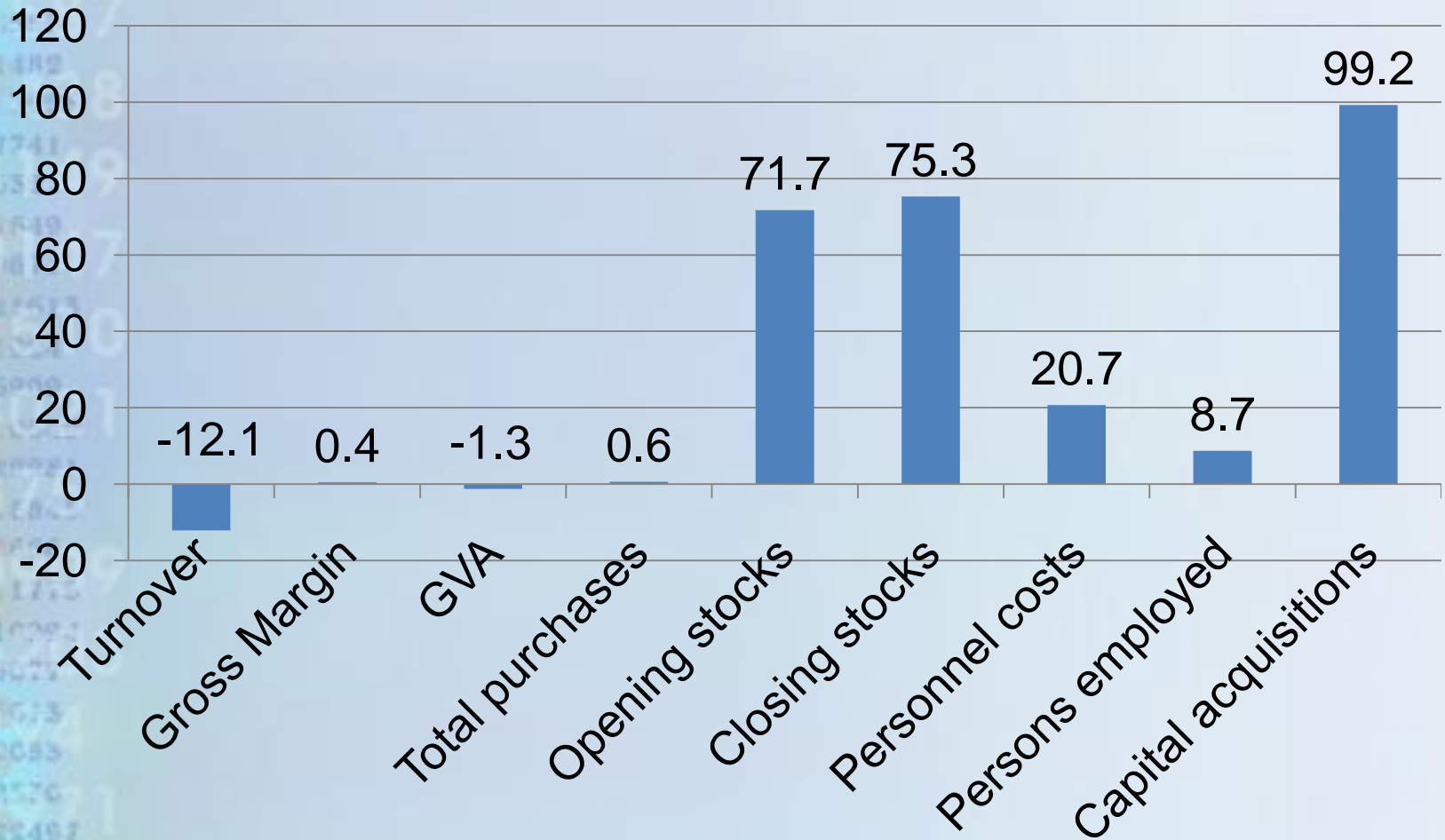
# Predictor variables

Variable from administrative source	Used to predict
Wages and salaries	Labour costs variables
Cost of sales	Stocks and purchases for direct resale
Gross profit	Other purchases
Turnover	Breakdown of turnover - retail, wholesale etc..

# % of state totals accounted for by single person enterprises



# % difference between administrative estimates and estimates from ASI



# Conclusions

- Impact on overall results should be low
- Agreed that model has potential for future use
- Similar modelling also done for industrial sectors
- May be possible to use model to impute for non-response for larger enterprises
- Why don't we use administrative data for all enterprises?