

USING ADMINISTRATIVE DATA TO GAIN INSIGHTS INTO MICRO-DRIVERS OF PRODUCTIVITY

CHIEN-HUNG CHIEN 

(Received 24 July 2021; first published online 13 September 2021)

2020 *Mathematics subject classification*: primary 62P20; secondary 91B38.

Keywords and phrases: administrative data, business network, productivity, synthetic data.

Since the late 1990s, the information technology revolution has improved our ability to process, store and analyse large administrative data sets. The idea of making better use of administrative data for social and economic research is not new. While researchers recognise that there is rich information contained in administrative data, they also acknowledge the need to address the methodological challenges of analysing it for research. This research explores methods to address some of these challenges to understand the micro-drivers of productivity. Productivity growth is important for firms so that they can either produce more at the same costs or produce the same amount at a lower cost. This gives firms the ability to increase wages, decrease prices (without reducing profits or wages) and increase profits. All of these have a positive impact on the economy, delivering higher living standards in the long run.

The thesis first addresses the challenge of handling large data sizes while estimating models to study the contributions from firm dynamics (that is, firm entry and exit, within-firm growth and reallocation) to aggregate productivity. The Australian Bureau of Statistics experimental integrated administrative data set contains more than 10 million workers across 1.5 million firms. The thesis uses a preconditioned conjugate gradient algorithm to solve a large sparse linear system. However, the solutions provided by the preconditioned conjugate gradient algorithm are not unique. The thesis shows how to impose appropriate restrictions to identify unique worker- and firm-specific effects. I use the estimated labour component to address a well-documented endogeneity problem in productivity analysis, to study contributions of firm dynamics to Australia's productivity growth in the period from 2002–03 to 2012–13 across 18 industries. The paper [1] shows that firms entering and exiting the market play a key role in improving productivity and therefore policies should aim to encourage competition (rather than provide advantages to continuing firms).

Thesis submitted to the Australian National University in February 2021; degree approved on 3 March 2021; principal supervisor Alan Welsh, associate supervisors Robert Breunig and Anton Westveld.

© 2021 Australian Mathematical Publishing Association Inc.

Firm dynamics are not the only micro-drivers of productivity. I analyse another experimental data set, integrating data from the Australian Bureau of Statistics, Intellectual Property Australia and the Australian Stock Exchange to understand the relationship between firm participation in business networks and firm performance. The thesis explores three types of business networks (research and development, commercial and shared directors) under three different assumptions. I use negative binomial models to account for over-dispersion observed in counts of patent and/or trademark applications. The paper [2] shows that, in general, there are positive associations between firm performance and these three types of business networks.

Another challenge with using large administrative data sets is applying methods to extract relevant information and use statistical models that do not assume the independence of observations to understand the complex relationships between firms. I use a semantic web approach to extract complex business network information from integrated administrative data. I combine the business network and firm information in exponential random graph models and latent space models to describe the factors contributing to firm participation in Australian business networks. The paper [4] shows that larger firms are more likely to form business networks than small and medium-sized firms. This may suggest a need for policies that encourage stronger collaboration between small and medium-sized firms.

Finally, there is a strong public good argument to make administrative data accessible for research while maintaining confidentiality. This is a particular challenge in the Australian context as some industries are characterised by an oligopoly or duopoly, making some firms easily identifiable. Typical data protection techniques such as information reduction may not be as effective for business microdata. I explore the use of synthetic data to make business microdata more accessible while maintaining confidentiality. The paper [3] shows that synthetic data can work as an alternative approach to maintaining confidentiality while making business microdata more accessible for research.

References

- [1] C.-H. Chien, A. H. Welsh and R. V. Breunig, 'Approaches to analysing micro-drivers of aggregate productivity', Research Paper 1351.0.55.164, Australian Bureau of Statistics, 2019. <https://www.abs.gov.au/ausstats/abs@.nsf/mf/1351.0.55.164>.
- [2] C.-H. Chien, A. H. Welsh and R. V. Breunig, 'Firm performance and business networks', Research Paper 1351.0.55.165, Australian Bureau of Statistics, 2019. <https://www.abs.gov.au/ausstats/abs@.nsf/mf/1351.0.55.165>.
- [3] C.-H. Chien, A. H. Welsh and J. D. Moore, 'Synthetic business microdata: an Australian example', *J. Priv. Confid.* **10**(2) (2020), 21 pages.
- [4] C.-H. Chien, A. H. Welsh and A. H. Westveld, 'Australian business networks', Research Paper 1351.0.55.063, Australian Bureau of Statistics, 2019. <https://www.abs.gov.au/ausstats/abs@.nsf/mf/1351.0.55.063>.

CHIEN-HUNG CHIEN, Mathematical Sciences Institute,
Australian National University, Canberra, ACT 2600, Australia
e-mail: joseph.chien@abs.gov.au