

# **How do we measure firm performance?**

## **A review of issues facing entrepreneurship researchers**

Josh SIEPEL\* and Marcus DEJARDIN\*\*

\* Science Policy Research Unit (SPRU), University of Sussex, UK; J.Siepel@sussex.ac.uk

\*\* DeFiPP – CERPE, Université de Namur & LIDAM – CIRTES,  
Université catholique de Louvain, Belgium; marcus.dejardin@unamur.be

May 2020

### **Abstract**

This paper aims to provide a succinct overview of the important challenges facing researchers seeking to perform firm level research, along with an overview of the different data sources that may be used, and some techniques that can be employed to ensure that data is robust. An emphasis is put on the linked importance of research design and choice of data. We discuss quantitative data and, more specifically, the measures used to observe firm performance, and present different types of data sources that researchers may use when studying firm level data, i.e. self-report data, official statistics, commercial data, combinations of data, and Big Data. We examine potential problems with data, from measurement to respondent and researcher errors. Finally, some key points and some avenues for future research are briefly reviewed.

**Keywords:** Firm Growth; Firm Performance; Methodology; Data Sources; Self-report Data; Official Data; Big Data.

**JEL-codes:** L26, L25, C18, C80.

## **1. Firm performance as a multifaceted phenomenon**

When researchers wish to study entrepreneurship, the choice of measures used to consider a particular phenomenon is of vital importance. Poor choices in relation to data can render the treatment of interesting research questions invalid and unpublishable. This paper aims to provide a succinct overview of the challenges facing researchers seeking to perform firm-level research, along with an outline of the different data sources that may be used, and some techniques that can be employed to ensure that data are robust.

While the nominal focus of ‘entrepreneurship’ as a field is the entrepreneur, the unit of analysis used in academic research in the field varies widely. Quantitative methods are widely applied to the study of individuals, teams, enterprises or geographical regions. Our focus in this paper is the ‘firm’, but as we discuss below this concept itself may be fluid depending on the research question. Fundamentally with different levels of analysis and different research questions, the demands placed upon the data will be necessarily different. A researcher trying to study the impact of founder attitudes on a firm’s innovative performance will require different data, and will face different challenges, from a researcher exploring the firm’s earnings volatility and survival.

Before we begin discussing key issues, it is important to set out some key definitions and exclusions regarding the issues that we are not considering in this case. Our primary focus here is firm-level research. We define a firm by focusing on the aggregate of the productive organisational structure rather than the production function (Mäntysaari 2012), with the difficulty that the limits of the organisation become more and more blurred, as Zingales (2000) has pointed out already, twenty years ago. In light of this definition, we do not consider research at the level of the individual, even if those individuals might be self-employed. Consequently, we do not directly address the issue of owners’ and workers’ earnings or well-being. We do not consider issues around the analysis of data (e.g., regression bias). While this represents a huge issue, there is an enormous body of literature on research methods around data analysis and interested readers are advised to refer to that separately. Our focus is on making sure that the data used are appropriate for a particular case.

One point that we wish to emphasise here is the linked importance of research design and choice of data. With this said, we do not explicitly discuss technical problems with research design, in as far as it relates to the design of instruments such as surveys. While we will talk about researcher error (in section 4 below), there is again a substantial literature on research design, and on instruments too. However, we remind the reader that a good research design is

vital for avoiding issues that may emerge later on in the research process (e.g., designing questionnaires that will allow researchers to address questions about endogeneity).

With the above caveats in place, we can begin to consider our discussion of quantitative data, the problems associated with them and how these difficulties may be addressed. The next section discusses measures used to observe firm performance. Section 3 then presents an overview of the different types of data sources that researchers may use when studying firm-level data. Section 4 discusses potential problems with data. Finally, the paper concludes with a brief discussion of some key points and some avenues for future research.

## **2. Firm performance: What are we measuring?**

What is a firm? While we happily leave this particular question to the literature on theory of the firm (see, e.g., Nightingale 2008), it is an important starting place in our discussion of firm performance. More specifically, there is an important disconnect between what firms *do* (produce and sell goods or services) and what we as researchers *observe* (and, for that matter, *can observe*). Quantitative research at the firm level is generally interested in the factors, all other things being equal, that cause firms to perform differently – better or worse than other firms. However, the measures that we use to observe performance may differ from those that are important to entrepreneurs: for instance, a scholar may choose to study employment growth, whereas an entrepreneur might be seeking to drive growth in profits instead. Consequently, we must be aware that firms, like the story about the blind men and the elephant,<sup>1</sup> look very different under the lens of different measures.

This leads to another important point about firm performance: it has a strong temporal component. This in turn leads to a significant observation to make about the nature of firm performance: as argued by Chandler et al. (2009), different performance measures are correlated, but the correlation is not strong. A company that is profitable may not have significant employment, and a company that is highly innovative may not have a high level of sales. Moreover, growth (as will be discussed below) does not happen at once across different measures but tends to happen in a sequence (Coad et al. 2017). So for these reasons it is

---

<sup>1</sup> This refers to an ancient parable, originating on the Indian subcontinent, of blind men who encounter an elephant and, having never encountered an elephant, disagree on what the elephant is like based on touching it. Each blind man feels part of the elephant – for instance, the trunk, a tusk, its skin, the tail – and generalises based on his observation. The moral of the parable is therefore that people tend to judge complex phenomena by limited experiences.

important to emphasise that what we capture as researchers may not necessarily be an ‘absolute’ picture of a firm; it is more akin to a snapshot. To extend the ‘gazelle’ analogy from the literature on high-growth firms, gazelles may run, but they also sleep, and a photo of a gazelle mid-stride will look different from one in which the gazelle is feeding, or asleep. Similarly, data on a firm over different periods may paint very different pictures. With this proviso in place, let us consider some of the widely used measures for firm performance.

## **2.1 Measures widely used in firm performance**

In this section, we will discuss some of the measures used by researchers to study firm performance, including their relative strengths and weaknesses.

### **2.1.1 Employment**

The number of people employed by a firm is a standard control and represents a core metric for understanding the firm’s size. Measures of employment growth are also relatively common, either on their own or as a metric supporting other measures such as high-growth firms. One common caricature is that employment growth is a preferred metric for economists and policymakers, who study employment growth as a means of exploring job creation. By contrast, entrepreneurs are less likely to view employment growth as a measure of success, as an increase in employee numbers drives up staff costs and means (as has been described by entrepreneurs) ‘simply more mouths to feed’. This illustrates the point made above regarding the different measures and judgements made on the part of the firm, and emphasises the relationship between metrics. For instance, the literature on the sequencing of growth provides some mixed findings on the role of employment growth. In general, employment growth is found to drive further sales and profit growth (Coad and Rao 2009; Coad 2010; Coad et al. 2014). However, in high-growth firms, employment growth often lags behind as one of the last steps in the growth process (Coad et al. 2017).

### **2.1.2 Turnover**

Following the caricature presented above, if employment growth is favoured by economists and policymakers, turnover (and derived measures) are more common in the management literature, as this aspect represents a growth in sales and better measures a firm’s performance, as opposed to its size. The literature on firm growth (see above) suggests that for most firms

turnover growth follows employment growth, but for high-growth firms turnover growth can precede employment growth.

### **2.1.3 Profits/Profitability**

A firm's ability to generate superior profits is a clear element of the recognition of entrepreneurial opportunity, and as such profits are another important measure of firm performance. Both Chandler et al. (2009) and Coad et al. (2017) (the latter only for high-growth firms) point to the importance of identifying profits for driving future growth. There are a number of ways of measuring profits, from direct measures as reported on financial statements to the financial ratios normally used in the finance literature (e.g., return on assets, etc.). These latter areas are less commonly used, which is typically a function of the availability of data, but they do occasionally appear in the literature.

### **2.1.4 Productivity**

Measured at the firm level, productivity is an important indicator of the firm's (in)efficiency in using production factors. Productivity can refer to labour productivity – for instance, value-added per worker or per worker hour – or less frequently to capital productivity, such as value-added per unit of fixed capital stock, whose value is closely linked to the types of technology that are used, usually labour- or capital-intensive technology. Another important productivity measure is the firm-level total factor productivity (TFP) (Gal 2013). Note that, when output is not available, a proxy for output might be deflated firm sales, but the quality of this proxy will drastically depend on the quality of the deflator and therefore the homogeneity of the produced good (Melitz 2000). Important antecedents of productivity are the size of the firms, capturing increasing return to scale; process as well as organisational innovation.

### **2.1.5 R&D/Innovation**

There are a number of links between the entrepreneurship and innovation literatures, and consequently there are a large number of studies in the entrepreneurship literature that focus on the innovative activities of firms. Innovation is a difficult topic to measure, primarily because studies of innovation are limited to inputs, such as research and development (R&D), staff or financial resources; intermediate outputs, such as patents; and outputs such as sales of new products (Hopkins and Siepel 2013). Consequently, much of the literature in

entrepreneurship that draws upon innovation measures will either rely on clearly available metrics such as R&D or patents, or on measures derived from survey instruments such as the Community Innovation Surveys in Europe, for instance the share of turnover derived from new-to-market products.

### **2.1.6 Firm survival**

The length of time a firm is active on a competitive market can be interpreted as a measure of its ability to compete. As such, the survival of the company is a performance measure. However, the survival approach of firm performance poses several problems. Measuring survival means being able to calculate the duration (usually in months or years) between firm entry into and exit from the market. A first batch of issues arises when you have to define 'entry' and 'exit'. The entry on a product market of an existing firm represents a different reality from the entry of a new firm. Regarding exit, this cannot be necessarily assimilated with failure and/or bankruptcy. The activity can indeed cease for many other, positive or more dramatic, reasons that are unrelated to business (e.g., the retirement or death of the owner). Exit of the firm does not correspond necessarily to business closure. The business activity can indeed be competitive and just taken over by another firm (Parker 2009, chapter 14). A second batch of issues relates to duration (see below, section 4).

## **3. Sources of data**

Following the discussion above, now let us consider some of the different classes of data that may be used to measure and observe firm performance. As there is an enormous range of datasets, this is only a rough typology. To illustrate the challenges associated with these types of data, we draw primarily on examples from the UK, though other cases are also discussed as these have some relevance.

### **3.1 Self-report data**

Within the academic field of entrepreneurship, a large majority of all publications draw upon self-report data (Chandler and Lyon 2001; Crook et al. 2010). These are data collected through either an interview or a questionnaire in which the respondent gives his or her answers to a series of questions. The potential of self-report data to show biases has already been widely discussed in the literature, particularly in fields such as psychology and organisational behaviour. For instance, Chan (2009) identifies two distinct problems with

self-report data: the validity of an individual self-reported variable (also known as construct validity); and the combining of two or more self-reported variables (resulting in ‘common method variance’).

Major datasets, well known by scholars in the entrepreneurship field and derived from questionnaires fulfilled with self-reported answers, are those of the international Global Entrepreneurship Monitor and the US Panel Study of Entrepreneurial Dynamics (PSED) I & II. They have been developed with much attention for constructs and measurement validity. The case of the PSED, being a panel with up to six follow-up interviews for PSED II, represents undoubtedly an asset for research, in particular for new firms’ processual development and survival analysis.

Whatever their limits, self-report data remain a major source of information for research. They open very largely the possibility for scholars to collect original data about personal or organisational characteristics that are barely available, if not at all, in official statistics. For example, this type of data allows for the identification of specific characteristics of the business owner, her personality traits or her personal value system; and with regard to the firm, a description of the firm’s strategy and so on.

While bespoke, hand-collected data are valuable, one limitation of collecting these data is that they are likely to be cross-sectional, which can limit the scope of the analysis that can be conducted, unless the researcher has the time and resources to collect multiple waves of data. One way of addressing this, if possible when defining one’s research design, is to combine self-report data with data coming from other, such as official or commercial, data sources. This would require a key allowing one to bridge datasets, like the firm identification number, but it can yield a range of different insights without the cost of additional data collection.

### **3.2 Official statistics**

An opposite to the collection of self-response data is the use of official, or government-collected or -sanctioned data. A vast amount of information is collected by the public sector, but firm-level studies may broadly be categorised in three ways: official surveys, mandated disclosure and administrative data.

Governments may undertake surveys for any number of reasons, but the resulting survey instruments will often have been carried out using a standard practice of national statistics offices in terms of survey design, sampling, data collection and so on. Likely the most widely used survey of this kind is the Community Innovation Survey (CIS), which is carried in a

(broadly) standardised form across the European Union and elsewhere in the world. The use of these surveys is beneficial in that they are widely accepted as a valid source for study and many publications have been derived from the analysis. As they are publicly available (albeit with varying access arrangements and data security restrictions by country), the core datasets are standardised and lend themselves to replication and the sharing of code. One drawback of these instruments is that they are subject to similar issues around response bias discussed above, but at least the researcher may know that she or he is using a widely used dataset that has been accepted by journal reviewers before. Another drawback of using these instruments is that it may be difficult to match the research question to the available data: for instance, researchers using the CIS studies who wish to study intellectual property will find a common set of binary variables that capture the quantity of different types of intellectual property and would need to look elsewhere.

The French Insee<sup>2</sup> SINE (Système d'information sur les nouvelles entreprises, New Enterprises Information System), which was launched in 1994, is another example of surveys being conducted by the national statistics offices. The programme is dedicated to the observation of new French firms and collects self-reported information about the business owner, the firm start-up, development and strategy, the firm relations with other firms and so on. A sample of new firms, representing a cohort of new firms created within the first half of the year, is selected every four years according to a sample selection method, and surveyed a few months after the firms' creation. The surviving firms of the cohort sample are surveyed again after three years and five years. Not all new firms are surveyed, but the statistical sampling method that is used by Insee provides a guarantee about the representativeness of the sample at both national and regional levels.

An alternative source of data is mandated disclosure, for example via company accounts or securities filings. The quantity, quality and level of detail required will vary by country, as will the availability of the data. In general, larger companies, such as those listed on stock markets, will have substantial amounts of data disclosed as terms of their stock market membership. However, for smaller firms, the availability of these data is variable. For instance, in the UK, all firms must make annual disclosures to Companies House, and all filings made by these companies are then publicly accessible on the Companies House website. However, as a means of data collection this website is not necessarily efficient, and

---

<sup>2</sup> Institut national de la statistique et des études économiques (National Institute of Statistics and Economics Studies).



commercial databases such as those of Bureau van Dijk (discussed below) can often provide the data disclosed by these firms without the challenges of data collection. With this said, when the production of performance figures is a matter of compliance, there is considerably less room for either intentional or accidental misstatement, as intentional misstatement is subject to legal proceedings.

An alternative approach to use of public data comes from the use of administrative data, which are collected by the government for its own purposes, typically taxation or generation of official statistics. The types of data available vary by country but often will contain detailed geographical information as well as information on employment and sales. In addition, other figures may be available, relating to workforce, taxation, innovation and any number of other areas. A significant benefit of these data sources is that the risk of measurement error associated with self-report data is substantially lower, as in many cases respondents could be criminally responsible for incorrect data provided. Alternatively, the data may be passively collected by the government (e.g., via the tax system), so there is little response at all. Access to these data is generally subject to extensive data protection regulation and researcher clearance, but in return the researcher has access to an entire universe of firms. Examples of these datasets include the Business Structure Database in the UK, the Central Balance Sheet Office at the National Bank of Belgium, which collects and makes publicly available the annual accounts of almost all commercial firms active in Belgium, and more. One challenge of using these forms of data is that elements of measurement error might be introduced that are not immediately obvious to the researcher. For instance, the Business Structure Database in the UK is based on a ‘snapshot’ of the live database containing information of employment and turnover for the population of UK firms on a particular date each year. Any subsequent changes are not captured until the next ‘snapshot’ is taken the following year. Similarly, the turnover figures are derived from the value-added tax figures, which may not capture a firm’s true performance if the company has sought to avoid paying tax. A final limitation is that the variables available are limited to those collected by the government, so other measures a researcher might want could very well not be included (as an example, a researcher wishing to study productivity may not find all the variables required to study TFP in a dataset that was not explicitly designed to do so). Despite some limitations, these datasets represent a powerful tool for the study of performance by providing validated, reliable information on a universe of firms.

### 3.3 Commercial data

Datasets that are commercially available may address some of the weaknesses described above in self-report and public data. Although they are themselves often based on official datasets, these commercial datasets may be relatively easy to access (at least for researchers at institutions who already have memberships) and will have a number of strengths. For this case, let us consider the datasets of Bureau van Dijk (BvD), which provide a good example. BvD supplies a number of datasets covering a range of geographical areas, including Orbis (worldwide), Amadeus (continental Europe) and FAME (UK). The general rationale of these datasets is bringing together data from securities or publicly accessible tax and registration filings for the firms in the areas in question, and augmenting these with other information where available. These datasets provide an impressive array of variables, including a range of detailed financial variables, geography, board members, R&D and product information. As such, they represent a very useful measure for researchers who wish to measure a range of metrics across a large number of firms. The data provided are robust, being generally derived from securities filings, and as they are commercial products researchers have access to customer support. Perhaps the major challenge of using these datasets is missing data. Datasets such as these rely on publicly available data, but in many cases a firm may be unable or unwilling to file the relevant documentation, resulting in substantial missing data problems. These issues are particularly pronounced for smaller firms. To give an example of where these data are used and why, the FAME dataset was used in Coad et al. (2017) to investigate the growth patterns of high-growth firms, considering measures of employment, turnover, profits and assets (which are not otherwise available in the UK via public or private datasets).

Another form of commercial data is generated as a by-product of a business's activities. An example here is the stream of work using data from Barclays bank by Storey, Coad, Frankish and colleagues (Frankish et al. 2013; Coad et al. 2013). These papers use proprietary data tracking the business activities of firms that opened bank accounts with Barclays in one quarter. The data allow the avoidance of particular biases (especially around survival), thus making them useful above all for tracking a range of issues around firm performance and demographics. Other uses of commercial by-product data could include 'scraped' web data or data accessed via the application programming interfaces (APIs) of web platforms. These forms of data have considerable potential for identifying specific answers to specific questions.

### **3.4 Combinations of data; employer–employee data; multilevel data; ‘Big Data’**

As already mentioned, when we were introducing self-report data, research can benefit tremendously from the combination of different datasets, implying the use of key variables to link datasets. Sometimes, the research questions that are raised make this combination simply necessary because no dataset exists that can cover all the variables involved in the analysis.

Regarding the explanation of firm performance, new Linked Employer–Employee Data (LEED) have made possible investigations that previously were unattainable. Employer–employee datasets include information about the company and its employees. They usually, but not exclusively, combine administrative data. The information can appear relatively basic (such as information on the origin of the employees before entering the company, or their destination when they leave), or more sophisticated. As an example, using employer–employee data for private-for-profit businesses active in New Zealand, Kirker and Sanderson (2018) were able to test the relation between firm productivity growth and the quality of their new workers, as well as the knowledge new workers may have acquired working for previous employers; the authors were able to explore whether companies benefit from knowledge brought to them by new employees.

Multilevel data are used in the estimation of linear mixed models – that is, hierarchical linear models and multilevel models (Garson, 2013). Suppose that you suspect that the performance of firms can be determined by their contexts, and contexts differ from each other with respect to some characteristics potentially entering into the explanation of firm performance. For example, you expect that the local cultural dimension, which differs within a given State according to local entities, could explain firm performance. By neglecting the existence of clusters (or groups) of firms sharing a given context that differs itself from other contexts, the scholar could neglect that the assumption of independent residuals is not respected. In our example, multilevel data including a measurement for culture at the local level and firm-level data would allow, conditionally to a sufficient number of observations for each level, the use and estimation of linear mixed models (Snijders and Bosker 2011; Maas and Hox 2005).

The collection, storage and provision of gigantic databases open up new, unsuspected opportunities for economics, management and econometrics research (see Einav and Levin 2014; George et al. 2014; Varian 2014).<sup>3</sup> Regarding empirical studies in firm performance,

---

<sup>3</sup> Coincidentally, while this paper was being written, a call for papers was launched by *Small Business Economics: An Entrepreneurship Journal* for a special issue entitled ‘Rethinking the entrepreneurial

big data represent an opportunity to open very widely what is still largely appearing as a black box and to address new, or neglected due to lack of data, research questions. For the sake of illustration, what is still very much a potentiality would consist in gathering individual (personal-)level data of structural and behavioural types within and outside the firm listed with reference to the firm and therefore connected with firm-level data, the whole forming a multilevel set of data.<sup>4</sup> However, it has to be made clear that what appear to be new and tremendous research avenues should not be allowed to evade crucial ethical and regulation issues to which the researcher has, or will have, the obligation to respond.

## **4. Types of problems**

### **4.1 Measurement error**

A host of issues can lead to incorrect measurement. These can consist of gaps between the measures used and the ‘facts’, as well as sample selection and biases. There is a substantial literature on measurement error in this context (cf. Bertrand and Mullainathan 2001; Bound et al. 2001), but here let us point out a few salient examples.

The use of measurement constructs is vital for the validation of new concepts, but there are substantial challenges in this area. Chandler and Lyon (2001) and Crook et al. (2010) summarise some of the key issues with research design and construct measurement by conducting both a meta-analysis and survey of experts in the entrepreneurship field. Their conclusion is that, while in its early days the field of entrepreneurship was characterised by weaker use of constructs, the field has developed substantially since then and the studies being published in top journals are becoming more robust. To our knowledge, a further updating of these studies has not yet taken place, but it could provide a good opportunity for further research.

Beyond the constructs used, there is a range of potential biases that may impact the analysis. For instance, nonresponse bias (i.e., those respondents who are invited to participate in the survey but do not participate) can be a source of a range of biases. Recent work by Rutherford et al. (2017) suggests that these issues are important, but there is little evidence to suggest that this form of bias is distorting results in entrepreneurship research more broadly.

---

(research) process: Opportunities and challenges of Big Data and Artificial Intelligence for entrepreneurship research’, with Martin Obschonka (Australian Centre for Entrepreneurship Research, QUT, Australia) and David B. Audretsch (Indiana University, USA) acting as guest editors.

<sup>4</sup> Contextual data would allow furthermore the consideration of an additional, upper level.

Beyond the more mechanical issues around samples, the biases among respondents represent another significant source of management error. Zhang and Cueto (2017) discuss a number of biases in entrepreneurship research, highlighting how these might impact results and potential avenues for future research. An example of these types of biases is the stream of research on entrepreneurial overconfidence and overoptimism, which leads entrepreneurs to overstate their abilities and prospects, respectively.

## **4.2 Missing data**

Missing data is a frequent problem in much empirical research, particularly research considering firms, as has been discussed above with regard to datasets. Missing data may happen for any number of reasons, but it may be treated in a number of ways, as outlined in Schafer and Graham (2002). There are some key issues that indicate how researchers deal with missingness in data. One key question is whether data are missing at random. If this is the case, these missing data may be convincingly ignored as they do not substantially impact or bias the results. By contrast, if data are missing but the missing data are not random, this represents an important bias. Fortunately, there are a number of statistical techniques that may be used to address missing data problems, as discussed in Schafer and Graham (2002), Allison (2001), Honaker and King (2010) and others.

The problem of censored and truncated data may be considered alongside the missing data issue. Censored data corresponds to a situation of a complete sample of individuals, but for some individuals the actual value of the variable of interest is unknown (e.g., the largest values in a distribution are grouped at a boundary). The presence of truncated data means that the sample is incomplete, with specific individuals excluded from the sample dataset leading to a truncated distribution compared with the population distribution (e.g., firms with a number of employees fewer than 10 are not surveyed). Although censoring and truncation problems might be pervasive, they are particularly notable in survival analysis. With censoring, the survival of a sample of firms is observed for the duration of the study; surviving firms at the end of the study will all be grouped at a boundary, though all will die one day. With truncation, some of the firms that are part of a cohort had already died when the study started and cannot be included in the sample. Luckily here too, statistical techniques have been developed to address censored and truncated data problems (see i.a., Klein and Moeschberger 2003).

### **4.3 Common method variance**

Among the types of problems or biases associated with entrepreneurship research, one frequent concern is common method variance, which relates to the method of measurement driving variance rather than the construct being studied. This may take several forms, for instance where the same respondent provides both the predicting and the dependent variable. A full list of forms of common method variance is found in the succinct and highly cited review by Podsakoff et al. (2003), but one theme that emerges from these biases is that the type of studies most subject to this form of variance comprises the cross-sectional, uncorroborated, self-reported datasets frequently found in the entrepreneurship literature. The solution to these problems is, as in many cases, use of more robust data: for example, panels, independent data in conjunction with self-response data and good survey design practice can help to address some of these issues. A reasonable summary of how authors and reviewers can deal with some of these issues is provided in Conway and Lance (2010). It is worth mentioning that, while statistical techniques are available to address this form of variance, good instrument design, as discussed above, is generally preferable to these post hoc statistical methods.

### **4.4 Incorrect data and respondent error**

One considerably more difficult problem to address is that of data that are present, but incorrect. This might take several forms. For instance, in survey data respondents may not have particular incentives to be truthful, particularly if responses are commercially sensitive. The question of whether self-response data may be considered reliable is an important one that has perhaps not been considered as widely as it could be. One germane study addressing many of the issues raised in this paper is the contribution of Fraser et al. (2007), who used panel survey data to explore recall bias among survey respondents. They found a close correlation between self-reported and administrative data, but identified a number of potential sources of bias. Among these, firms that have undergone changes, such as growth, contraction or relocation, are more likely to demonstrate bias. More recently, Bloom et al. (2019), working with the US Census, took advantage of an administrative error to resample a group of survey respondents, and found that measurement error accounted for 45 per cent of variation in responses.

Moreover, even if we assume the sincerity of responses made by respondents, simple mistakes such as data entry errors or improperly read instructions can prove to be problematic.

For instance, if a survey respondent is asked for his or her company's annual turnover, in thousands, a busy person who is not reading closely might enter £100,000 in the space, resulting in a data point where a company might have two employees and £100m in turnover. Some of these issues may be addressed in data cleaning, but they may prove to be considerably more difficult to address.

#### **4.5 Researcher error**

Further to these issues, another major issue is simply that of researcher error. Researchers, particularly those working with quantitative data, have the implicit responsibility to use the most appropriate and robust approaches in addressing their research question. Yet there is a wide variety of errors or misjudgements that can arise through inappropriate use of econometric or data management techniques. Addressing this vast array of issues is beyond the scope of this paper, but the interested reader could do worse than consulting Kennedy (2003) or any of the more practitioner-oriented econometrics texts that have emerged over the years.

Finally, under the category of researcher error, we must also include those sources of bias whose intent is not innocent but represents wilful misrepresentation of results or data. The associated literature on academic misconduct is substantial and growing (cf. O'Boyle et al. 2017) due to the unfortunate proliferation of these issues, but researchers, particularly new ones, are advised to acquaint themselves with good practice.

#### **5. Conclusion**

This paper has aimed to identify some key issues around the measurement of firm performance and the data sources that researchers can use to study it. In conclusion, we wish to highlight a few additional points that may be of use to readers. Firstly, it is important to realise that advances in both computing power and econometric and other analytical techniques have meant that researchers have the ability to answer questions that would have been virtually impossible even ten years ago. The direction of travel in academia is toward the acceptance of these new tools and techniques, and junior researchers in particular will do themselves well to become familiar with these advanced techniques, which unlock huge opportunities in terms of analysis.

Another point to highlight is the importance of research design. It may seem trite to emphasise that the measures being studied must match the theories or concepts being

explored, but this remains a huge issue. There is a range of approaches to research design that allows for the identification of very interesting findings, and a student or junior researcher is well advised to familiarise themselves with these techniques. For instance, while we have not discussed the rise of quasi-experimental data or discontinuity studies, these are the types of research designs that are enabling ever-bolder statements to be made about entrepreneurship and firm performance more broadly.

In conclusion, a vast array of data resources is available to researchers wishing to study the drivers of firm performance. However, the careful alignment of measures, data sources and empirical approaches is required in order to generate convincing findings that will allow for the development of robust explanations for the variations in firm performance.

## References

- Allison, P.D. (2001). *Missing Data: Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: SAGE Publications.
- Bertrand, M. and Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67–72.
- Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I. and Van Reenen, J. (2019). *American Economic Review*, 109(5), 1648–1683.
- Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of Econometrics*, 5, 3705–3843.
- Chandler, G.N. and Lyon, D.W. (2001). Issues of research design and construct measurement in entrepreneurship research: the past decade. *Entrepreneurship Theory and Practice*, 25(4), 101–113.
- Chandler, G.N., McKelvie, A. and Davidsson, P. (2009). Asset specificity and behavioral uncertainty as moderators of the sales growth: Employment growth relationship in emerging ventures. *Journal of Business Venturing*, 24(4), 373–387.
- Coad, A. (2010). Exploring the processes of firm growth: Evidence from a vector autoregression. *Industrial and Corporate Change*, 19(6), 1677–1703.
- Coad, A. and Rao, R. (2009). Firm growth and R&D expenditure. *Economics of Innovation and New Technology*, 19(2), 127–145.
- Coad, A., Daunfeldt, S.O., Holzl, W., Johansson, D. and Nightingale, P. (2014). High-growth firms: Introduction to the special issue. *Industrial and Corporate Change*, 23(1), 91–112.



- Coad, A., Siepel, J. and Cowling, M. (2017). Growth processes of high-growth firms as a four-dimensional chicken and egg. *Industrial and Corporate Change*, 26(4), 537–554.
- Conway, J.M. and Lance, C.E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, 25(3), 325–334.
- Crook, T.R., Shook, C. Morris, M.L., and Madden, T. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. *Organizational Research Methods*, 13(1): 192–206.
- Einav, L. and Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210), 715–721.
- Frankish, J.S., Roberts, R.G., Coad, A., Spears, T.C. and Storey, D.J. (2013). Do entrepreneurs really learn? Or do they just tell us that they do? *Industrial and Corporate Change*, 22(1), 73–106.
- Fraser, S., Green F.J. and Mole, K. (2007). Sources of bias in the recall of self-generated data: the role of anchoring. *British Journal of Management*, 18(2), 192–208.
- Gal, P. (2013). Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS, OECD Economics Department Working Papers, No. 1049, OECD Publishing, Paris.
- George, G., Haas, M.R. and Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–332.
- Honaker, J. and King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2), 561–581.
- Hopkins, M.M. and Siepel, J. (2013). Just how difficult can it be counting up R&D funding for emerging technologies (and is tech mining with proxy measures going to be any better)? *Technology Analysis & Strategic Management*, 25(6), 655–685.
- Kirker, M. and Sanderson, L. (2018). Firm Productivity Growth and Its Relationship to the Knowledge of New Workers, 15 February, Mimeo.
- Klein, J.P. and Moeschberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd edition. Heidelberg: Springer-Verlag.
- Mäntysaari, P. (2012). *Organising the Firm: Theories of Commercial Law, Corporate Governance and Corporate Law*. Heidelberg: Springer.
- O’Boyle, E., Banks, G. and Gonzalez-Mule, E. (2017). The chrysalis effect: how ugly initial results metamorphosize into beautiful articles. *Journal of Management*, 43(2) 376–399.

- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Rutherford, M.W., O’Boyle, E.H., Miao, C., Goering, D. and Coombs, J.E. (2017). Do response rates matter in entrepreneurship research? *Journal of Business Venturing Insights*, 8, 93–98.
- Schafer, J.L. and Graham, J.W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2), 147–177.
- Varian, H.R. (2014). Big data: new tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28.
- Zhang, S.X. and Cueto, J. (2017). The study of bias in entrepreneurship. *Entrepreneurship Theory and Practice*, 41(3), 419–454.
- Zingales, L. (2000). In search of new foundations. *The Journal of Finance*, 55, 1623–1653.

## Websites

- GEM: <http://www.gemconsortium.org/> (accessed January 2020).
- PSED I: <https://doi.org/10.3886/ICPSR37203.v1> (accessed January 2020).
- PSED II: <https://doi.org/10.3886/ICPSR37202.v1> (accessed January 2020).
- SINE: <https://www.insee.fr/en/metadonnees/source/s1108> (accessed January 2020).