INTERNATIONAL DATA ON MEASURING MANAGEMENT PRACTICES

By Nicholas Bloom, Renata Lemos, Raffaella Sadun, Daniela Scur and John Van Reenen*

Rapid advances in computer power and increased openness of national statistical agencies have led to unprecedented availability of large datasets. Consider three types of firm datasets. First, governments collect administrative data on firms: information on jobs, investment and output has long been collected to calculate national, industrial and regional statistics. In recent years, the underlying micro-data, typically at the establishment level, have become widely available to researchers in many countries.

Second, there has been an explosion of Big Data - various forms of data typically created for business purposes. Although data scraped from the internet, video and other media is more often discussed, the most common form of Big Data for researchers is firm accounts. In most developed countries there is a legal duty to publish basic annual accounts for the protection of investors (even if this is only a name, address and owner), and these have been digitized by private sector firms like Bureau van Dijk (BvD).[1] Products like ORBIS contain over 50 million firms from almost every country in the world and can be used to address many questions. Another example is Compustat, which contains extensive data for about 6,000 listed US firms but excludes the other 99% of private firms. We focus on a third type of international firm data, which is collected from surveys. In an age of rich administrative and Big Data why bother with such surveys? The primary

* Bloom, Stanford University, 579 Serra Mall, Stanford CA 94305, nbloom@stanford.edu, Lemos, London School of Economics, Houghton Street, London, WC2A 2AE, r.lemos@lse.ac.uk; Sadun, Harvard University, Morgan Hall, Harvard Business School, rsadun@hbs.edu, Scur, University of Oxford, Nuffield College, Oxford, daniela.scur@economics.ox.ac.uk, Van Reenen, London School of Economics, Houghton Street, London, WC2A 2AE j.vanreenen@lse.ac.uk,

[1] Usually aimed at investors like banks who are interested in particular firms or asset classes.

reason is that many important social science concepts such as management and organizational practices are not well measured in other types of data. Perhaps the best data that currently exists on this is the identity and history of senior managers that is available for some high profile firms, e.g. publicly traded US firms. While this is useful for certain questions, such as the link between managerial style and company performance (e.g. Bertrand and Schoar 2003), there is evidence that suggests a company is far more than simply the identity of its most senior employee (Bender et al, 2015). Moreover, this type of data on its own does not tell much about how firms are managed or organized. Therefore, over the last decade we have been working to fill this gap in data by collecting comprehensive information on management practices (see Bloom et al. 2014).

I Measuring Management

We began with an interest in trying to account for the very large and persistent productivity differences between firms even within narrow industries. After discussions with management consultants and industry participants, we focused on three broad areas that were generally agreed to be important for firm productivity. These were (A) performance monitoring (information collection and analysis), (B) target setting (using and stretching short and long run targets), and (C) performance incentives (rewarding high-performing employees, and retraining or moving underperformers).

The measures we focus on were designed to capture shifters of Total Factor Productivity (TFP). How could a firm produce the good or service it currently made in a more efficient way? This leaves out, of course, many important aspects of management – mergers and acquisitions, innovation, pricing, investment, leadership. We focus on the operational and Human Resource side of the business because this is an area where there is the most consensus over what constitutes a

"best practice". In principle, other areas of "strategic management" may also be open to the methods we describe here.

The two methods for gathering management data that we have focused on are (i) Open Ended questions (those with a wide variety of possible answers) and (ii) Closed Ended questions (those with a list of potential answers like "Yes or No"). We will use the World Management Survey (WMS) as our example of open interviews and the Management and Organizational Practices Surveys (MOPS) as our example of closed interviews.

Open Ended Questions: World Management Survey (WMS)

The WMS approach is modelled on what leading management consulting firms do when interviewing client firms in consulting engagements. We first implemented this in 2004 in a survey developed jointly with the consulting firm McKinsey & Co. (Bloom and Van Reenen, 2007). We used open questions to collect information. For example, on monitoring, we begin with asking the open question "*can you tell me how you monitor your production process?*". We continued with open questions focusing on actual practices and examples until the interviewer can make an accurate assessment of the firm's practices. For example, the second question on that monitoring dimension is "*what kinds of measures would you use to track performance?*" and the third is "*if I walked round your factory could I tell how each person was performing?*". These open questions are designed to mimize the chance we steer respondents to a particular answer

We target production plant managers using a 'double-blind' technique. One part of this technique is that managers were not told in advance they were being scored or shown the scoring grid. They were only told they were being "*interviewed about management practices for a piece of work.*"

(we avoid the words "survey" or "research" because of connotations with market research). The other side of the technique is that interviewers were not told in advance about the firm's performance. They were only provided with the company name, telephone number and industry. Since the survey requires some degree of business acumen and knowledge, we hired skilled interviewers – usually graduate students with business qualifications to run interviews. This double-blind approach tries to prevent firms from biasing their responses towards higher-scores, and interviewers from biasing their scores based on knowledge of the firm's performance.

To score these interview responses we had a grid for each question running on a scale from 1 to 5, where for example on the monitoring question discussed above a score of 1 was defined as "*Measures tracked do not indicate directly if overall business objectives are being met. Tracking is an ad-hoc process (certain processes aren't tracked at all)*" while a score of 5 was defined as "*Performance is continuously tracked and communicated, both formally and informally, to all staff using a range of visual management tools*". From this example it is clear that designing these surveys take some expertise in terms of selecting questions and response grids, and our experience was that this is an iterative process involving repeated rounds of testing and refinement. The full questionnaire is available on www.worldmanagementsurvey.com.

Finally, these surveys have to be run as an interactive conversation, which we did over the telephone to reduce travel time and ensure consistency. We obtained response rates of about 40%, interviewing managers for around 45 minutes. We provided one week of intense training combined with daily coaching and monitoring for our interview team.

Response rates to surveys in general have been falling in the US and other countries over time. For these type of surveys, private sector companies often only have response rates of 5-10% and although attempts are made to balance these on observables such as size, industry and geography there is an obvious concern over selection on unobservables. The much higher response rates achieved by the WMS are partly due to interviewer persistence, as senior managers are hard to reach and convince to take part on our interviews, but also because the survey itself is very interactive and thus more enjoyable for managers than simply being "pumped for information."

We also use endorsement letters from senior officials from respected institutions such as the Central Bank, Finance Ministry and Employers Federation. Given the high overhead costs to administer these surveys, each interview is budgeted at between US$400 and US$500.

Close Ended Questions: Management and Organizational Practices Survey (MOPS)

Closed ended surveys allow respondents to choose from a menu of answers, so the survey does not need an interviewer to run it over the telephone or face-to-face. The MOPS was designed in collaboration with the US Census Bureau to be comparable to the WMS questions. For example, in the monitoring section we asked how frequently were performance indicators tracked at the establishment, with options ranging from "*hourly", "daily", "weekly", "monthly, "quarterly", "yearly"* to *"never"*. The targets section asked about the design, integration and realism of production targets and the incentives section asked about non-managerial and managerial bonus, promotion and reassignment/dismissal practices. The full questionnaire is available on http://bhs.econ.census.gov/bhs/mops/form.html.

MOPS was sent to respondents by mail or e-mail as a mandatory supplement of the Annual Survey of Manufacturing (ASM) and followed up with multiple rounds of mail and e-mail for non-respondents, achieving a response rate of 80%. This survey mode is far cheaper to run: the 2010 US MOPS cost around $1.2m for 35,000 responses, that is, about $35 per response.

II Comparison of Open vs Closed Ended Surveys

Table 1 compares the two approaches on a number of dimensions. No one method clearly dominates the other, with the WMS vs MOPS a quality-cost and flexibility-scale tradeoff. In summary, the WMS approach likely elicits more accurate responses as respondents can be probed more deeply and asked for examples. It also can be run without any government support and still achieve reasonable response rates. However, the WMS has the disadvantage that it requires trained highly quality interviewers, which is expensive and harder to organize.

For the closed approach, collaborating with national statistical agencies like the US Census Bureau is a major advantage. First, it is possible to leverage off the sampling frames of existing surveys like the ASM. Second, it makes it easier to link to data on productivity from these surveys. Third and most importantly, if it goes out as a mandatory survey alongside the standard official surveys, response rates can be much higher (around 80% in the case of MOPS) and the survey can be administered at a larger scale. Overall, the WMS method has the advantage of accuracy, but the MOPS has the advantage of lower per-survey cost.

The WMS randomly samples medium-sized manufacturing firms (employing between 50 and 5,000 workers). Our initial view was that in smaller firms formal management practices may be less valuable. In very large firms we worried that one plant-interview would be too limited to

evaluate the whole firm. By contrast, in MOPS, we covered the entire firm size distribution using plant-level interviews. Although it was true that large firms were more likely to have higher management scores, we found that the link with performance extended throughout the size distribution, similar to McKenzie and Woodruff (2015) who find an important role for management in micro-firms in developing countries.

Both surveys were targeted at plant managers, who are senior enough to have an overview of management practices but not so senior as to be detached from day-to-day operations. We also collected a series of "noise controls" on the interview process itself – such duration, as the time of day, day of the week, characteristics of the interviewee and the identity of the interviewer. Including these in our regression analysis typically helps to improve our estimation precision by stripping out some of the measurement error. We have focused on management, but similar issues arise when measuring other aspects of firm organization. For example, in WMS and MOPS we collect data on the decentralization between the Head Quarters and plant managers over investment, hiring, sales and innovation decisions.

Measurement Error. Measurement error is endemic to all surveys, but may be a particular concern for more "subjective" management question than for questions like the number of employees. There are many ways to examine this issue. For example, the noise controls mentioned above and the correlation of management with external measures of firm performance provide useful checks. In the WMS we also re-interviewed 222 firms using both a different interviewer and a different plant manager at the same firm, finding scores have a correlation of 0.51 (p-value <0.001). In MOPS about 1,000 interviews were sent to the same plant twice, with around 500 being completed by different managers.

Table 1: Strengths and Weaknesses in two ways to collect management and organization data

Aspect	Open (e.g. WMS)	Closed (e.g. MOPs)
Accuracy of responses	High: Interview format gives opportunity to probe and ask for examples. Possible to implement "double blind" method to reduce preconception bias.	Medium: Harder to elicit truthful answers if respondents have preconceptions. Greater risk that respondents might misinterpret questions or rush through the survey.
Cost per survey	High: High quality trained interviewers needed to run survey. Training includes one week initial training and ongoing debriefing and calibration. Interviewers' time primarily spent recruiting managers to take part in the survey (rather than just running interviews).	Low: Initial design and execution costs, but this fixed cost can be spread over a very large number of respondents, so cost per survey is low. Costs can be higher in poorer countries where enumerators administer surveys onsite because of unreliable mail and e-mail networks.
Response rates	Medium: Interview is interactive and managers more engaged. We obtained an average response rate of 40%.	High: Co-operation with a National Statistics Agency can enable the survey to be mandatory. Given this response rates of around 80%. Without such co-operation, response rates will be low.
Replicability	Medium: Training needed to ensure the survey is delivered in same way. Useful to have some individuals who have worked in previous survey waves as trainers for other surveys foster comparability. Training and survey material is available on-line.	High: Questionnaire essentially the same across countries and already available pre-tested from by US Census Bureau.
International comparability	High: Multiple countries can be interviewed from same location. Using bilingual interviewers means makes it is easier to cross-check responses.	Medium/High: Easier to implement but there is a risk of differential interpretation if this is not carefully translated across languages.
Speed of delivery	High: Can complete a full survey wave in about 10 weeks. So including recruitment and set-up time possible to complete a survey wave from scratch in about 4 months.	Medium: Involves cooperation with national statistical agencies, so more planning work in advance. The survey period typically is around 3 months plus 1 to 3 months of data cleaning.

We found these different answers from the same plant had a correlation of 0.5 across respondents, suggesting that about half the management score is measurement error (and about 50% signal). In Bloom et al. (2013) we suggest this measurement error is about equal in magnitude to that of TFP, and found our management score and TFP have similar predictive power for future plant performance.

Hybrid Approaches. We have used Open vs Closed ended surveys as two binary altenatives but hybrid approaches are also possible. WMS has to involve a discussion with the manager and we typically have delivered this over the telephone. These are cheaper than face to face (as travel costs are saved) and are more comparable (the interviews are all conducted from a single location with common training and calibration) which is important when running international surveys. An alternative is to conduct the interviews face to face as we did in the 2010 Management, Organizational and Innovation (MOI) survey in Eastern Europe (see Bloom et al, 2012). This was delivered by a private survey firm (TNS) running face-to-face interviews across different countries, which made the survey execution relatively easy but increased the challenges of comparing scores across countries as different teams ran the surveys.[2] Similarly, the Mexican and Pakistan Statistical Agencys' own MOPS surveys – which are comparable to the self-reported US MOPS - were run face to face to increase response rates, due to the difficulties of contacting firms by e-mail or mail.

[2] Likewise, Lemos and Scur (2015) have a "Development WMS" for the public sector that has a much more detailed scoring grid that enables less skilled enumerators to administer it. It has been successfully used in India, Mexico and Colombia.

Conclusion: The impressive growth in the availability of detailed datasets over the past decades has greatly enhanced the scope of research opportunity. In this paper, we have presented an overview of the datasets we have been involved in creating over the past 12 years, along with a summary of the methodology behind them. There are exciting times ahead as more data is collected and becomes widely available.

REFERENCES:

Bertrand, Marianne, and Antoinette Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies" *Quarterly Journal of Economics*, 118 (4): 1169-1208.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter. 2015. "Of managers and management: evidence from employer-employee data", LSE mimeo.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen. 2014. "The new empirical economics of management", *Journal of the European Economic Association*, 12(4), pp.835-876.

Bloom, Nick Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten and John Van Reenen. 2013. "Management in America" CEP Discussion Paper 1258

Bloom, Nicholas, Helena Schweiger and John Van Reenen. 2012. "The land that Lean manufacturing forgot? Management practices in transition countries" *Economics of Transition* 20(4), 569-785

Lemos, Renata, and Daniela Scur. 2015. "Developing Management: an expanded evaluation tool for developing countries" CEP-LSE Mimeo

David McKenzie and Christopher Woodruff. 2015. "Business Practices in Small Firms in Developing Countries" NBER Working Papers 21505.