CENTRE for ECONOMIC PERFORMANCE

Occasional paper

41

The New Empirical Economics of Management

Nicholas Bloom
Renata Lemos
Raffaella Sadun
Daniela Scur
John Van Reenen

April 2014

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
Over the last decade the World Management Survey (WMS) has collected firm-level management practices data across multiple sectors and countries. We developed the survey to try to explain the large and persistent TFP differences across firms and countries. This review paper discusses what has been learned empirically and theoretically from the WMS and other recent work on management practices. Our preliminary results suggest that about a quarter of cross-country and within-country TFP gaps can be accounted for by management practices. Management seems to matter both qualitatively and quantitatively. Competition, governance, human capital and informational frictions help account for the variation in management. We make some suggestions for both policy and future research.

JEL classification: L2, M2, O14, O32, O33
Key words: Management, organization, productivity

This paper was produced as part of the Centre's Productivity and Innovation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

This paper is based on Van Reenen's FFBVA lectures in the 2013 San Diego AEA and in Madrid. Comments from seminar participants and an anonymous referee have helped improve the paper. WMS has received financial support from the Economic and Social Research Council, the International Growth Center, PEDL, the UK Department of Business and the National Science Foundation. Our partnership with McKinsey and Company (from whom we received no funding) has been essential for the project, in particular Pedro Castro, Stephen Dorgan, John Dowdy and Dennis Layton.

Nicholas Bloom is an Associate of the Centre for Economic Performance, London School of Economics and Political Science. He is also a Professor of Economics at Stanford University. Renata Lemos is a research project director for the World Management Survey within CEP's productivity and innovation programme, and a PhD candidate at Cambridge University. Raffaella Sadun is a CEP Associate and an Assistant Professor of Economics at Harvard Business School. Daniela Scur is a research project director for the World Management Survey within CEP's productivity and innovation programme, and a PhD candidate at Oxford University. John Van Reenen is the Director of the Centre for Economic Performance and Professor of Economics, London School of Economics and Political Science.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

N. Bloom, R. Lemos, R. Sadun, D. Scur, J. Van Reenen, submitted 2014

1. Introduction

The enormous variation in firm and establishment performance has become a focus of empirical and theoretical interest throughout the social sciences, including economics. The opening up of business micro-data by national statistical agencies and vast improvement in computer power to store and analyze very large and complex datasets have facilitated the careful documentation of this first order economic fact.

A decade ago we began a project now called the World Management Survey[1] (WMS) which sought to address the issue of whether management practices were an important factor in understanding the heterogeneity of firm productivity. Many theories put entrepreneurial or managerial ability at the heart of this issue, but until recently there was precious little large-scale quantitative data across firms, industries and countries to empirically investigate these claims.

This review paper seeks to draw together what has been learned from the research program in measuring and understanding management practices. It is an exciting research area, and there is a huge amount remaining to be done.

In short, there do appear to be methodologically robust ways of measuring core management practices. These do not cover every aspect of management – for example, we explicitly leave out more "strategic" aspects of management relating to innovation, marketing and finance. Nevertheless, the practices identified in our survey - monitoring, targets and incentives - appear to be informative for organizational performance across disparate sectors such as manufacturing, hospitals, schools and retail stores. Further, the small Randomized Control Trial (RCT) evidence does suggest a causal impact of "high dosage" management on productivity. In summary, management does indeed appear to be important in accounting for the large differences in cross-country Total Factor Productivity (TFP) as well as within-country differences.

[1] The WMS website (http://worldmanagementsurvey.org) has details on datasets, methods (with training materials on how to run your own survey), reports, papers and an online benchmarking tool for firms.

Our knowledge about why there are such large variations in management is still rudimentary. Competitive intensity is one important and robust factor in raising management quality, as is ownership and governance (e.g. family firms appear to have weak management on average). But empirical work examining other potentially fundamental factors such as information and co-ordination frictions is almost non-existent.

The structure of the paper is as follows. Section 2 looks at productivity variation across firms and countries both cross-sectionally and over time. Section 3 describes the WMS methodology, gives some results, and responds to criticisms. Section 4 examines the impact of management on performance and section 5 discusses some theoretical models of management. Section 6 offers some brief remarks on the causes of the variation of management and section 7 concludes.

2. Productivity variation

We begin by documenting the different types of productivity variation across countries, firms and time.

2.1 Aggregate time series

Solow (1957) found that a large fraction (87.5%) of the growth of output per worker in the US was due to growth in TFP rather than capital accumulation. The finding that TFP is at least as important as observable factors of production in such growth accounting exercises has been replicated for numerous countries. It is easy to forget that there was initially much skepticism over this result with many attempts to statistically explain away residual TFP as due to standard mis-measurement of capital or labor services.[2] The growth literature[3] has generally understood TFP to be due to the generation and diffusion of "hard" technological innovations such as hybrid corn, beta-blockers and information and

[2] Jorgenson and Griliches (1967) argued that when aggregation was done properly and inputs and outputs were correctly measured US TFP growth was negligible between 1945 and 1965. Griliches (1996) was later to revise his views.

[3] For example, see the Aghion and Howitt (2009) textbook.

communication technologies (ICT). Another important factor, however, could be "soft" technologies such as the management practices of Taylor's Scientific Management or Toyota's Lean Manufacturing.[4] Indeed, in Solow's original article he emphasized that TFP meant "*any kind of shift* in the production function" (emphasis in original).

2.2 Industry level time series

With the advent of better micro-economic data on plant and firm[5] productivity it became possible to decompose the growth of TFP into a "within-firm" and "between-firm" component. The traditional view is that the economy can be summarized by a representative firm, implying that productivity growth is *within-firm*. This could be from innovation expanding the technological frontier outward or from the adoption of existing ideas by incumbent firms.

The Schumpeterian tradition, however, has long emphasized the *between-firm* component. Much of aggregate productivity growth is from the reallocation of output away from less productive firms towards more productive firms. This reallocation can take place on the *extensive* margin as less productive firms exit and more productive firms enter.[6] This is the traditional notion of creative destruction, which is a Darwinian force of natural selection. But reallocation can also take place on the *intensive* margin as market shares get reallocated among incumbents away from the least efficient and towards the more efficient firms. In either case these are between-firm effects that are distinct from the traditional within-firm effects.

[4] See Alexopoulos and Tombe (2012) for a systematic analysis of these at the macro-economic level.

[5] We will tend to use plant and firm interchangeably for expositional ease, although obviously they differ in interesting ways. A firm can increase productivity by shrinking/shutting down its less efficient plants and growing/entering more efficient ones. This appears to be an important channel in the retail sector (e.g. Foster, Haltiwanger, and Krizan, 2006). The WMS data is collected at the establishment level, but most of the manufacturing firms have only one or two plants.

[6] Analysis of entrants has found that their measured productivity is surprisingly low, usually no better than incumbents. However, this appears to be due to an overestimation of their output price, because firm specific prices are usually unobserved and researchers use industry-wide price deflator instead. Foster, Haltiwanger and Syverson (2008) show in industries where plant-specific prices are observed entrants typically price below the average incumbent, so revenues deflated by industry prices will lead to an underestimate of entrant output and therefore also their productivity.

Bailey, Hulten and Campbell (1992) analyzed data from US manufacturing plants and argued that over a five year period about half of a typical industry's TFP growth was due to the reallocation of output between plants rather than ongoing incumbent within plant productivity growth. There are multiple ways in which to perform such statistical decompositions of industry productivity growth into within and between components - see, for example, Olley and Pakes (1996), Foster, Haltiwanger, and Krizan (2001) and more recently, Melitz and Polanec (2013). Whichever way this is performed there is almost always a substantial between-firm component.

2.3 TFP differences between cross sections of countries

Figure 1 shows the correlation between GDP per capita and TFP for a large number of countries (Jones and Romer, 2010). It is clear that those countries with high TFP are also the countries with high GDP per capita, suggesting that TFP is important for understanding cross-country success. Development accounting (e.g. Caselli, 2005) focuses on how to account for these large cross-sectional differences across countries. It is the cross sectional analog of the Solow growth accounting approach. As with the time series, a puzzle remains that observables such as human and non-human capital seem unable to account for the large GDP per capita differences observed across countries.[7]

Aggregate TFP differences across countries are also influenced by how different economies allocate output to plants of heterogeneous productivity levels. For example, Figure 2 shows the estimated productivity distribution of the manufacturing sectors in the US and India (Hsieh and Klenow, 2009). Compared to the US, India appears to have a much longer left tail of low productivity plants. This suggests that there is something about the structure of the Indian economy that allows less productive plants to survive more easily than they do in the US. A large number of possible explanations present themselves that we will later examine, such as competitive intensity in the product market, labor market frictions, size-related regulations and other distortions due to corruption and tax.

[7] Gennaioli, LaPorta, Lopez-de-Silvanes and Shleifer (2013) perform development accounting using cross sectional data from the regions within a large number of countries. They argue that an expanded view of human capital (which includes managerial/entrepreneurial skills) can account for most of the TFP differences. We pursue the managerial idea in this paper.

Hsieh and Klenow (2009) estimate that Indian manufacturing TFP would be 40-60% higher if misallocation was reduced to US levels.

These lines of research show that productivity dispersion at the micro-economic level is fundamental to understanding the macro-economic patterns. But this only pushes the question one level deeper: what causes firm heterogeneity?

2.4 Firm heterogeneity within countries

Firm heterogeneity has a long history in social science (see Syverson, 2011, for an economics perspective). Today we are lucky to live in a world of large-scale (frequently near population) on firms. These are usually from national statistical agencies that collect micro-data primarily to build aggregate information either at the industry or macro level. Increasingly, researchers have been granted confidential access to such government data (e.g. the Longitudinal Business Database of US establishments). A second source is from the private sector. Companies such as Bureau Van Dijk have collated firm-level accounting panel data from almost every country in the world, for both publicly listed and private companies. These sources have enabled researchers to look at a wide range of variables including employment, output and productivity. Liberalization of administrative data and rapid increases in computer power have enormously enhanced our capacity to store and interrogate micro-data. Bartelsman, Haltiwanger and Scarpetta (2013) detail many examples of the cross-country micro-datasets now being used for productivity analysis.

The first systematic empirical analysis focused on the firm size distribution measured by employment, sales or assets. Gibrat (1931) characterized the size distribution as approximately log normal and sought to explain this with reference to simple statistical models of growth. In fact, the firm size distribution is closer to a Pareto distribution, and this power law is now well documented in every country in the world were data is available and is a central topic of the field of "econo-physics" (e.g. Hernández-Pérez, Angulo-Browna and Tun, 2006 or Axtell, 2011).

For example, using data on the population of manufacturing firms for France (following Gibrat) and the US we plot the firm size distribution in Figure 3. The power law implies that in log-log space there is a negative linear relationship between firm size and density, which is what we observe for the US data, except for the far right tail. France looks similar except for a break at 50 employees, which is an important regulatory threshold for labor laws. Garicano, Lelarge and Van Reenen (2013) discuss how the presence of many regulatory "taxes" that begin when the firm reaches 50 employees implies a broken power law exactly as described by the data.[8]

As noted in the introduction, one of the robust facts emerging from the analysis of large-scale firm-level databases is the very high degree of heterogeneity between business units (see Bartelsman and Doms, 2000). For example, Syverson (2004a) analyses labor productivity (value-added per worker) in US manufacturing establishments in the 1997 Economic Census and shows that on average, a plant at the 90th percentile of the labor productivity distribution has four times higher labor productivity than a plant at the 10th percentile in the same four digit sector. Similarly, Criscuolo, Haskel and Martin (2003) show that there is a fivefold difference in labor productivity between these deciles in 2000 in the UK.

What could explain these differences in productivity, and how can they persist in a competitive industry? One explanation is that if we accounted properly for the different inputs in the production function there would be little residual productivity differences. It is certainly true that moving from labor productivity to total factor productivity (TFP) reduces the scale of the difference. For example, in Syverson (2004a) the 90-10 productivity difference falls from a factor of 4 to a factor of 1.9, a smaller but still substantial difference.

These productivity differences show up clearly even for quite homogeneous goods. An early example is Salter (1960) who studied the British pig iron industry between 1911 and

[8] The welfare losses of such regulations could be substantial – up to 5% of GDP according to Garicano et al (2013).

1926. He showed that the best practice factory produced nearly twice as many tons per hour as the average factory. A major problem in measuring productivity is the fact that researchers rarely observe plant level prices so an industry price deflator is usually used. Consequently, measured TFP typically includes an element of the firm-specific price-cost margin (e.g. Klette and Griliches, 1996; De Loecker and Goldberg, 2014). Foster, Haltiwanger and Syverson (2008) study 11 seven-digit homogeneous goods (including block ice, white pan bread, cardboard boxes and carbon black) where they have access to plant specific output prices. They find that conventionally measured revenue based TFP ("TFPR") numbers actually *understate* the degree of "true" quantity-based productivity dispersion ("TFPQ") especially for newer firms as the more productive firms typically have lower prices and are relatively larger.[9]

Higher TFP is positively related to firm size, growth and survival probabilities. Further, Bartelsman and Dhrymes (1998, Table A.7) show that over a five-year period around one third of plants stay in their productivity quintile. This suggests that productivity differences are not purely transitory but significant component of persistence.

The evidence of substantial TFPQ dispersion found in Foster et al (2008) and in other studies that have tried to control for firm-specific prices implies that observed productivity heterogeneity is not all simply attributable to temporary fluctuations. For example, one could imagine a model where firms have homogenous productivity but are subject to heterogeneous price shocks. This would show up in variations of measured TFPR but not in TFPQ. Of course, there may well be adjustment costs and other frictions that cause a deviation between market-wide factor prices and their marginal revenue products. This will show up in variations of TFPR, and such deviations are also indicators of misallocation. In Hsieh and Klenow (2009) intra-industry variation in TFPR is due to distortions as firms face different unobserved input prices (due to subsidies and political connections for example).

[9] Foster et al (2009) show that measured revenue TFP will in general be correlated with true TFP but also with the firm specific price shocks. Hsieh and Klenow (2009) detail a model where heterogeneous TFPQ produces no difference in TFPR because the more productive firms grow larger and have lower prices, thus equalizing TFPR.

In summary, there is a substantial body of evidence of persistent firm-level heterogeneity in firm productivity (and other dimensions of performance) in narrowly defined industries in many countries and time periods. Differential observable inputs, heterogeneous prices and idiosyncratic stochastic shocks are not able to adequately account for the remarkable dispersion of productivity. So what else could account for these persistent productivity differences?

There are two levels to addressing this issue. One level is the proximate causes of the differences and the second is more fundamental causes. This is like peeling the layers of an onion. If we discovered that all labour productivity differences were due to fixed capital like plant and machinery (i.e. no TFP differences) we would then have to address the question of why these differed. But at least observable capital would give us a proximate explanation. Consider one of the possible proximate causes of productivity differences - hard technologies. The generation of new technologies (as proxied by measures of R&D or citation-weighted patents) or the adoption of technologies (as proxied by such things as hybrid corn, new drugs or information and communication technologies) would therefore be the things to focus on. There is a huge literature on such observable measures of innovation and diffusion.

Differences in hard technologies, however, are not able to fully account for productivity spreads for at least two reasons. First, even after controlling for a host of observable technology measures there remains a very large TFP residual. Second, the impact of observable technologies seems to vary systematically with the management and organization of the firm. This has most clearly been seen in studies of the effect of ICT on productivity (e.g. Bresnahan, Brynjolfsson and Hitt, 2002). There is a very wide range of effects of ICT on productivity and the impact seems to be much higher when firms are more decentralized and have stronger "people management" practices – structured policies over hiring and a strong emphasis on ability and effort when determining promotion, dealing with underperformance and pay (Bloom, Sadun and Van Reenen, 2012a).

Since technologies have been much more of a focus of empirical economic research, we will focus on management practices for the rest of this paper.

3. Management and the World Management Survey

Francis Walker, the first president of the American Economic Association wrote an 1887 paper published in the first volume of the *Quarterly Journal of Economics* that argued for the primacy of managerial ability in understanding the phenomenon of firm heterogeneity

> "***This excess of produce*** *[TFP] has not, speaking broadly, been generated by any greater strain upon the nervous or muscular power. Indeed, it may, as a rule, be confidently stated that, in works controlled* ***by men who have a high power of administration and a marked degree of executive ability****, where everything goes smoothly and swiftly forward to its end, where emergencies are long foreseen and unfavorable contingencies are carefully guarded against, where no steps have to be retraced, and where nothing ever comes out wrong end foremost, there is much less nervous and muscular wear and tear than in works under* ***inferior management****" (our bolding)*

Walker's observations were based on his experience running the 1870 US Census, and this emphasis on management has been taken up wholeheartedly by business schools. But as the survey by Syverson (2011) remarks, "*no potential driving factor of productivity has seen a higher ratio of speculation to empirical study*". There are a huge number of case studies discussing the importance of management, mostly focusing on CEOs of top corporations. Much can be learned from case studies in the formulation of hypotheses and the understanding of theories and mechanisms. They are wonderful tools for teaching, but they are poor tools for hypothesis testing.

The typical case study has a sample size of one. Even more problematic is the fact that the sample is highly non-random. Indeed, it is selected precisely to illustrate a point rather than being something that could test a theory. In the late 1990s, there were numerous case

studies and books praising a highly successful firm with a dynamic CEO that had a relentless emphasis on talent, aggressively promoting and paying smart young professionals with freshly minted MBAs from top US Business Schools. Everything possible was outsourced, the organization was extremely flat and innovation prized over dull experience. This company was called Enron (see Gladwell, 2002). When the firm collapsed due to extensive accounting frauds and huge losses, the case studies switched from the strategy sequence to the ethics sequence in the Business School curriculum.[10]

Thus, for an informative discussion on the importance of management in driving productivity, we needed to collect systematic data on representative samples of firms to empirically test our hypothesis.

3.1 How Can Management Practices Be Measured?

To measure management practices, we developed a new survey methodology first described in Bloom and Van Reenen (2007). In summary, we use an interview-based evaluation tool that defines and scores from 1 ("worst practice") to 5 ("best practice") across 18 key management practices. Appendix Table 1 lists the management questions for manufacturing, and it also gives some sense of how each is mapped onto the scoring grid.[11]

As mentioned, this evaluation tool attempts to measure management practices in three key areas. First, *monitoring*: How well do organizations monitor what goes on inside the firm, and use this information for continuous improvement? Second, *targets*: Do organizations set the right targets, track the right outcomes, and take appropriate action if the two are inconsistent? Third, *incentives/people management*: Are organizations promoting and

[10] A similar but slightly more subtle phenomenon is the "Halo effect" (Rosenzweig, 2007). This is the psychological tendency to try and backwardly induct causal factors in success from the characteristics of those who are successful. The fundamental problem is the lack of a clear counterfactual group that we need to compare to the successful group. This is a problem with approaches such as Peters and Waterman (1982) who interviewed 43 US firms who had out-performed their peers over a number of years and detailed eight themes that they claimed determined success.

[11] For the full set of questions for each sector (manufacturing, retail, schools and hospitals) see www.worldmanagementsurvey.org. The difference in the survey questions across industries primarily reflect different organizational structures—for example, using the words "nurse manager" and "unit" in hospitals as compared to "plant manager" and "factory" in manufacturing firms.

rewarding employees based on performance, prioritizing careful hiring, and trying to keep their best employees?[12]

Our methodology gives a firm a low score if it that fails to track performance, has no effective targets, does not take ability and effort into account when deciding on promotions (e.g. completely tenure-based) and has no system to address persistent employee underperformance. In contrast, a high scoring organization frequently monitors and tries to improve its processes, sets comprehensive and stretching targets, promotes high-performing employees and fixes (by re-training/rotating and, if unsuccessful, exit) underperforming employees.

To collect the data, we hired and trained teams of MBA-type students who had some business experience to conduct the telephone interviews. These students were all from the countries we surveyed (and so could interview managers in their native languages) and were studying at top U.S. or European universities. The survey was completed by plant managers in manufacturing, retail store managers, clinical service leads in hospitals, and school principals or headmasters. This level of middle managers was purposely selected: they were senior enough to have an overview of management practices but not so senior as to be detached from day-to-day operations.

We interviewed these managers using a double-blind survey technique. The first part of this double-blind technique was that managers were not told they were being scored or shown the scoring grid. They were told only that they were being "interviewed about their day-to-day management practices." To do this, we asked open-ended questions in the survey. For example, on the first monitoring dimension in the manufacturing survey, we start by asking the open question "Could you please tell me about how you monitor your production process?" rather than closed questions such as "Do you monitor your production daily [yes/no]?"

[12] These practices are similar to those emphasized in earlier work on management practices, by, for example, Osterman (1994), Ichniowski, Shaw, and Prennushi (1997) and Black and Lynch (2001).

We continue with open questions focusing on actual practices and always elicit 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?" rather than "Do you track your performance?" and the third is "If I walked around your factory what could I tell about how each person was performing?" The combined responses to the questions within this dimension are scored against a grid that goes from 1, which is 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),*" to 5, which is defined as "*Performance is continuously tracked and communicated, both formally and informally, to all staff using a range of visual management tools.*"

The other side of our double-blind approach was that our interviewers were not told in advance anything about the organization's performance; they were provided only with the organization's name, telephone number, and industry. We randomly sampled medium-sized firms (employing between 50 and 5,000 workers) in manufacturing and retail, hospitals that deliver acute care, and schools that offered education to 15-year-olds (which corresponds to high schools in most of the countries we surveyed). The formal practices we focus on are not likely to be relevant for very small organizations with few employees.[13]

We used a variety of procedures to obtain a high success rate and to remove potential sources of bias from our estimates. First, we obtained government endorsements for the surveys in most countries and industries.[14] Second, we never asked interviewees for performance or financial data; instead, we obtained such data from independent sources such as company accounts or hospital and school league tables. Third, the interviewers were encouraged to be persistent; they ran about two interviews, lasting 45 minutes each on average, per day, with the rest of the time spent contacting managers to schedule interviews. We also ran interviews in the managers' native languages to make the process

[13] In the MOPS survey we survey firms of all size classes in the US and confirm this intuition.
[14] We positioned the surveys as "an interview on management," never using the word "survey" or "research," as telephone operators usually block surveys and market research.

as comfortable as possible. These steps helped yield a response rate of about 50% across industries, which was uncorrelated with the (independently collected) performance measures for the firm—thus, we were not disproportionately interviewing successful or failing organizations.

We also collected a series of "noise controls" on the interview process itself (such as the time of day and the day of the week), characteristics of the interviewee (such as tenure in firm), and the identity of the interviewer (so we could include a full set of dummy variables for the interviewer to deal with interviewer bias). Including these in our regression analysis typically helps to improve the precision of our estimates by stripping out some of the measurement error.

3.2 Validating the Management Practices Data

To accurately validate the data we took several steps. First, for almost three quarters of all interviews we had a second person listening in on a phone extension as a "silent monitor" to independently score the interview. For these double-scored interviews we found the correlation across scores was 0.887, which shows that two interviewers typically gave the same score to the same interview.

Second, we also ran repeat interviews on 222 firms from our manufacturing sample, using a different interviewer and a second plant manager within the same firm. This helped to evaluate how consistently we were measuring management practices within firms by interviewing one manager. We found that the correlation between our independently run first and second interview scores was 0.51 (p-value< 0.001). Part of this difference across plants within the same firms is likely to be real internal variations in management practices; no two plants within the same firm will have identical management practices. The rest of this difference across plants within firms reflects measurement error in the survey process. Nevertheless, while our management score is clearly noisy, it picks up significant

management differences across firms. Similar high correlations are found in the hospital surveys.[15]

3.3 Some basic descriptive statistics on the WMS Management data

Manufacturing

The median firm is privately owned, employs around 300 workers, and operates two production plants. Initially, we take the simple average across the 18 questions, but we discuss more sophisticated methods of aggregating individual management scores. Figure 4 presents the average management practice score across countries. The US has the highest average management score followed by Japan, Germany and Sweden. Half way down the table are Southern European countries like Portugal and Greece, followed by emerging economies like India and China. African countries come at the bottom of the table. This cross-country ranking is perhaps not surprising, since it approximates the cross-country productivity and income rankings. Figure 5 plots the management scores against GDP per capita which has a reasonably tight fit.

We plot a firm-level histogram of the distribution of management practices *within* countries in Figure 6. There is a wide variation everywhere, just like the productivity distribution. One of the features distinguishing the US is not just that the mean of the distribution is to the right of other countries, but also that there is an unusually thin left tail of very badly managed firms. This is suggestive of harsher forces of selection in the US that could be related to tougher competition, a theme we return to below.

Figure 7 shows the average management scores broken down by country and whether it is an affiliate of a foreign multinational or a (non-multinational) domestically owned firm. The domestic firms dominate the overall sample so the light bars look like Figure 4. By

[15] Further evidence of the consistency of the management scores is in Grous (2011). He conducted extensive factory visits of 23 British aerospace firms, administering both the WMS telephone survey on the plant manager and face-to-face interviews with up to three other employees (the CEO/Managing Director, a manager and a shopfloor worker). The management scores from his site visits were highly associated with the scores from the telephone interviews (the correlation coefficient was 0.89). Similar results were found in the India project (Bloom et al. 2013, footnote 11) where the management scores were compared to factory visits practice evaluations run by Accenture, with a correlation (p-value) of 0.404 (0.077).

contrast, the subsidiaries of foreign multinationals score highly regardless of which country they operate in. This is not just a feature of size as the multinational premium on management persists after controlling for firm size. It is consistent with the idea that multinationals are able to spread better practices across the countries that they work in. We also found that multinationals transplant other features of their organizational form overseas, such as the average degree of decentralization (Bloom, Sadun, and Van Reenen, 2012b). The higher people management scores of US multinational subsidiaries in Europe, for example, helps explain the greater association of their IT with productivity (Bloom, Sadun, and Van Reenen, 2012a).

Outside manufacturing: Hospitals, Schools, Retail and beyond

In Figure 8, we report management scores for almost 2,000 hospitals (Bloom, Sadun and Van Reenen, 2013b). US hospitals and retailers are again the best managed across our international sample and emerging economies like India and Brazil the worst. The ranking is similar in the retail sector.[16] Figure 9 reports a similar exercise for over 1,800 schools (Bloom, Lemos, Sadun and Van Reenen, 2014). Whether or not we control for observable characteristics, US schools are more in the middle of the pack with UK and Swedish schools topping the ranks. One reason for this may be that both UK and Swedish schools have undergone a series of reforms in the last decade to improve management (see McNally, 2010).

As in manufacturing, we observe a wide spread of management practices within countries.[17] To illustrate this, Figure 10 plots the distributions of management scores for hospitals, schools, and manufacturing firms in the US for the 16 questions that are identical across the surveys. Figure 10 also highlights that average management scores for manufacturing are higher than for hospitals that are, in turn, higher than for schools.

[16] Bloom, Genakos, Sadun, and Van Reenen (2012) show that the US tops the cross country ranking followed by Canada and then the UK.

[17] These spreads in management practices appear to mimic the wide dispersions in performance in these sectors as reported in, for example, Skinner and Staiger (2009) for hospitals, Foster, Haltiwanger, and Krizan (2006) for retail, and Hoxby (2000) for schools.

One possible reason for the difference is that schools are dominated by the public sector compared to manufacturing, with hospitals in between. In each individual sector (manufacturing, hospitals and schools), government owned organizations have lower average management scores than the non-government owned. This is true even after controlling for size, country and other factors. The main reason government owned organizations have lower scores is that they have weaker people management practices. In particular, promotion is often based on time served, and persistent underperformers are not retrained or moved to different positions. Interestingly, it is not the profit motive that matters. 'Not for profit' hospitals and more autonomous public schools[18] look as good as, or better than their private counterparts in terms of management. This suggests that it may be the lack of managerial autonomy, the power of unions or the unobserved characteristics of public sector employees that matter more than public ownership *per se*.

Other research teams have also used our management scoring method to study other sectors. For example, Delfgaauw et al. (2011) look at fostering, adoption, and nursing homes; Dohrmann and Pinshaw (2009) survey various tax agencies in OECD countries; Homkes (2011) studied global public-private partnerships; McConnell et al. (2009) examine substance abuse clinics; McCormack et al (2013) examine UK university departments; McKinsey (2009) studied Irish tradable service firms; and Rasul and Rogger (2013) look at Nigerian Civil Servants. In every case the researchers found extremely wide variations in management practices across the organizations studied.

3.4 Some drawbacks of the World Management Survey

Many important aspects of management are left out

The focus of the WMS questions are on practices that are likely to be associated with delivering existing goods or services more efficiently. We think there is some consensus over better or worse practices in this regard. By contrast, we are not measuring "strategic" aspects of management such as innovation, pricing, advertising, M&A, leadership, the

[18] Such as magnet or charters in the US; academies, foundations, and voluntary-aided schools in England or free schools in Sweden.

decision whether to enter new markets, shut down existing operations, etc. These are important, no doubt, but we do not feel confident of judging anything to be on average better or worse in this regard.

It may be that a firm that scores highly on the WMS metrics may systematically also score *badly* on these other unobserved dimensions of management. For example, some firms may specialize in creativity rather than operational efficiency (a "high quality instead of low cost" strategy). Trying to improve our notion of management practices may dull the creative spark. We see this view as an interesting hypothesis over whether our measures of management are substitutes or complements with other strategic aspects. *A priori* one could equally well make a case that the WMS management scores correlate positively with these other dimensions (rather than substitute for them). For example, if a company's R&D lab is run efficiently with good collection of data, value mapping and strong incentives, then it may be better at producing innovations.[19] In the data there is a positive correlation between the management practice scores and measures of R&D, patenting and technology adoption.

A related concern is that the measures we focus on may be beneficial for productivity, but they come at the expense of making life miserable for workers or the environment. Again, we cannot rule this out, but the simple correlations go in the opposite direction. Measures of work-life balance and family friendly policies are positively correlated with the WMS management measures (Bloom, Kretchmer and Van Reenen, 2011), as are measures of energy efficiency (Bloom, Genakos, Martin and Sadun, 2010).

Are the WMS questions culturally biased?

Another concern with the questions is that they are picking up "Anglo Saxon" practices (or the ability of a manager to talk them up) rather than something that is genuinely related to better performance. Although we were very concerned about this when we started the project, we do not think this is a major concern. First, the methodology is expressly

[19] For example, see Levy (2011, Chapter 3) for a description of how Google uses the kind of people management practices over performance to determine employee rewards. These were brought in by CEO Eric Schmidt when Google's size had become too large to manage in the informal way they began with in their early start-up period.

designed to mitigate this problem – we focus on practices rather than aspirations, what is happening on the ground rather than what the firm claims are its formal policies. Second, each interviewer is multilingual. They interview managers in their own language, but they can all speak English so they also interview another English speaking country, typically the US or UK. Because we have many interviews of US firms, we can check whether the US firms score higher regardless of the nationality of the person doing the interview (they do). Third, we test whether the association between productivity and management is stronger in the US than elsewhere. It is not: the correlation is similar across all countries. Fourthly, multinationals from every country appear to adopt these management practices in every country around the world (recall Figure 7), suggesting these monitoring, targets and incentives practices are seen as the global basics of good management.

Naturally, in some countries these types of practices will be adopted less frequently because of differences in regulations, legal system, culture and other factors. For example, in Italy there are strong regulations on dismissing long-tenured employees which make it hard to have the highest scoring practices for dealing with under-performers.[20] However, we still want to measure the adoption of these management practices in Italy, so we can potentially examine the impact of these regulations on the use of this and other forms of management practice (rather than simply assume these regulations bind and change the survey technique on a country-specific basis). Hence, using the same instrument across countries is essential for the evaluation of the adoption of a core set of management practices across firms and countries.

WMS interviews are expensive

The methodology involves the recruitment, paying and training of high human capital interviewers. They have to be able to ask open questions intelligently and press for examples as well as use their judgment to score. We deliver the surveys from the same location so we can do intensive calibration of scores and de-briefings. These all make the

[20] Question 15 in Appendix Table 1. Note that the highest score here is awarded to firms who actively identify and try to fix under-performers. Usually, this will initially be through moving position or re-training. Only if these first type of interventions fail, will the high scoring firm try to move the employee out of the company.

interviews relatively expensive: including fixed cost, each interview costs in the order of $400. By contrast, a more traditional survey approach with a fixed script and closed answers requires minimal training and is easier to administer.

We have therefore experimented with lower cost versions of the management survey. First, we switch to a completely traditional "tick box" approach with closed questions called the Management, Organizational Practices Survey (MOPS). We added these questions to the US Annual Survey of Manufacturers (ASM) with support from the Census Bureau. Because answering the survey is mandatory (like the ASM) we obtained an 85% response rate from about 40,000 plants. Preliminary results from MOPs are contained in Bloom, Brynjolfsson, Foster, Jarmin, Saporta and Van Reenen (2012). In MOPS, as with the WMS, we scored highly responses indicating heavy data collection and monitoring, extensive and stretching targets, and aggressive performance incentives. Doing this yields qualitatively similar patterns of results to our standard WMS management data - TFP and management are positively correlated both in the cross section and time series with (we asked retrospective questions in MOPS). Furthermore, because we have many multi-plant firms with TFP and management from each plant, we can include firm level effects in a cross sectional plant level production function. Even in this demanding specification, there is a significant and positive association between TFP and management.[21]

What is the right unit of normalization?

The WMS attempts to measure practices, so it is like a test score. In principle the test is administered in the same way to all firms who take it and attempts to be a cardinal measure. But there is no natural unit of measurement. Hence, we tend to discuss changes in terms of standard deviations of the management score (transformed from the support between the minimum of one and the maximum of five). In regressions we usually z-score each

[21] An intermediate approach between MOPS and WMS is a face-to-face interview with closed questions (MOPS is answered remotely on the internet or by filling in a questionnaire and mailing it back to the Census). Bloom, Schweiger and Van Reenen (2012) implemented this in a large number of East European and Eurasian nations. The results were again broadly comparable to what emerged from the other methodologies, but with measurement error a greater problem than with WMS. Our sense is that the MOPS approach is better value for money, at least in countries where there is already a reasonably good Census data infrastructure system.

individual question, average across all 18 questions and take the z-score of the resulting index (taking the first principal component from a factor analysis yields a very similar result).

In terms of what this means economically, we tend to then look at the association of the management scores with some other cardinal outcome such as productivity or profitability (see Section 4). An attractive alternative would be to measure the time cost of building up managerial capital in a similar way in which we would look at human capital (the time spent in education such as years of schooling) or physical capital (the depreciated sum of past investment spending). This is the approach of authors in the growth accounting literature who seek to measure intangible capital in an analogous way to other forms of capital and build these into the national accounts. In principle, one could do this by looking at spending on (external and internal) consultants and the time managers spend in building such capital. This often comes under the category of "economic competencies", for example in the work of Corrado, Hulten and Sichel (2009). Currently this is done crudely by assuming some fraction of time of high human capital workers is spent in management. But such estimates could be refined, for example with time use surveys of senior managers on the lines of Bandiera, Prat and Sadun (2013). Combining such managerial capital measures with those in WMS would be a major advance in measuring management.

4. The influence of management on organizational performance

There is a large literature examining the effects of management on firm performance. We survey this in Bloom and Van Reenen (2011) with an emphasis on human resource management (such as incentive pay). The vast majority of these studies are not Randomized Control Trials (RCTs), but non-experimental regressions in either a cross sectional or panel data setting. In personnel economics there is a tradition of exploiting changes in firm policies initiated by a CEO (a natural experiment such as Lazear, 2000) or engineered by the team of researchers (e.g. the fruit farm field experiments summarized by Bandiera, Barankay and Rasul, 2011). Consequently, sub-section 4.1 focuses on non-experimental work using our management data and sub-section 4.2 discuss the more sparse RCT evidence.

4.1 Non-experimental Evidence

Performance and Management in Manufacturing

A simple way to summarize the management practices of the firm is to use the same summary management quality measure underlying Figure 6 and correlate this with various firm performance outcomes. For example, Figure 11 shows the local linear regression of ln(firm sales) on the management score. Since we would expect the better-managed firms to capture a larger fraction of sales, the positive and monotonic relationship is consistent with this prediction. Figure 12 repeats this analysis for plant size (left panel) and firm size (right panel) but using the establishments in the US MOPS instead. A similar positive and monotone relationship is revealed.

Higher management scores are positively and significantly associated with higher productivity, firm size, profitability, sales growth, market value and survival. For example, Bloom, Sadun and Van Reenen (2013a) estimate production functions where they regress real firm sales on the management score controlling for conventional inputs (e.g. labor, capital, employee education) and other covariates (e.g. firm age, noise controls, industry, country and year dummies). In the cross section their results show that a one standard deviation increase in management is associated with an increase in TFP of 15%. An example of such a result is in Figure 13, which shows the results of a local linear regression of estimated TFP on the management score. The relationship is monotonically increasing over the support of the distribution. The figure also has a hint of convexity towards the top end of the management distribution (scores above about 4.2) suggesting that introducing many top practices simultaneously has an especially large correlation with productivity. Meagher and Strachan (2013) apply Bayesian techniques to the Bloom and Van Reenen (2007) data over four countries and also find that there is some convexity for high scores. They interpret this as consistent with the idea that there is complementarity between multiple managerial practices (as in Gibbons and Henderson, 2013; Milgrom and Roberts, 1990).[22]

[22] Meagher and Strachan (2013) also find that the relationship between management and performance is flat and decreasing for lower levels of management (between average scores of 1 and 2). But it is unclear whether

The panel dimension of the management data allows more sophisticated ways of estimating the performance-management relationship. Bloom, Sadun and Van Reenen (2013a) show that Olley-Pakes (1996) style estimates are similar results to levels OLS. Fixed effects estimates of the management coefficient are also positive and significant, although the magnitude of the association is much smaller. These within group estimates may bias downwards the coefficient on the management score due to attenuation and/or because of unobservable shocks. For example, firms may only upgrade their management significantly when they face a crisis.

The relationship between productivity and management is also robust to different ways of combining the management questions. For example, factor analysis on our 18 management questions reveals one principal component that loads positively on all questions and explains most of the variance. This reflects a common factor of "good management": if a firm is strong on one managerial question it will tend to be strong on all of them. Replacing our management score with this factor in a regression yields very similar results. There is also a second factor that explains only a small amount of the data. This loads positively on the monitoring and targets questions and negatively on incentives, suggesting that some firms specialize more in monitoring (often those from Germany, Sweden, and Japan) and other firms specialize more in incentives (often those from Anglo-Saxon countries). One explanation of these cross patterns is that countries with weaker labor regulations will tend to have relatively better incentives management compared to monitoring and targets management.

Performance and Management outside Manufacturing

The association of management with organizational performance is also clear in other sectors outside manufacturing. Bloom, Propper, Seiler, and Van Reenen (2010) interviewed managers and physicians in the orthopedic and cardiology departments of 100 UK hospitals. They found that management scores were significantly associated with better

this is particular to the smaller sample or estimation technique. The more recent data in Figure 13 looks like a clear and string positive relationship for lower scores.

patient outcomes (as indicated by survival rates from emergency heart attacks and general surgery) as well as other productivity indicators (such as average length of stay and finished consultant episodes per patient). For example, a one standard deviation increase in management is associated with about a one percentage point fall in the risk-adjusted mortality rate (say from the sample mean of 17% to 16%).

Chandra, Finkelstein, Sacarny, and Syverson (2013) show that there is also a positive association between case mix adjusted AMI (heart attack) survival rates and management scores among US hospitals. In subsequent work, Bloom, Sadun and Van Reenen (2013b) show that this positive relationship between patient outcomes and management holds in other countries. In Figure 14 we report their regression of AMI mortality rates in four countries (the UK, US and Sweden and Canada). Columns (1) and (2) show that with or without general controls higher management scores are associated with lower death rates. Breaking this down by country in the last four columns shows that there is a significant relationship in all countries.

Figure 15 examines the relationship between pupil outcomes as measured by test results and our management scores in schools in six countries where we can obtain school level pupil outcome data (UK, US, Sweden, Brazil, India and Canada). There is again a positive and monotonic relationship between pupil test scores and management.

Most of the other studies also find that the management scores are positively associated with measures of organizational performance. For example, McCormack, Propper and Smith (2013) examine UK university departments and also find that better managed departments appear to have higher scores in teaching and research. Chong et al (2013) find that the management score is correlated with postal service efficiency at the country level. Delfgaauw et al (2011) find a positive and significant association amongst for-profit nursing homes (but no significant relationship for the not-for profits).

One exception, however, is the Rasul and Rogger (2013) study of the Nigerian civil service. They have information on the success rates of 4,721 projects such as plans to build bore

holes, dams and roads. After implementing the WMS method they found that contrary to the other studies, organizations with high management scores were *less* likely to successfully complete projects. By contrast, decentralization was found to be associated with a greater likelihood of project success. The authors' preferred explanation of this is that the greater monitoring associated with higher management scores crowds out the intrinsic motivation of the public servants.

4.2 Management and Performance: RCT Evidence

A problem with the non-experimental evidence is that management is likely to be endogenous. Even in the panel estimates, there may be time-varying unobservables correlated with both management and performance. There may also be reverse causality: perhaps better performing firms can employ superior management consultants, for example. In recent years there has been an emphasis on Randomized Control Trial (RCTs) evidence to obtain causal estimates.

Bloom, Eifert, Mahajan, McKenzie and Roberts (2013) provided free management consulting to a set of randomly selected textile plants outside Mumbai to help them adopt the kind of modern management practices measured by WMS and compared their performance to another randomly chosen set of control plants. The Indian experiment revealed that the adoption of these management practices leads to large increases in productivity. This took several months to occur as the firms slowly adopted modern management practices. As shown in Figure 16 there was an improvement of about 20% in productivity from an increase in the management score which the authors show is equivalent to twice the initial cross-sectional standard deviation. This implies that a one standard deviation increase in the management score would increase productivity by 10%: a figure lying between the OLS levels cross sectional and within groups panel estimates in Bloom, Sadun and Van Reenen (2013a). Profits increased on average by $325,000 in the first year which compared to a market cost of the intervention of $200,000. So the intervention more than paid for itself in the first year – the returns would be even higher to the extent the improvements persisted, which they appear to do.

Interestingly, the Indian experiment also found that the adoption of these types of practices were more likely to occur when production conditions were bad. When facing tough times, firms were more likely to try to upgrade their management practices. In contrast, when conditions were better, firms were reluctant to change or adjust management practices. If this type of endogeneity was common, it would lead to systematic underestimation of the impact of management on performance, especially in panel data estimates that rely on changes in performance following changes in management.

There are also a growing number of RCTs on other management interventions in developing countries in micro-enterprises (single or few person firms) many of which are still in the field. The results of these are much more ambiguous than the Indian textile experiment which focused on large (several hundred employee) firms.

Karlan, Knight and Udry (2012) survey 11 studies of managerial interventions.[23] Several of these find positive effects on profits like the Indian textile RCT such as Mano, Iddrisu, Yoshino and Sonobe (2011) on sub-Saharan Africa, Valdivia (2012) in Peru, Bruhn, Karlan and Schoar (2012) and Calderon, Cunha and De Giorgi (2013) in Mexico. Others find insignificant or mixed results – Berge, Oppedal, Bjorvatn, and Tungodden (2011) for example find positive effects for men but negative effects for women. Some other studies find negative effects – such as Giné and Mansuri (2011) or Drexler, Fisher and Schoar's (2011) basic accounting training. Karlan, Knight and Udry (2012) run their own experiment providing consulting advice (and later, free cash) for tailors in Ghana. Surprisingly, they found that neither the business advice not the cash infusion raised firm profits. They interpret these results in the context of a model whereby such interventions enable entrepreneurs to take more risks and increase the probability of extreme positive draws.

Why does the wider literature not find uniformly strong and positive effects like Bloom et al's (2013) RCTs? There are several possibilities. First, this intervention (like the WMS)

[23] See also the McKenzie and Woodruff (2012) survey of managerial RCT interventions which has similar qualitative conclusions.

emphasizes formal systems for monitoring output, inputs and defects, setting short and long-run targets, and rigorous employee appraisal systems. These are less likely to be important for the micro and mini-enterprises – mostly single person firms - that the rest of the literature focuses on. The Indian textile RCTs (and the WMS survey) explicitly target medium sized firms with several hundred employees spread across multiple factories. Secondly, the firms who deliver the management consultancy services in the wider literature are usually local firms, unlike Accenture who delivered the services for the Indian experiment. Local consultancy firms will struggle to deliver the quality of intervention of that global consultancy firms can. Thirdly, the type of management training differs substantially. The WMS method focuses on operational improvements whereas many of the treatments have a focus on "strategic management" such as improved marketing and pricing.

This is an emerging field and hence there is unlikely to be a quick consensus. The RCT closest to the WMS approach does find causal effects that are consistent with the non-experimental work. Understanding the heterogeneity of the effects across different RCTs is an important area for future research. This is a general lesson for economics. RCTs are a much more credible way of establishing causal relations between variables than conventional approaches, but it is still necessary to relate these treatment effects to "deep" parameters in order to understand the world and make robust policy recommendations.

4.3 How much of TFP spread can management account for?

If we take some of the effects of management on TFP seriously, how important are they from a macro-economic perspective? In the spirit of development accounting, Bloom, Sadun and Van Reenen (2013a) estimate that management accounts for (on average) a quarter of the TFP gaps between the US and other countries. To do this they use: (i) the size-weighted average management scores by country, (ii) an average treatment effect of a 10% increase in TFP from a one standard deviation increase in management; and (iii) the

cross country TFP differences from Jones and Romer (2010).[24] For some Southern European countries such as Portugal and Italy, management accounts for half of the TFP gap with US, whereas for other nations like Japan or Sweden the fraction is only one tenth. Management can potentially account for a great deal of the TFP spread within countries. In the US and UK about a third of the 90-10 difference in TFP can be related to management practices.

Although these estimates highlight that there are many other things apart from management that are important for TFP and are very crude, they do imply that management is potentially quantitatively as well as qualitatively important in explaining TFP differences between and within countries.

5. Models of management

When considering how to interpret management scores there are two broad approaches in economics. First, we can consider higher scores as reflecting at least in part better management quality (vertical dimension). Our work, following *inter alia* Walker (1887) has focused on this perspective and we discuss this further below. An alternative view is that no one practice is on average better than another (the horizontal dimension of management). We call this the "Design" perspective because here all practices are designed to be adapted to the idiosyncratic local environment and do not systematically reflect any better or worse management quality. In management science this is called "contingency theory" (Woodward, 1958).

5.1 The Design Perspective on management

To fix ideas, consider an example around promotion practices. In WMS we measure whether or not firms take effort and ability into account when making promotion decisions or whether (for example) they simply promote on tenure regardless of performance. We suspect that ignoring any measure of ability or effort is not, on average, a wise policy. The

[24] For example, there is about a 1.65 standard deviation gap in the management score between Greece and the US and a TFP gap of 51%. This implies management "accounts" for a third (= (1.65*0.1)/0.5) of Greece's TFP deficit with the US.

design view, however, would emphasize that basing promotions on objective performance measures has many well-studied problems such as multi-tasking (Holmstrom and Milgrom, 1991). Management involves many hard to measure tasks. In the face of high-powered incentives around promotions, managers may focus only on activities which are easily observable (such as output production) and ignore those which are hard to observe (like safety maintenance). What about subjective performance metrics? Such measures may be subject to influence activities (Milgrom and Roberts, 1988) with workers investing time (and bribes) to get a better grading from their supervisors rather than focusing on production. If these problems were severe enough, then even if collecting performance information were monetarily costless it may be optimal to ignore all performance metrics and promote purely on, say, tenure. Consequently, firms with low WMS scores on this question are in no sense "worse managed". Forcing them to promote partly on performance would cause a reduction in the firm's productivity and value. Similar remarks could be made on all other questions in the WMS survey since organizational economics has a wealth of theories why seemingly inefficient management practices may, in fact, be profitable.

We do find some evidence for the Design perspective. As noted above factor analysis suggests a second factor of some firms specializing in targets and monitoring and others in incentives/people management. Some countries (like the US and UK) have a much higher relative score of people compared to monitoring/targets management than others (such as Germany and Japan). This appears related to labor regulations. Across sectors, we systematically relate these relative scores to industry characteristics. Industries that are more innovative (e.g. with higher R&D and patents) and that have more human capital tend to focus relatively more on people management. By contrast, industries that are more physical capital intensive tend to be relatively stronger in monitoring/targets.

Despite important elements of contingency, the main aspect of the WMS data does appear to be more tightly linked to firm performance. If management was all by design it is unclear why higher scores should be systematically linked to higher performance. The evidence from Figure 16 is that the Indian textile firms who were "forced" to increase their management scores became more productive and profitable. if the Design perspective was the whole story their performance should have deteriorated.

5.2 Management vs. Managers: Management as entrepreneurial talent

One popular approach is to consider management as a matter of the talent of the CEO and an interpretation of the WMS scores is that they simply reflect entrepreneurial ability. The Lucas (1978) model is in this spirit whereby individuals are endowed with some element of managerial talent. Those individuals with the highest level of talent will run the largest firms and earn rents from their talent as the residual claimants on a firm's profits. Managerial overload means that there are diminishing returns so that the best manager does not take over the entire economy. In equilibrium, there will be a cut-off below which and individual will optimally decide to be a worker rather than a firm. A power law in managerial ability then gives rise to the power law in the firm size distribution (Garicano, Lelarge and Van Reenen, 2013).

There is empirical evidence that CEOs do matter. For example, Bertrand and Schoar (2003) use CEO data from ExecuComp and Forbes matched to publicly listed US firms to show that there are important CEO fixed effects in corporate strategy and performance. These fixed effects are also systematically correlated with observables. CEOs from later cohorts and those with an MBA tend to be more aggressive in leveraging up debt and have a closer correlation between investment and stock market value (as measured by Tobin's average Q).[25]

More broadly, the skills of all managers (indeed, all employees) are important for a firm's performance.[26] If these are all available in perfectly competitive factor markets then they should be measured as factor inputs and not part of TFP. Of course, this is unlikely to happen in practice if management has an element of intangible capital as discussed in the next sub-section.

[25] Benmelech and Frydman (2014) show that CEOs with a military background have more conservative investment policies and perform better during industry downturns. They also survey the literature showing how CEO characteristics affect company performance.

[26] Lazear, Shaw and Stanton (2012) show the importance of supervisor specific effects in a large IT firm.

Despite the appeal of the managerial talent approach, it seems to us likely that management is a broader concept than simply adding up the atoms of human capital of the entrepreneur and all employees. Some firms seem to be able to obtain more productivity from the same group of employees than other firms, which is likely to relate to the deep-seated organization of firms.[27] The CEO and founder will have a large influence on this corporate culture, but the culture may persist after the departure of the CEO or founder. Toyota would be such an example. For example, in the Indian RCTs, the managers largely stayed in place, but the productivity of the firm improved enormously.

5.3 Management as intangible capital

Walker's (1887) emphasis on managerial ability as the source of firm heterogeneity was met with a response by Alfred Marshall (1887) in the next edition of the Quarterly Journal of Economics. Marshall wrote *"I am very nearly in agreement with General Walker's Theory of profits….the earnings of management of a manufacturer represents the value of the addition which his work makes to the total produce of capital and industry…."*. In other words, management should be thought of as endogenously chosen by a firm and paid a wage consummate with its contribution to marginal productivity.

The intangible capital approach to management is to treat it as another factor of production. Its level can be altered at some cost. One could think of this as purchasing advice from consultants, hiring new managers or diverting some current time by employees into building managerial capital.

Formally, consider the production function where the value added, Y, of firm *i* depends on TFP (A), labor (L), non-managerial capital (K) and managerial capital (M).

$$Y_i = A_i K_i^{\alpha} \, \mathrm{L}_i^{\beta} \, M_i^{\gamma} \tag{1}$$

Since managerial capital is usually not measured, it will typically be picked up by residual TFP. For example, if management was upgraded through the purchase of consultancy

[27] This is why wage decompositions of matched worker-firm data typically find that firm effects are as important as individual effects (and more important than worker-firm match effects). See Abowd, Kramarz and Margolis (1999) and Card, Henning and Kline (2013).

services, these would normally be charged as intermediate inputs and not be counted in Y. However, these purchases would actually be serving to increase M, managerial capital.

5.4 Management as a technology

In Bloom, Sadun and Van Reenen (2013a) we try to combine some of these ideas into a simple model that we call "Management as a Technology". The production function is the same as equation (1). We add a monopolistic competition demand side (a greater demand elasticity will index higher competition). Profits will depend on revenues less a fixed cost and adjustment cost for changing managerial or non-managerial capita. When firms enter the market they pay a sunk cost and take a draw from a known distribution of managerial talent. This is a reduced form way of capturing many factors that may influence management such as entrepreneurial talent, but also the informational, incentive and co-ordination problems that cause the underlying heterogeneity of management practices (see next section).

In this set-up management differs from tangible capital because there is initial heterogeneity in the distribution of management at entry, which will be a cause of (quasi) persistent firm differences. We allow other sources of idiosyncratic heterogeneity as the firm also has a draw of the TFP distribution and the distortions distribution (frictions that act like taxes to reduce revenue[28]). What distinguishes management from these other two sources of initial heterogeneity is that it can be altered endogenously. We allow the economy to evolve as firms are subject to idiosyncratic TFP shocks which will cause it to want to alter its optimal level of management and other factors of production. Adjustment costs mean that the shocks will cause a firm to choose a different path for investment in managerial (and non-managerial) capital. Consistent with the panel data evidence that management practices change very slowly we assume that management has higher adjustment costs than capital. But we allow firms to alter the level of management over time (at a cost) rather than assume it is completely fixed in an individual like Lucas (1978) or in a firm like Melitz (2003).

[28] These are a way of capturing corruption, size contingent regulations, etc. following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009).

Although very stylized, we find that this simple model does a reasonable job of describing some of the first order facts of the data. First, there should be a positive covariance between management and performance. As discussed above, this seems to be a robust stylized fact. Second, tougher competition should increase average management quality through selecting out the badly managed firms. Since the average surviving firm will also tend to grow as the market is larger (lower prices) and with fewer firms, there will also be an increase in management endogenously on the intensive margin. We will discuss this prediction in the next section and find it receives considerable empirical support. Thirdly, when the price of management falls, its level should increase. Proxying the cost of managerial skills by the geographic proximity of universities and business schools, we also find evidence for this proposition.

A fourth, and more subtle implication of the management as a technology model, is that in economies where there are greater distortions (arbitrary taxation, many size-contingent bribes and regulations, etc.) the covariance between management and size should be lower. In other words, even firms with very high managerial quality will struggle to grow and reach a large size in an economy like India compared to one like the US. Bloom, Sadun and Van Reenen (2013a) find considerable evidence in support of this. The covariance between management and size is stronger in the US than other countries. Furthermore, in environments where explicit policy indicators of distortions (such as industry-country specific tariffs, trade costs or labor regulations) are worse, the management-size covariance is weaker. This finding is consistent with Bartelsman, Haltiwanger and Scarpetta (2013). Across countries they found that the TFP-size covariance was much stronger in the US than other countries and that this covariance grew stronger in Eastern European countries that moved towards a market economy.

6. What causes the heterogeneity in Management Practices?

The theories discussed in the previous section have different implications over what causes the (large) heterogeneity that we observe in management in just about every sector. If the persistent performance differentials observed are more than design, then they do pose an

important question as to why seemingly profit-enhancing practices are not universally adopted.

At a high level we can follow Rivkin (2000) and distinguish between four reasons (four 'shuns'). First there is *perception*. The CEO may simply not even know her firm is badly run – there are many other excuses for under-performance. Further, even if she does know the firm is poorly managed she may not know how to change it (*inspiration*). Both these relate to informational frictions (and CEO human capital), but even if the CEO is fully informed she may lack *motivation* to adopt best practices. This is where economists have focused on problems of incentives due to weak competition in product markets, agency and governance problems, etc. Finally, even if the CEO is fully informed and well incentivized, she may still not do the right thing as changing the firm is a complex matter of *persuasion* (Gibbons and Henderson, 2013). It is not just a decision-theoretic problem: it requires a coalition of many agents in the organization to introduce major changes. Even if the change increases the size of the surplus many powerful insiders may receive a smaller slice, thus generating resistance to change.[29] This is the meat and blood of organizational economics (Gibbons and Roberts, 2013) and the political economy approach to firms.

6.1 Product Market Competition

The broader productivity literature has found an important role for competition in raising productivity (Van Reenen, 2011) and argue that one mechanism through which this happens has been through improving management quality. The management as a technology model of sub-section 5.5 has this property, for example.

In our work we have consistently found that greater levels of competition in the product market are associated with higher management scores, both in the cross section and in the panel dimension. There is no one ideal measure of competition that can be used across all sectors. In Bloom and Van Reenen (2007) we used three indicators: (i) trade openness as

[29] In a Coasian world, losers could be bought off with appropriate side payments. But such credible promises are extremely difficult in the real world of incomplete contracts. Relational contracts may be critical in generating a high trust environment that enables change to occur.

measured by the ratio of imports to production measured in the industry-country pair; (ii) the Lerner index which is a proxy for super-normal profits also measured at the industry-country pair level[30] and (iii) the perceived number of competitors faced as declared by the plant manager. In all three cases we found that higher levels of competition were associated with higher management quality, both with and without firm fixed effects.

A concern with all of the associations is that they are not necessarily causal. We can try to tackle this by exploiting quasi-experiments. The number of rivals, often weighed by sales as in a Herfindahl Index, is often used as an indicator of competition in the hospital sector as patients dislike being treated far from where they live. In our hospital data we also tend to find a positive correlation between competition as measured by the number of rivals and our management scores. Bloom, Propper, Seiler, and Van Reenen (2010) exploit the fact that in the UK politicians control exit and entry. They keep hospitals open in politically marginal districts as shutting down all or part of a hospital is incredibly unpopular and loses large numbers of votes in elections. This creates some quasi-random variation in the number of hospitals across different areas. We use political marginality as an instrumental variable for market structure (more hospitals in marginal areas) and find that the positive causal effect of competition on management (and better clinical outcomes) is *stronger* than the simple correlation would suggest.

Returning to the manufacturing sector, Bloom, Draca and Van Reenen (2011) use the growth of Chinese imports as a quasi-experiment for import competition. The trade success of China has come from the ongoing liberalization process begun by Deng Xiapong since the early 1980s - an exogenous supply shock. The impact has also varied between industries, with China's import threat being much greater in low-wage industries. This was particularly strong around the time of China's Accession to the World Trade Organization in December 2001 in the textile and apparel industries. Using the differential exposure across industries to quotas as an instrumental variable (e.g. the differential effect of the

[30] We defined the Lerner index as 1 minus the average profits/sales ratio of all other firms in the country industry cell over the past five years. High values suggest low long-run profits, suggestive of tough competition. When we used this and the import measure data we added country and industry dummies to control for factors like country size and different reporting requirements.

abolition of Western quotas against Chinese goods due to the WTO) the authors find a positive causal impact of competition on management (as well as innovation and productivity).

6.2 Ownership and Governance

We can divide the firms by ultimate ownership. One interesting group that emerges is family firms, which our research defines as firms owned by the descendants of the founder – usually their sons, grandsons, etc. Those firms that are family owned *and* family managed have on average much worse management scores, while the family owned but externally managed look much better. The negative effect of family firms holds up after controlling for a host of factors such as age and size (see Bloom and Van Reenen, 2007). The reason appears to be that many family firms adopt a rule of primogeniture (the eldest son becomes the next boss regardless of merit).

These results are consistent with the negative effect of family firms on performance as shown by Perez-Gonzalez (2006) on US data. Bennesden et al (2007) show that this result is even stronger when using the gender of the founder's first born child as an instrument (using Danish data). The reason is family firms appear to get in outside managers during a crisis, leading to an underestimate of the negative impact of family management on performance in standard OLS regressions. Hence, OLS regressions of changes in performance on changes in management may underestimate its importance.

Many governments around the world also provide tax subsidies for family firms. For example, the UK has many more family-run and -owned firms than the US, which is likely to be related to the more generous treatment of estate tax exemption for inherited business in the UK.[31]

[31] Since family firms typically have less debt, product market competition may not be as effective in driving them out of business if they are badly managed. Without debt firms have to cover operating costs (e.g., salaries and wages) but not capital costs (e.g., the rent on property or equipment since these were typically bought outright many years ago). Hence, family firms can continue to generate positive cash flow while generating economic losses because their family owners are subsidizing them through cheap capital.

6.3 Human Capital

The human capital of managers as measured by the proportion who have college degrees is strongly positively associated with management scores. It is interesting that this is also true for non-managers which suggests that having workers who are smart enough to respond to continuous improvement initiatives, for example. It is possible to geocode the locations of all the WMS plants and calculate the drivetimes to the closest university and business schools. Conditional on other local characteristics like population density, being close to a university is significantly correlated with better management scores (Feng, 2013). This is consistent with the idea that reductions in the costs of management cause firms to increase managerial investment.

6.4 Information

Although lack of knowledge is frequently mentioned as a constraint on the adoption of managerial practices, hard evidence is difficult to come by. This was frequently mentioned in the Indian textile experiment for example. Some suggestive evidence is contained in a question we ask at the end of the management survey "*Excluding yourself, how well managed would you say your firm is on a scale of 1 to 10, where 1 is worst practice, 5 is average and 10 is best practice*". Unlike the management score, this is a purely subjective question capturing how the managers' perceive the management quality in their firms.

Figure 17 plots these scores against labor productivity. Unlike the management scores in Figure 13 there is no relationship at all. Many good managers underestimate their firm's quality whereas many poor managers over-estimate it. This illustrates the challenge facing firms in how to upgrade their practices: there is much ignorance in simply knowing how well managed a firm is. This is why when a consultancy attempts an operational transformation, the first thing they typically do to perform a "diagnostic" to evaluate the performance and practices of the firm. Having some objective sense of strength and weakness is the first step to improvement.

7. Conclusions

Studying the causes and implications of variation in productivity across firms has become an important theme in social science. While several fields have been studying management for many decades, empirical economists after some early focus in the 19th century under Francis Walker and Alfred Marshall, have more recently mostly ignored management as a factor behind differences in productivity. We believe the discipline would benefit from more interaction with management. We have started to bridge this gap by developing a simple methodology to quantify some basic aspects of management practices across sectors and countries, and using experiments to identify causal impacts. These are hard, but not impossible, to measure, and we hope the methodology we have developed will be refined and used by other researchers to help draw the international map of management in finer detail in additional countries, industries, and practices.

The patterns we find lead us to believe that an important explanation for the substantial differences in productivity among firms and countries are variations in management practices. Preliminary estimates suggest around a quarter to a third of cross-country and within-country TFP gaps appear to be management related.

From a policy perspective, several factors seem important in influencing management quality b**y** thinning the ranks of the badly managed and incentivizing the survivors to improve. One reason for higher average management scores in the United States is that better managed firms appear to be rewarded more quickly with greater market share and the worse managed forced to rapidly shrink and exit. Avoiding regulatory barriers to entry, protection of inefficient incumbents and having a vigorous competition policy are to be recommended. Tax incentives to protect family firms, onerous regulations to slow reallocation and barriers to skill acquisition are also to be avoided.

It is also likely that there is a role for reducing informational frictions. There is no reason to generally subsidize management consulting, but reducing barriers to the market for advice should be high on the policy agenda. The creation of better benchmarks, advice shops and management demonstration projects, especially for smaller firms could be

beneficial. A plethora of these business support policies exist, but they are never credibly evaluated. Rigorous RCTs and other evaluations would both help governments determine "what works" and also shed light on the fundamental drivers of firm heterogeneity.

From a research perspective, understanding the causes of the variation in management is a key issue. As economists we have focused a lot on human capital, incentives and selection through market competition. But it is likely that informational constraints and within firm co-ordination are equally important, but even harder to measure. Understanding these factors will help us advance the field and develop better policies for improving management and productivity.

REFERENCES

Abowd, John, Francis Kramarz, and David Margolis (1999) "High Wage Workers and High Wage Firms" *Econometrica*, 67(2) 251–333

Aghion, Philippe and Howitt, Peter (2009) *The Economics of Growth*, Cambridge: MIT Press.

Alexopoulos, Michelle and Trevor Tombe "Management Matters" (2012) *Journal of Monetary Economics* 59, 269–285

Baily, Martin, Charles Hulten and David Campbell "Productivity Dynamics in Manufacturing Plants" (1992) *Brookings Papers on Economic Activity: Microeconomics*, 187-267

Bandiera, Oriana, Iwan Barankay and Imran Rasul (2011) **"**Field Experiments With Firms" *Journal of Economic Perspectives*, 25 (3)

Bandiera, Oriana, Andrea Prat, Raffaella Sadun (2013) "Managing the Family Firm: Evidence from CEOs at Work" CEP Discussion Paper 1250

Bartelsman, Eric and Phoebus Dhrymes (1998) "Productivity Dynamics U.S. Manufacturing Plants 1972-1986", *Journal of Productivity Analysis*, 9(1), 5-33.

Bartelsman, Erik, Haltiwanger, John and Scarpetta, Stefano (2013) "Cross Country Differences in Productivity: The Role of Allocation and Selection" *American Economic Review*, 103(1) 305-334

Benmelech, Efraim and Antoinette Schoar (2014) "Military CEOs" NBER Working Paper 19782

Bennedsden, Morten, Kasper Nielsen, Francisco Perez-Gonzalez and Daniel Wolfenzon (2007) "Inside the family firm: the role of families in succession decisions and performance", *Quarterly Journal of Economics*, 647-691.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden (2011) "Human and financial capital for microenterprise development: Evidence from a field and lab experiment" CMI Working Papers 1

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

Black, Sandra and Lisa Lynch (2001) "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity", *Review of Economics and Statistics*, 83(3), 434–445

Bloom, Nick, Christos Genakos, Ralf Martin and Raffaella Sadun, 2010. 'Modern Management: Good for the Environment or Just Hot Air', *Economic Journal*, May.

Bloom, Nicholas and John Van Reenen (2007) "Measuring and explaining management practices across firms and countries" *Quarterly Journal of Economics*, 122(4), 1351–1408.

Bloom, Nicholas and John Van Reenen (2011) "Human Resource Management and Productivity", *Handbook of Labor Economics Volume 4B* in Orley Ashenfelter and David Card (eds), Chapter 19 1697-1769.

Bloom, Nicholas, Mirko Draca and John Van Reenen (2011) "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity" Centre for Economic Performance Discussion Paper No. 1000.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012a) "Americans do I.T. better: US multinationals and the productivity miracle" *American Economic Review* 102 (1), 167-201

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012b) "The organization of firms across countries" *Quarterly Journal of Economics* 127(4): 1663-1705

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2013a) "Management as a Technology", LSE mimeo http://cep.lse.ac.uk/textonly/_new/staff/vanreenen/pdf/mat_2013dec1.pdf

Bloom, Nicholas, Raffaella Sadun and John Van Reenen (2013b) "Does Management Matter in Healthcare" LSE mimeo

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

Bloom, Nicholas, Carol Propper, Stephan Seiler and John Van Reenen (2010) "The impact of competition on management quality: Evidence from public hospitals" Centre for Economic Performance Discussion Paper No. 983.

Bloom, Nick, Tobias Kretschmer and John Van Reenen (2011) "Are Family Friendly Workplace Practices a valuable firm resource?" *Strategic Management Journal* 32(4) 343-367

Bloom, Nicholas, Renata Lemos, Raffaela Sadun and John Van Reenen (2014) "Does management matter in schools?" LSE mimeo

Bloom, Nicholas, Ben Eifert, Abrijit Mahajan, David McKenzie and John Roberts (2013) "Does management matter? Evidence from India" *Quarterly Journal of Economics* 128 (1): 1-51.

Bloom, Nick Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Itay Saporta-Eksten and John Van Reenen (2012) "Management in America" Census Bureau Centre for Economic Studies Working Paper 13-01

Bresnahan, Timothy, Erik Brynjolfsson and Lorin Hitt (2002) "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence", *Quarterly Journal of Economics*, 117(1), 339-376.

Bruhn, Miriam and Bilal Zia (2011) "Stimulating managerial capital in emerging markets: the impact of business and financial literacy for young entrepreneurs" Policy Research Working Paper Series 5642, The World Bank.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2012) "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico," Yale Economics Department Working Paper No. 100

Calderon, Gabriela, Jesse M. Cunha and Giacomo De Giorgi (2013) "Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico" NBER Working Paper 19740

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson (2013) "Healthcare Exceptionalism? Productivity and Allocation in the U.S. Healthcare Sector", NBER Working Paper 19200.

Corrada, Carol, Charles Hulten and Daniel Sichel (2009) "Intangible Capital And U.S. Economic Growth" *Review of Income and Wealth*, 55(3), pages 661-685, 09.

Card, David, Joerg Heining and Patrick Kline (2013) "Workplace Heterogeneity and the Rise of West German Wage Inequality" *Quarterly Journal of Economics*, 128, 967-1015.

Caselli, Francesco (2005) "Accounting for Cross Country Income Differences" in *Handbook of Economic Growth*, Philippe Aghion and Stephen Durlauf (eds) North-Holland: Elsevier.

Chong, Alberto Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shleifer (2012) "Letter Grading Government Efficiency" NBER Working Paper 18268 http://www.nber.org/papers/w18268

De Loecker, Jan and Penelopi Goldberg (2014) "Firm Performance in a Global Market" forthcoming, *Annual Review*

Delfgaauw, Josse, Dur, Robert, Proper, Carol and Smith, Sarah (2011), "Management practices: are not for profits different?", CMPO working paper 11/263.

Dohrmann, Thomas and Gary Pinshaw. 2009. *The road to improved compliance: A McKinsey benchmarking study of tax administrations 2008-2009*. New York: McKinsey and Company.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar (2011). "Keeping it Simple: Financial Literacy and Rules of Thumb," CEPR Discussion Paper No. DP7994.

Feng, Andrew (2013) "Essays in management and human capital", LSE mimeo

Foster, Lucia, John Haltiwanger, and C. J. Krizan (2001). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence" in *New Developments in Productivity Analysis* University of Chicago Press

Foster, Lucia, John Haltiwanger, and C. J. Krizan (2006). Market selection, reallocation, and restructuring in the U.S. retail trade sector in the 1990s. *Review of Economics and Statistics*, *88*(4), 748–758.

Foster, L., Haltiwanger, J., and Syverson, C. (2008) "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review*, *98*(1), 394–425.

Garicano, Luis, Claire Lelarge and John Van Reenen, J. (2013) "Firm Size Distortions and the Productivity Distribution: Evidence from France" NBER Working Paper 18841

Gibrat, Robert, 1931. *Les Inegalites Economiques*, Sirey, Paris.

Gibbons, Robert and John Roberts (2013) "Introduction" *Handbook of the Economics of Organization*, Princeton: Princeton University Press
Gibbons, Robert and Rebecca Henderson (2012) "Relational Contracts and Organizational Capabilities" *Organizational Science*, 23 (5), 1350-1364
Giné, Xavier, and Ghazala Mansuri (2011) "Money or Ideas? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan," mimeo.
Gladwell, Malcolm (2002) "The Talent Myth" *New Yorker* July http://www.newyorker.com/archive/2002/07/22/020722fa_fact
Griliches, Zvi (1996) "The Discovery of the Residual: A Historical Note", *Journal of Economic Literature*, 34(3), 1324-1330.
Grous, Alex (2011) "Management Practices in the UK Aerospace Sector", LSE mimeo
Gennaioli, Nicola, Rafael LaPorta, Florencio Lopez-de-Silanes, and Andrei Shleifer. (2013) "Human Capital and Regional Development" *Quarterly Journal of Economics* 128(1) 105-164
Homkes, Rebecca (2011) "The missing management link: Why Management Matters in global public-private partnerships" LSE mimeo
Hoxby, Caroline (2000). "Does competition among public schools benefit students and taxpayers?" *American Economic Review, 90*(5), 1209–1238.
Hsieh Chang-Tai and Peter Klenow, 2009. 'Misallocation and Manufacturing TFP in China and India', *Quarterly Journal of Economics*, CXXIV (4).
Hernández-Pérez, Ricardo, Federico Angulo-Browna and Dionisio Tun (2006) "Company size distribution for developing countries", *Physica A*, 359, 607-618
Holmstrom, Bengt and Paul Milgrom (1991) "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design" *Journal of Law, Economics and Organization* 7, 24-52
Ichniowski, Casey, Katheryn Shaw and Prennushi, Giovanni. (1997) "The effects of human resource management practices on productivity: A study of steel finishing lines" *The American Economic Review, 87*(3), 291–313.
Jones, Charles and Paul M. Romer (2010) "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital" *American Economic Journal: Macroeconomics*, 2(1), 224-45,
Jorgenson, Dale W and Griliches, Zvi (1967) "The Explanation of Productivity Change" *Review of Economic Studies*, 34(3), 249-83.
Karlan, Dean, Ryan Knight and Christopher Udry (2012) "Hoping to Win, Expected to Lose: Theory and Lessons on Micro Enterprise Develop ment" NBER Working Paper 18325
Karlan, Dean and Martin Valdivia (2011) "Teaching Entrepreneurship: Impact Of Business Training On Microfinance Clients and Institutions," *Review of Economics and Statistics*,93(2):510-527.
Klette, Tor and Zvi Griliches, 1996. "The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous", *Journal of Applied Econometrics*, 11, 343-361.
Lazear, Edward (2000) "Performance Pay and Productivity", *American Economic Review* 90(5), 1346-1361.
Lazear, Edward P. Kathryn L. Shaw and Christopher T. Stanton (2012) "The Value of Bosses" NBER Discussion Paper 18317

Levy, S. (2011) *In The Plex*: *How Google Thinks, Works, and Shapes Our Lives*
Lucas, Robert (1978) 'On the Size Distribution of Business Firms', *Bell Journal of Economics*, 9, 508-523
Mano, Yukichi, Alhassan Iddrisu, Yutaka Yoshino, and Tetsushi Sonobe (2011). "How Can Micro and Small Enterprises in Sub-Saharan Africa Become More Productive? The Impacts of Experimental Basic Managerial Training," Policy Research Working Paper Series 5755, World Bank.
Marshall, Alfred (1887) "The Theory of Business Profits: Reply" *Quarterly Journal of Economics* 1 (4): 477-481
McCormack, Carol Propper and Sarah Smith (2013) "Herding cats? Management and university performance", forthcoming *Economic Journal*
McConnell, John, Hoffman, Kim, Quanbeck., Andrew, and McCarty, Dennis. (2009) "Management practices in substance abuse treatment programs" *Journal of Substance Abuse*, 37(1), 79-89.
McNally, S. (2010) "Evaluating education policies: The evidence from economic research" *CEP Election Analysis*, http://cep.lse.ac.uk/pubs/download/ea008.pdf
McKinsey and Company (2009) "Management Matters in Northern Ireland and the Republic of Ireland", http://www.delni.gov.uk/managementmatters
McKenzie, David, and Christopher Woodruff (2012) "What are we learning from business training and entrepreneurship evaluations around the developing world?" working paper.
Meagher, K. and Strachan, R. (2013) "Evidence on the Non-Linear Impact of Management" Australian National University mimeo
Melitz, Marc (2003) "The impact of trade on intra-industry reallocations and aggregate productivity growth" *Econometrica*, 71, 1695-1725
Melitz, Marc, and Saso Polanec (2013) "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit", Harvard mimeo
Milgrom, Paul and John Roberts, (1988) "An Economic Approach to Influence Activities and Organizational Responses" *American Journal of Sociology* 96: S154–S179.
Milgrom, Paul and John Roberts (1990) "The Economics of Modern Manufacturing: Technology, Strategy, and Organization", *American Economic Review*, 80(3), 511–28.
Olley, Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica* 64(6), 1263–1298.
Osterman, Paul (1994) "How common is workplace transformation and who adopts it?" *Industrial and Labor Relations Review, 47*(2), 173–188.
Peters, Tom and Robert Waterman (1982) *In Search of Excellence*, New York: Harper business
Perez-Gonzalez, Francisco (2006), "Inherited Control and Firm Performance," *American Economic Review*, 1559–1588.
Rasul, Imran and Daniel Rogger (2013) "Management of Bureaucrats and Public Service Delivery" UCL mimeo
Rivkin, Jan (2000) "Imitation of Complex Strategies." *Science* 46(6) 824–844.
Rosenzweig, Phillip (2007) *The Halo Effect and the Eight Other Business Delusions that Deceive Managers* New York: Free Press

Restuccia, Diego and Richard Rogerson (2008) "Policy Distortions and Aggregate Productivity with Heterogeneous Plants" *Review of Economic Dynamics*, 11(4), 707-720

Salter, W G., (1960) *Productivity and Technical Change.* Cambridge: Cambridge University Press.

Skinner, J., and Staiger, D. (2009) "Technology diffusion and productivity growth in health care" NBER Working Paper No. 14865

Solow, Robert (1957) "Technical Change and the Aggregate Production Function" Review of Economics and Statistics, 39(3) 312-320

Syverson, Chad (2004a) "Product Substitutability and Productivity Dispersion" *Review of Economics and Statistics* 86(2), 534-550.

Syverson, Chad (2011) "What determines productivity?" *Journal of Economic Literature*, 49(2) 326–365

Valdivia, Martin (2012) "Training or Technical Assistance for Female Entrepreneurship? Evidence from a Field Experiment in Peru," GRADE working paper.

Van Reenen, John (2011) "Does competition raise productivity through improving management practices?" *International Journal of Industrial Organization*, (2011) 9(3), 306-317

Walker, Francis W. (1887) "The Source of Business Profits" *Quarterly Journal of Economics* 1(3): 265-288

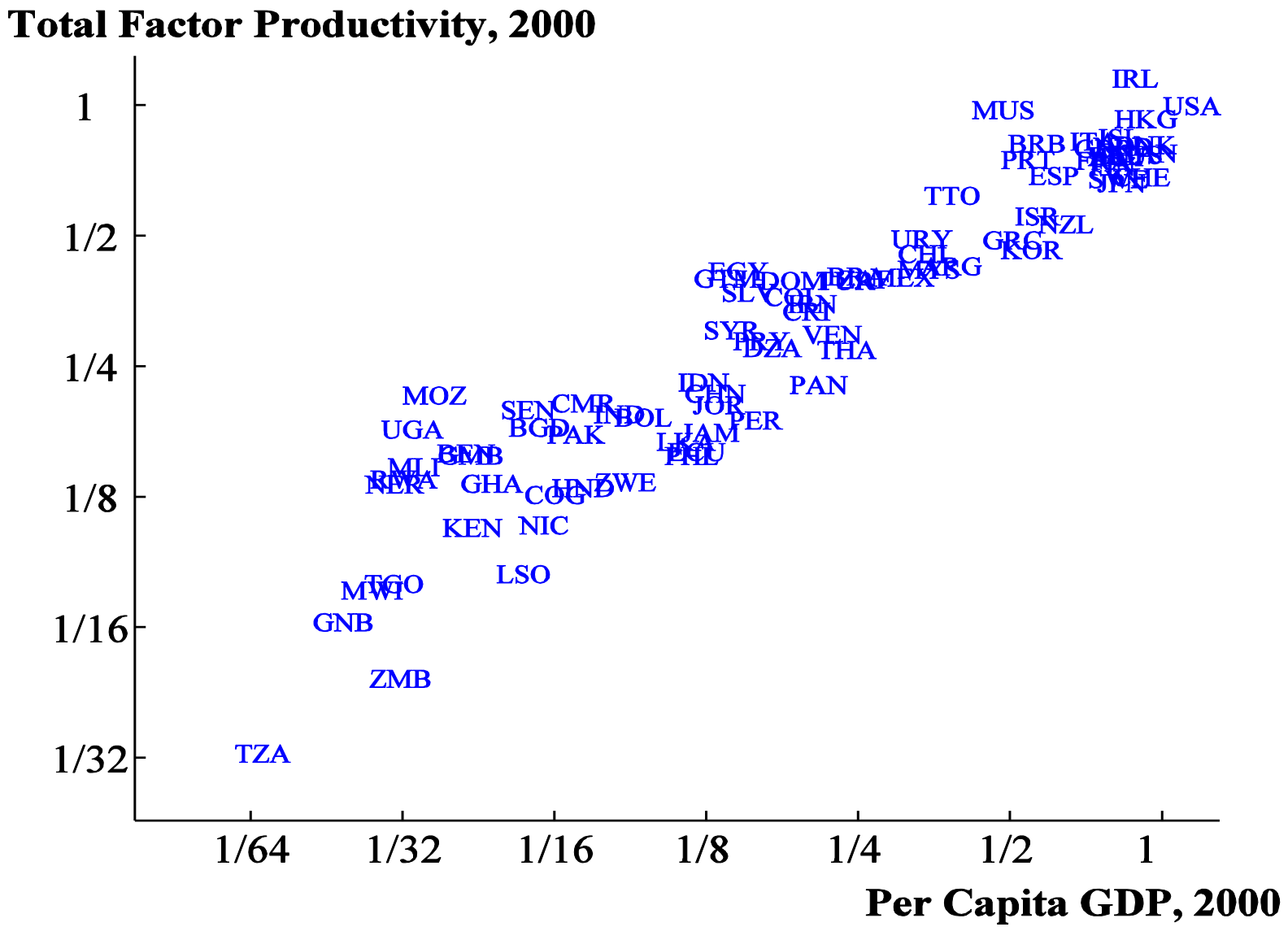
Woodward, Joan (1958) *Management and technology*. London: H. M. Stationery Office.

Appendix Table 1. Management Practice Questions[32]

Categories	Score from 1 to 5 based on:
1) Introduction of modern manufacturing techniques	What aspects of manufacturing have been formally introduced, including just-in-time delivery from suppliers, automation, flexible manpower, support systems, attitudes, and behavior?
2) Rationale for introduction of modern manufacturing techniques	Were modern manufacturing techniques adopted just because others were using them, or are they linked to meeting business objectives like reducing costs and improving quality?
3) Process problem documentation	Are process improvements made only when problems arise, or are they actively sought out for continuous improvement as part of normal business processes?
4) Performance tracking	Is tracking ad hoc and incomplete, or is performance continually tracked and communicated to all staff?
5) Performance review	Is performance reviewed infrequently and only on a success/failure scale, or is performance reviewed continually with an expectation of continuous improvement?
6) Performance dialogue	In review/performance conversations, to what extent are the purpose, data, agenda, and follow-up steps (like coaching) clear to all parties?
7) Consequence management	To what extent does failure to achieve agreed objectives carry consequences, which can include retraining or reassignment to other jobs?
8) Target balance	Are the goals exclusively financial, or is there a balance of financial and nonfinancial targets?
9) Target interconnection	Are goals based on accounting value, or are they based on shareholder value in a way that works through business units and ultimately is connected to individual performance expectations?
10) Target time horizon	Does top management focus mainly on the short term, or does it visualize short-term targets as a "staircase" toward the main focus on long-term goals?
11) Target stretching	Are goals too easy to achieve, especially for some "sacred cow" areas of the firm, or are goals demanding but attainable for all parts of the firm?
12) Performance clarity	Are performance measures ill-defined, poorly understood, and private, or are they well-defined, clearly communicated, and made public?
13) Managing human capital	To what extent are senior managers evaluated and held accountable for attracting, retaining, and developing talent throughout the organization?
14) Rewarding high performance	To what extent are people in the firm rewarded equally irrespective of performance level, or is performance clearly related to accountability and rewards?
15) Fixing poor performers	Are poor performers rarely removed, or are they retrained and/or moved into different roles or out of the company as soon as the weakness is identified?
16) Promoting high performers	Are people promoted mainly on the basis of tenure, or does the firm actively identify, develop, and promote its top performers?
17) Attracting human capital	Do competitors offer stronger reasons for talented people to join their companies, or does a firm provide a wide range of reasons to encourage talented people to join?
18) Retaining human capital	Does the firm do relatively little to retain top talent, or does it do whatever it takes to retain top talent when they look likely to leave?

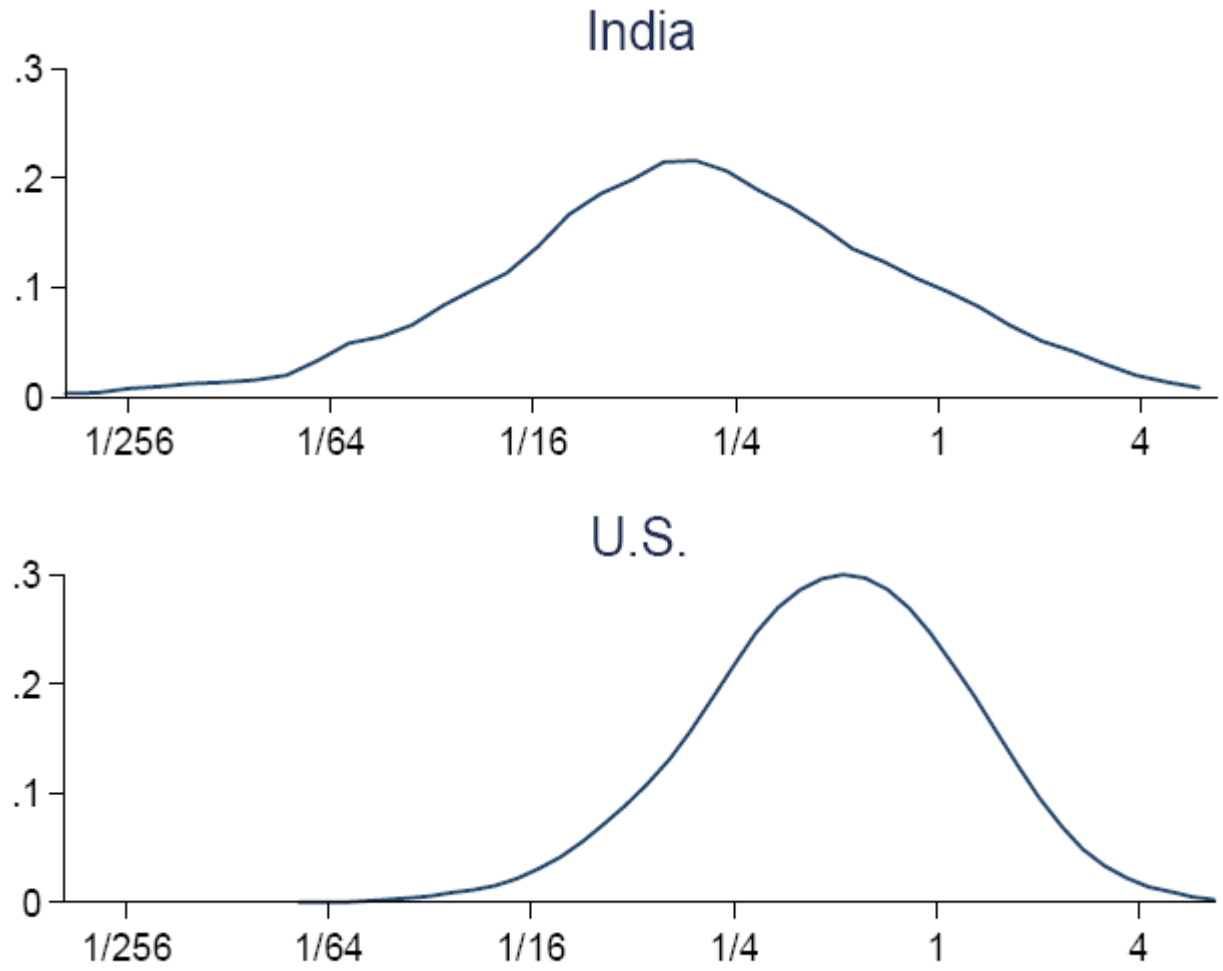
[32] The full set of questions that are asked to score each dimension are included in Bloom and Van Reenen (2007) and also at www.worldmanagementsurvey.com.

FIGURE 1: PER CAPITA & TFP BETWEEN COUNTRIES



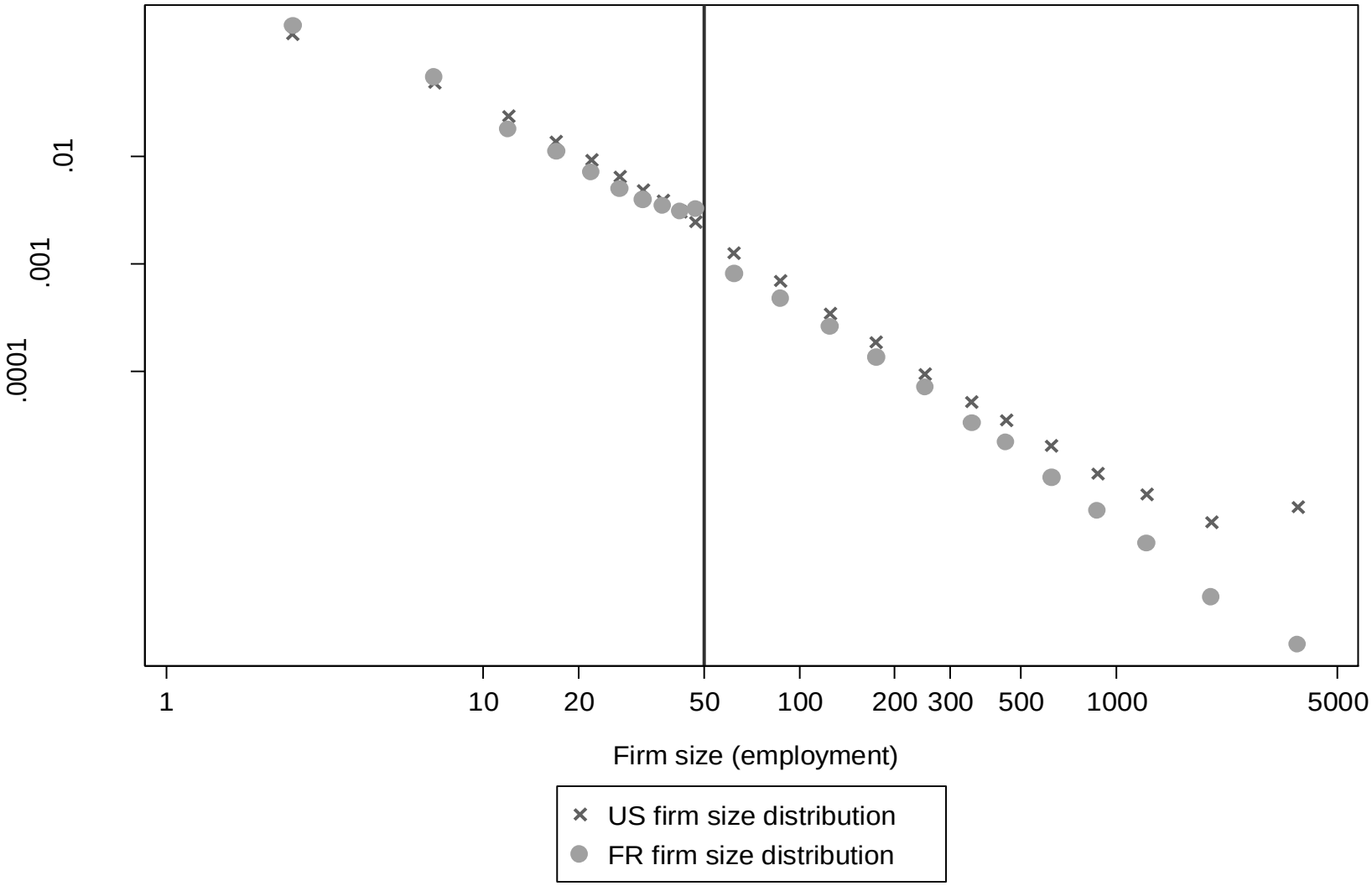
Source: Jones and Romer (2010). US=1

FIGURE 2: DISTRIBUTION OF TFP: US HAS MUCH SMALLER "LEFT TAIL" OF LESS PRODUCTIVE PLANTS THAN INDIA



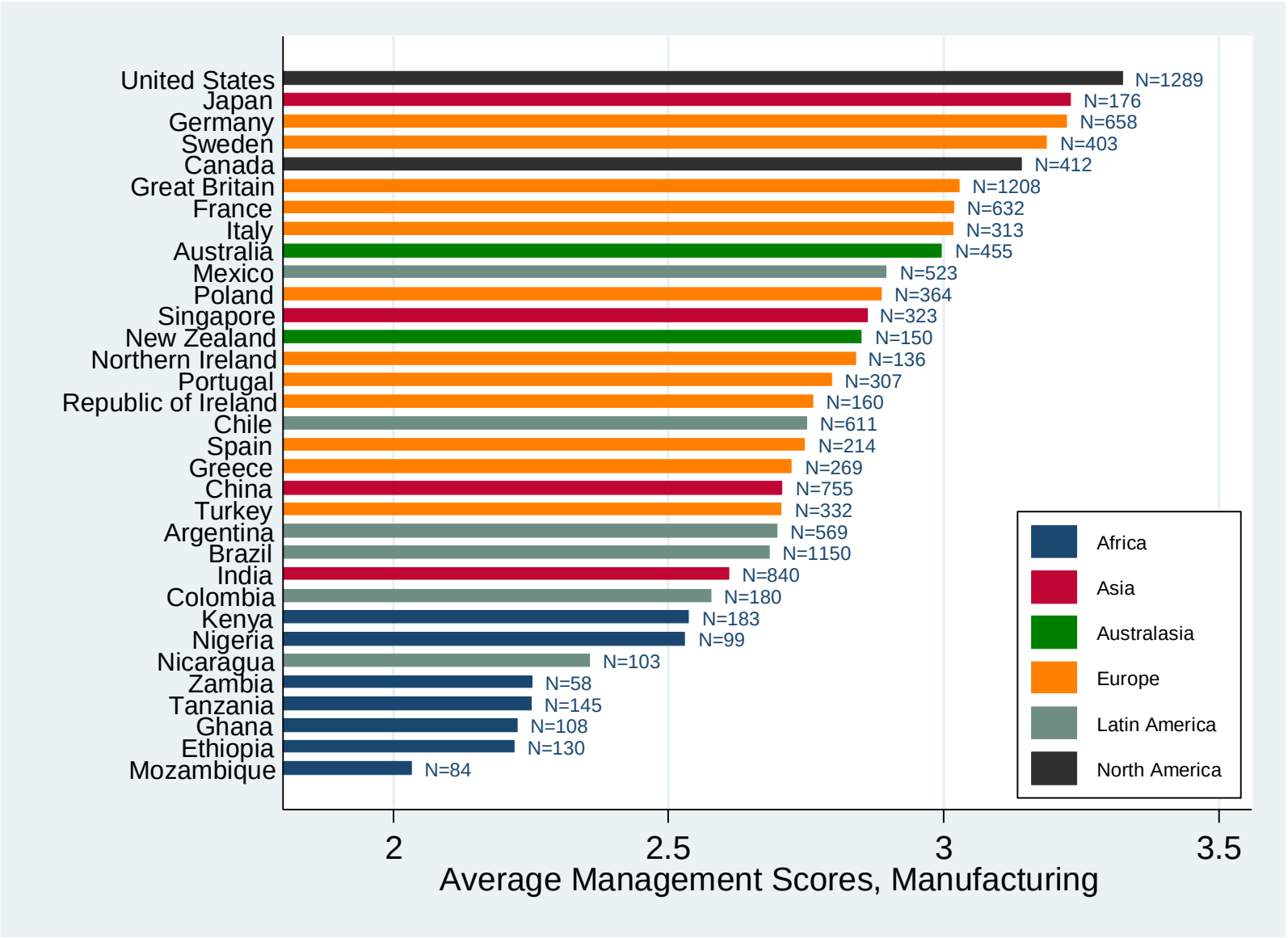
Source: Hsieh and Klenow (2009); mean=1, manufacturing plants

FIGURE 3: FIRM SIZE DISTRIBUTION IN THE US & FRANCE



Source: Garicano, Lelarge and Van Reenen (2013)
Notes: 2007 data. FICUS from France and Census of Manufacturing for US

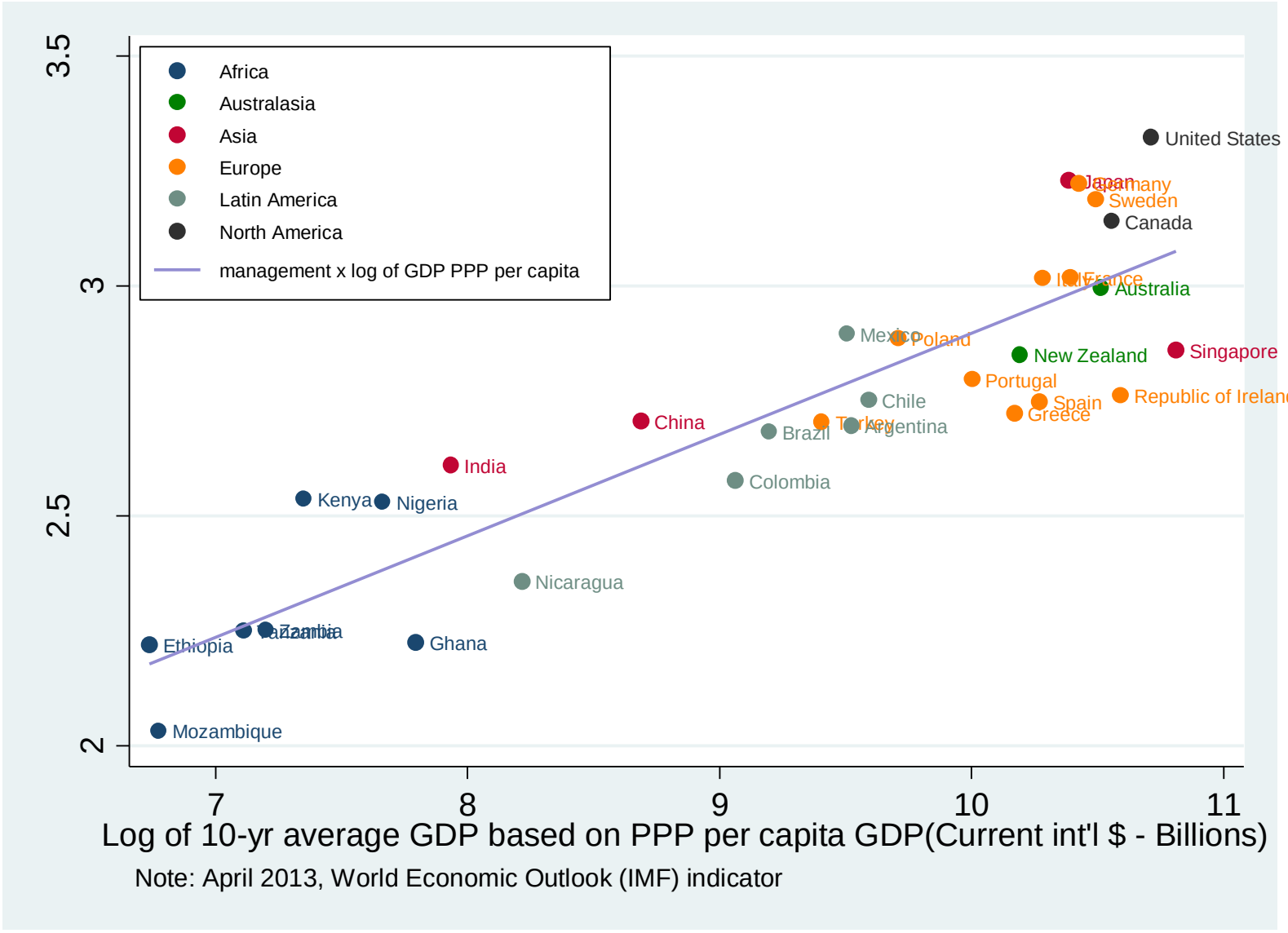
FIGURE 4: AVERAGE MANAGEMENT SCORES BY COUNTRY



Source: Bloom, Sadun and Van Reenen (2013a)

Notes: Data includes 2013/14 survey wave as of April 14th 2014

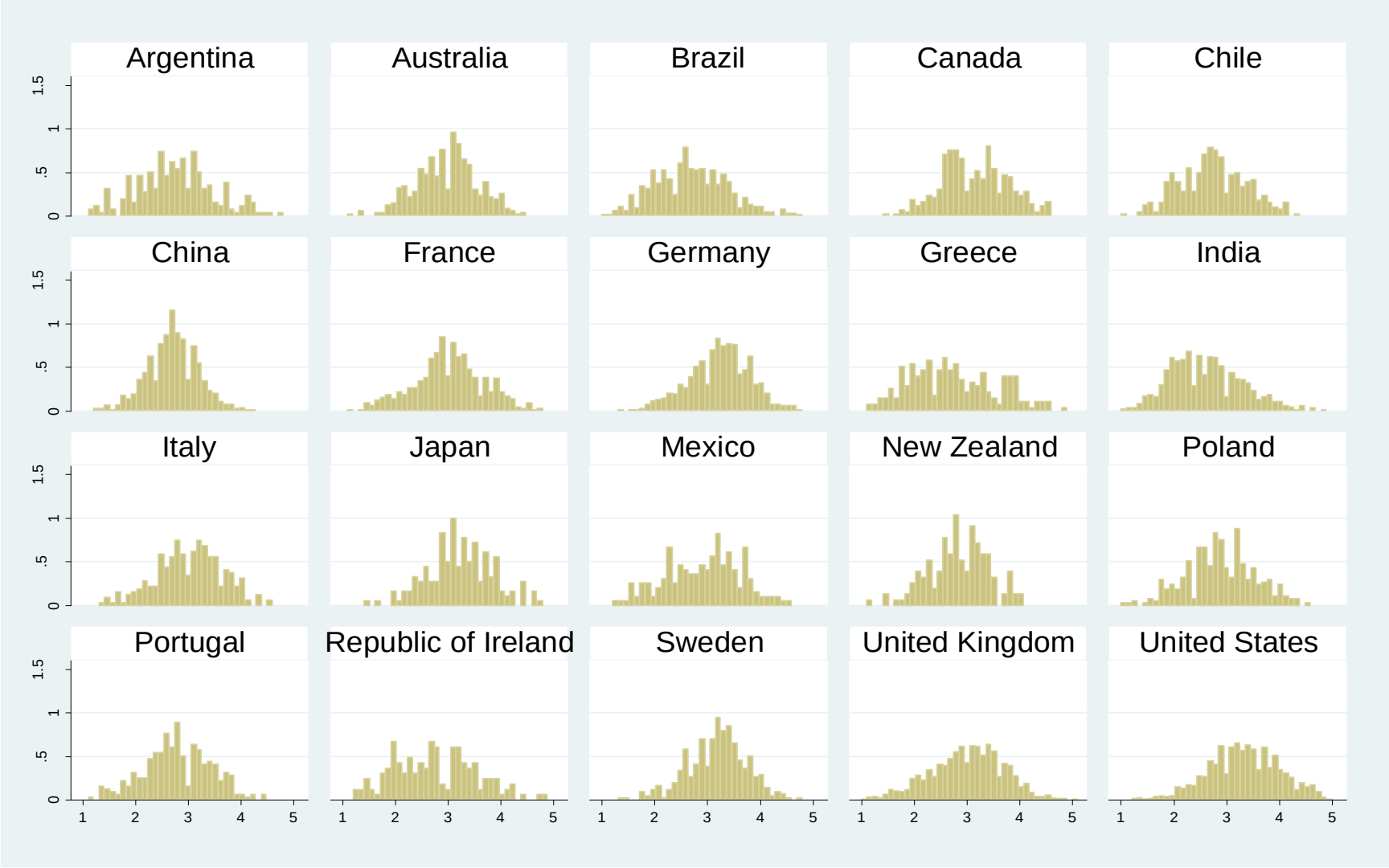
FIGURE 5: AVERAGE MANAGEMENT SCORES ACROSS COUNTRIES ARE STRONGLY CORRELATED WITH GDP PER CAPITA



Source: Bloom, Sadun and Van Reenen (2013a)

Notes: Data includes 2013/14 survey wave as of April 14th 2014

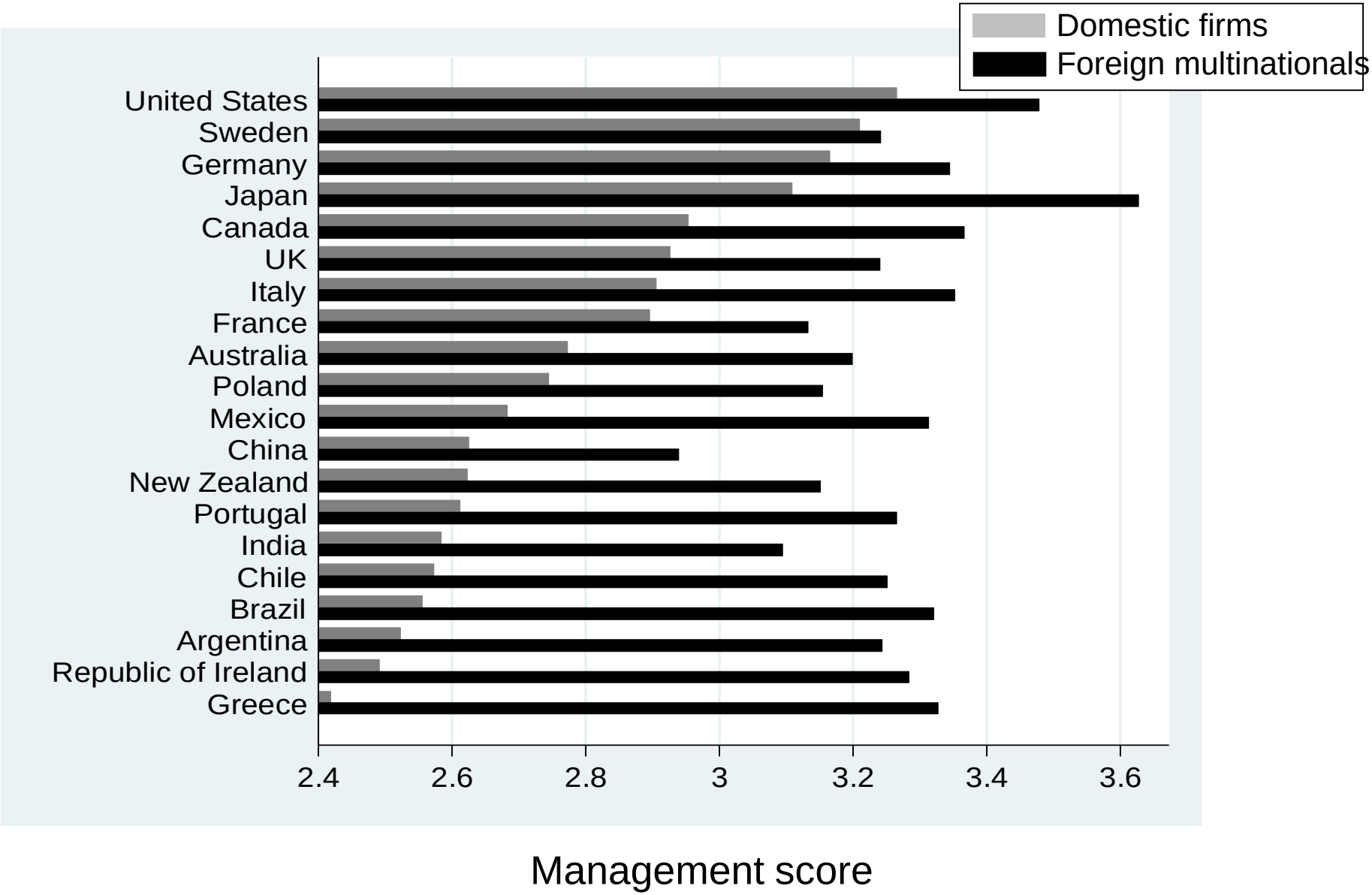
FIGURE 6: LARGE VARIATION IN MANAGEMENT SCORES ACROSS FIRMS WITHIN COUNTRIES



Source: Bloom, Sadun and Van Reenen (2013a)

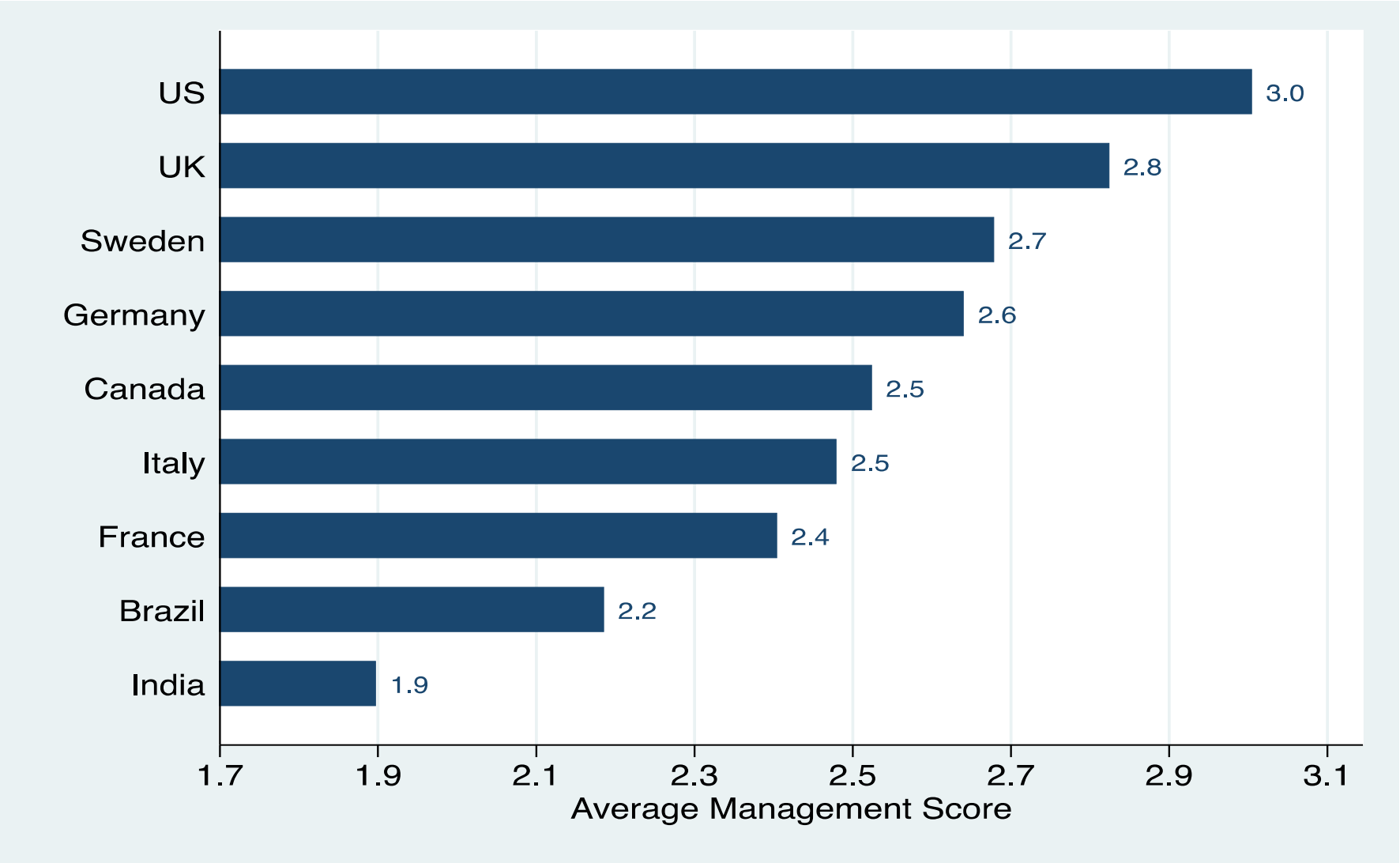
Notes: Data includes 2013 survey wave as of Oct 4th 2013, Bars are the histogram of the actual density. Scores from 9,995 management interviews across 20 countries.

FIGURE 7: MULTINATIONALS ACHIEVE HIGH MANAGEMENT SCORES WHEREVER THEY LOCATE



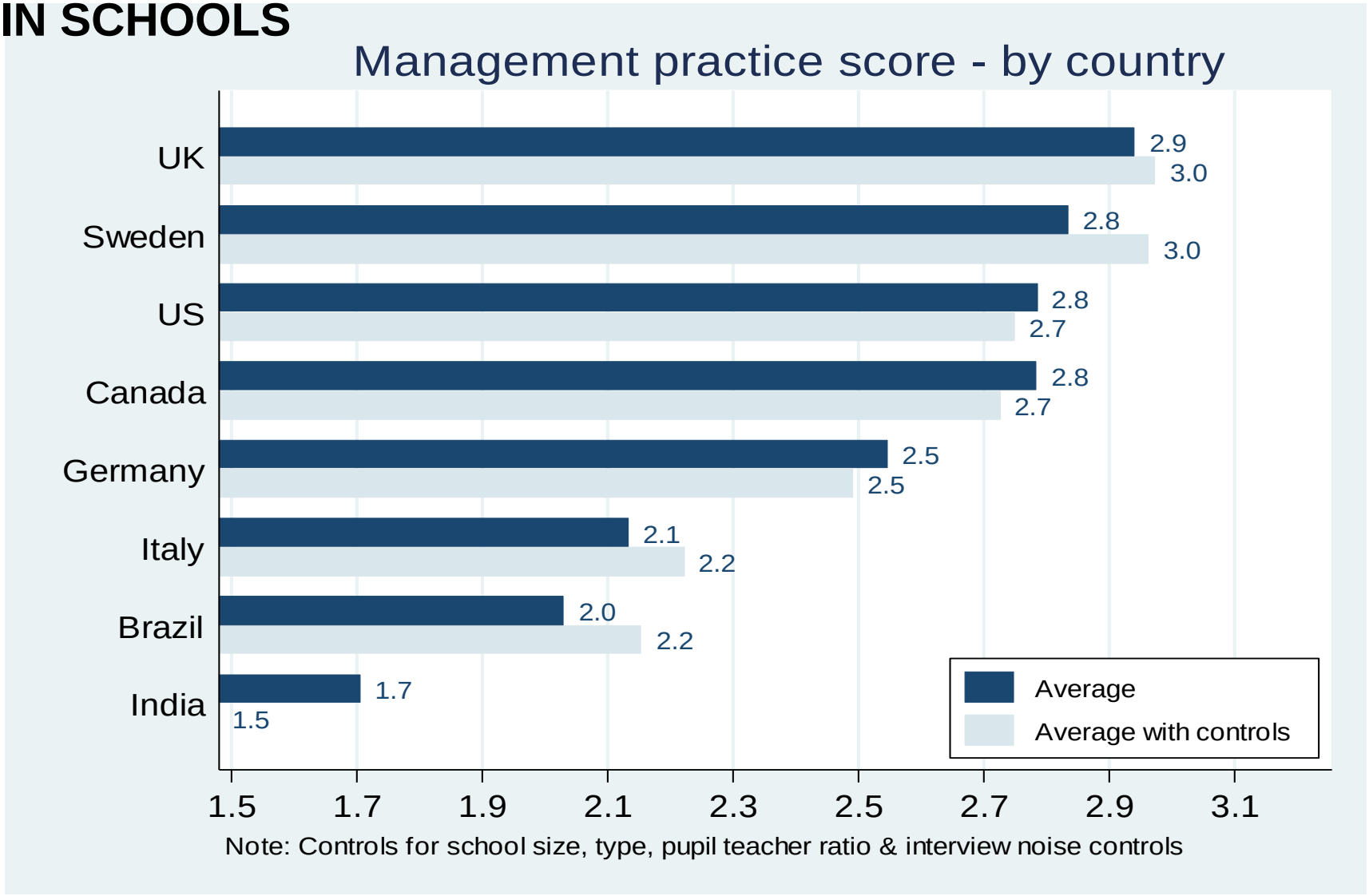
Notes: Sample of 7,303 manufacturing firms, of which 4,926 are purely domestic and 2,377 are foreign multinationals. Domestic multinationals are excluded – that is the domestic subsidiaries of multinational firms (like a Toyota subsidiary in Japan).

FIGURE 8: CROSS COUNTRY AVERAGE MANAGEMENT SCORES IN HOSPITALS



Source: Bloom, Sadun & Van Reenen (2013b)
Notes: 1,971 acute care hospitals with a cardiology and orthopedics department

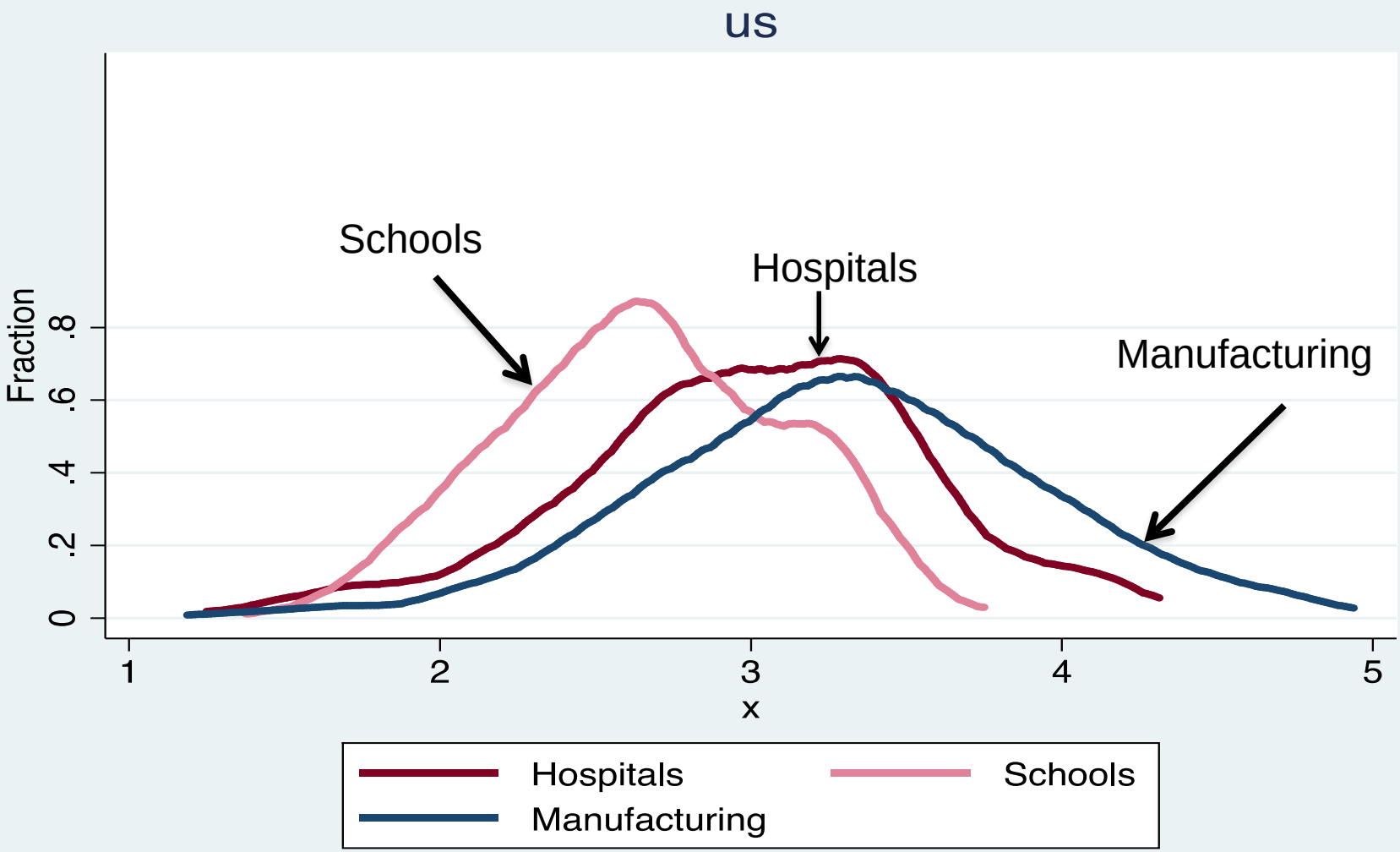
FIGURE 9: CROSS COUNTRY AVERAGE MANAGEMENT SCORES IN SCHOOLS



Source: Bloom, Lemos, Sadun and Van Reenen (2014)

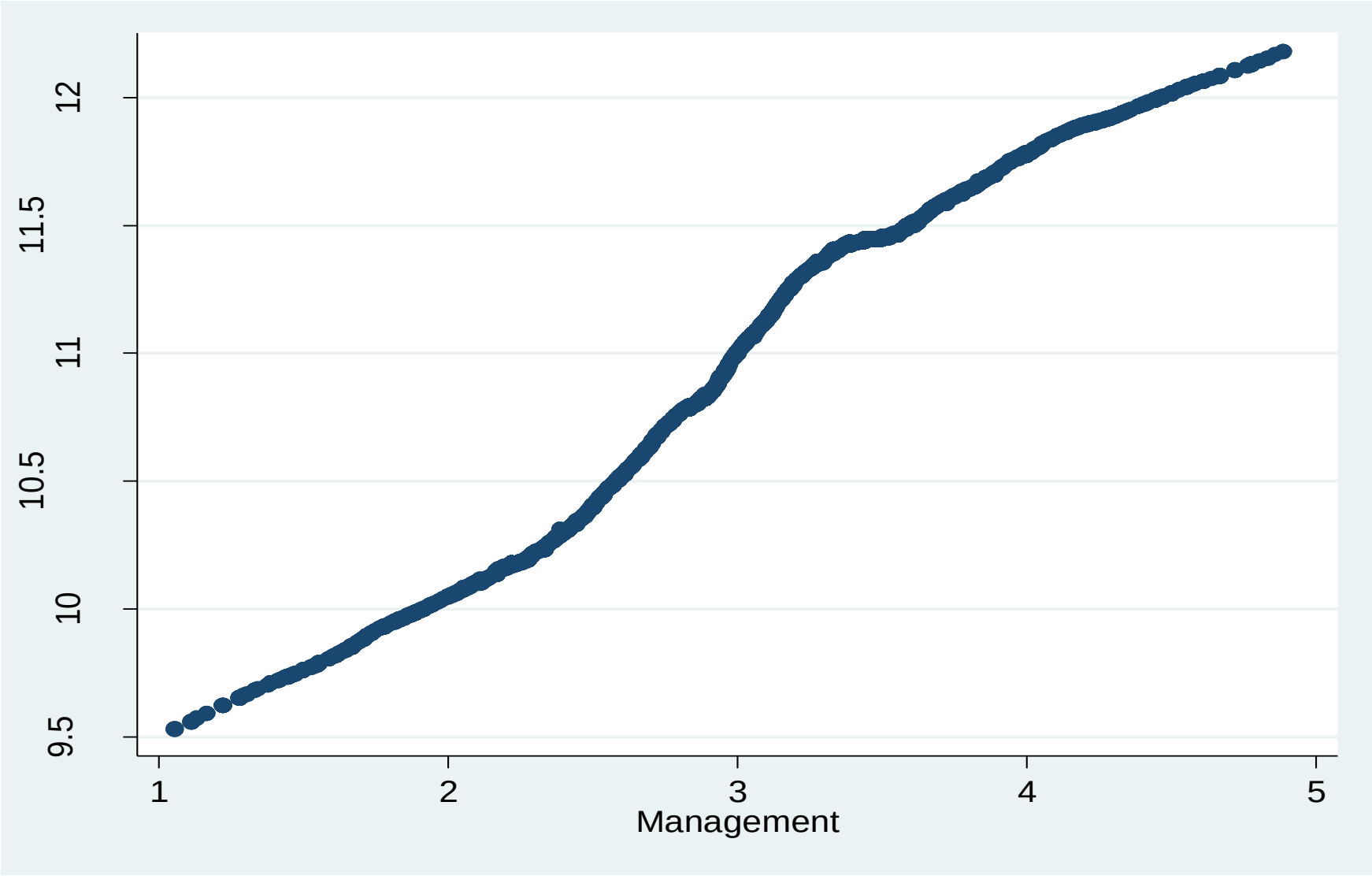
Notes: 1,851 secondary schools (defined as those teaching 15 year olds); Controls include # students, pupil/teacher ratio, school type (autonomous government, private, regular government); curriculum type (academic, vocational), noise controls

FIGURE 10: COMPARISON OF MANAGEMENT SCORES ACROSS THREE SECTORS (IN US)



Source: Bloom, Lemos, Sadun and Van Reenen (2014)
Notes: Management kernel density plot

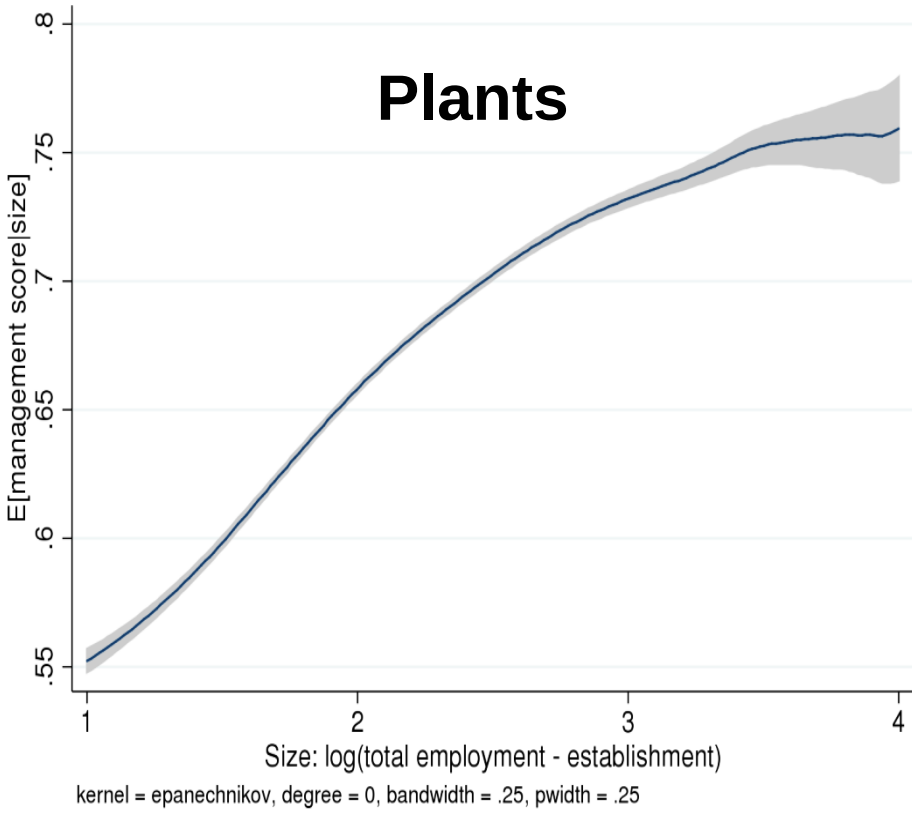
FIGURE 11: FIRM SIZE IS INCREASING IN MANAGEMENT SCORE

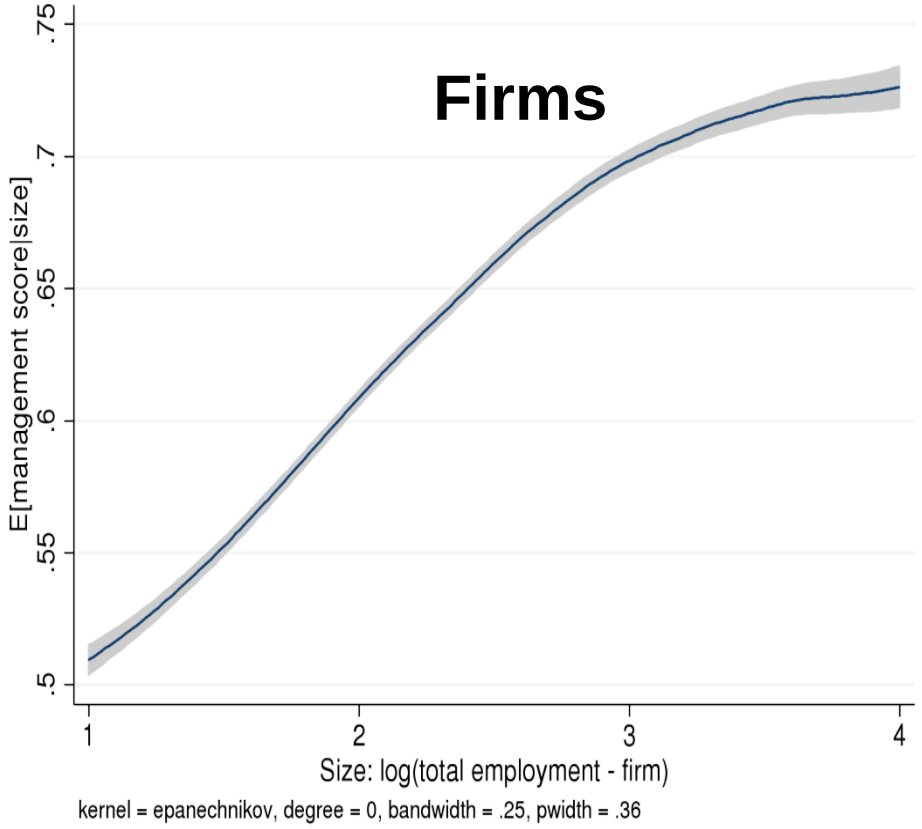


Source: Bloom, Sadun and Van Reenen (2013a)

Notes: WMS: Management is average of 18 questions. Sales is log(sales) in US$. N=10,197

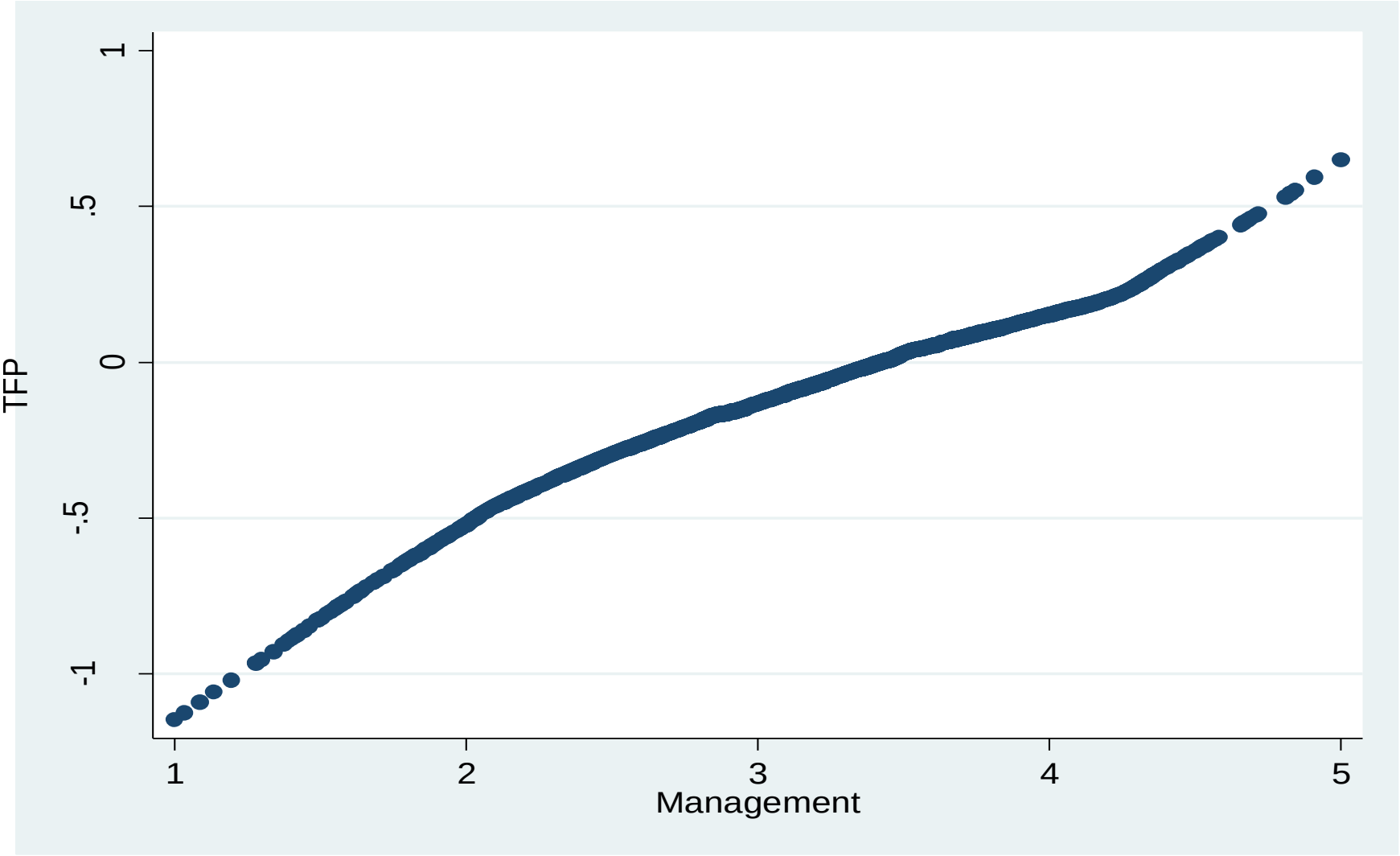
FIGURE 12: BETTER MANAGED FIRMS AND PLANTS ARE LARGER





Notes: Management data from the 2011 MOPS, Management and Organizational Performance Survey (US Census data on 32,000 Annual Survey of Manufacturing establishments).
Source: Bloom, Brynjolfsson, Foster, Jarmin, Saporta and Van Reenen (2012)

FIGURE 13: TFP IS INCREASING IN MANAGEMENT



Source: Bloom, Sadun and Van Reenen (2013a)
Notes: Management is an average of all 18 questions (set to sd=1). TFP residuals of sales on capital, labor, skills controls plus a full set of SIC-3 industry, country and year dummies controls. N=8,314

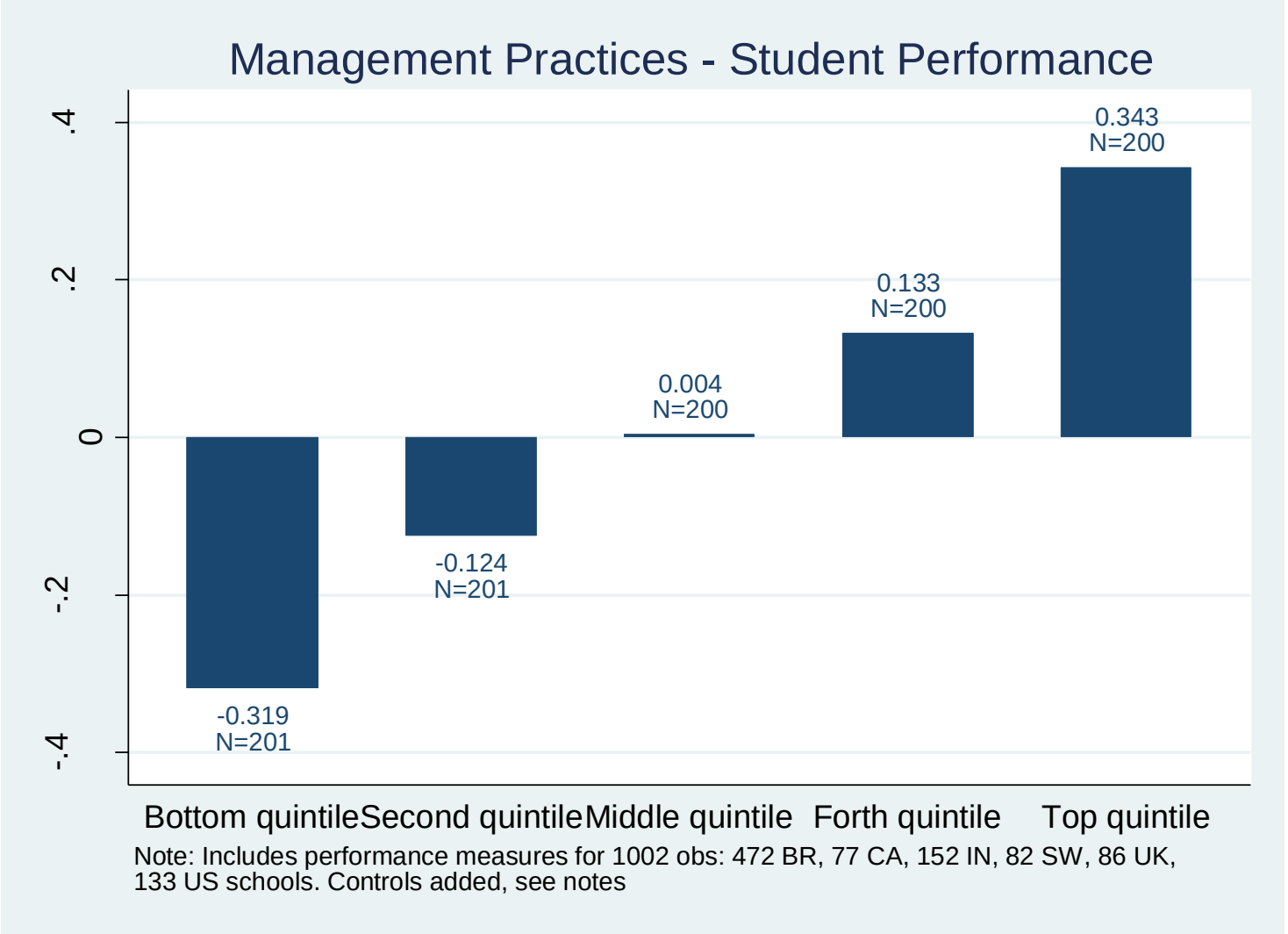
FIGURE 14: PATIENT OUTCOMES BETTER WHEN MANAGEMENT SCORES HIGHER (US, UK, CANADA & SWEDEN)

Dependent Variable: Case mix adjusted AMI 30 days mortality rates (z-scored by country)						
Countries	**All**		**US**	**UK**	**Canada**	**Sweden**
Management (z-score)	-0.162***	-0.246***	-0.211**	-0.416*	-0.717**	-0.543***
	(0.056)	(0.075)	(0.100)	(0.224)	(0.316)	(0.193)
Observations	324	324	178	74	24	48
Country dummies	y	y				
Hospital controls	y	y	y	y	y	y
Region & noise controls		y	y	y	y	y

Notes. OLS estimates with standard errors clustered by hospital. Hospital controls: size, specialty, % managers with a clinical degree. “Noise controls”: interviewer dummies, seniority & tenure of manager; interview duration, reliability indicator, interviewee type (nurse, doctor or manager).

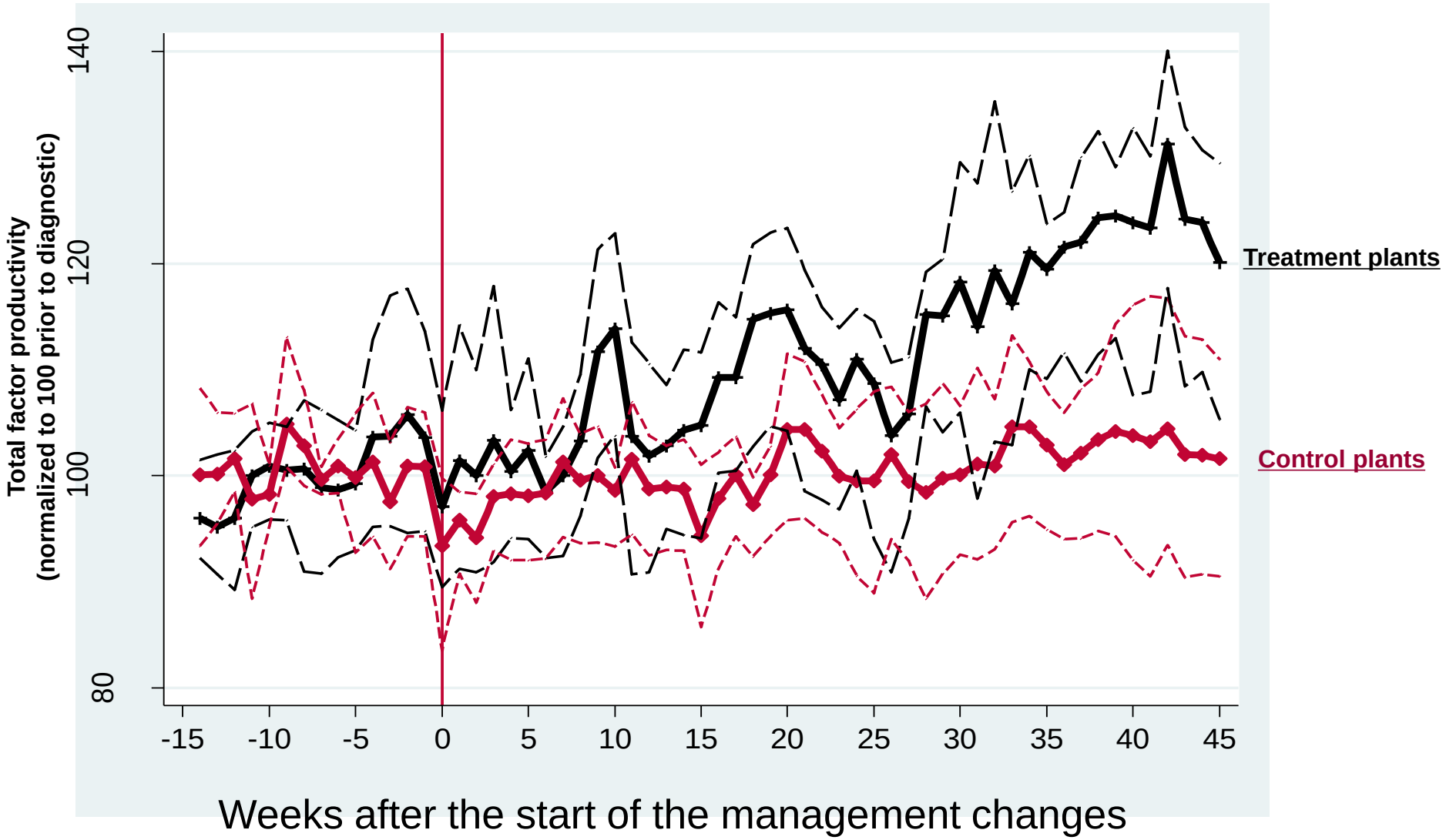
Source: Bloom, Sadun & Van Reenen (2013b)

FIGURE 15: PUPIL TEST SCORES CORRELATED WITH HIGHER MANAGEMENT SCORES



Source: Bloom, Lemos, Sadun and Van Reenen (2014) **Notes:** We use the math exam pass rate from HSEEs in US public schools, GCSE score in UK, the school-level rating produced by the Fraser Institute in Canada, the 9th grade GPA in Sweden, the school-level average in maths in the High School National Exam in Brazil, and the X Standards Math Score in India. We z-score the student achievement data within country to take into account differences in school performance measures. Regional dummies and school-level controls for number of students, pupil/teacher ratio, school type dummies and noise controls included.

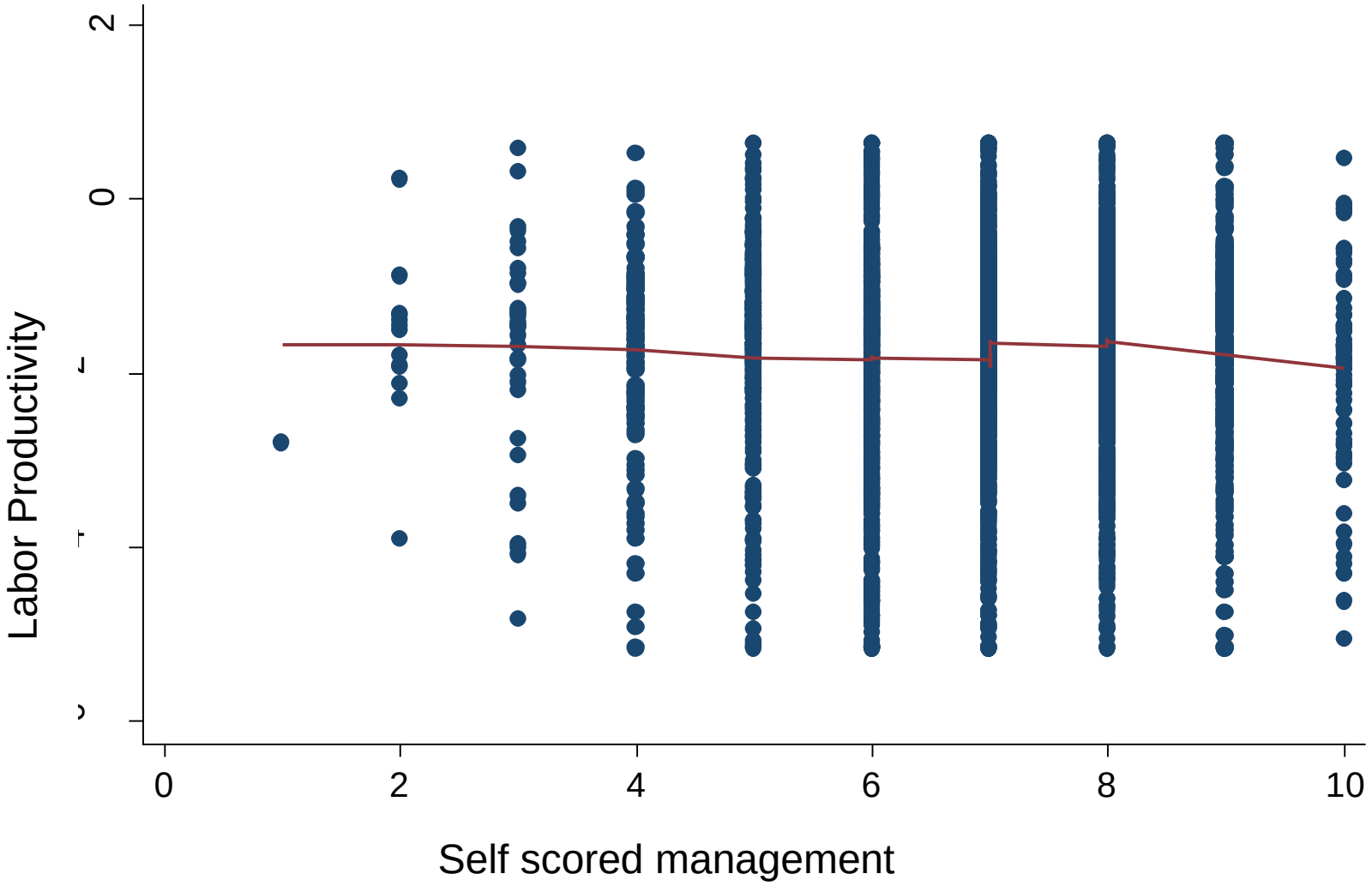
FIGURE 16: PRODUCTIVITY IMPROVEMENTS IN RCT ON ADOPTION OF MANAGEMENT PRACTICES



Notes: Weekly average total factor productivity for 14 treatment & 6 control plants. All plants make cotton fabric near Mumbai, India, with between 100 and 1000 employees. Values normalized so both series have an average of 100 prior to the start of the intervention. Confidence intervals bootstrapped over firms.

Source: Bloom, Eifert, Mahajan, McKenzie & Roberts (2013).

FIGURE 17: SELF-SCORED MANAGEMENT UNCORRELATED WITH PRODUCTIVITY



Source: Bloom, Sadun & Van Reenen (2013b)
Note: Insignificant 0.03 correlation with labor productivity

CENTRE FOR ECONOMIC PERFORMANCE
Occasional Papers

40	Ghazala Azmat Barbara Petrongolo	Gender and the Labor Market: What Have We Learned from Field and Lab Experiments?
39	Christopher J. Boyce Alex M. Wood James Banks Andrew E. Clark Gordon D.A. Brown	Money, Well-Being, and Loss Aversion: Does an Income Loss Have a Greater Effect on Well-Being than an Equivalent Income Gain?
38	Nicholas Bloom	Fluctuations in Uncertainty
37	Nicholas Oulton	Medium and Long Run Prospects for UK Growth in the Aftermath of the Financial Crisis
36	Phillipe Aghion Nicholas Bloom John Van Reenen	Incomplete Contracts and the Internal Organisation of Firms
35	Brian Bell John Van Reenen	Bankers and Their Bonuses
34	Brian Bell John Van Reenen	Extreme Wage Inequality: Pay at the Very Top
33	Nicholas Oulton	Has the Growth of Real GDP in the UK been Overstated Because of Mis-Measurement of Banking Output?
32	Mariano Bosch Marco Manacorda	Social Policies and Labor Market Outcomes in Latin America and the Caribbean: A Review of the Existing Evidence
31	Alex Bryson John Forth Minghai Zhou	What Do We Know About China's CEO's? Evidence from Across the Whole Economy
30	Nicholas Oulton	Horray for GDP

29	Stephen Machin	Houses and Schools: Valuation of School Quality through then Housing Market – EALE 2010 Presidential Address
28	John Van Reenen	Wage Inequality, Technology and Trade: 21st Century Evidence
27	Barry Anderson Jörg Leib Ralf Martin Marty McGuigan Mirabelle Muûls Laure de Preux Ulrich J. Wagner	Climate Change Policy and Business in Europe Evidence from Interviewing Managers
26	Nicholas Bloom John Van Reenen	Why do Management Practices Differ Across Firms and Countries?
25	Paul Gregg Jonathan Wadsworth	The UK Labour Market and the 2008-2009 Recession
24	Nick Bloom Raffaella Sadun John Van Reenen	Do Private Equity Owned Firms Have Better Management Practices?
23	Richard Dickens Abigail McKnight	The Impact of Policy Change on Job Retention and Advancement
22	Richard Dickens Abigail McKnight	Assimilation of Migrants into the British Labour Market
21	Richard Dickens Abigail McKnight	Changes in Earnings Inequality and Mobility in Great Britain 1978/9-2005/6
20	Christopher Pissarides	Lisbon Five Years Later: *What future for European Employment and Growth?*
19	Richard Layard	Good Jobs and Bad Jobs

For more information please contact the Publications Unit
Tel: +44 (0)20 7955 7673; Fax: +44 (0)20 7955 7595; Email: cep.info@lse.ac.uk
Website http://cep.lse.ac.uk