

Coad, Alex; Mathew, Nanditha; Pugliese, Emanuele

Working Paper

Positioning firms along the capabilities ladder

UNU-MERIT Working Papers, No. 2021-031

Provided in Cooperation with:

Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT),
United Nations University (UNU)

Suggested Citation: Coad, Alex; Mathew, Nanditha; Pugliese, Emanuele (2021) : Positioning firms along the capabilities ladder, UNU-MERIT Working Papers, No. 2021-031, United Nations University (UNU), Maastricht Economic and Social Research Institute on Innovation and Technology (UNU-MERIT), Maastricht

This Version is available at:

<https://hdl.handle.net/10419/326790>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>



**UNITED NATIONS
UNIVERSITY**

UNU-MERIT

Working Paper Series

#2021-031

Positioning firms along the capabilities ladder

Alex Coad, Nanditha Mathew and Emanuele Pugliese

Published 13 August 2021

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)

email: info@merit.unu.edu | website: <http://www.merit.unu.edu>

Boschstraat 24, 6211 AX Maastricht, The Netherlands

Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers

ISSN 1871-9872

**Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT**

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT to stimulate discussion on the issues raised.



Positioning firms along the capabilities ladder *

Alex Coad^{†1}, Nanditha Mathew², and Emanuele Pugliese³

¹Waseda Business School, Waseda University, Tokyo, Japan

²UNU-MERIT, Maastricht, Netherlands

³European Commission, Joint Research Centre (JRC), Seville, Spain

Abstract

We develop and apply a novel methodology for quantifying the capability development of firms, and putting these capabilities (and hence also the firms) in a hierarchy, that we refer to as their position on the capabilities ladder. Our nestedness algorithm, inspired by biology and network science, defines a capability as complex if it is performed by only a few firms at the upper rungs of the ladder. We analyze balance sheet and innovation data of almost 40'000 Indian firms for the time period 1988-2015, and observe significant nestedness. Lower rungs of the capabilities ladder correspond to basic managerial and production capabilities. Mid-level rungs correspond to internationalization and acquiring absorptive capacity. Higher level rungs are more related to M&A and innovation. ICT capabilities have become more fundamental lower-level rungs on the capabilities ladder in recent years. We find that capability ranking can explain future growth patterns and survival probability of firms, summing up in one number their future potential trajectories.

Keywords: Capabilities, Competences, Complexity, Balance sheet data, Resources

JEL codes: L10, D22, O12

*Authors are listed alphabetically. The content of this article does not reflect the official opinion of the European Union. Responsibility for the information and views expressed therein lies entirely with the authors.

[†]*Corresponding author:* Waseda Business School, Waseda University, Bldg. 11 1-6-1 Nishi-Waseda, Shinjuku, Tokyo 169-8050, Japan. ORCID: 0000-0002-5706-0528. Email: alex.coad@waseda.jp

1 Introduction

Strategy scholars have long emphasized that superior firm-level performance refers to the competitive advantage that is conferred by firm-specific capabilities and resources (Penrose, 1959; Barney, 1991). Naturally, this has led to interest in how such superior capabilities can be developed (Teece et al., 1997; Dosi et al., 2000; Helfat and Peteraf, 2003; Ethiraj et al., 2005). While strategy scholars generally agree on the importance of the development of capabilities for superior performance, nevertheless progress in the area is being held back by difficulties in operationalizing the capabilities construct in empirical research. This is a pity, because empirical progress can be expected to lead to further theoretical refinement. This was emphasized in a recent survey of the strategic management literature on capabilities: “Organizational capability is a central construct in theories of competitive advantage, strategic choice, firm boundaries, learning and adaptation, and technological change. Yet, if organizational capabilities are to fulfill their potential to illuminate these phenomena, there is an urgent need to surmount the problems of identification and measurement that continue to impede empirical research.” (Grant and Verona, 2015, p.72).

This article therefore seeks to contribute to the measurement of capabilities and the dynamic processes of capability development. Research into capability development and capability accumulation faces (at least) two major challenges, however. A first challenge relates to how capabilities are operationalized and linked to real-world management phenomena. The lack of precision and specificity with which capabilities are identified in some studies is a problem for knowledge accumulation and subsequent theory-building (Grant and Verona, 2015, p.66). Failure to connect abstract constructs like “capabilities” and “resources” to real-world phenomena risks leading to tautological statements (Arend, 2006), for example if (unobserved) valuable capabilities are invoked to explain the performance of successful firms. We address this concern by complementing the existing literature with a new approach that uses objective data on firms’ reported expenditures.

A second challenge relates to how these capabilities can be arranged (ideally, in quantitative terms) into a hierarchy, such that a distinction can be made between advanced and lower-level capabilities. Positioning capabilities into a hierarchy or a ladder is useful for the purposes of making specific normative recommendations to firms regarding which steps to take next, i.e. the paths that should be taken by firms to develop their capabilities and progress to higher-value markets and activities. In this way, we seek to contribute to our understanding of the dynamics of capability

development. In the development economics literature, Cirera and Maloney (2017) discuss the concept of a “capabilities escalator” according to which capabilities are in a tripartite hierarchy of basic production/management, technology adoption, and invention/innovation capabilities. We draw on this intuition to attempt to develop a hierarchy of capabilities based on company expenditures in 47 different dimensions, sorting these capabilities into a hierarchical order by applying network science algorithms (Tacchella et al., 2012).

We therefore contribute to the literature by proposing a new method for measuring capabilities and arranging routines and capabilities (and hence firms) within a hierarchy, that we refer to as the ‘capabilities ladder’. Our measure of capabilities has a number of advantages. While some previous research has made valuable contributions to the measurement of capabilities using subjectively-reported indicators of capabilities (e.g. Bloom and Van Reenen, 2010; Danneels, 2016), nevertheless subjectively-reported indicators have some drawbacks, such as a possible lack of detailed knowledge on the part of the respondent, perceptual and cognitive distortions (such as recall bias, social desirability bias, etc), and the possibility of deliberate misreporting (Grant and Verona, 2015). The costs of administering large-scale questionnaire studies may also discourage empirical investigations of capabilities. Our approach to measuring capabilities refers to concrete observations of firm-level expenditures in various areas. While subjectively-reported data carries the risk of ‘cheap talk’, our indicators of activity require the actual commitment of resources. Although some previous research sought to analyze detailed data on capabilities from specific industries (e.g. Ethiraj et al 2005 for the software services industry; Pandza 2011 for a pharmaceutical firm), our approach is general in the sense that it refers to firm activities that are observed across sectors. Our approach requires detailed balance sheet data for firms, but this type of information is usually readily available across sectors and countries, therefore making our approach amenable to replication and extension in a large range of contexts.

Our analysis shows that the lower rungs of the capabilities ladder correspond to basic managerial and production capabilities, such as borrowing money, owning land and building assets, and expenses due to outsourcing professional jobs. Mid-level rungs correspond to internationalization and acquiring absorptive capacity. Higher-level rungs in the capability ladder are less clearly ordered into a hierarchy. Overall, therefore, the path for capability development is more clearly defined at lower and mid-level rungs than at the top. Focusing on changes in the structure of the ladder over time, ICT capabilities have become more present at the fundamental lower-level rungs on the capabilities

ladder in recent years. Our capabilities ranking is correlated with firm size for small firms but not for large firms, indicating that firm growth is associated with capability development in small firms but not for larger firms.

The paper unfolds as follows. Section 2 introduces our statistical methodology and links it to theoretical issues in the strategic management literature. Section 3 presents our data on 39,992 Indian firms for the period 1988-2015. Section 4 contains our analysis and results, first sorting the capabilities into a hierarchy according to our algorithm, and then investigating dynamics in the hierarchy of the ladder. Overall scores for capabilities are compared across broad industry groups and for firm-level variables (firm size and survival). Section 5 concludes.

2 Methodology

2.1 Theoretical background on the measurement of capabilities

“The core problem of empirical research into organizational capability is that organizational capabilities are latent constructs that are inherently unobservable.” (Grant and Verona, 2015, p.61).

The unobservable nature of capabilities is one of their core characteristics, because it is this tacit and ambiguous aspect of capabilities that allows firms to provide competitive advantage without being imitated by rivals (Hilliard and Goldstein, 2019). However, the unobservable nature of capabilities is a headache for empirical researchers. As a consequence, indirect approaches to measuring capabilities are required. A common approach in the previous literature involved the analysis of questionnaire or telephone survey responses from firms to obtain subjective assessments of their firms’ capabilities (e.g. Danneels, 2016; Bloom et al., 2013; Costa et al., 2021). Surveys of firms can shed valuable light on phenomena that are not detectable using administrative sources only (e.g. Costa et al., 2021). Subjectively-reported indicators of capabilities have drawbacks however. Grant and Verona (2015) identify three sources of error in self-reported assessments of capabilities: deliberate misreporting, perceptual and cognitive distortions (leading to unintended biases in responses), and a lack of information on the part of the respondent. These biases can be attenuated, for example by checking the background and expertise of individual respondents (Danneels, 2016), or by performing double-blind interviews (Bloom and Van Reenen, 2010) but this is time-consuming and costly. We do not seek to diminish the commendable achievements of previous research that obtains subjectively-reported measurements of capabilities, but we seek to complement the existing

literature with a new technique for the development of capabilities based on factual observations in company balance sheets.

This paper therefore measures capabilities using objective data, although (given the unobserved nature of capabilities) this is done in an indirect way. Our approach draws on a widely-held definition of a capability as the ability to perform an activity or function using a set of tangible resources (Amit and Schoemaker, 1993; Helfat and Peteraf, 2003; Danneels, 2016). (Capabilities are treated here as being synonymous with competences.) An important point is that a capability has material as well as cognitive components (Danneels, 2016). While the “ability to perform” is fundamentally unobservable, nevertheless this ability can only manifest itself through the deployment of material resources. As such, an organizational capability is only observed when it is performed through the application of a skill involving tangible resources (Grant and Verona, 2015).

An example of an (individual-level) capability is given in Grant and Verona (2015, p.61):¹ playing the violin. Here, the capability of playing a violin is only observed when someone actually starts to play the violin, and the capability remains latent or dormant until a violin is physically involved. In the context of our approach, buying a violin would be taken as evidence of a capability for playing violin. Furthermore, we do not distinguish between the quantity and quality of violins being purchased, instead we analyze binary variables regarding objectively-reported expenditures in various areas to deduce the state of a firms’ set of capabilities. Applying the analogy to our approach, we are not concerned with whether some violinists are better than others, because our main concern is whether an agent is able to carry out a specific activity (such as playing the violin) purposefully, repeatedly, reliably, and in a minimally satisfactory manner (Helfat and Winter, 2011; Danneels, 2016). In this way, firms’ expenditures in various areas can be considered to leave a paper trail of the performance of tasks that would only be undertaken if firms have capabilities in certain areas that are (given the current competitive environment) considered to be sufficiently developed that they are consistent with an overall profit-seeking strategy.

A similar approach has been taken in previous papers that took R&D expenditures as an indicator of (unobserved) technological capability (Helfat, 1997; Rothaermel and Hess, 2007). In these papers, the fact that firms consider it to be in their interests to invest in R&D is taken as a signal that these firms have sophisticated routines and management practices in place to undertake complex activities relating to R&D. Our paper extends this line of investigation, because the dimension of

¹A violin analogy also appears in the discussion of capabilities in Costa et al. (2021, p.5).

R&D investment is only one out of 47 areas of expenditure that are analyzed to assess a firm’s position in capability space.

An influential stream of research in the measurement of capabilities has emerged from the World Management Survey (e.g. Bloom and Van Reenen, 2010; Sadun et al., 2017). Our approach is similar to theirs, in that we measure and quantitatively analyze capabilities in a variety of sectors and for a potentially large number of firms. However, our approach is different from theirs in that our approach can be applied at low cost to a population of firms that report detailed balance sheet data. Furthermore, our focus is not on management capabilities but on capabilities across various functions, such as marketing, financial practices, human resource practices, IT strategy, internationalization and innovation. Importantly, these capabilities are reflected not in self-reported data but in how firms actually spend their money in various activities and areas. Another important difference is that analysis of the World Management Survey does not seek to arrange capabilities in a hierarchy (but instead calculate an aggregate sum as an indicator of management capabilities, implicitly giving equal weight to each dimension of management capability). In contrast, our analysis yields is a natural ordering of firms’ activities and capabilities into a hierarchy, i.e. a map of which activities generally appear to be precursors or prerequisites for others. We then investigate how this natural ordering relates to actual firms’ performance in terms of growth and survival.

2.2 Empirical methodology

The biology literature introduced the concept of “nestedness” to explain that in many systems the capabilities required to deal with specific advanced ecological niches were nested in those required for more basic ones. Similarly, in economic systems, those countries that were able to export advanced products were also exporting basic ones (Bustos et al., 2012). This intuition can be formalized through algorithms inspired by network science (Hidalgo and Hausmann, 2009; Tacchella et al., 2012) to provide a natural ranking of countries and products. Domínguez-García and Munoz (2015) shows that Tacchella et al. (2012)’s algorithm works best for recovering the rankings in nested biological systems.

Here, the same algorithm analyzes activities of around 40’000 Indian firms, to explore patterns in the following types of areas: Is a firm investing in R&D? Applying for patents? Using external know-how? Exporting? Each activity is the fingerprint of advanced production and technological capabilities, and the algorithm uses this data to simultaneously measure both the rankings of the

most advanced activities, i.e. the activities that are not accessible to firms that are not high enough in the capability ladder, and the position of the firms in the ladder. This allows us to give a quantitative definition to concepts that were previously either described qualitatively or measured by ad hoc case studies.

To give an illustration: research has found that innovative firms are more dependent on external finance (Brown et al., 2017). How are these two capabilities (innovation and seeking external finance) arranged in the capabilities ladder? A simple approach would involve comparing the conditional probabilities of innovating if a firm has access to external finance, and vice versa. The algorithm we will use instead does not look at the two capabilities in a vacuum, but takes an eco-systemic approach by looking at all the activities of all firms to evaluate the relative positioning of the two activities. At the same time it defines the ladder – i.e. the ranking of capabilities – and the position on the ladder of the firms, by defining a firm as high up the ladder if it is able to perform complex activities, and defining a capability as complex if those firms that are at the lower rungs of the ladder are not observed to perform it. The algorithm is validated against a null model in which firms’ activities are reshuffled between firms to show that nestedness is indeed a key force shaping firms’ behavior. More details on the algorithm are reported in Appendix C.

3 Data

Some influential empirical investigations of capabilities have previously focused on Indian firms (Ethiraj et al., 2005; Bloom et al., 2013). A drawback of these two previous studies, however, is that they focused on small sample sizes. We use PROWESS, from the Centre for Monitoring the Indian Economy (CMIE). Data comes from the annual reports of companies (both public-listed and unlisted), which include balance sheets and profit and loss accounts. PROWESS is the most comprehensive database available for Indian firms. The companies covered account for about 70% of industrial output, 75% of corporate taxes, and over 95% of excise taxes. We also merge with Patstat for patent data.

The database we use has 39,992 firms in the period 1988 to 2015, for a total of 348,422 observations. In Table 1 we show the general characteristics of the firms in our sample. Sales demonstrates considerable skew, and therefore we take the natural logarithm of sales as a measure of size, to avoid outliers in the regressions.

Table 1: Mean, Standard Deviation, Minimum, and Maximum of the main descriptive variables in our sample.

Variable	Mean	St.Dev.	Min	Max
Sales (million rupees)	3826	52569	0	5034766
Age	23.1	18.4	0	60
Growth Rate	+11.1%	86.1%	-1071%	+1258%

Firm activities reflect whether a firm has a certain capability (e.g. investing in R&D; exporting; obtaining access to external finance; owning machinery). The underlying assumption is that a firm would not invest in a certain activity (e.g. investing in R&D) if it did not seriously consider itself to possess the required capabilities (e.g. skilled researchers, sufficient knowledge base, and innovation management capabilities required for investing in R&D).

A binary variable indicates whether or not the firm is investing in any particular area. The amount of spending is therefore considered to be less important than whether or not a firm actually spends anything in this area (following on from e.g. Crépon et al., 1998; Bartz-Zuccala et al., 2018; Coad et al., 2020), because what matters is the binary variable of reaching a “rung” on the capabilities ladder, not the scale of activity on that particular rung. In this way, small firms that spend less on certain activities (e.g. R&D) receive the same ‘score’ regarding their capability in this area as large firms that spend larger amounts on the same activities, because both firms appear to possess the required capability.

We divide 47 firm activities (signalling underlying capabilities) into 7 groups: 1) Production Capabilities; 2) Technological Capabilities; 3) Innovation; 4) Basic Managerial & Organizational Capabilities; 5) Internationalization; 6) Environment and Social Amenities; 7) I.T. and Communication.² In Appendix B we describe the activities used in our analysis. This binary dataset is the starting point of our analysis, represented in Figure 1 in matrix form for one year (2012): each row corresponds to a firm and each column corresponds to an activity.

²Previous research into capabilities has distinguish between client-specific capabilities and project management capabilities (Ethiraj et al., 2005), or between technological and R&D capabilities, and customer and marketing capabilities (Danneels, 2016).

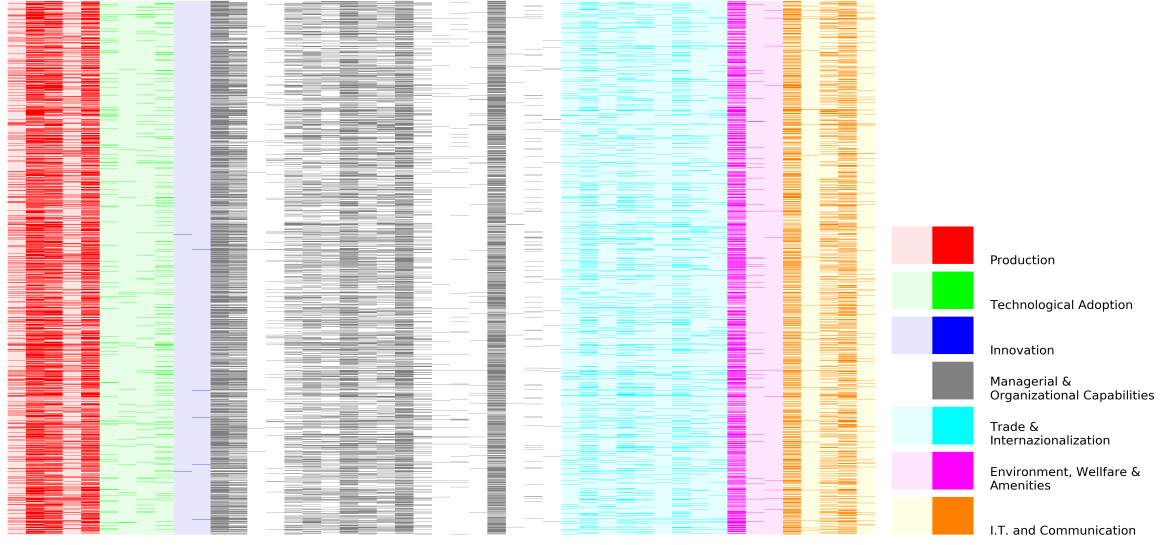


Figure 1: Matrix of activities by firm for year 2012. 15797 firms active in that year in our database are represented by rows and the 47 firms' activities are represented by columns.

4 Results

4.1 The capability ladder

Having constructed the matrix of activities in Figure 1, our algorithm then checks for nestedness, i.e. whether activities can be ranked to highlight that firms doing complex activities are also likely to do simpler ones (Nestedness and the ranking algorithm are presented in Appendices C and D.) Our results matrix clearly displays nested behaviour (see Figure 2). Indeed, it is possible to rearrange rows and columns to cluster the present and absent values much better than in a random matrix with similar characteristics. This means that it is meaningful to sort activities into a hierarchy. We observe that, as a first step, firms need to have basic managerial capabilities and production capabilities. Internet access and basic communication needs arise at about the same time. At a second step, firms engage in internationalization and gaining absorptive capacity for future innovation. At a third step, firms have more idiosyncratic patterns of engaging in product innovation or patenting, merging, environmental policies, and so on.

The optimal rearranging of columns and rows can be also considered as a natural ranking of activities and firms. In the next section we will look at the ranking of firms, and its significance in

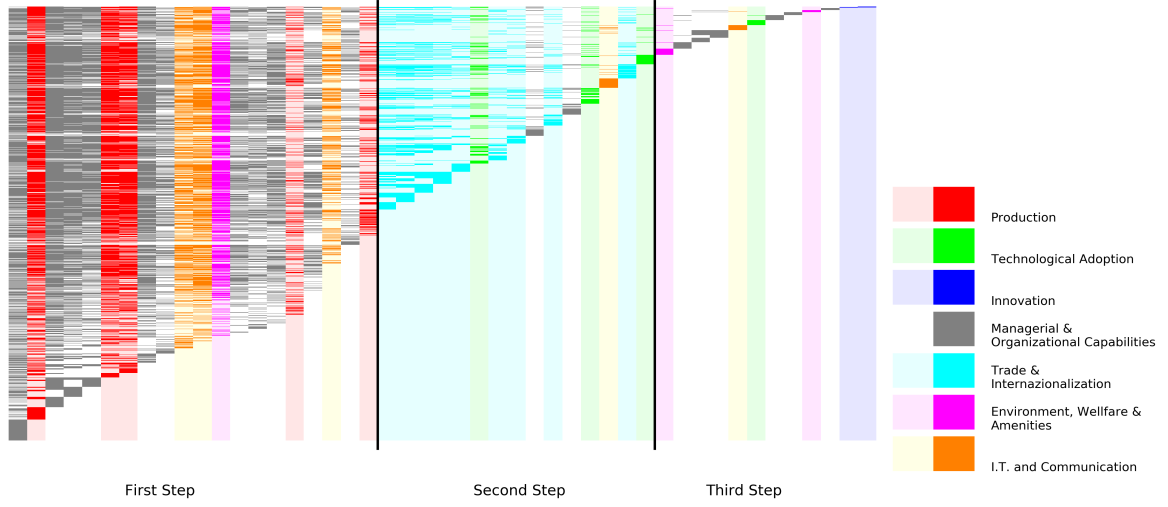


Figure 2: Matrix of activities by firm for year 2012. 15797 firms active in that year in our database are represented by rows and the 47 firms' activities are represented by columns. Same as Figure 1, but columns and rows are reordered to highlight nestedness.

measuring firms' capabilities. Focusing now on the 47 activities, we see that the first three rungs of the ladder are the ability to borrow, to own land & building assets, and to outsource professional employment. The three most advanced are product development, patenting, and external specialist consultants.

Figure 3 shows the dynamics of the hierarchy of activities over the available time window (2001-2014). As expected, this hierarchy of activities is quite stable over this period. The main exception is ICT related capabilities. While ICT capabilities were the footprint of an advanced organization in the early 2000s, they later became a standardized and necessary step for climbing the capabilities ladder. In appendix A we show the evolution of the single activities during the time period of our analysis.

4.2 Populating the ladder

In the previous section, we defined and described the ladder as it emerged from the algorithm. This process defined the landscape where firms coevolve. We will now describe the firms present at different levels of the capabilities ladder. In Figure 4 we show the same matrix as Figure 2, however

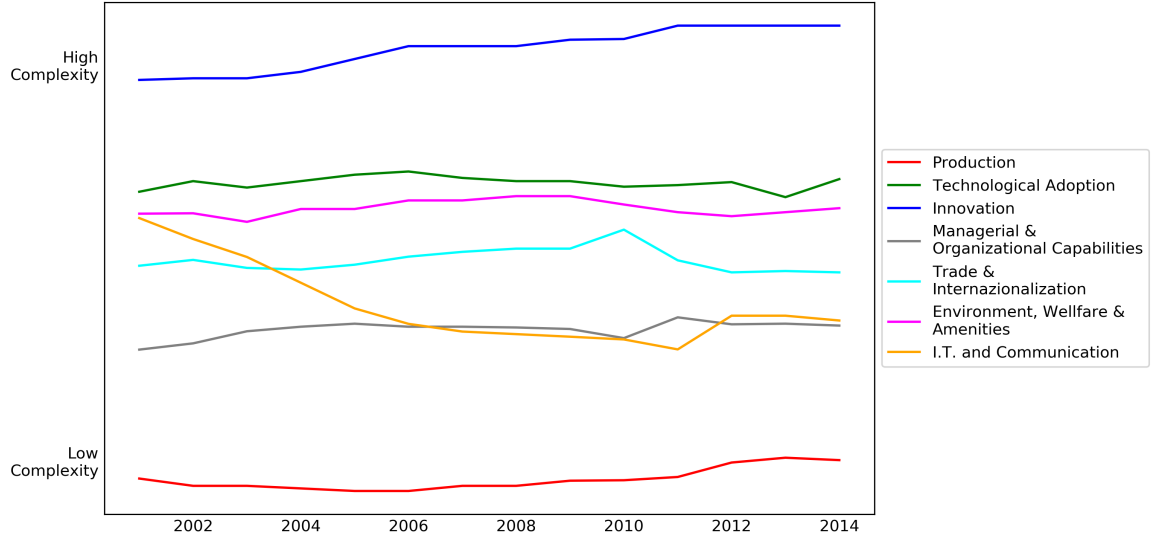


Figure 3: Dynamics of the hierarchy of activities over the period 2001-2014.

we now highlight the industry of the firm (different rows) instead of focusing on the activities (different columns).

Figure 4 shows several descriptive aspects of the Indian firms' ecosystem and how they are positioned on the ladder. First, it shows that, at each rung, there is a mix of firms of different sectors. However, this mix is not random: firms in the financial and real estate (K-L), trade (G-I), professional scientific and technical sectors (M), as well as energy (D-F) and other services (N-S) tend to be present along all the spectrum of capabilities, while ICT service sectors and manufacturing firms are mostly located in the two highest steps. This is shown by the average position of firms in the global ranking (emerging in Figure 4) by industry, as plotted in Figure 5.

Figure 5 clearly shows that the industrial sectors are not equally populated, or homogeneous. One of the biggest (more than 6000 firms out of 16000 in 2006) and most heterogeneous sectors is the manufacturing sector. We further disaggregate it using Pavitt's taxonomy (Pavitt, 1984) to split the manufacturing firms into four main taxa, as reported in Figure 6. We see that while firms in supplier dominated sectors do not have average capabilities higher than a service-sector firm, taxa characterized by more reliance on technological capabilities (Specialized Suppliers and Science Based sectors) show a significantly higher position in the ranking, similar to ICT service firms.

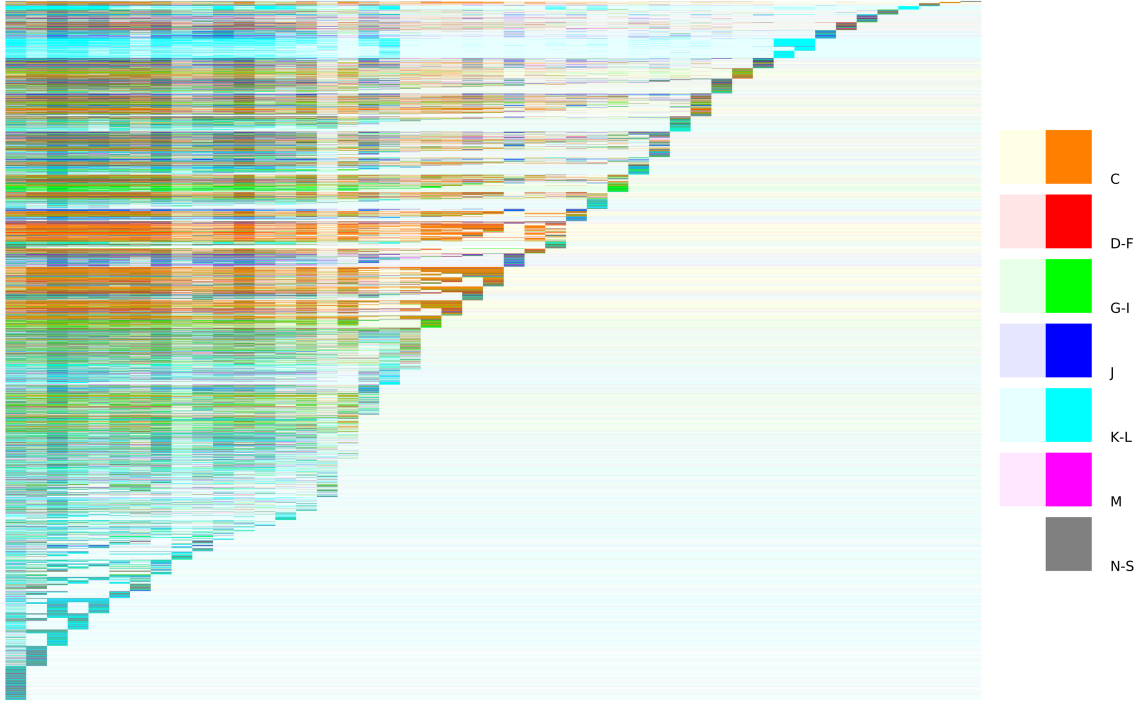


Figure 4: Matrix of firms and activities with columns and rows ordered to highlight nestedness. Rows are highlighted to show different industrial sectors: C- Manufacturing; D-F: Energy, Water, Construction; G-I: Wholesale and retail trade, deposit and distribution, food and accommodation; J- Information and Communication; K-L- Financial and real estate; M- Professional, Scientific and Technical Activities; N-S: Other services (Administration, Health, Education, Other).

4.3 Climbing the ladder

The firm growth process involves enlargement in quantitative terms (regarding indicators such as revenues, employees, assets, profits) as well as episodic reorganization and restructuring events (linked to the strategic development of capabilities) that are more qualitative in nature. Figure 7 shows that these two processes (firm size and capabilities) are indeed correlated. Firms increase in size while climbing the capabilities ladder in a coevolving process between capability accumulation and firm size. Interestingly, however, the positive correlation fades in the upper half of the distribution, roughly corresponding to firms that passed the first step highlighted in section 4.1. While a certain size is necessary to support advanced capabilities, the relationship disappears above $\log(\text{sales}) \approx 6$, equivalent to sales of about 40 crore rupees, or approximately 5 million US\$. Overall, the relationship arises from two different dynamics: the growth dynamics and the survival probability.

To better understand the dynamic coevolution of sales and capabilities, we now look at the growth

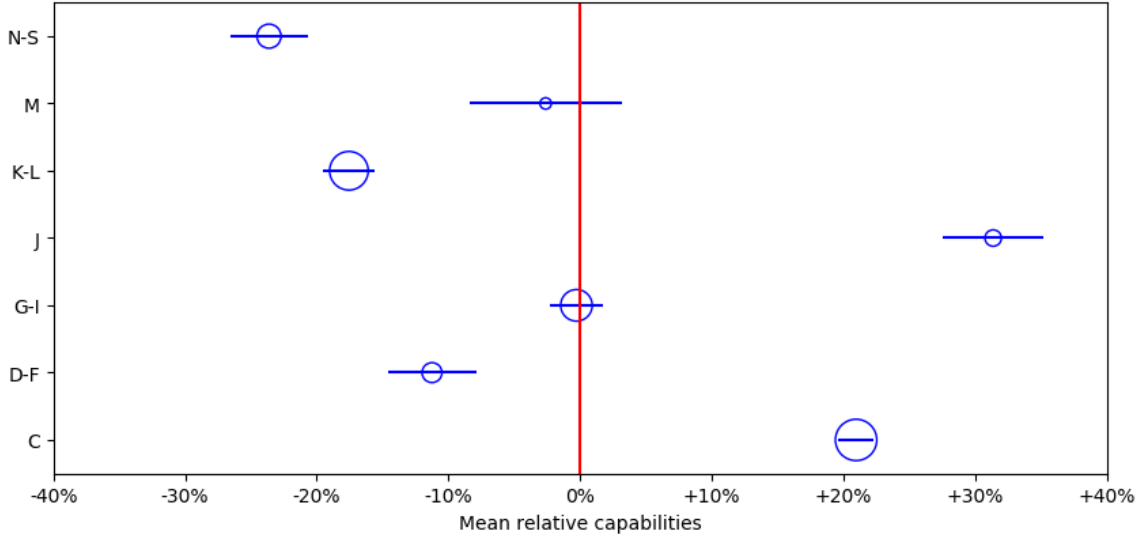


Figure 5: Relative average position in the capabilities ladder with respect to the average of all Indian firms of the firms of a specific industry. Year 2012. Error bars represents 2 standard errors – i.e. $\approx 95\%$ confidence interval. Size of the circle is proportional to the number of firms in the sector. Industries: C- Manufacturing; D-F: Energy, Water, Construction; G-I: Wholesale and retail trade, deposit and distribution, food and accommodation; J- Information and Communication; K-L- Financial and real estate; M- Professional, Scientific and Technical Activities; N-S: Other services (Administration, Health, Education, Other).

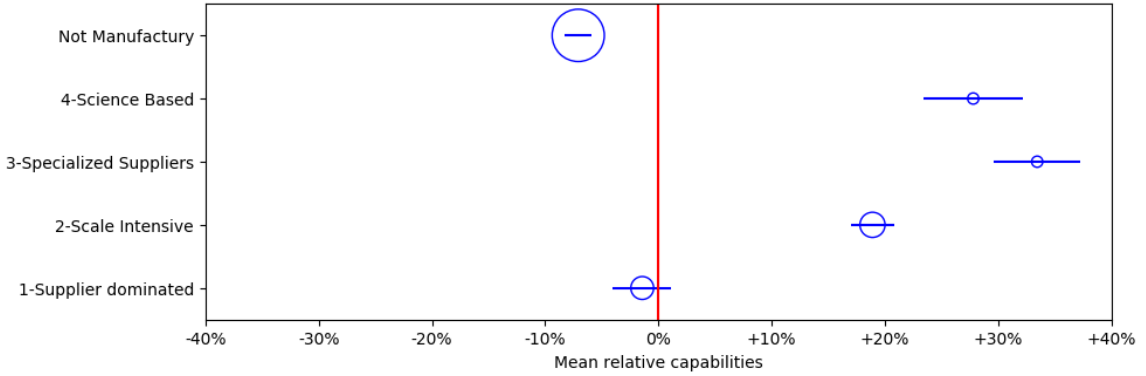


Figure 6: Relative average position in the capabilities ladder of manufacturing firms of a specific Pavitt taxum with respect to the average of all Indian firms. Year 2012. Error bars represents 2 standard errors – i.e. $\approx 95\%$ confidence interval. Size of the circle is proportional to the number of firms in the sector.

and survival of firms depending on their capabilities and sales. In Figures 8 and 9 respectively, we look at the expected growth rates (relative to the growth rate of that size class) and survival probability of firms after 7 years³ depending on their initial size and capability class. Firms are

³The general pattern for different delta years is similar. The same figures for different delta years are provided

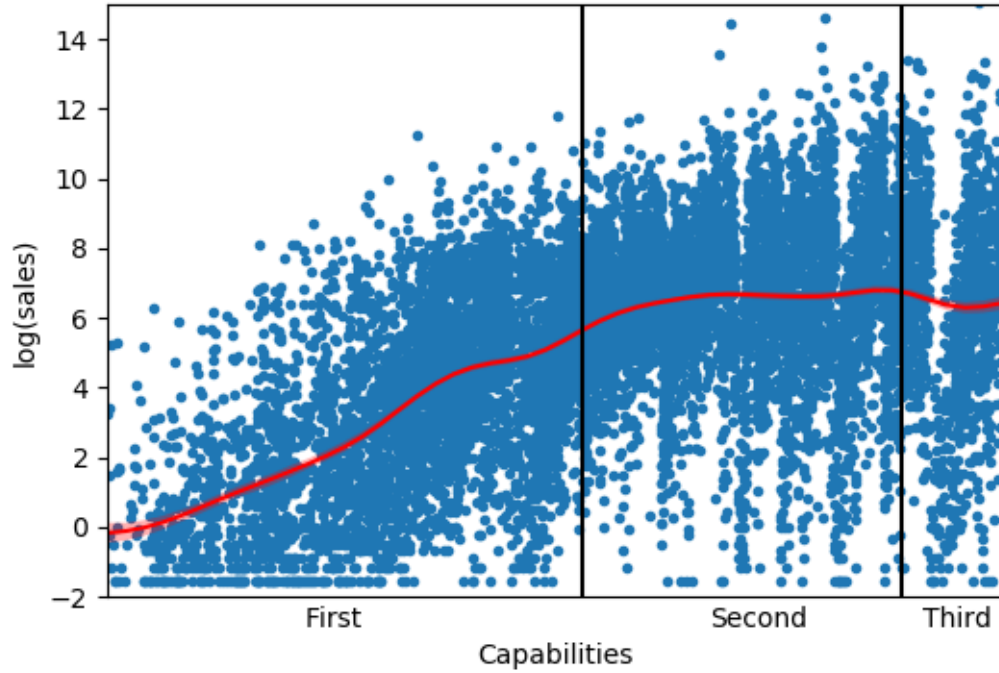


Figure 7: Size of the firms (in terms of the natural logarithm of sales) and its (inverse) position in the ranking, its ‘capabilities’. The two vertical black lines separates the three steps. The red line is a non-parametric Nadaraya-Watson kernel regression.

divided by size class in four quartiles (approximately less than 5.6 crores rupees, between 5.6 and 23.7 crores rupees, between 23.7 and 81.8 crores rupees, and over 81.8 crores rupees), and in terms of the capability ladder we split the firms in the first step (the most populated step) into two subclasses, to get a total of four capability classes. Once we observe the two processes separately, we better understand the dynamics leading to Figure 7. First, the datapoints in Figure 8 shows a general trend from bottom-left to top-right, indicating that, for firms in every size class, higher capabilities are generally associated with higher average growth. This positive relationship between capabilities and growth seems stronger after the first step. Second, by looking at the survival rate (Figure 9), capability development improves survival outcomes from step 1A to 1B, although subsequent development of capabilities has no clear survival benefits for firms. A possible explanation is that higher positions in the capability ladder correspond to entering into more risky activities (exporting, patenting) and sustaining higher fixed costs.⁴ This is reminiscent of Arora and Nandkumar (2011),

upon request by the authors.

⁴This result contrasts with Bloom et al. (2019, Figure 2, Panel C, p.1659), who observe higher survival rates for firms

who observe that capable entrepreneurs with high opportunity costs will take more risks and place less value on the option of survival. Another possible explanation is that some of these cases of non-survival correspond to lucrative exits via acquisition or trade sale. While firms in ‘Size class 0’ have maximum survival probability at step I.B, climbing up the ladder of production and managerial capabilities without engaging in R&D and exporting activities, firms in higher size classes have their maximum survival probability while remaining at step II of the ladder, avoiding only the riskiest activities (patenting, introduction of new products, etc).

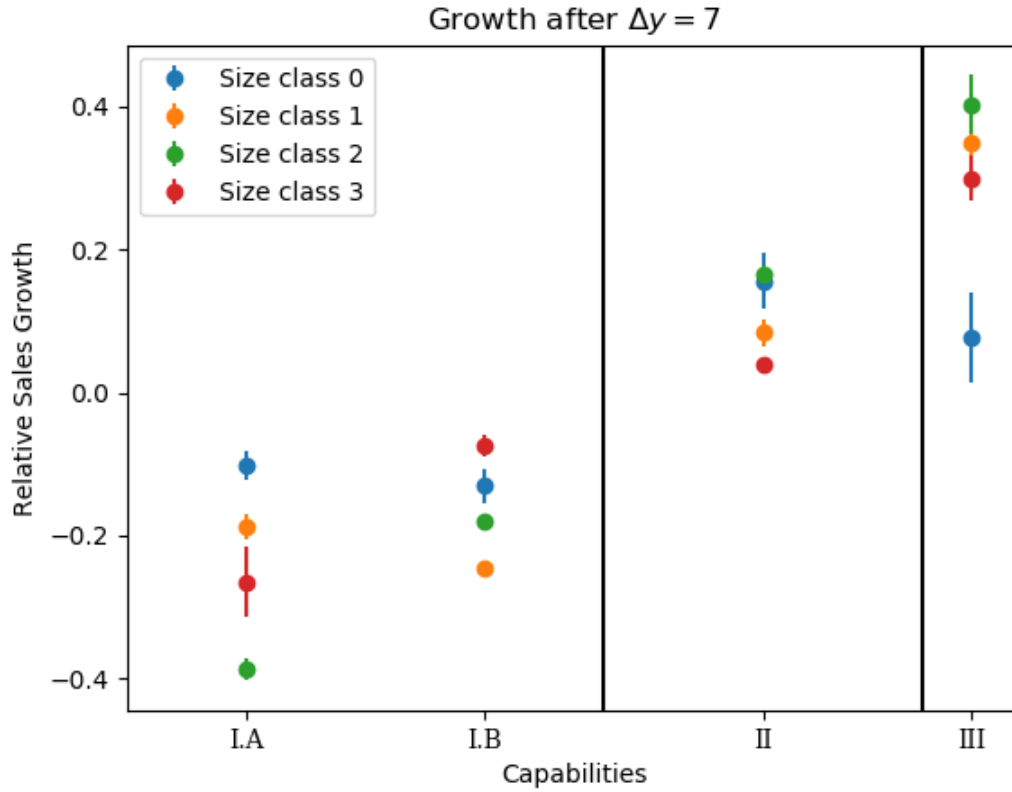


Figure 8: Average 7 years sales growth relative to the size class for different size and capability classes. Bars represents two standard errors, $\approx 95\%$ confidence interval.

Assuming a linear behaviour for the relationship between growth and capabilities (as suggested by Figure 8) we can add further controls to better understand the growth process by using parametric with higher-scoring management capabilities. This is presumably because of the differences in the operationalizations of ‘capabilities’ across the two studies, given that the capabilities indicators in Bloom et al. (2019) correspond to relatively low-risk activities such as frequency of tracking performance indicators and procedures for setting production targets. This underscores the distinct and complementary contribution of our approach.

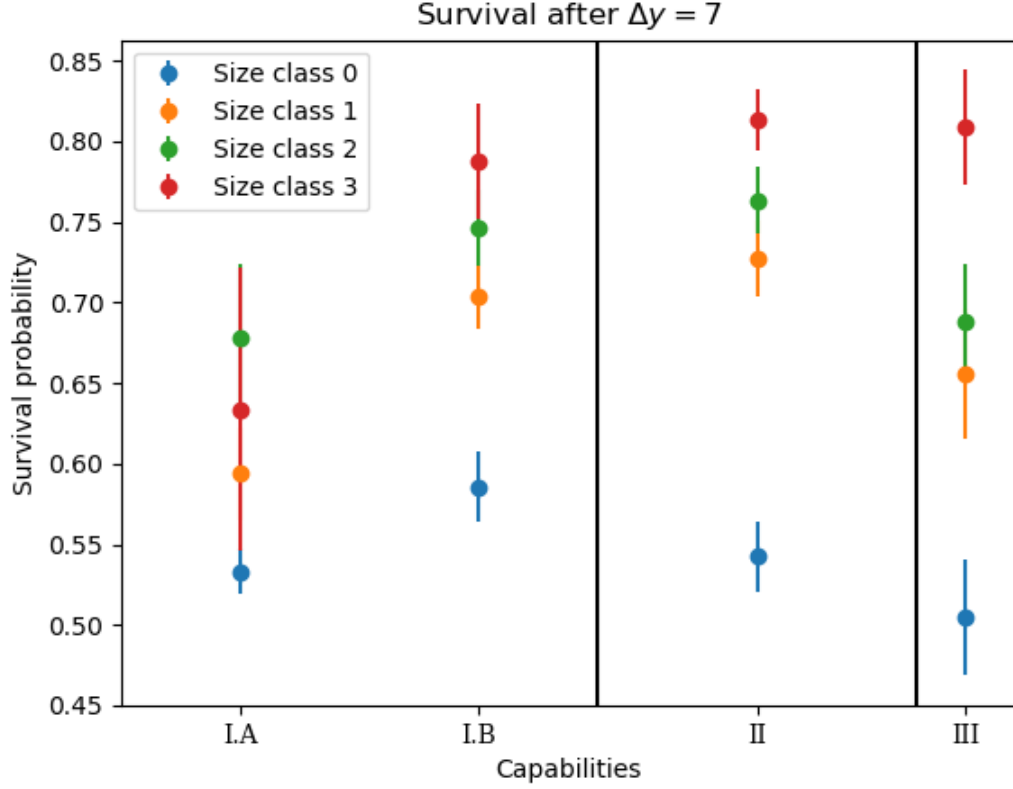


Figure 9: Survival probability after 7 years for different size and capability classes. Bars represents two standard errors, $\approx 95\%$ confidence interval.

regressions. The relationship between firm size and capabilities is confirmed by looking at the future growth of the firms that are higher in capabilities than their size would indicate, i.e. firms that are below the red line in Figure 7. Table 2 shows the results of a Fixed Effects regression of firm growth for different sectors:

$$Growth_{i,t} = \alpha + \beta_1 Size_{i,t-1} + \beta_2 Age_{i,t-1} + \beta_3 Capabilities_{i,t-1} + v_i + \gamma_t + \epsilon_{i,t} \quad (1)$$

where firm growth is defined in terms of log-differences of sales, i.e. $Growth(i,t) = \log(sales(i,t)) - \log(sales(i,t-1))$ for firm i in year t . The controls include size (measured by log sales), log of age of the firm, year dummies (γ_t) and two-digit sector dummies (v_i). Interestingly, there is a clear threshold effect: the larger the firm is, the higher on the ladder it should be placed to be able to grow.

Table 2: Fixed Effect regression of firm growth with respect to lagged size, age and the position in the ladder of the firm ('capabilities') normalized from 0 (beginning of the ladder) to 1 (top of the ladder). Column 1 and 2 involve manufacturing firms (industry C), columns 3 and 4 involves ICT firms (industry J), while columns 5 and 6 involves all firms in the sample.

	(1)	(2)	(3)	(4)	(5)	(6)
Size	-0.2493*** (0.0070)	-0.2513*** (0.0072)	-0.3022*** (0.0171)	-0.3094*** (0.0174)	-0.2850*** (0.0048)	-0.2874*** (0.0048)
Age	-0.2847*** (0.0160)	-0.2969*** (0.0163)	-0.3079*** (0.0586)	-0.2934*** (0.0588)	-0.2931*** (0.0139)	-0.2971*** (0.0140)
Capabilities		0.1234*** (0.0188)		0.2640*** (0.0873)		0.1969*** (0.0171)
Observations	96793	92315	7776	7681	173782	167740
Firm clusters	11032	10583	1120	1112	23892	23192

Notes: Standard errors in parentheses.

Key to significance stars: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notice that this effect is specific for firm growth: other performance variables, like profitability, are not affected by the spread of firms in different activities (see Appendix B). Indeed, a lean operation not involved in R&D or internationalization could well have a higher profitability than a large corporation could. An explanation could be that the gains from capability development are not clearly visible in the 'bottom line' of final profits, because they are shared with stakeholders such as skill-upgrading employees. Increasing the operational complexity of the firm only seems necessary for increasing the scale of operations, not for boosting profitability or enhancing survival.

5 Concluding Discussion

We contribute to the strategic management literature by introducing and applying a novel methodology for quantifying firms in terms of their capabilities, and putting these capabilities (and hence also the firms) in a hierarchy that we refer to as the capabilities ladder. These operational routines correspond to a variety of functional areas, such as marketing, financial practices, human resource practices, internationalization, and innovation. Our nestedness algorithm, inspired by biology and network science, defines a capability as complex if few firms at lower rungs of the ladder are observed to perform it. We analyse balance sheet and innovation data of almost 40'000 Indian firms 1988-2015, and observe significant nestedness, as compared to a random allocation of activities in firms. Lower rungs of the capabilities ladder correspond to basic managerial and production capabilities. Mid-level rungs correspond to internationalization and acquiring absorptive capacity. Higher-level

rungs are less clearly ordered into a hierarchy. Higher-level rungs are more focused on one side with environmental and welfare, and the other side with innovation in terms of patenting and introducing new products. Further analysis showed some interesting changes in the ordering of the rungs over time. For example, ICT capabilities – once rare – have become more fundamental lower-level rungs on the capabilities ladder in recent years, as ICT capabilities have become more widely used and have permeated a wider range of processes.

Interestingly, we saw that firms with higher capabilities have higher growth rates, but when the firm is involved in too many advanced activities with respect to their size it lowers survival probability. As a first step, firms need to have basic managerial capabilities and production capabilities. Internet access and basic communication needs to arise at about the same time. At a second step, firms engage in internationalization and gaining absorptive capacity for future innovation. Subsequent rungs of the capabilities ladder are less precisely arranged in a hierarchy (i.e. nestedness is weaker), as firms with adequate capabilities have more latitude to climb their own customized ladders. For lower capability firms, though, the capability ladder is set out quite clearly. These combined dynamics generate an equilibrium distribution where capabilities highly relate with firm size for small firms with few capabilities, while for higher capability levels there is a different regime, likely depending on the firm sector, age, and some idiosyncratic firm’s characteristics, breaking the observed relationship between size and capabilities.

Some limitations of our work (and empirical work in this area more generally) can be mentioned. First, we use a dummy variable for capability, for reasons discussed in Section 2, which could be a limitation because this operationalization does not take into account how much a firm spends in a certain area, or how skilful the capability is (e.g. some violinists have better capabilities than others). Second, and relatedly, we take an inputs-based indicator of capabilities, rather than an output-based indicator. A problem could be that inefficient and optimistic firms may lack the capabilities to effectively create value from expenditures in these areas (e.g. to benefit from R&D investment).

Some policy implications can be derived from our analysis. First, a number of policy initiatives seek to provide support to firms facing particular challenges, such as taking first steps into export markets. Our results help to better understand the precursors and prerequisites for certain types of activities, which could make policy targeting more effective if firms become eligible for policy support conditional on their existing capabilities. For example, publicly-funded export assistance could be

made conditional on having a website; or eligibility for R&D grants could be made conditional on patent applications or having certain technological search routines or skilled employees. There may be drawbacks to making policy support conditional on capabilities, however. Grazzi et al. (2021) argue that some firms might accelerate their movement up the capabilities ladder (e.g. prematurely enter export markets) in order to become eligible for export promotion policies, which leads to a misallocation of policy support towards ill-prepared firms. Another possible drawback of making policy support conditional on existing capabilities (according to our methodology) would be that firms could learn how to “game the system” and imitate high-capability firms by spending in areas where they have no capabilities, in an attempt to appear sophisticated.

Another thought is that some advanced capabilities (e.g. internationalization, R&D investment) are possessed by firms from birth (i.e. the case of born globals). Such cases are indeed impressive and seem worthy of particular research focus. For example, it could be interesting to investigate how severely firms are penalized (if at all) for moving up the rungs in an order that is different from the commonly-observed order of rungs.

To summarize, therefore, we present a novel methodology for the measurement and sorting of capabilities within firms, and hope that our methodology and results will help to improve our ability to measure capabilities as well as helping to refine the theory of capabilities and capability development.

References

- ALMEIDA-NETO, M., P. GUIMARAES, P. R. GUIMARAES JR, R. D. LOYOLA, AND W. ULRICH (2008): “A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement,” *Oikos*, 117, 1227–1239.
- AMIT, R. AND P. J. SCHOEMAKER (1993): “Strategic Assets and Organizational Rent,” *Strategic Management Journal*, 14, 33–46.
- AREND, R. J. (2006): “Tests of the resource-based view: do the empirics have any clothes?” *Strategic Organization*, 4, 409–421.
- ARORA, A. AND A. NANDKUMAR (2011): “Cash-out or flameout! Opportunity cost and entrepreneurial strategy: Theory, and evidence from the information security industry,” *Management Science*, 57, 1844–1860.

- BARNEY, J. (1991): “Firm resources and sustained competitive advantage,” *Journal of management*, 17, 99–120.
- BARTZ-ZUCCALA, W., P. MOHNEN, AND H. SCHWEIGER (2018): “The role of innovation and management practices in determining firm productivity,” *Comparative Economic Studies*, 60, 502–530.
- BLOOM, N., E. BRYNJOLFSSON, L. FOSTER, R. JARMIN, M. PATNAIK, I. SAPORTA-EKSTEN, AND J. VAN REENEN (2019): “What Drives Differences in Management Practices?” *American Economic Review*.
- BLOOM, N., B. EIFERT, A. MAHAJAN, D. MCKENZIE, AND J. ROBERTS (2013): “Does management matter? Evidence from India,” *The Quarterly Journal of Economics*, 128, 1–51.
- BLOOM, N., C. GENAKOS, R. SADUN, AND J. VAN REENEN (2012): “Management practices across firms and countries,” *Academy of management perspectives*, 26, 12–33.
- BLOOM, N. AND J. VAN REENEN (2010): “Why do management practices differ across firms and countries?” *Journal of economic perspectives*, 24, 203–24.
- BROWN, J. R., G. MARTINSSON, AND B. C. PETERSEN (2017): “What promotes R&D? Comparative evidence from around the world,” *Research Policy*, 46, 447–462.
- BRUNO, M., F. SARACCO, D. GARLASCHELLI, C. J. TESSONE, AND G. CALDARELLI (2020): “The ambiguity of nestedness under soft and hard constraints,” *Scientific reports*, 10, 1–13.
- BUSTOS, S., C. GOMEZ, R. HAUSMANN, AND C. A. HIDALGO (2012): “The dynamics of nestedness predicts the evolution of industrial ecosystems,” *PloS one*, 7, e49393.
- CIRERA, X. AND W. F. MALONEY (2017): *The innovation paradox: Developing-country capabilities and the unrealized promise of technological catch-up*, The World Bank.
- COAD, A., N. MATHEW, AND E. PUGLIESE (2020): “What’s good for the goose ain’t good for the gander: heterogeneous innovation capabilities and the performance effects of R&D,” *Industrial and Corporate Change*, 29, 621–644.
- COSTA, S., S. DE SANTIS, G. DOSI, R. MONDUCCI, A. SBARDELLA, AND M. E. VIRGILLITO (2021): “From organizational capabilities to corporate performances: at the roots of productivity slowdown,” Tech. rep., Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies. LEM working paper 2021/21.

- CRÉPON, B., E. DUGUET, AND J. MAIRESSE (1998): “Research, Innovation And Productivity: An Econometric Analysis At The Firm Level,” *Economics of Innovation and new Technology*, 7, 115–158.
- DANNEELS, E. (2016): “Survey measures of first-and second-order competences,” *Strategic Management Journal*, 37, 2174–2188.
- DOMÍNGUEZ-GARCÍA, V. AND M. A. MUNOZ (2015): “Ranking species in mutualistic networks,” *Scientific reports*, 5, 1–7.
- DOSI, G., R. R. NELSON, AND S. WINTER, eds. (2000): *The nature and dynamics of organizational capabilities*, Oxford University Press: Oxford.
- ETHIRAJ, S. K., P. KALE, M. S. KRISHNAN, AND J. V. SINGH (2005): “Where do capabilities come from and how do they matter? A study in the software services industry,” *Strategic Management Journal*, 26, 25–45.
- GRANT, R. M. AND G. VERONA (2015): “What’s holding back empirical research into organizational capabilities? Remedies for common problems,” *Strategic Organization*, 13, 61–74.
- GRAZZI, M., N. MATHEW, AND D. MOSCHELLA (2021): “Making one’s own way: jumping ahead in the capability space and exporting among Indian firms,” *Journal of Evolutionary Economics*, 1–27.
- HELFAT, C. E. (1997): “Know-how and asset complementarity and dynamic capability accumulation: the case of R&D,” *Strategic Management Journal*, 18, 339–360.
- HELFAT, C. E. AND M. A. PETERAF (2003): “The dynamic resource-based view: Capability lifecycles,” *Strategic Management Journal*, 24, 997–1010.
- HELFAT, C. E. AND S. G. WINTER (2011): “Untangling dynamic and operational capabilities: Strategy for the (N)ever-changing world,” *Strategic Management Journal*, 32, 1243–1250.
- HIDALGO, C. A. AND R. HAUSMANN (2009): “The building blocks of economic complexity,” *Proceedings of the national academy of sciences*, 106, 10570–10575.
- HILLIARD, R. AND D. GOLDSTEIN (2019): “Identifying and measuring dynamic capability using search routines,” *Strategic Organization*, 17, 210–240.
- PANDZA, K. (2011): “Why and how will a group act autonomously to make an impact on the development of organizational capabilities?” *Journal of Management Studies*, 48, 1015–1043.
- PENROSE, E. (1959): *The Theory of the Growth of the Firm*, Oxford University Press, 3rd ed.

- ROTHAERMEL, F. T. AND A. M. HESS (2007): “Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects,” *Organization science*, 18, 898–921.
- SADUN, R., N. BLOOM, AND J. VAN REENEN (2017): “Why Do We Undervalue Competent Management? Great Leadership and Brilliant Strategy Won’t Succeed Without Operational Excellence,” *Harvard Business Review*, 95, 120–127.
- SARACCO, F., R. DI CLEMENTE, A. GABRIELLI, AND T. SQUARTINI (2015): “Randomizing bipartite networks: the case of the World Trade Web,” *Scientific reports*, 5, 1–18.
- TACCHELLA, A., M. CRISTELLI, G. CALDARELLI, A. GABRIELLI, AND L. PIETRONERO (2012): “A new metrics for countries’ fitness and products’ complexity,” *Scientific reports*, 2, 1–7.
- TEECE, D. J., G. PISANO, AND A. SHUEN (1997): “Dynamic Capabilities and Strategic Management,” *Strategic Management Journal*, 18, 509–533.

A Ranking of the capability variables in time

In this appendix, we list the individual firm activities used in the analysis (Table A.1) and we show their individual evolution in time in terms of complexity (Figure A.1).

Table A.1: Definition of variables

Variable name	Set	Definition
Dpatents	Innovation	Patents
Dsa_prod_dev_amort	Innovation	Product Development Expenses
Dsa_consult_fees_auditors	Managerial & Organizational	Expenses: Consultancy fees to auditors
Dno_branches	Environment, Welfare & Amenities	No of branches
Dno_segments	Managerial & Organizational	No of segments or business units
Dsa_lt_loan_advance_nbfcs	Managerial & Organizational	Long term loans and advances
Dsa_st_loan_advance_nbfcs	Managerial & Organizational	Short term loans and advances
Dsa_tech_knowhow	Technological Adoption	Technical knowhow fees paid to external parties
Dsa_internet_serv	I.T. and Communication	Expenses on internet services
Dmergers	Managerial & Organizational	Merged with another company
Dsa_environment_related	Environment, Welfare & Amenities	Environment and pollution control related expenses
Dsa_leased_in_ast	Managerial & Organizational	Leased in assets
Dsa_license_fees	Technological Adoption	License fees (for use of intellectual property) paid
Dsa_comm_equip	I.T. and Communication	Assets – communication equipment
Dshareholder_outsideIn	Internationalization	Shareholders outside India
Dsa_ITES	Managerial & Organizational	Expenses – Outsourced IT enabled services and other professional jobs
Dsa_staff_training	Technological Adoption	Staff Training Expenses
Dsa_import_fg	Internationalization	Import of finished goods
Dsa_rnd	Technological Adoption	R&D expenses
Dsa_lt_invest_abroad	Internationalization	Long-term investment abroad
Dsa_export_serv	Internationalization	Export of Services
Dforeignborrowings	Internationalization	Foreign currency borrowings
Dacquisition	Managerial & Organizational	Acquirer
Dsa_import_stores_spares	Internationalization	Import of stores and spares
Dsa_import_cap_goods	Internationalization	Import of capital goods
Dsa_import_rawmat	Internationalization	Import of raw materials
Dsa_export_goods	Internationalization	Export of goods
Dsa_outsourced_mfg_jobs	Production Capabilities	Expenses: Outsourced manufacturing jobs
Dsa_total_provisions	Managerial & Organizational	Provisions for i) bad and doubtful advances, ii) debts & debtors, diminution in investments and iii) unspecified contingencies
Dsa_distribution_exp	Managerial & Organizational	Distribution expenses
Dsa_elec_install_fitting	Production Capabilities	Assets – Electrical installations
Dsa_software	I.T. and Communication	Assets - Software
Dsa_marketing	Managerial & Organizational	Expenses on market research, branding, product designing for targeted markets
Dsa_printing_stationery	Managerial & Organizational	Printing and Stationery Expenses
Dsa_staff_welfare	Environment, Welfare & Amenities	Expenses on staff welfare
Dsa_advertising	Managerial & Organizational	Advertising expenses (Media campaign etc)
Dsa_communications	I.T. and Communication	Communication expenses (telephone, postage etc)
Dsa_net_computer_it	I.T. and Communication	Assets - Computers and IT systems
Dsa_fee_based_fin_serv_exp	Managerial & Organizational	Fee based financial services expenses (Bank charges etc)
Dsa_travel_exp	Managerial & Organizational	Travel expenses
Dsa_transport_vehicles	Production Capabilities	Assets – transport vehicles
Dsa_plant	Production Capabilities	Assets – plant and machinery
Dsa_fund_based_fin_serv_exp	Managerial & Organizational	Fund based financial services expenses (Interest paid on loans)
Dsa_auditor_fees	Managerial & Organizational	Fees paid to auditor associated with the company (internal)
Dsa_outsourced_professional_job	Managerial & Organizational	Expenses: Outsourced professional jobs
Dsa_land_building	Production Capabilities	Assets – land and building
Dborrowings	Managerial & Organizational	Borrowings

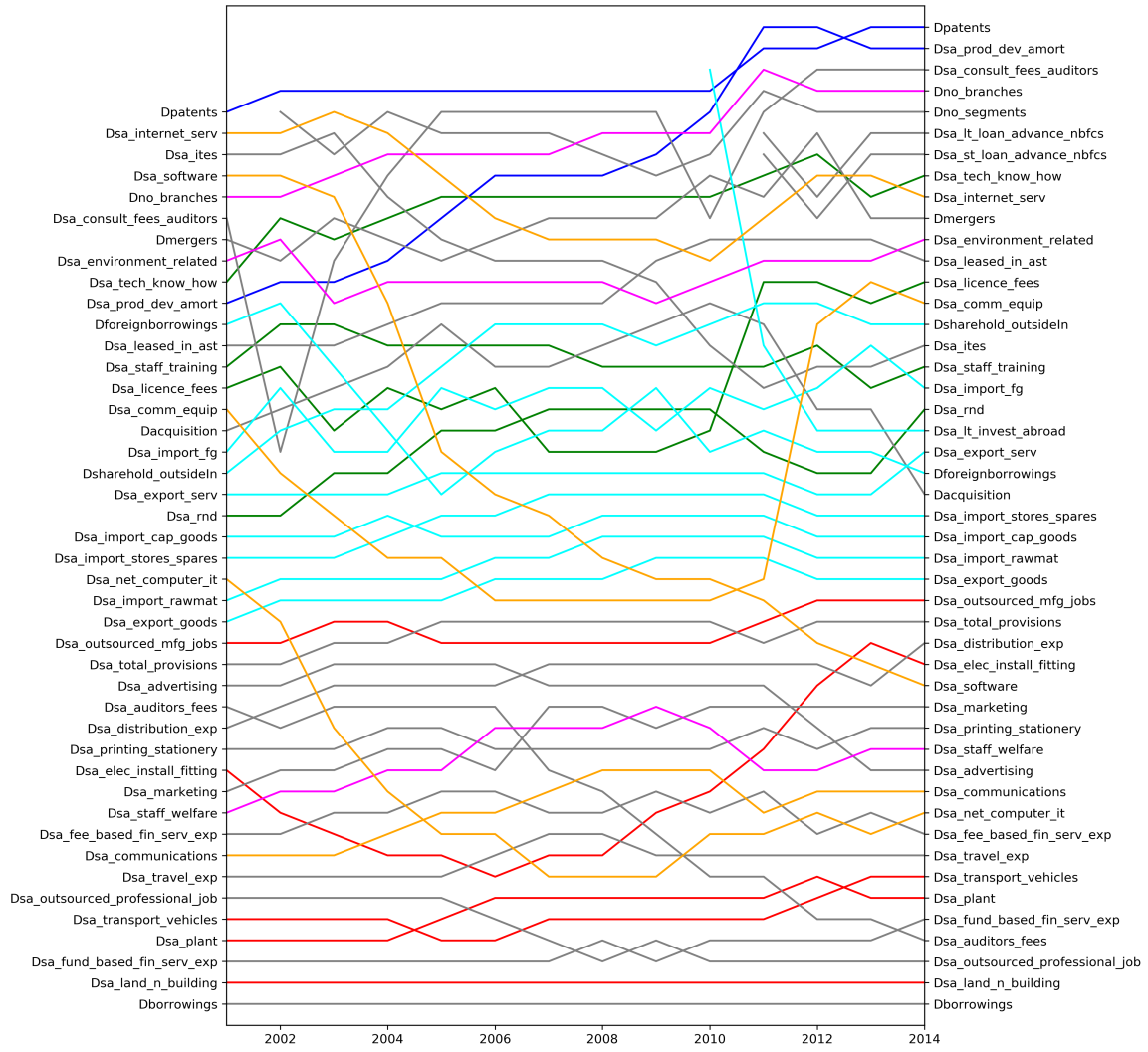


Figure A.1: Dynamics of the hierarchy of activities over the period 2001-2014.

B Relative Profitability

In this appendix, we show the results of running a regression looking at the relative profitability of the firm compared to the other firms in its 4-digits sector.

	(1)	(2)
Size	0.0021 (0.0012)	0.0020 (0.0013)
Age	-0.0241*** (0.0045)	-0.219*** (0.0045)
Capabilities		-0.0033 (0.0038)
Observations	27411	26104
Firm clusters	4289	4094

Table B.1: Capabilities and firm profitability: Fixed effects estimation

C The ranking algorithm

In this type of analysis, the basic input is the matrix $M_{f,c}$ for a given year, whose entries take the value 1 if the firm f is active in the activity revealing the capability c and 0 otherwise. For example, the input matrix $M_{f,c}$ in our case for year 2012 is shown in Figure 1. From this starting input we try to infer at the same time the complexity of the activities and the “fitness” of the firms, i.e. their ability to perform complex activities. To task we use the Tacchella et al. (2012) algorithm, as it has been shown to be the best algorithm to highlight nestedness in biological systems (Domínguez-García and Munoz, 2015).

In particular, to calculate the fitness of firms and the complexity of the activities we iterate upon convergence the following set of non linear coupled equations:

$$\left\{ \begin{array}{l} \tilde{F}_f^{(n)} = \sum_c M_{fc} Q_c^{(n-1)} \\ \tilde{Q}_c^{(n)} = \frac{1}{\sum_f M_{fc} \frac{1}{F_f^{(n-1)}}} \end{array} \right\} \left\{ \begin{array}{l} F_f^{(n)} = \frac{\tilde{F}_f^{(n)}}{\langle \tilde{F}_f^{(n)} \rangle} \\ Q_c^{(n)} = \frac{\tilde{Q}_c^{(n)}}{\langle \tilde{Q}_c^{(n)} \rangle} \end{array} \right. \quad (C.1)$$

where the normalization of the intermediate tilded variables is made as a second step and n is the iteration index. Equations C.1 can be explained as following. F_f , the fitness of firm f , is high if the firm is able to perform complex activities. The complexity of activity c , Q_c , is high if those firms that *do not have* high fitness are *not able* to perform it. Simpler linear algorithms (for example Hidalgo and Hausmann, 2009) would see an activity as complex if advanced firms perform it. However, this is clearly not as informative: if one small firm lacking any other capability is able

to perform the task, this has clearly more weight on our assessment than the fact that also a big multinational is able to perform it, as those advanced firms are involved in both simple and advanced tasks.

The iteration of these coupled equations until a fixed point is reached allow us to rank firms and activities in a unique way. These rankings, as shown in Figure 2, allow reorganizing the matrix to highlight its nestedness. Indeed it has been shown in Domínguez-García and Munoz (2015) that equations C.1 are the best equations to find the ranking representing the importance of species in nested systems: once the species are ordered according to equations C.1, the nestedness of the system is the highest with respect to ordering of rows and columns based on different centrality measures. Equations C.1 will also be the rankings we use in our computations when referring to the complexity of capabilities (Q_c) or firms (F_f).

D Measuring nestedness

A crucial part of this work is assessing the nestedness of the matrix, i.e. the fact that firms able to perform activities classified as complex, are also able to perform simpler activities more than what we would expect if firms were selecting activities randomly. While this looks intuitively obvious, we first need to clarify two points: how are we going to measure the nestedness, and what “randomly” means in this context. First, as a measure of nestedness we use ‘Nestedness as a measure of Overlap and Decreasing Fill’ (NODF), (Almeida-Neto et al., 2008), a measure of the actual overlap between firms’ activities. The measure gives a number between 0 (no overlap) and 1 (perfectly triangular matrix). According to this measure, our matrix for 2012 has a NODF of **0.710**. To be able to compare this value with what we would expect from a random matrix, we have to clarify how a random assignation of activities to firms would occur.

The easiest process we can imagine is a process where each firm-activity link has the same probability. As the density of our matrix is 0.216 — firms perform on average 21.6% of the activities — if this is our null model each firm will have a probability 0.216 of being active in any given activity. The distribution of NODF for matrices produced by this null model is represented in Figure D.1. Clearly, the empirical NODF is much higher than the NODF measured in the matrices produced by this null model.

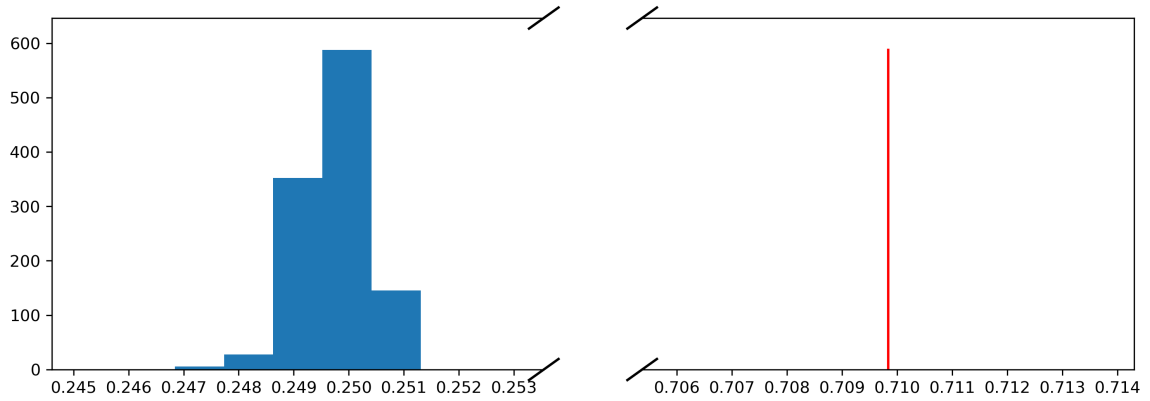


Figure D.1: Probability distribution of the NODF of matrices generated by a homogeneous null model (blue area), compared with the NODF of the empirical matrix (red line).

Of course, this does not seem particularly surprising: we know that there are firms more diversified in different activities, and that this is not the outcome of a random process but of different firm capabilities. We can therefore change our null model, and assume now that each firm has the

same expected diversification as in the empirical matrix, but with random selection regarding which specific activities it is doing. If a firm is empirically active in 50% of the activities, the corresponding random firm will have a 0.50 probability of being active in each activity. The distribution of NODF for matrices produced by this null model is represented in Figure D.2. Again, also this null model cannot produce matrices with a NODF as high as the empirical matrix.

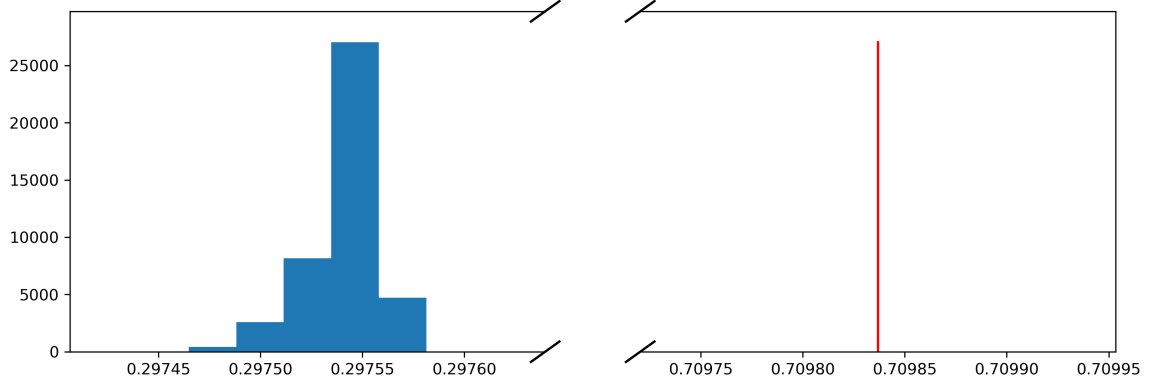


Figure D.2: Probability distribution of the NODF of matrices generated by a null model preserving heterogeneous expected firms diversification (blue area), compared with the NODF of the empirical matrix (red line).

There is a further objection: we know that activities are not all the same and some activities are less ubiquitous than others. This is not the outcome of a random process but of higher capabilities required. We can propose therefore one more null model, and assume now that each activity has the same expected ubiquity as in the empirical matrix, but which specific firms are doing that activity is randomly selected. If an activity is performed by 30% of the firms, each firm will have 0.30 probability of being active in that activity. The distribution of NODF for matrices produced by this null model is represented in Figure D.3. Also this null model cannot produce matrices with a NODF as high as the empirical matrix.

Finally, in the empirical matrix are present at the same time both the constraints tackled in the previous two paragraphs. Therefore, we want a null model representing random matrices such that both of the following are satisfied: each firm has its expected diversification and each activity has its expected ubiquity. To solve both constraints at the same time we need to solve a system of equations numerically with techniques inspired by the maximization of entropy (Saracco et al., 2015). The distribution of NODF for matrices produced by this last null model is represented in Figure D.4. Even this last null model cannot produce matrices with a NODF as high as the empirical matrix.

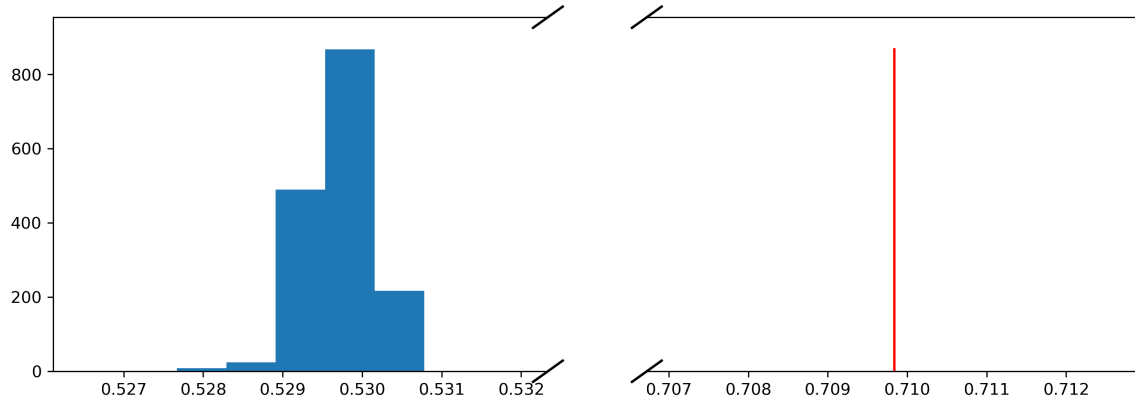


Figure D.3: Probability distribution of the NODF of matrices generated by a null model preserving heterogeneous expected activities ubiquities (blue area), compared with the NODF of the empirical matrix (red line).

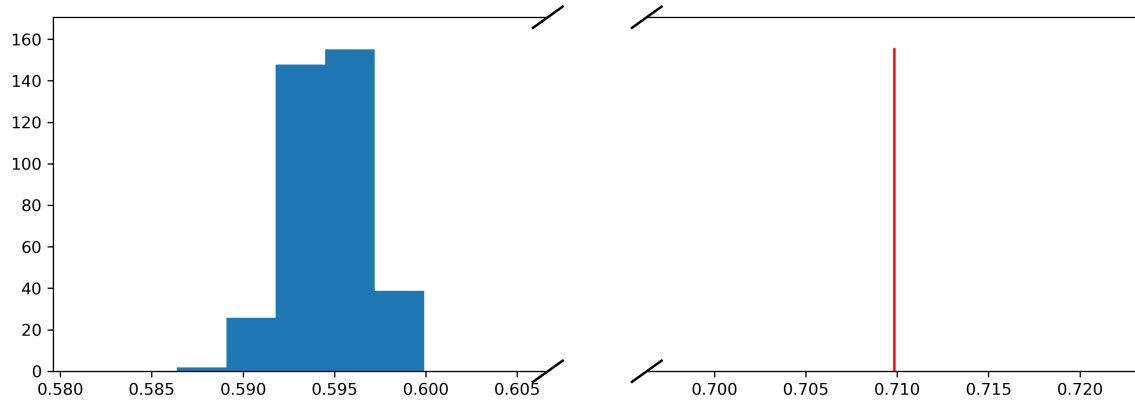


Figure D.4: Probability distribution of the NODF of matrices generated by a null model preserving heterogeneous expected activities ubiquities and expected firms diversifications (blue area), compared with the NODF of the empirical matrix (red line).

In table D.1 we show the statistics of the distribution of NODF for matrices generated with the four models, and the Z-score of the empirically observed values of NODF for the matrix assuming the four different null hypotheses. An interesting result is that, even if no null model approaches the empirical value (0.710), we can see how the heterogeneous ubiquity of activities is a crucial ingredient to generate a nested matrix, more than the diversification of firms.

Overall, the comparison between the nestedness of an empirical matrix with a random matrix is not a trivial task, as it is not clear which constraints are actually to be discounted and which are into the definition of nestedness itself (Bruno et al., 2020). In our case however, the nestedness observed, measured as NODF, is above the random case even discounting every possible constraint.

Null model	Mean	St.Dev.	Zscore
Homogeneous probability - D.1	0.249	5.8510^{-4}	11110.8
Heterogeneous firms diversification - D.2	0.298	1.5610^{-5}	263039.9
Heterogeneous activities ubiquity - D.3	0.529	4.0310^{-4}	6310.0
Bilateral configuration model - D.4	0.593	2.0510^{-3}	803.2

Table D.1: Mean and Standard deviation of the NODF for matrices generated with different null model, and T-statistics of the empirical value (0.710). All the corresponding pvalues are below machine precision.

To perform advanced activities, a firm has to climb the ladder.

Notice that the large values of the Z-score in table D.1 should not surprise the reader: having many firms in our sample, even small signals are highly statistically significant. As a comparison, it is instructive to look at the case of survey data. Applying the same computations to data from the World Management Survey⁵ (employed by several studies, among others, Bloom et al. (2012)), we can define firm capabilities from managers' answers to the survey questions. For instance, to the question "Can you describe the production process for me?", a manager could answer with different scores, where score 1 would imply the introduction of few modern manufacturing techniques in an ad-hoc manner, score 3 would imply introduction of modern manufacturing techniques in an informal or isolated manner and score 5 would imply that all major aspects of lean manufacturing have been introduced formally. The observed nestedness in the empirical matrix of capabilities extracted by this survey data, as measured by NODF, is 0.299. While it is *statistically* significantly different from the strict null model, with a Z-score of ≈ 10 from the null model average of 0.253, it is not *economically* significantly different. As NODF spans from 0 to 1, the signal with respect to random noise is just 4.5% more. More striking, even with the looser null model devoid of any constraint, a complete random rendition of a matrix of a certain density, the null model average is 0.218.

E Geographical characterization of firms

We can look at the position of firms in the ladder from a development angle, by looking at the average capabilities of firms in different Indian states exploiting their vast heterogeneity in terms of industrial capabilities of Indian states. Notice that these results are suggestive, but they do not allow for a clear interpretation. On one side, here we do not take into account the potentially confounding role of variables such as urban density, sectoral specialization, firm size, regional infrastructure, and

⁵<https://worldmanagementsurvey.org/>

so on. On the other side, the average capability of the firms in the state should not to be seen as a measure of excellence: a state with many young and active start-ups would have low average firm capabilities. With these caveats, in Figure E.1 we show the average capabilities per state. Haryana and Kerala firms have on average the most developed capabilities, while West Bengal and NCT of Delhi have the least developed capabilities. In Figure E.2 we show, for the four states with the most firms in our sample, the distribution of firms at different steps of the ladder.

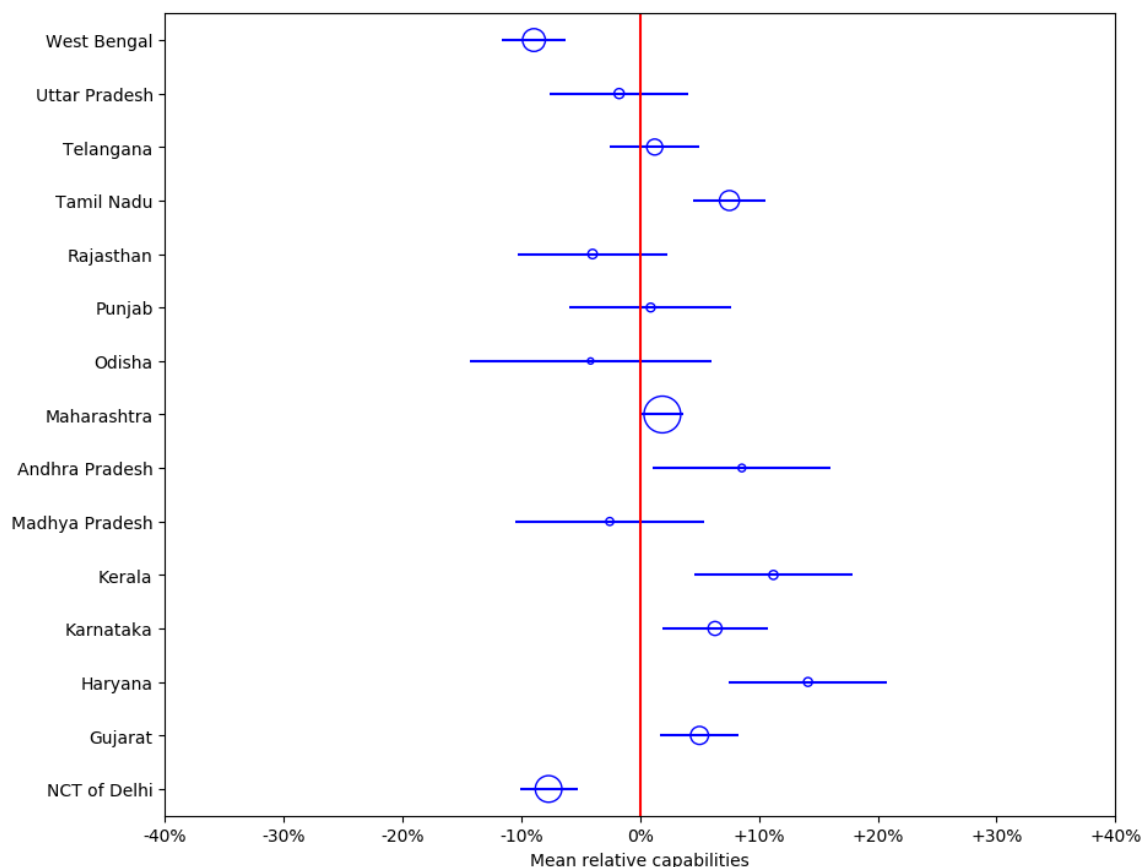


Figure E.1: Relative average position in the capabilities ladder of firms in different Indian states with respect to the average of all Indian firms, Year 2012. Error bars represents 2 standard errors – i.e. $\approx 95\%$ confidence interval. Size of the circle is proportional to the number of firms in the sector.

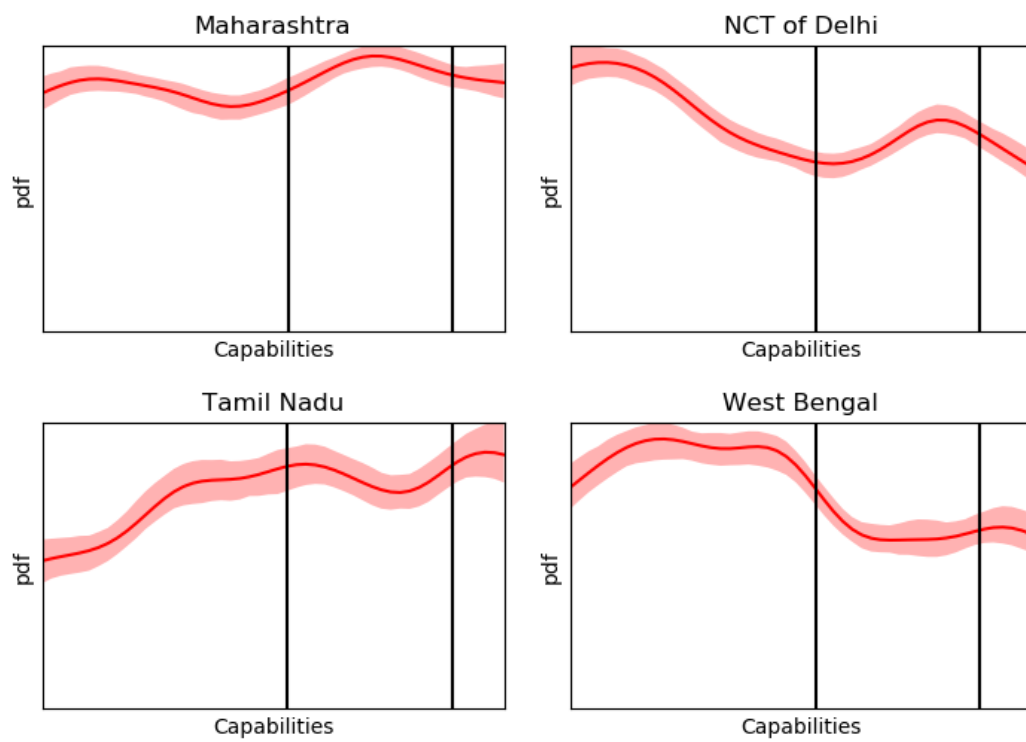


Figure E.2: Probability distribution of firms of the state by capability (inverse) ranking among Indian firms, for the four states with the most firms in our database. The two vertical black lines divide the three steps.

The UNU-MERIT WORKING Paper Series

- 2021-01 *Transformation towards sustainable development goals: Role of innovation* by Michiko Iizuka and Gerald Hane
- 2021-02 *Access to social protection for platform and other non-standard workers: A literature review* by Tamara A. Kool, Giulio Bordon and Franziska Gassmann
- 2021-03 *Labour-augmenting technical change data for alternative elasticities of substitution, growth, slowdown, and distribution dynamics* by Thomas Ziesemer
- 2021-04 *Democracy and COVID-19 Outcomes* by Gokhan Karabulut, Klaus F. Zimmermann, Mehmet Huseyin Bilgin and Asli Cansin Doker
- 2021-05 *Patent landscaping using 'green' technological trajectories* by Önder Nomaler & Bart Verspagen
- 2021-06 *Macroeconomic stimulus packages and income inequality in developing countries: Lessons from the 2007-9 Great Recession for the Covid-19 crisis in South Africa* by Alexis Habiyaemye, Peter Jacobs, Olebogeng Molewa and Pelontle Lekomanyane
- 2021-07 *Deepening or delinking? Innovative capacity and global value chain participation in the ICT sectors* by Rasmus Lema, Carlo Pietrobelli, Roberta Rabellotti and Antonio Vezzani
- 2021-08 *COVID-19 policy responses, mobility, and food prices: Evidence from local markets in 47 low to middle income countries* by Stephan Dietrich, Valerio Giuffrida, Bruno Martorano and Georg Schmöder
- 2021-09 *Attitudes towards inequality in Brazil: An analysis of a highly unequal country* by Cintia Denise Granja and Ana Maria Carneiro
- 2021-10 *Mobile phones and HIV testing: Multi-country evidence from sub-Saharan Africa* by Francesco Iacolla and Nyasha Tirivayi
- 2021-11 *What makes a productive Ph.D. student?* by Alberto Corsini, Michele Pezzoni and Fabiana Visentin
- 2021-12 *Do institutions and ideology matter for economic growth in Latin America in the first two decades of the 21st century?* by Pamela L. Navarrete Gallo and Jo Ritzen
- 2021-13 *How social assistance affects subjective Well-being: Lessons from Kyrgyzstan* by Franziska Gassmann, Bruno Martorano and Jennifer Waidler
- 2021-14 *Do pandemics lead to rebellion? Policy responses to COVID-19, inequality, and protests in the USA* by Francesco Iacolla, Patricia Justino and Bruno Martorano
- 2021-15 *Foreign R&D spillovers to the USA and strategic reactions* by Thomas H.W. Ziesemer
- 2021-16 *Will the AI revolution be labour-friendly? Some micro evidence from the supply side* by G. Damioli, V. Van Roy, D. Vertesy and M. Vivarelli
- 2021-17 *The influence of value-chain governance on innovation performance: A study of Italian suppliers* by Emanuele Brancati, Carlo Pietrobelli and Caio Torres Mazzi
- 2021-18 *Does entrepreneurship increase the chances of the poor?* by Omar Rodríguez Torres
- 2021-19 *How different are necessity and opportunity firms? Evidence from a quantile analysis of the Colombian microenterprise sector* by Omar Rodríguez Torres
- 2021-20 *A taxonomy of European innovation clubs* by Ariel L. Wirkierman, Tommaso Ciarli and Maria Savona
- 2021-21 *Supporting academic advising through self-directed, blended learning* by Mindel van de Laar, Katerina N. Triantos and Lutz F. Krebs
- 2021-22 *Powering structural transformation and productivity gains in Africa: The role of global value chains and resource endowments* by Solomon Owusu

- 2021-23 *A map of the fractal structure of high-tech dynamics across EU regions* by Ariel L. Wirkierman, Tommaso Ciarli and Maria Savona
- 2021-24 *How does market competition affect firm innovation incentives in emerging countries? Evidence from Latin American firms* by Jose Miguel Benavente and Pluvia Zuniga
- 2021-25 *The effectiveness of innovation policy and the moderating role of market competition: Evidence from Latin American firms* by Jose Miguel Benavente and Pluvia Zuniga
- 2021-26 *Knowledge transfer activities and conditions for impact in Bulgarian public research institutions: A survey-based diagnostic* by Anwar Aridi, Daniel Querejazu and Pluvia Zuniga
- 2021-27 *Religiosity, smoking and other addictive behaviours* by Monica Roman, Klaus F. Zimmermann and Aurelian-Petrus Plopeanu
- 2021-28 *Women's empowerment in Colombia: A multidimensional approach* by Esmeralda Rodríguez Guerrero, Zina Nimeh and Andrea Franco Correa
- 2021-29 *Simulating the impact of a raise in education salaries on economic growth in Peru* by Luciano Paredes Chervellini and Jo Ritzen
- 2021-30 *Social networks and agricultural performance: A multiplex analysis of interactions among Indian rice farmers* by Bruhan Konda, Mario González-Sauri, Robin Cowan, Yashodha Yashodha and Prakashan Chellattan Veettill
- 2021-31 *Positioning firms along the capabilities ladder* by Alex Coad, Nanditha Mathew and Emanuele Pugliese