

# The Catch-22 of External Validity in the Context of Constraints to Firm Growth

By GREG FISCHER AND DEAN KARLAN\*

\*Fischer: London School of Economics, IPA, MIT J-PAL, 32 Lincoln's Inn Fields, 3.09, London WC2A 2AE (email: [g.fischer@lse.ac.uk](mailto:g.fischer@lse.ac.uk)), Karlan: Yale University, IPA, MIT J-PAL and NBER, 27 Hillhouse Avenue, New Haven, CT 06511 (email: [dean.karlan@yale.edu](mailto:dean.karlan@yale.edu)). The authors thank Tavneet Suri for comments, and acknowledge support from the Templeton Foundation. We thank Sarah Craig and Else van Schijndel, who provided outstanding research support and project management, and we are grateful to Ron Chua, Mau Bolante and Ricky Lim at the Asian Institute of Management for their partnership.

There is a disconnect between academic economists' search for individual mechanisms that constrain firm growth and the more complex reality facing firms and policymakers aiming to alleviate these constraints. The comprehensive, some would say scattershot, approaches that are common in practice are considered challenging for evaluators because of the difficulty in identifying any particular causal mechanism. More targeted attempts to improve business performance typically generate mixed performance (McKenzie and Woodruff 2012) or do not seem to scale either in the market or with public support.

With that in mind, we partnered with the Asian Institute of Management, a leading Philippine business school, to launch a class-based program that had MBA students providing consulting services for local small and medium enterprises. We had three goals, spanning policy, research and teaching: to pilot a potentially scalable approach to improving management practices for small businesses; to better understand the complex

set of constraints facing individual small businesses; and, to test a hands-on, multi-skill teaching approach for MBA students.

We began with the administrative list of all tax-registered businesses in Makati City, Manila, where AIM is located. For our pilot, we restricted our attention to businesses in operation for at least two years; reporting revenues in 2010 between 1 and 15 million Philippines Pesos (PHP)<sup>1</sup>; and in industries where general consulting was feasible (e.g., we excluded foreign exchange services). We attempted to visit all 4,212 eligible businesses. Nearly 40% were not reached because they had changed address, closed, or otherwise could not be located. We explained (but did not promise) the consulting program to the 2,533 businesses that were reached. Ultimately, only 177 interviews were completed, as many owners or managers were either too busy to complete the interview, not interested in participating, or repeatedly out of

<sup>1</sup> Approximately US\$ 23,600 – 354,600 at the mid-2011 exchange rate of 42.3 PHP/USD.

the office. Of the 177 business owners interviewed, 142 upon completion of the survey expressed interest in receiving free consulting from AIM students. We completed detailed qualitative and quantitative surveys with 95 of these businesses. Given the structure of our sample, we cannot argue that it is representative of small and medium enterprises in urban and peri-urban Manila. However, we note one key observation that has implications for both research and practice: most firms have a complex set of constraints, many of which are interconnected.

The presence of multiple and varied constraints to firm growth is an emerging theme. For example, the World Management Survey (Bloom, Genakos, et al. 2012) shows that poorly managed firms have a number of weaknesses rather than problems clustering in any particular area. The World Bank's Enterprise Survey shows a similar pattern in the external business environment. The median firm lists three significant obstacles in the business environment.<sup>2</sup> As with management practices, these challenges are diffuse: after demeaning at the country level,

<sup>2</sup> Represents obstacles considered major or severe. Other categories comprise none, minor, moderate, not applicable and don't know. Authors calculations from 15 potential obstacles in combined WBES data: electricity, transportation, customs & trade regulation, informal sector competitors, access to land, crime & disorder, access to finance, tax rates, tax administration, business licensing & permits, political instability, corruption, courts, labor regulations, and inadequately educated workforce.

the first principal component of the constraints matrix explains 52% of the variation with similar weight on all obstacles.

The dataset from our project in the Philippines is smaller and more selected (those willing to participate in a consulting program), but provides richer information, more focused on perceived constraints as well as detailed quantitative and qualitative information about what is happening inside the firms. Figure 1 shows the histogram of obstacles identified per firm, grouping detailed items such as employee retention into common themes such as human resource management. Even after grouping, the median number of constraints is two out of a possible five.<sup>3</sup>

Moreover, these constraints are quite varied and consistent with an overall observation of missing "managerial" capital (Bruhn, Karlan, and Schoar 2010). Figure 2 shows the share of firms in our sample identified as facing constraints in a particular area. Within the sample for which we have detailed, qualitative data, there are two clusters. Nearly 70% of firms require some form of assistance on sales and marketing. Another 42% need assistance with accounting or cash flow management. No

<sup>3</sup> The modal constraint reported by respondents is competition; however, detailed information in our baseline survey of 95 businesses and from the consultants' engagements with 26 treatment firms suggests that this is almost universally price or quality competition in competitive markets.

other category is preponderant; however, even these groupings belie significant variation in the detailed needs of firms.

For example, among the firms needing sales and marketing assistance, one struggles to recruit sales personnel with the technical competence to accurately describe the product. Another firm struggles with marketing staff turnover and an inability to generate sales beyond the owner's personal contacts. While there are some common themes in the challenges these firms face, little suggests a one-size-fits-all training program would appeal to or benefit these firms.

The Online Appendix<sup>4</sup> presents summary statistics from the baseline, and then richer qualitative information on the constraints of the 26 firms that participated in the AIMS consulting project. These data show somewhat more clustering with respect to financial management. Several demonstrate a need for better inventory and cash flow management, and specific recommendations in these areas by the consultants were particularly well received by management.

The stories though are quite varied, demonstrating our main point. We also believe the stories are a first step towards more detailed ethnographic research, which could

be helpful for economists in forming hypotheses on constraints to firm growth.

A similar pattern of varied constraints to firm growth and performance is evident in other data. In the WMS, a small majority of firms are classified as poor performers (score less than three) for management practices related to human resources (rewarding high performers, getting rid of poor performers, performance clarity, and retaining human capital); however, this may reflect both internal and external constraints. No other practice has more than 40%.

Taken together, these results not only point to a weakness in providing one-size-fits-all business training interventions. They also present a challenge for academic economists looking to identify mechanisms through which training programs may affect business outcomes. When there is significant diversity in the obstacles faced by firms, it may simply be mechanically difficult to identify the particular channel through which a program or policy may work without large samples and detailed baseline diagnostics with which to test interactions. For example, marketing training is unlikely to benefit the 40% of firms that do not appear to have any problems with marketing. Moreover, even a well-structured marketing course may fail to address the specific needs of firms struggling in this area.

<sup>4</sup> Available at <http://karlan.yale.edu>.

Firms may be aware of these challenges and therefore rationally choose not to attend such trainings, a possibility consistent with low observed attendance of even subsidized training programs.

Based on the qualitative evidence, we posit that identifying any one mechanism through which firm performance and growth could be improved may be hindered by a dismal application of Kremer's O-ring theory (Kremer 1993). Each reported obstacle or poor management practice is a failed O-ring. Removing one obstacle would not improve outcomes because several others still persist.

There are two implications of this pattern. First, it presents a challenge to the evaluation space but not an insurmountable one. One option would be to begin with larger sample frames, run detailed diagnostics prior to treatment assignment, and then put forward trainings that target the identified constraints (firms diagnosed with financial problems get finance advice, those with human resources problems get human resources advice, etc.). This approach is not without challenges. It tests not the training alone, but training preceded by the diagnostic process, which itself may yield false positives or negatives. Moreover, the diagnostic process itself may be a treatment, changing the firms' behavior. An evaluation can only assess the impact of

training over and beyond the diagnostics. If the diagnostics are not part of everyday business, which typically they would not be, then estimating the treatment effect of the training in this setting is useful for cost-benefit analysis of the training but not as much for learning why firms are constrained.

Alternatively, one could design multi-arm experiments that randomly assign firms to different training modules: some receive all, some receive a random subset, etc. The permutations here are practically unbounded, and the sample size required to identify economically meaningful effects quickly gets out of control when one considers the variation in constraints across firms.<sup>5</sup> Naturally any result on the relative magnitude of treatment effects of one treatment arm to another will be dependent on the characteristics of the sample frame. Thus the highly-self-selective nature of many business training programs suggests that such exercises are useful for learning about the relative treatment effects but less useful for making grand statements about the constraints to growth for firms overall in that market.

Regardless of the approach, this issue also makes clear the need for monitoring and process data to help shed light on which

<sup>5</sup> See Anderson-Macdonald and Chandy (2014) for an example of this approach for just marketing and finance training.

components of training are actually adopted. This is critical for helping to track the theory of change of a training program: first, measure actual activities, the teaching activities, number of hours of meetings, etc.; then measure whether participants' knowledge increases on the specific topics taught; then measure whether participant behavior and choices change; then measure whether business outcomes change, as well as overall aspirations, motivation and "entrepreneurial spirit" of the business owners.

The second implication of firms facing multiple and complex constraints is that more tailored consulting or mentoring programs may be more appropriate for improving firm performance (e.g., see Bloom, Eifert, et al. 2012; Bruhn, Karlan, and Schoar 2013 for examples of successful consulting interventions; and Karlan, Knight, and Udry 2013 for an example of an unsuccessful consulting intervention, on smaller firms). The intervention that we piloted with AIM was designed to test a potentially scalable approach to do just that for small businesses in low- and medium-income countries. Such approaches face their own set of challenges. For example, providing effective business consulting in the face of multiple, diverse, and unpredictable challenges may require a level of expertise that would be unrealistic to expect

of students or other low-cost providers. We are, however, encouraged by the fact that even in the face of these challenges, 60% of student teams made recommendations that were implemented by clients and subjectively judged effective. When we restrict our attention to those teams independently judged as delivering satisfactory work—as we know, not every student assignment will receive full effort—that rises to over 70%.

We return to our opening point. External validity is typically discussed in two ways: with theory and with empirics. As Deaton (2010) argues and we agree, external validity is gained by having an empirically-validated theory of why something is working, and that theory ought to include relevant contextual factors. External validity is also discussed empirically: a result from one sample frame at one point in time can be used to predict results elsewhere? The less selected the sample frame and context – i.e. the more representative it is of a defined population – the more convincingly one can translate results to elsewhere.

These two aspirations are at odds with each other. The quest for theory and cleanly identified mechanisms calls for narrow, highly-selected sample frames. Yet given the complexity of constraints to firm growth, any successful attempt to identify a particular

mechanism would likely require narrowing one's sample frame to highly specific firms that may not even be representative of other firms in the same market. In contrast, the desire for empirical breadth and representativeness pushes towards sample frames that will be full of complex, changing and amorphous issues with no readily identifiable mechanisms.

We need both to move forward. Through iteration and extension—using one approach to inform the other, back and forth—we can arrive at a mosaic understanding of the constraints to firm growth. But to get there we need to lower our expectations for what we can learn from any one study.

## REFERENCES

Anderson-Macdonald, Stephen, Rajesh Chandy, and Bilal Zia. 2014. "The Impact of Marketing (versus Finance) Skills on Firm Performance." *Working Paper*.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2012. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics*, November, qjs044. doi:10.1093/qje/qjs044.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen. 2012. "Management Practices Across Firms and

Countries." *The Academy of Management Perspectives* 26 (1): 12–33. doi:10.5465/amp.2011.0077.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2010. "What Capital Is Missing in Developing Countries?" *American Economic Review* 100 (2): 629–33. doi:10.1257/aer.100.2.629.

———. 2013. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Working Paper*, March.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55. doi:10.1257/jel.48.2.424.

Karlan, Dean, Ryan Knight, and Christopher Udry. 2013. "Consulting and Capital Experiments with Micro and Small Tailoring Enterprises in Ghana." *Working Paper*.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *The Quarterly Journal of Economics* 108 (3): 551–75. doi:10.2307/2118400.

McKenzie, David, and Christopher Woodruff. 2012. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *World Bank Policy Research Working Paper* 6202 (September).

Figure 1: Histogram of number of constraints (primary categories only) identified by each firm (N=177), from qualitative interviews.

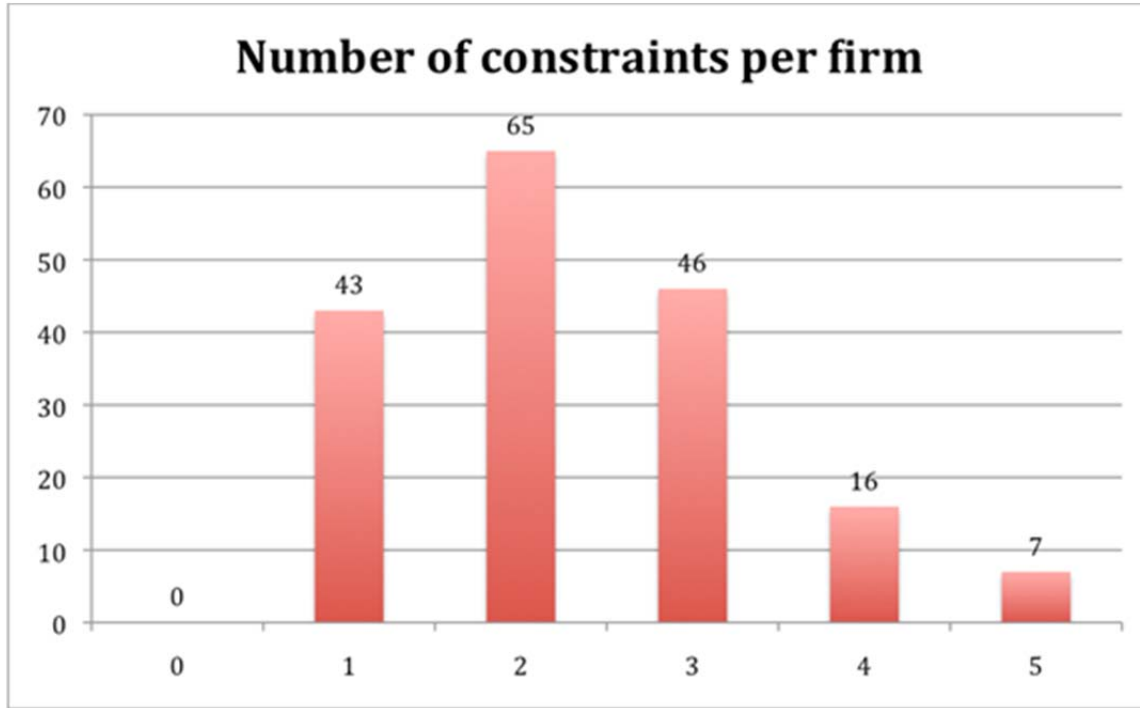


Figure 2: Histogram of the number of constraints identified per main category across all firms completing qualitative interviews (N=177). Includes firms identifying at least one constraint in a given sub-category.

