
What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?

David McKenzie and Christopher Woodruff

Business training programs are a popular policy option to improve the performance of enterprises around the world, and the number of rigorous impact evaluations of these programs is growing. A critical review reveals that many evaluations suffer from small sample sizes, measure impacts only within a year of training, and experience problems with survey attrition and measurement that limit the conclusions one can draw. Over these short time horizons, there are relatively modest effects of training on the survivorship of existing firms. However, there is stronger evidence that training programs help prospective owners launch new businesses more quickly. Most studies find that existing firm owners implement some of the practices taught in training, but the magnitudes of the improvement to practices is often modest. Few studies find significant impacts on profits or sales, although some studies with greater statistical power have done so. There is little evidence to guide policymakers regarding whether any identified effects are due to trained firms drawing sales from competing businesses rather than through productivity improvements or to guide the development of the provision of training at market prices. We conclude by summarizing some directions and key questions for future studies. JEL codes: O12, J16, L26, M53

Walk into a typical micro or small business in a developing country and spend a few minutes talking with the owner, and it often becomes clear that owners are not implementing many of the business practices that are standard in most small businesses in developed countries. Formal records are not kept, and household and business

The World Bank Research Observer

© The Author 2013. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
doi:10.1093/wbro/lkt007 Advance Access publication July 22, 2013 29:48–82

finances are combined. Marketing efforts are sporadic and rudimentary. Some inventory sits on shelves for years at a time, whereas more popular items are frequently out of stock. Few owners have financial targets or goals that they regularly monitor and act to achieve. The picture is not much better in some medium and large firms: few firms use quality control systems, reward workers with performance-based pay, or adopt many other practices that are typical of well-managed firms in developed countries.

It is small wonder, then, that business training is one of the most common forms of active support provided to small firms around the world. There are a number of programs offered by governments, microfinance organizations, and NGOs in many countries around the world. Perhaps the mostly widely implemented training program is the International Labor Organization's Start and Improve Your Business program. Established in 1977, the program claims more than 4.5 million trainees with implementation in more than 100 countries.¹ Other widely used programs include the GTZ/CEFE program, the UNCTAD/EMPRETEC program, business plan competitions and training run by Technoserve, content for microfinance clients developed by Freedom from Hunger, and the IFC's Business Edge and SME Toolkit programs.

Until recently, however, there has been very little rigorous evidence on the impacts of these programs. Overviews of evidence from mostly nonexperimental evaluations of programs that focus on training for the unemployed in developed countries (Dar and Tzannatos 1999) and developing and transition countries (Betcherman et al. 2004) have found the existing evidence to be mixed, at best. A 2009 overview of impact evaluations in finance and private sector development found very little work on business training (McKenzie 2010). The last three years have seen a rapid increase in attention to the idea that "managerial capital" or poor management is a constraint to production in developing countries (Bruhn et al. 2010; Bloom and Van Reenen 2010) as well as the emergence of a number of impact evaluations of business training programs. This paper provides a critical overview of lessons from these evaluations for both policy and the next generation of research.

We use a variety of methods to identify all published studies and recent working papers that examine the impacts of business training in developing countries. These include an Econlit search for published studies, Google Scholar searches of papers that cite these published studies or other working papers, our contacts with scholars working in this field, input from recent training program inventory exercises, and knowledge of papers presented in recent seminars or conferences. We restrict our attention to papers with a clear impact evaluation design that address the selection of both observable and unobservable characteristics of business owners and that focus on enterprise management rather than solely on technical or vocational training.

We begin by assessing the comparability of these programs in terms of their course content and participants. We find considerable variation across studies in terms of the participants and the length and content of the training provided, although a number of core topics are covered in most training sessions. Next, we discuss a number of challenges faced by researchers when measuring impact. Critically, most of the existing studies measure impacts on relatively small samples of very heterogeneous firms. In addition, many existing studies only consider impacts within a year of training, a period that is too short to detect some changes. Many studies also experience problems with attrition, selective survival and start-up, and nonresponses for sensitive outcomes such as profits and revenues. A final concern is that training may change the measurement of outcomes even if it does not change the outcomes themselves. We discuss several studies' attempts to show that their results are robust to reporting issues.

With these issues in mind, we assess what we have learned about the impacts of different programs on business survivorship and start-up, business practices, and profitability and enterprise growth. Among the minority of studies that have examined the effects of these programs on the survivorship of existing businesses, there is some weak evidence for a positive effect for male-owned businesses. However, for female-owned businesses, training is found to have either no effect or a slightly negative effect on survivorship. Stronger results have been found with respect to the impacts of training programs on new business start-ups. All of the studied training programs that include content specifically intended to help people start new businesses have found that training helps in starting firms, although there is some evidence that training merely hastens the entry of firms that would enter anyway and potentially changes the selection of which firms enter.

Almost all training programs find that treated firms implement some of the business practices taught in the training. However, the magnitude of the impact is small in many cases; a typical change is 0.1 or 0.2 standard deviations, or 5 to 10 percentage points. The combination of relatively small changes in business practices and low statistical power means that few studies find effects of training on sales or profitability, although a few studies find some positive short-term effects. Studies of microfinance clients find some evidence that training changes the rates of client retention and the characteristics of loan applicants. Finally, the three studies that examine the impact of individualized consulting provided to larger firms find evidence that consulting services can improve the performance of firms, including those with multiple plants and more than 200 workers.

Before concluding, we discuss several important issues for which existing studies provide very little evidence but which are crucial for the development of policy recommendations. These issues include whether gains from training are long lasting and whether these gains result from competing away sales from untrained firms or through other channels. We also discuss the need to address the heterogeneity of

training content and participants and to identify the market failures that may prevent firms from investing in training that may be beneficial. We conclude with recommendations for future work in this area.

What Does a Typical Business Training Program Involve?

Attempts to measure the impact of “business training” face multiple challenges that complicate comparisons across studies. The first challenge is that business training varies in what is offered and how it is offered across different locations and organizations. These differences in content are likely to be important, and they induce much more variation into the treatment of business training than exists in other firm interventions, such as access to capital through credit or grants. A second challenge (common to most evaluations) is that the impact of training is likely to differ depending on who receives the training. Thus, even if we compare the same training content in different locations, differences in the characteristics of the individuals receiving the training may result in different measured impacts. Therefore, it is important to carefully examine who participates and what is offered before making comparisons among studies.

Who Participates in Business Training Experiments?

Table 1 summarizes the key characteristics of the participants in recent business training evaluation studies. Classroom-based training offered by microfinance organizations or banks to their clients is the most common modality among these studies. This approach is particularly common for training offered to female micro-enterprise owners because the majority of microfinance clients are women. A second strategy is to offer training to firms in a particular industry or industrial cluster (Mano et al. 2011; Sonobe et al. 2011). A third strategy is for individuals to apply to participate in training as part of a competition, as Technoserve does (Klinger and Schündeln 2011), to be screened for interest in participating (Valdivia 2012), or for students to apply to participate in an entrepreneurship course (Premand et al. 2012). All of these approaches result in a selected sample of firms, which may differ from the general population, making it difficult to generalize their findings to an average firm. A final approach, used only by de Mel et al. (2012) and Calderon et al. (2012), is to draw a representative sample of the microenterprise population of interest and then offer the training to a random subsample of this population.

Most evaluations focus on existing businesses. Exceptions include studies in which many of the microfinance clients are borrowing or saving for household purposes but do not necessarily have an enterprise (Field et al. 2010) and studies based

Table 1. Who Are the Participants in Business Training Evaluations?

<i>Study</i>	<i>Country</i>	<i>Existing Businesses?</i>	<i>All microfinance/ bank clients?</i>	<i>Rural or Urban</i>	<i>Business Sector</i>	<i>Selected on interest in training?</i>	<i>Mean Age</i>	<i>% Female</i>
Berge et al. (2011)	Tanzania	Existing	Yes	Urban	Many	No	38	65
Bruhn and Zia (2012)	Bosnia-Herzegovina	67% existing	Yes	Urban	Many	Yes	28	35
Calderon et al. (2012)	Mexico	Existing	No	Rural	Many	No	46	100
De Mel et al. (2012)	Sri Lanka	50% existing	No	Urban	Many	No	34-36	100
Drexler et al. (2012)	Dominican Republic	Existing (a)	Yes	Urban	Many	No	40	90
Field et al. (2010)	India	24% existing	Yes	Urban	Many	No	32.4	100
Giné and Mansuri (2011)	Pakistan	61% existing	Yes	Rural	Many	No	37.6	49
Glaub et al. (2012)	Uganda	Existing	No	Urban	Many	Yes	39	49
Karlan and Valdivia (2011)	Peru	Existing	Yes	Both	Many	No	n.r.	96
Klinger and Schündeln (2011)	El Salvador, Guatemala, Nicaragua	39% existing	No	n.r.	Many	Yes	36	28
Mano et al. (2012)	Ghana	Existing	No	Urban	Metalwork	No	45	0
Premand et al. (2012)	Tunisia	No	No	Urban	Many	Yes	23	67
Sonobe et al. (2011)	Tanzania	Existing	No	Urban	Garments	No	45	85
	Ethiopia	Existing	No	Urban	Metalwork	No	44	4
	Vietnam	Existing	No	Urban	Rolled Steel	No	40	55
	Vietnam	Existing	No	Urban	Knitwear	No	41	66
Valdivia (2012)	Peru	Existing	No	Urban	Many	Yes	43	100

Note: n.r. denotes not reported.

(a) 78 percent of sample is existing businesses, and study does not look at business outcomes for those who were not existing at baseline.

on competitions or training of new businesses (Klinger and Schündeln 2011; Premand et al. 2012; de Mel et al. 2012). The majority of the evaluations to date have focused on urban clients, which likely reflects the greater density of businesses and training providers in urban areas.

The average age of a participant in a typical study is 35 to 45 years, although two studies focus on young entrepreneurs (Bruhn and Zia 2012; Premand et al. 2012). Some studies focus entirely on female business owners, and others focus on male owners. Relatively few studies have sufficient numbers of both genders to compare impacts separately. Finally, there is substantial heterogeneity in the education levels of participants, with averages as low as 2.5 years of schooling for females and 5.7 years for males in the study of rural Pakistan by Giné and Mansuri (2011) and as high as university level in the study by Premand et al. (2012).

Table 2 shows the degree of heterogeneity in firm size at baseline among studies that include existing firms. At the low end are subsistence firms run by women in Giné and Mansuri (2011) and de Mel et al. (2012), where 95 percent of the firms

Table 2. Heterogeneity in Baseline Size of Firms Participating in Business Training Experiments

Study	% with zero employees	Mean Employees	Monthly Profits (USD)		Monthly Revenues (USD)	
			Mean	S.D.	Mean	S.D.
Berge et al. (2011)	n.r.	1.08	480	384	2102	3083
Males	n.r.	1.18	528	432	2586	2876
Females	n.r.	1.03	455	354	1847	3160
Bruhn and Zia (2012)	n.r.	2.08	700	n.r.	n.r.	n.r.
Calderon et al. (2012)	60	1.6	121	183	398	610
De Mel et al. (2012)	95	0.06	35	17	109	99
Drexler et al. (2012)	60	n.r.	n.r.	n.r.	747	1215
Giné and Mansuri (2011)	90	2.43	n.r.	n.r.	n.r.	n.r.
Males	86	2.51	n.r.	n.r.	380	n.r.
Females	95	2.34	n.r.	n.r.	80	n.r.
Glaub et al. (2012)	n.r.	1.5	n.r.	n.r.	100	n.r.
Karlan and Valdivia (2011)	n.r.	0.22	-165	4118	534	1230
Klinger and Schündeln (2011)	n.r.	8	n.r.	n.r.	6916	17333
Mano et al. (2012)	n.r.	n.r.	2200	2700	4717	5658
Sonobe et al. (2011)						
Tanzania	n.r.	5	530	1056	866	1393
Ethiopia	n.r.	33	19599	38048	142311	354163
Vietnam - Steel	n.r.	17	2627	4181	105787	98526
Vietnam - Knitwear	n.r.	20	-888	7234	7055	16509
Valdivia (2012)	n.r.	0.23	n.r.	n.r.	740	1696

Note: n.r. denotes not reported.

have no paid employees, average monthly revenues are only \$80–100² at market exchange rates, and profits are approximately \$1 per day. Most of the rest of the studies focus on microenterprises, albeit ones with slightly larger revenues and potentially one or two employees. The main exceptions are the firms chosen from industrial clusters (Mano et al. 2011 and Sonobe et al. 2012), in which the firms are SMEs with five to 50 workers and monthly revenues of \$5,000 or more (and in some cases, more than \$100,000).

Training Delivery and Costs

All of the training courses reviewed here are classroom-based courses delivered to groups of individuals, although several of the programs provide additional one-on-one follow-up training, which we will discuss later. Table 3 provides key characteristics of the training delivery in the different studies. A first point is that many of the studies test content that is modified or developed specifically for the study of interest rather than content that has been taught for years. This situation may be significant if it takes time to adapt particular content to a local context or for instructors to become familiar with new material.

The length of the training course also varies substantially across studies. The shortest courses are two days or two half-days (Bruhn and Zia 2012; Field et al. 2010), whereas other courses are full time and last one week or more (de Mel et al. 2012; Sonobe et al. 2011). In most cases, the training is concentrated in a relatively short period, but in some of the cases, especially where training takes place in microfinance group meetings, it is spread over many months in blocks as short as half an hour (Karlan and Valdivia 2011). Longer full-time courses allow more content to be taught, but they are more costly and require business owners to be away from their businesses for longer.

In all of these experiments, training is offered for free. In addition, some studies have provided small supplements for travel or food or have offered the prospect of grants as an additional incentive. The training costs per person range from as little as \$21 in Drexler et al. (2011), where training was conducted by local instructors once per week over five or six weeks in local schools, to more than \$400 per firm in Sonobe et al. (2011), where instructor costs and venue rental costs per person for 15 days were relatively high. One argument for subsidizing costs is that many business owners have little perception of how badly managed their firms are. To these owners, training is a new and unproven concept with uncertain payoffs. Even those who are not liquidity constrained may be reluctant to pay, and training providers may find it costly and difficult to credibly signal quality. Two studies find evidence to support the idea that individuals who are the least interested potentially have the most to gain from training (Bjorvatn and Tungodden 2010; Karlan and Valdivia

Table 3. Key Characteristics of Training Delivery

<i>Study</i>	<i>Training Provider</i>	<i>Training content new or established?</i>	<i>Course Length (hours)</i>	<i>Participant Cost (USD)</i>	<i>Actual Cost (USD)</i>	<i>Attendance Rate</i>
Berge et al. (2011)	Training professionals	New	15.75	0	\$70	83%
Bruhn and Zia (2012)	Training organization	New	6	0	\$245	39%
Calderon et al. (2012)	Professors & Students	New	48	0	n.r.	65%
De Mel et al. (2012)	Training organization	Established (ILO)	49-63	0	\$126-140	70-71%
Drexler et al. (2012)						
"Standard"	Local instructors	New	18	0 or \$6	\$21	50%
"Rule-of-thumb"	Local instructors	New	15	0 or \$6	\$21	48%
Field et al. (2010)	Microfinance credit officers	New (a)	2 days	0	\$3	71%
Giné and Mansuri (2011)	Microfinance credit officers	New (b)	46	0	n.r.	50%
Glaub et al. (2012)	Professor	New	3 days	0	\$60	84%
Karlan and Valdivia (2011)	Microfinance credit officers	Established (FFH)	8.5-22 (c)	0	n.r.	76-88%
Klinger and Schündeln (2011)	Training professionals	Established (Empretec)	7 days	0	n.r.	n.r.
Mano et al. (2012)	Local instructors	New (d)	37.5	0	\$740	87%
Premand et al. (2012)	Govt. office staff	New	20 days +	0	n.r.	59-67%
Sonobe et al. (2011)						
Tanzania	Training professionals	New (d)	20 days	0	>\$400	92%
Ethiopia	Training professionals	New (d)	15 days	0		75%
Vietnam - Steel	Training professionals	New (d)	15 days	0		39%
Vietnam - Knitwear	Training professionals	New (d)	15 days	0		59%
Valdivia (2012)	Training professionals	New	108 (e)	0	\$337 (f)	51%

Note: FFH denotes Freedom from Hunger; ILO denotes the International Labor Organization.

(a) Shortened version of existing program + new content on aspirations added.

(b) Adapted from ILO's Know About Business modules.

(c) Training sessions were each 30 minutes to 1 hours, and up to 22 sessions occurred, but only half had done 17 sessions over 24 months.

(d) Based in part on ILO content + Japanese Kaizen content.

(e) Although only 42 percent of those attending completed at least 20/36 sessions, and only 28 percent attended 30 sessions or more.

(f) The basic training cost \$337, while the technical assistance plus basic training cost \$674.

2011). We will return to a discussion of market failures and subsidies later in the paper.

Although training is offered for free, the average participation rate across the different studies for individuals who are offered training is only about 65 percent. Low take-up rates make it difficult to measure impacts; decreasing the take-up rate from 100 percent to 65 percent increases the required sample size by 2.4 times. One would expect take-up rates to be highest when training occurs in the context of regular group meetings organized by microfinance organizations, but even in the “mandatory” treatment of Karlan and Valdivia (2011), attendance rates are only 88 percent. Screening for initial interest in training does not guarantee high take-up rates either. Bruhn and Zia (2012) and Valdivia (2012) focus on samples that had initially expressed interest in attending a training course, but they still only obtain attendance rates of 39 percent and 51 percent, respectively. In most short courses, there is very little drop out conditional on attending the first session of the course, but longer courses experience more drop out over time.

Training Content

Table 4 summarizes the key topics taught in the different courses. All of the studies focus on general business skills that should be broadly applicable to most businesses rather than technical knowledge or sector-specific content. However, there is significant variation in the depth and breadth of topics. The most common set of topics focuses on maintaining business records and encouraging small business owners to separate household and business finances. Many courses, especially those targeted to potential rather than existing business owners, focus on generating a product idea and the steps needed to take the product to market. A core set of topics for attempting to grow existing businesses includes marketing, pricing and costing, inventory management, customer service, and financial planning. Because few microenterprises have employees, employee management is not a significant part of most courses. Courses that focus on larger firms include content on quality management, lean production, or Kaizen and 5S techniques³ for continuous production improvement. Finally, in addition to targeting improvements in business practices, some courses attempt to change entrepreneurial attitudes or aspirations. The amount of time devoted to attitudes has been relatively low in the courses studied by economists, but Glaub and Frese (2011) review a number of nonexperimental studies of training programs in developing countries that focus on strengthening psychological factors. Glaub et al. (2012) provide an example of a three-day course focused on personal initiative training, a psychological intervention aimed at making business owners more proactive and self-starting with respect to new ideas and opportunities and more persistent in overcoming barriers.

Table 4. Training Content

<i>Study</i>	<i>Separating household and business finances</i>	<i>Accounting</i>	<i>Financial Planning</i>	<i>Product ideas</i>	<i>Marketing</i>	<i>Pricing and Costing</i>	<i>Inventory Management</i>	<i>Customer Service</i>	<i>Business Investment & Growth Strategies</i>	<i>Employee Management</i>	<i>Savings</i>	<i>Debt</i>	<i>Using Banks</i>	<i>Quality</i>	<i>Kaizen/ 5S/ Lean</i>	<i>Aspirations/ Self-esteem</i>
Berge et al. (2011)	X	X	X	X	X	X	X	X		X	X	X	X	X		X
Bruhn and Zia (2012)	X	X	X		X				X							
Calderon et al. (2012)		X		X	X	X		X	X							
De Mel et al. (2012)	X	X	X	X	X	X	X	X								
Drexler et al. (2012)																
"Standard"	X	X	X				X				X	X				
"Rule-of-thumb"	X	X									X	X				
Field et al. (2010)	X	X	X													X
Giné and Mansuri (2011)	X	X		X		X										
Glaub et al. (2012)																X
Karlan and Valdivia (2011)	X	X	X	X	X	X										

Continued

Table 4.. *Continued*

<i>Study</i>	<i>Separating household and business finances</i>	<i>Accounting</i>	<i>Financial Planning</i>	<i>Product ideas</i>	<i>Marketing</i>	<i>Pricing and Costing</i>	<i>Inventory Management</i>	<i>Customer Service</i>	<i>Business Investment & Growth Strategies</i>	<i>Employee Management</i>	<i>Savings</i>	<i>Debt</i>	<i>Using Banks</i>	<i>Quality</i>	<i>Kaizen/ 5S/ Lean</i>	<i>Aspirations/ Self-esteem</i>
Klinger and Schündeln (2011)		X	X		X	X	X	X							X	X
Mano et al. (2012)	X	X			X	X	X						X	X	X	
Premand et al. (2012)		X		X	X	X										X
Sonobe et al. (2011)																
Tanzania		X			X	X	X						X	X	X	
Valdivia (2012)		X		X	X	X		X						X		X

Note: Based on training descriptions provided in research studies.

The different types of content may affect business performance in different ways. Simple accounting practices and financial literacy training may give business owners a better understanding of the profitability of their business but may have little immediate effect on sales or profit levels. However, in the longer term, better accounting practices may enable owners to reinvest more in their firms because of higher savings or to put more effort on product lines that are more profitable. In contrast, some other practices may show impacts more quickly. For example, better marketing and customer service may directly increase sales, whereas costing and quality control practices may lead to reduced costs and increased profits. The development of a new product idea may have rapid and long-lasting benefits even if no other additional practices are introduced. Changes in entrepreneurial attitudes may affect how hard the owner works and the way the owner thinks about various business decisions. However, because all of the available training experiments contain a mixture of different content, existing studies are unable to determine which components of training are most important.

Challenges in Measuring Impact

Impact evaluations that measure the effects of business training programs on business performance rely primarily on survey data to measure outcomes. To obtain credible and useful estimates, studies must have sufficient statistical power, measure impacts over an appropriate time horizon, address survey attrition and the selective survival and start-up of firms, and address the possibility that training changes how firms report business outcomes even if it does not change those outcomes. We discuss each of these challenges and assess how well existing studies have met them.

Power

The power of a statistical test is the probability that it will reject a null hypothesis given that the null hypothesis is false. A starting point for most business training evaluations is to test the null hypothesis that the intervention had no effect, so the power of the experiment is a measure of the ability to detect an effect of training if such an effect does exist. The key determinants of the power of a study are the size of the sample, the amount of heterogeneity in the sample (the more diverse the set of firms, the more difficult it is to measure change in them), whether the intervention occurs at an individual or group level (power is lower for a given sample size when treatments are allocated at the group level), and the size of the treatment effect. Low take-up rates dilute the treatment effect, reducing power.

Table 5 compares studies in terms of these components of power.⁴ A typical study involves approximately 200 to 400 individuals or groups in each of the treatment

Table 5. Power of Studies to Detect Increases in Profits or Sales

Study	Group or Individual Randomization?	Sample Sizes in Treatment (T) and Control (C) Groups	C.V. Profits	C.V. Revenues	Attendance Rate	25% in Profits	Power to Detect Increase of:		
							50% in Profits	25% in Revenues	50% in Revenues
Berge et al. (2011)	Group	119 (T), 116 (C) groups (a)	0.80	1.47	83%	0.631-0.842	0.996-1.000	0.239-0.365	0.705-0.897
Bruhn and Zia (2012)	Individual	297 (T), 148 (C)	2.69	n.a.	39%	0.070	0.132	n.a.	n.a.
Calderon et al. (2012)	Two-stage	164 (T), 711 (C) (c)	1.51	1.53	65%	0.263 (b)	0.754 (b)	0.257 (b)	0.743 (b)
De Mel et al. (2012)	Individual	200 (T1), 200 (T2), 228 (C)	0.49	0.91	70%	0.990	1.000	0.632	0.994
Drexler et al. (2012)	Individual	402 (T1), 404 (T2), 387 (C)	n.a.	1.63	49%	n.a.	n.a.	0.231	0.686
Giné and Mansuri (2011)	Group	373 (T), 374 (C) groups	n.a.	n.a.	50%	n.a.	n.a.	n.a.	n.a.
Glaub et al. (2012)	Individual	56 (T), 53 (C)	n.a.	n.a.	84%	n.a.	n.a.	n.a.	n.a.
Karlan and Valdivia (2011)	Group	138 (T), 101 (C) groups	-24.96	2.30	80%	0.057 (b)	0.078 (b)	0.120-0.757	0.335-1.000
Klinger and Schündeln (2011)	Individual RD	377 (T), 278 (C)	n.a.	2.51	n.a.	n.a.	n.a.	0.259 (d)	0.746 (d)
Mano et al. (2012)	Individual	47 (T), 66 (C) (b)	1.23	1.20	87%	0.188	0.571	0.195	0.592
Sonobe et al. (2011)									
Tanzania	Individual	53 (T), 59 (C)	1.99	1.61	92%	0.109	0.292	0.141	0.414
Ethiopia	Individual	56 (T), 47 (C)	1.94	2.49	75%	0.087	0.204	0.072	0.142
Vietnam - Steel	Individual	110 (T), 70 (C)	1.59	0.93	39%	0.075	0.153	0.124	0.353
Vietnam - Knitwear	Individual	91 (T), 70 (C)	-8.15	2.34	59%	0.052	0.058	0.074	0.150
Valdivia (2012)	Individual	709 (T1), 709 (T2), 565 (C)	n.a.	2.29	51%	n.a.	n.a.	0.207	0.626

Notes: n.a. denotes not available, either because the study did not report this outcome, or because it didn't report the coefficient of variation (C.V).

Personal correspondence with authors used to obtain C.V.s from studies which only report sample means and not standard deviations.

Where range is shown, first number is power if intra-cluster correlation is one, second is power if intra-cluster correlation is zero.

(a) Numbers in control and training only groups - the study also includes groups with grants. Power calculations based on random assignment to groups, which is the working assumption of the paper, although in practice true random assignment only occurred at the branch-day of the week level, in which case power is zero.

(b) Power calculation assuming randomization was at the individual level. Actual power will be lower once group-level randomization is accounted for.

(c) Assignment first at the village level to 7 treated villages and 10 control villages, then assignment within village to treatment and control.

(d) Study does not examine revenue as an outcome, since some data is collected retrospectively.

Power calculations ignore survey attrition, which would further lower power. They also assume entire sample are existing enterprises.

Attendance rate for Klinger and Schündeln (2011) assumed to be 90 percent for purpose of power calculations.

Power calculations assume one baseline and one post-treatment survey, with an autocorrelation in the outcome variable of 0.5, and ANCOVA estimation.

and control groups, although sample sizes have been smaller for studies based on specific industrial clusters (Mano et al. 2012; Sonobe et al. 2011). A useful summary statistic of the cross-sectional heterogeneity in baseline firms is the coefficient of variation of profits or revenues, which is the ratio of the standard deviation and the mean. The two studies with the lowest coefficients of variation are both studies that restrict the heterogeneity in firms eligible for the study. De Mel et al. (2009) required firms to have baseline profits below Rs 5,000 per month (\$43), whereas Berge et al. (2011) restricted training to firms with loan sizes in a narrowly defined range. In contrast, most studies contain a much wider mix of firms, resulting in coefficients of variation exceeding two or more. The more heterogeneous the firms are, the more difficult it is to detect changes in their average outcomes arising from treatment.

Many funding agencies consider 80 percent to 90 percent power an appropriate target (Duflo et al. 2008), and power of 80 percent or more is the standard in medical trials (Schultz and Grimes 2005). Table 5 shows that many—indeed, most—business training experiments fall well below these levels in terms of power to detect a 25 percent or even 50 percent increase in profits or revenues. For a microenterprise earning \$25 per month (about \$1 per day), a 25 percent increase in profits would be \$75 per year, or about 75 percent of the direct costs of a typical microenterprise training program. Therefore, a reasonable assessment of impact should have the power to measure returns at least at this level. However, in fact, none of the studies achieves 80 percent power to detect a 25 percent increase in revenues, and only de Mel et al. (2012) and possibly Berge et al. (2011) exceed 80 percent power for a 25 percent increase in profits.⁵ Valdivia (2012) demonstrates the importance of heterogeneity and take-up. Although that study has the largest sample size of any individual experiment, high heterogeneity and a low 51 percent take-up rate (requiring four times the sample size to achieve a given power compared with a 100 percent take-up rate) yield very low power.

We should also note that power is generally much higher for detecting binary outcomes, such as whether a new business is started, whether a firm applies for a loan, or whether a firm implements a particular business practice. Therefore, studies with low power to inform about the impact of training on ultimate business outcomes may still be informative about other training impacts.

Timing of Effects

The short- and long-term impacts of many policies may differ substantially, so a key challenge for impact evaluation is determining when to measure outcomes (King and Behrman 2009). For business training, one might expect firms to make some changes relatively quickly after training. However, the full impact of training may take some time. Impacts on business survival may also take time to materialize.

Table 6. Follow-up Survey Timing of Different Studies

<i>Study</i>	<i>Number of Follow-up Surveys</i>	<i>Months since Intervention</i>	<i>Attrition rate</i>
Berge et al. (2011)	2	5 to 7, 29-31	13 to 18 (c)
Bruhn and Zia (2012)	1	5 to 6	11
Calderon et al. (2012)	2	8, 28	15-26 (b)
De Mel et al. (2012)	4	4, 8, 16, 25	6 to 8
Drexler et al. (2012)	1	12	13 to 46 (a)
Field et al. (2010)	1	4	5.3
Giné and Mansuri (2011)	1	19-22	16
Glaub et al. (2012)	2	5, 12	11
Karlan and Valdivia (2011)	1	12 to 24	24
Klinger and Schündeln (2011)	1	12	28
Mano et al. (2012)	1	12	17
Premand et al. (2012)	1	9 to 12	7.2
Valdivia (2012)	1	10	18

(a) Attrition rate is 46 percent for business outcomes like sales, 13 percent for business practices.

(b) Rates are for first and second follow-ups respectively. Additionally note that 21 (50) percent of non-attriters had closed down by the first (second) follow-up surveys, so profit and revenue outcomes are on smaller sample.

(c) Note the study only surveys 644 out of the 1164 clients, based on accessibility by phone.

However, firms may begin some practices and then drop them, so surveys that measure what occurs in the business only several years after training may miss the period of experimentation. Ideally, studies should trace the trajectories of impacts, measuring both short- and long-term effects.

Table 6 provides details on the number of follow-up surveys, their timing, and their attrition rate for the different studies. The majority of studies that we review use a single follow-up survey, providing a snapshot of information on the training impact but no details on the trajectory of impacts. Eight of the 13 studies are very short-term studies that examine impacts one year or less after training. De Mel et al. (2012) find that the impacts differ in the short and medium term in their study. For example, in their study, examining impacts within the first year shows that business training for women out of the labor force led to large increases in business entry, whereas surveys 16 and 25 months after training shows that the control group had caught up in terms of business ownership rates.

Survey Attrition and Selective Survival or Start-up

Survey attrition is another problem that complicates inference, especially if the reasons for attrition are business failure, refusal because of disappointment with the training effects, or successful business owners moving out of the area. Attrition

rates range from as low as 5.3 percent in Field et al. (2010) and 6 percent to 8 percent in de Mel et al. (2012) to 24 percent in Karlan and Valdivia (2011) 26 percent in Calderon et al. (2012), and 28 percent in Klinger and Schündeln (2011).

Attempts to examine the impacts of training on business outcomes face additional difficulties when training influences the rate of business survivorship or the likelihood of business start-up. If training leads to the survival of relatively unsuccessful firms that would otherwise have closed, then a straight comparison of profits or sales by treatment status will understate the impact of training. Note that even if training has no impact on the rate of business survivorship or start-up, it may still affect the characteristics of which firms survive, requiring authors to use nonexperimental methods to address this selectivity. For example, de Mel et al. (2012) find that training (and grants) leads to changes in the characteristics of who opens businesses, even though the rates of ownership do not differ in the treatment and control groups. They therefore use a generalized propensity score to reweight their regression estimates to correct for the selectivity they find on observables such as ability and wealth.

Measurement Changed by Training

A final challenge in measuring the impact of business training on business outcomes is measuring those outcomes. Start-up and survivorship are objective measures that can be verified, whereas business practices, profitability, and revenues are difficult to measure for most firms. Business practices (for example, keeping accounts, separating business and household expenses, advertising in the past month) are normally relatively easy concepts for firms to understand and are questions that firm owners are usually willing to answer. However, Drexler et al. (2012) note that treated individuals may report performing certain behaviors (for example, separating personal and business accounts) because the training told them this was important rather than because they actually perform the behavior.

Measuring profits and revenues poses further problems. Owners of the smallest businesses typically do not keep written records of these items, and owners of larger firms who do keep records may be reluctant to share them. De Mel et al. (2009a) study several approaches to obtaining profits from microenterprises and conclude that, in their context at least, a simple, direct question is more accurate and much less noisy than calculating profits from revenues and expenses. However, collecting profits has proved difficult for many studies, and several studies have not collected profit data at all (Valdivia 2012; Klinger and Schündeln 2011), have collected it but not used it because of too much noise (Drexler et al. 2012), or have collected only profit margins on the main product rather than overall profits (Karlan and Valdivia 2011). Most studies have collected revenue data, but some have struggled with

much lower response rates for revenues than for nonfinancial business questions (for example, Drexler et al. (2012) have a 46 percent attrition rate on revenues compared to 13 percent for their questionnaire as a whole).

Even when studies are able to obtain data on profits and sales, business training may change the reporting of this data irrespective of whether it actually changes profits and sales. This may occur because the practices taught in the training course lead to more accurate accounting or because training recipients are less likely to underreport profit and sales levels because, for example, they trust the enumerators more after being given the training.⁶ Few studies to date attempt to address this issue. Exceptions are Drexler et al. (2011), who examine reporting errors (for example, reporting profits higher than sales or bad week sales higher than average sales) to determine whether treatment reduces these reporting errors and the difference between self-reported profits and profits calculated as the difference between revenue and expenses; Berge et al. (2011), who compare self-reported profits to revenue minus expenses for treatment versus control groups; and de Mel et al. (2012), who do the same and who control for detailed measures of accounting practices as a further robustness check. De Mel et al. (2012) find little evidence that training has changed reporting, whereas Drexler et al. (2012) find that their rule-of-thumb training reduces the number of errors in reporting, and Berge et al. (2011) find that training increases the gap between self-reported profits and revenue minus expenses.

Impacts of Business Training Interventions

The previous section highlights issues with statistical power, timing of follow ups, attrition, and measurement that present challenges for interpreting the impacts identified in the different studies. With these caveats in mind, we examine the extent to which business training is found to impact business start-up and survivorship, business practices, business outcomes, and outcomes for microfinance lenders. Because studies of other microenterprise interventions (De Mel et al. 2009b) often find differences by gender, we separate results by gender to the extent possible.

Impacts on Start-up and Survivorship

Table 7 summarizes the impacts of different studies on business survivorship and new business start-ups. The coefficients are marginal effects on the probability of either outcome occurring, so a coefficient of 0.06 can be interpreted as a 6 percentage point increase. Consider first the impact on business survival. Survivorship is difficult to examine when attrition rates are high because closing is often a cause of

Table 7. Impacts of Business Training on Business Start-up and Survival

Study	Gender	Impact on Survival		Impact on Start-up		
		Point estimate	95% CI	Point estimate	95% CI	
Bruhn and Zia (2012)	Mixed	0.013	(-0.09, +0.10)	0	n.r.	
	Female	-0.125	n.r., not sig.	0	n.r.	
	Male	0.072	(-0.07, 0.21)	0	n.r.	
Calderon et al. (2012)	Female	-0.034	(-0.13, +0.06)	n.r.	n.r.	
De Mel et al. (2012)	Current Enterprises	Female	-0.026	(-0.102, +0.051)	n.r.	n.r.
	Potential Enterprises	Female	n.r.	n.r.	+0.09 (4 months)	(0, 0.18)
Giné and Mansuri (2011)	Mixed	0.034	(-0.021, 0.089)	-0.006	(-0.02, +0.01)	
	Male	0.061	(-0.012, 0.133)	-0.011	(-0.04, +0.01)	
	Female	0.001	n.r., not sig.	0.002	n.r., not sig.	
Glaub et al. (2012)	Mixed	0.05	n.r.	n.r.	n.r.	
Karlan and Valdivia (2011)	Female	n.r.	n.r.	-0.019	(-0.05, +0.01)	
Klinger and Schündeln (2011)	selected in first phase	Mixed	n.r.	n.r.	0.044	(-0.12, 0.21)
	selected in first phase	Female	n.r.	n.r.	-0.019	(-0.31, +0.27)
	getting trained in second phase	Mixed	n.r.	n.r.	0.465	(0.10, 0.82)
		Female	n.r.	n.r.	0.572	(0.04, 1.10)
Mano et al. (2012)	Male	0.095	(0.022, 0.167)	n.r.	n.r.	
Premand et al. (2012)	Mixed	n.r.	n.r.	0.04	(0.02, 0.06)	
	Male	n.r.	n.r.	0.06	(0.04, 0.08)	
	Female	n.r.	n.r.	0.03	(0.01, 0.05)	
Valdivia (2012)	General training	Female	-0.045	(-0.094, +0.004)	0.014	(-0.03, +0.06)
	Training + technical assistance	Female	0.021	(-0.014, +0.056)	-0.006	(-0.05, +0.04)

Notes: 95% CI denotes 95 percent confidence interval. Impacts significant at the 10 percent level or more reported in bold. n.r. denotes not reported. Not sig. denotes point estimate is not significantly different from zero.

Berge et al. (2011) and Drexler et al. (2012) do not report impacts on either survivorship or start-up.

Note Valdivia (2012) survival is based on whether they stopped any business in the past two years, while start-up is based on whether they started a new business in the last year.

attrition, and bounds that allow for attrition can be very wide. Because many studies examine impacts over only a short time, rates of business failure are often low. However, there are exceptions. [Bruhn and Zia \(2012\)](#) find that 36 percent of businesses close during their study period in Bosnia, a rate that is due in part to the downturn caused by the global economic crisis, while [Calderon et al. \(2012\)](#) find that 50 percent of the nonattriting businesses close by the time of their second follow-up survey 28 months after training.

The only study with a survival effect significant at the 5 percent level is [Mano et al. \(2012\)](#), which finds a 9 percentage point increase in the likelihood of survival 12 months after training. These authors do not provide bounds for this effect that control for survey attrition, but they note that none of the training participants had closed. [Giné and Mansuri](#) find a 6 percent increase in the likelihood of survival 18 to 22 months after training for the male owners in their sample, an effect that is significant at the 10 percent level, but no change for female owners, whereas [Valdivia \(2012\)](#) finds that training leads to a marginally significant reduction in the likelihood of survival for female firm owners. He attributes this phenomenon to the possibility that training teaches owners to close losing firms. The remaining studies that report survivorship find insignificant impacts but with confidence intervals that are wide enough to include at least a 5 percentage point increase or decrease.

Studies that focus on existing firm owners sometimes consider the start-up of a second business, but none has found significant impacts. However, studies that focus on training specifically tailored for starting new businesses have found some impacts. [Klinger and Schündeln \(2011\)](#) find very large point estimates for entry one year after participation in the second phase of Technoserve's business plan competition in which training occurs, although the confidence intervals are very wide, and this impact includes the joint impact of grants given to the winners. [Premand et al. \(2012\)](#) examine a sample of 1,500 youths and find that participation in an entrepreneurship track rather than an academic track in the final year of university leads to an increase in self-employment rates of 6 percent for males and 3 percent for females one year later.⁷ Four months after training, [Field et al. \(2010\)](#) examine whether women reported business income over the preceding week, which reflects a combination of an effect on business start-up and an effect on survival. They find that upper-caste Hindu women who took the training were 19 percentage points more likely to report income, whereas the training had no effect on lower-caste Hindu women or on Muslim women. They attribute the lack of impact on these groups to social restrictions, arguing that training helped women whose businesses had been limited by social restrictions, but women who faced more extreme restrictions could not respond to training.

Training therefore appears to generate some short-run impacts on business start-up. However, this effect does not necessarily increase employment among trainees, who may simply switch from wage work. [Premand et al. \(2012\)](#) and [de Mel et al.](#)

(2012) both find that short-run increases in self-employment from training are coupled with reductions in the likelihood of wage work, so net employment effects on trained individuals are insignificant. Moreover, it is unclear whether training merely speeds up the rate of entry or permanently increases it. De Mel et al. (2012) find that training alone increases the rate of business ownership among a group of women out of the labor force by 9 percentage points within four months of the training, and giving these women grants increases this effect to 20 percentage points. However, by 16 and 25 months after training, the control group catches up. Given the short time horizon of the other studies that have found start-up impacts, it is unclear whether they too would show these effects dissipating over longer time horizons.

Impacts on Business Practices

A first link in the causal chain from business training to business profitability and growth is that business training improves the knowledge and implementation of business practices by business owners. There may be other potential mechanisms through which training affects business outcomes (for example, changing attitudes or work hours). However, failure to find any change in practices should cast doubt on the ability of the training to improve firm outcomes.

Table 8 summarizes the impacts identified by various studies on business practices. Almost all studies find a positive effect of business training on business practices, although the effect is often not significant once the sample is divided by gender. Studies differ in what specific practices they measure, how comprehensively they measure them, and how (if at all) they aggregate them. Several studies measure only one to three basic practices, such as Calderon et al. (2012), who examine whether the firm uses formal accounting, and Mano et al. (2011) who record whether the firm keeps records, whether it analyzes them, and whether it visits customers. Others record a broader range of practices, including different types of record keeping, different marketing activities, and other specific practices taught in the training.

One common approach to aggregating different practices is to normalize each practice as a z score (subtracting the mean and dividing by the standard deviation) and then to average these z scores. A coefficient of 0.03, as in Karlan and Valdivia (2011), is interpreted as an impact of 3 percent of a standard deviation. This is useful for considering the magnitude of the increase in relative terms, but it does not provide much guidance regarding the absolute size of the effect. Alternatively, one can examine the percentage point increase in the likelihood that a particular practice will be implemented or the change in the number of practices implemented out of some total, both of which provide more guidance on the absolute magnitude of the increase.

Table 8. Impact of Business Training on Business Practices

Study	Gender	Units	Number of Practices		95% CI
				Point estimate	
Berge et al. (2011)	Male	p.p.	4 (a)	0.03-0.08	n.a.
	Female	p.p.	4 (a)	-0.02-0.00	n.a.
Bruhn and Zia (2012)	Mixed	s.d.	3	0.272	(0.03, +0.51)
	Male	s.d.	3	0.290	(0.01, 0.57)
	Female	s.d.	3	0.214	n.r.
Calderon et al. (2012) De Mel et al. (2012)	Female	p.p.	1	0.062	(-0.02, +0.14)
	Current Enterprises	Female	num	2.03	(1.27, 3.30)
Potential Enterprises	Female	num	29	0.87	(-0.23, +1.97)
Drexler et al. (2012)	"Rule-of-thumb"	Mostly Female	s.d.	0.14	(0.06, 0.22)
	"Standard"	Mostly Female	s.d.	0.07	(-0.03, 0.17)
Giné and Mansuri (2011)	Mixed	s.d.	3	0.131	(0.01, 0.25)
	Male	s.d.	3	0.114	(-0.05, 0.28)
	Female	s.d.	3	0.140	n.r.
Karlan and Valdivia (2011)	Mostly Female	s.d.	14	0.03	(0.00, 0.06)
Mano et al. (2012)	Male	p.p.	3	0.24-0.42	n.a.
Valdivia (2012)	General training	Female	s.d.	11	(-0.02, +0.04)
	Training + technical assistance	Female	s.d.	11	0.05

Notes: 95% CI denotes 95 percent confidence interval. Impacts significant at the 10 percent level or more reported in bold. Units for measuring practices are either standard deviations of a normalized aggregate (s.d.), percentage points (p.p.), or number of distinct practices improved (num).

Number of practices is the total number of practices measured.

When no aggregate measure is reported, the range of point estimates for individual practices is given. n.r. denotes not reported. n.a. denotes not applicable since range of estimates given.

(a) we include here their index of three marketing practices, plus their result on record-keeping. No aggregate measure is provided.

Many studies find baseline levels of business practices that are relatively low. For example, [Giné and Mansuri \(2011\)](#) report that only 18 percent of firms record money taken from the business, and only 18 percent record sales. Even among larger metalwork firms, [Mano et al. \(2012\)](#) report that only 27 percent of their sample keep business records, and only 20 percent visit customers at baseline. Although most studies find significant increases in the use of business practices taught during the training, the magnitude of these effects, although sometimes large in relative terms, is often small in absolute terms. For example, [Drexler et al. \(2012\)](#) find that rule-of-thumb training leads to an increase in individuals reporting that they separate personal and business expenses, keep accounting records, and calculate revenues formally, with each of these measures increasing 6 to 12 percentage points relative to the control group. In [Giné and Mansuri \(2011\)](#), treatment

impacts include a 6.6 percentage point increase in recording sales and a 7.6 percentage point increase in recording money taken for household needs. In [de Mel et al. \(2012\)](#), existing enterprises implement an additional two practices out of 29. [Mano et al. \(2012\)](#) are an exception in this regard: they find a 30 percentage point increase in the percentage of firms keeping records in the treatment versus the control group. However, in general, given that the magnitude of the changes in business practices is relatively small, we might expect it to be difficult to detect impacts of these changes on business outcomes.

Impacts on Business Profits and Sales

Ultimately, from the viewpoint of an individual firm owner, an investment in training is justified only if there is an increase in profits. However, as noted previously, many studies struggle to measure profits, so not all studies consider this as an outcome. Table 9 summarizes those studies that do, converting, where necessary, point estimates of profit or sales levels to percentage increases relative to the control group mean to enhance comparability across studies. Several studies examine gender heterogeneity by reporting a point estimate for males and then an interaction effect for females, but they do not test the overall impact on females. Therefore, the table sometimes shows confidence intervals for males but not for females. Often, studies have more than one specification for profits or revenues, with variation in whether they include different controls and whether they truncate or trim the data or take a log transformation. We report impacts on the measure that corresponds most closely to profits or sales in the previous month. The data shown in the table do not account for differential attrition, though some studies report bounds that adjust for attrition.

The table shows that few studies detect significant impacts of business training on business profits or sales, although the confidence intervals are very wide in many cases. The wide confidence intervals reflect the issue of statistical power discussed earlier. The studies that have the most power according to the calculations in table 5 are the ones that are most likely to show significant effects. [Berge et al. \(2011\)](#) find that training increases profits by 24 percent and sales by 29 percent for males in the short run (five to seven months posttraining), but the point estimate of the impact on profits drops to 5 percent and is statistically insignificant in their longer-term follow up (30 months posttraining). There is a continued and marginally significant impact on sales.⁸ Their point estimates are much closer to zero and statistically insignificant for women in both the short and medium term. [De Mel et al.'s \(2012\)](#) study also has enough power to detect reasonable changes in profits. They find no impact of training alone on profits of existing firms over either the short or medium run, but they do find significant impacts of the combination of training and a grant on short-run profits, with these gains dissipating over time. In

Table 9. Impacts on Business Profits and Sales

Study	Gender	Profits		Revenues	
		% increase	95% CI	% increase	95% CI
Berge et al. (2011)					
	Male	5.4%	(-20%, +38%)	31.0%	(-4%, +79%)
	Female	-3.0%	(-23%, +22%)	4.4%	(-23%, +22%)
Bruhn and Zia (2012)	Mixed	-15%	(-62%, +32%)	n.r.	n.r.
Calderon et al. (2012)	Female	24.4%	(-1%, 56%)	20.0%	(-2%, +47%)
De Mel et al. (2012)					
Current Enterprises	Female	-5.4%	(-44%, +33%)	-14.1%	(-68%, +40%)
Potential Enterprises	Female	43%	(+6%, +80%)	40.9%	(-6%, +87%)
Drexler et al. (2012)					
"Standard"	Mostly Female	n.r.	n.r.	-6.7%	(-24.5%, +11.2%)
"Rule-of-thumb"	Mostly Female	n.r.	n.r.	6.5%	(-11.4%, +24.4%)
Giné and Mansuri (2011)	Mixed	-11.4%	(-33%, +17%)	-2.3%	(-15%, +13%)
	Male	-4.3%	(-34%, +38%)	4.8%	(-14%, +27%)
	Female	n.r. (a)	n.r.	n.r. (a)	n.r.
Glaub et al. (2012)	Mixed	n.r.	n.r.	57.4% (c)	n.r.
Karlan and Valdivia (2011)	Mostly Female	17% (b)	(-25%, +59%)	1.9%	(-9.8%, +15.1%)
Mano et al. (2012)	Male	54%	(-47%, +82%)	22.7%	(-31%, +76%)
Valdivia (2012)					
General training	Female	n.r.	n.r.	9%	(-8%, +29%)
Training + technical assistance	Female	n.r.	n.r.	20.4%	(+6%, 37%)

Notes: 95% CI denotes 95 percent confidence interval. Impacts significant at the 10 percent level or more reported in **bold**. n.r. denotes not reported.

(a) They look at an aggregate sales and profitability measure and find no significant impact for either gender.

(b) Impact on profit from main product.

(c) Calculated as difference-in-difference calculation. Study reports difference in log sales is significant at the 1 percent level. Profit increases are scaled as a percentage of the control group mean to enable comparability.

When multiple rounds are used, longest-term impacts available are reported.

a separate sample of women who were out of the labor force at baseline, training increased the profits and sales of start-up businesses by a statistically significant 40 percent, although the confidence intervals around this level are wide.

Calderon et al. (2012) find a 24 percent increase in weekly profits and a 20 percent increase in weekly revenues, both significant at the 10 percent level. However, given that attrition is 26 percent by the second round survey and that 50 percent of the nonattritors have closed, there is reason to be cautious in interpreting this estimate of the impact on surviving nonattriting firms. The only other study to find significant impacts on revenues, Valdivia (2012), finds a 20 percent increase for the group that received both training and intensive one-on-one technical

assistance but no significant increase for training alone. Finally, [Glaub et al. \(2012\)](#) find a positive effect of personal initiative training on sales one year later, although they do not survey the noncompliers (individuals selected for training who do not attend), which is problematic if there is selective participation.

Several studies have emphasized the possibility that business training may have its strongest impact on sales during a bad month. The working paper version of [Karlan and Valdivia \(2011\)](#) stressed this avenue, noting that training might help clients identify strategies to reduce downward fluctuations in sales by considering diversifying the products that they offer and by being more proactive about alternative activities during slow months. The working paper estimate, which has gained some policy attention, showed a 30 percent increase in sales during a bad month. However, the published version of the paper deemphasizes this impact, noting that when an alternative (and now preferred) specification is used, the impact falls to an insignificant 5 percent to 7 percent increase. The possibility that training may be particularly valuable during bad times is also emphasized by [Drexler et al. \(2012\)](#), who find that their rule-of-thumb training leads to an increase in sales during bad weeks that is significant at the 10 percent level. However, Drexler et al. also ask firm owners to report sales in a bad month and find a very small and insignificant impact of training on this measure. [Giné and Mansuri \(2011\)](#), [de Mel et al. \(2012\)](#), and [Valdivia \(2012\)](#) find no significant impacts of training alone on sales during bad months. Viewing these studies together leads us to conclude that the evidence that training has particularly strong effects during bad periods is weak.

A microenterprise earning \$1 per day would need to see only a 13.7 percent increase in profits to recoup the cost of \$100 of training over two years. The confidence intervals for the studies that consider profits are almost all wide enough to include this level of return. For larger firms, the percentage increase in profits required to repay training costs is likely lower because the costs of training often increase more slowly than the size of the firm undertaking the training.⁹ For example, a firm with \$500 in monthly profits would only need a 2 percent increase in monthly profits to recoup \$250 worth of training costs over two years. The result is that training costs may be justified by increases in profits that are far too small for existing studies to detect.

Impacts on Employment

A further justification by policymakers for subsidizing business training is that business growth may have broader benefits for others in the community by increasing employment opportunities. For programs working with microenterprises, the most direct employment impacts are likely to be for the owner himself or herself, increasing employment by increasing the likelihood of starting a new business or reducing the chance of business failure.

The few studies using samples of microenterprises that report impacts on employment of other workers robustly show very small and statistically insignificant effects. Karlan and Valdivia (2011) find an increase of 0.02 workers, Valdivia (2012) finds a decrease of 0.06 workers from straight training and a similar decrease from training plus technical assistance, and Drexler et al. (2012) find an increase of 0.05 workers from standard training and a decrease of 0.02 workers from rule-of-thumb training. None of these impacts is statistically significant, but their point estimates suggest that no more than one in 20 microenterprises that take business training will hire an additional worker.

The one study to show a stronger employment effect is Glaub et al. (2012), which hints at the possibility of employment impacts when training larger firms. These authors find that employment in treated firms grows from 7.9 employees at baseline to 10.7 at follow up, whereas employment in control firms falls from 6.6 employees at baseline to 5.0 at follow up. This difference is significant at the 5 percent level. Their sample is small, and they drop noncompliers to their treatment, so this result is likely an overstatement of the effect. More studies with larger firms are needed.

Impacts on Microfinance Institution Outcomes

Because many of the studies work with microfinance clients, they also consider outcomes using administrative data from the microfinance organization. These data have the advantage of being available with less attrition and over longer periods, and they are useful for assessing whether offering training is cost effective for the microfinance organization. However, these data are less useful for explaining how such training affects firms. Karlan and Zinman (2011) find that training results in a 4 percentage point increase in client retention rates and a 2 to 3 percentage point increase in the likelihood of perfect repayment (although this is only marginally significant). However, they also note that some of the clients who leave cite the added length of the weekly meetings due to the training sessions as a factor in dropping out of the program. They note that these benefits appear to make the training profitable from the lender side. After their study, FINCA Peru implemented the mandatory version of their training in all village banks.

Giné and Mansuri (2011) find that training leads to a 16 percent increase in loan size for males, a reduction in loan size for females, and no change in repayment rates. They also find a change in the selection of who borrows; individuals with higher predicted probabilities of default are less likely to borrow after training. Field et al. (2010) find that upper-caste Hindu women are 13 percentage points more likely to borrow after training. In contrast, Drexler et al. (2012) and Bruhn and Zia (2012) find no significant impacts of training on the likelihood of taking loans or loan size, although Bruhn and Zia find an increase in loan duration and

more refinancing of loans. They attribute this finding to trained individuals making longer-term investments and being more aware of available interest rates.

Boosting the Intensity and Working with Larger Firms

Many of the training sessions are relatively brief, and the increase in business practices has been relatively small in a number of studies. One response to this phenomenon is that more in-depth and individualized follow ups on the training are needed, whereas another response is to focus on larger firms in which management practices may be of greater importance. We discuss the results of studies that have pursued these two approaches.

Individualized Follow Ups

Three of the business training evaluations had a treatment group that added individualized follow ups to the classroom training. In [Drexler et al. \(2012\)](#), trainers visited eight times over five months to answer queries, verify and encourage the use of accounting books, and correct any mistakes in completing books. These authors find no significant effects of this additional follow up. [Giné and Mansuri \(2011\)](#) added “hand-holding sessions” in half of the community organizations, with firms receiving visits one to two times per month for four months to discuss topics learned, answer questions, and suggest solutions to potential problems. They find that this hand holding had no effect on any of the aggregate outcomes for either men or women.

In both of these cases, the follow ups mostly reinforced the general business skills taught in training rather than providing firm-specific individualized advice. [Valdivia \(2012\)](#) examines more intensive follow up, with trainers providing specific technical assistance tailored to the needs of women’s businesses. The follow ups combined individual visits with group sessions among small groups of similar businesses during a three-month period. This component included 22 three-hour group sessions and five to six hours of individual sessions or visits. Valdivia finds some evidence to suggest this technical assistance helped firms; women assigned to receive the assistance experienced a 20 percent increase in revenue relative to the control group (significant at the 1 percent level) and showed more improvement in business practices than women who were assigned to only the basic training. This additional attention cost twice as much as the basic training alone.

Individual Consulting

A related body of literature examines the impact of providing consulting services on a one-on-one basis to firms to improve business and management practices. The

closest study to the business training experiments is the work of [Karlan et al. \(2012\)](#), who examine a mix of 160 male and female tailors in Ghana with five or fewer workers. Their study used local consultants from Ernst and Young in Ghana, who met with the tailors for 30 minutes to 1 hour several times a month over one year, with the average firm receiving 10 hours of consulting over a year at no cost to the firm. They find that some of the consultants' recommendations were adapted for some months but had been abandoned one year after training stopped. There is no significant impact of either treatment on profits or revenues, with some specifications showing negative effects in the short run, although the power is very low and confidence intervals are wide.

[Bruhn et al. \(2012\)](#) evaluate a state government program in Puebla, Mexico, that paired small businesses with a consultant from one of several local consulting firms. Consultants spent approximately four hours per week over a year assisting the firm in overcoming constraints to growth. A total of 432 firms applied to the program, and 150 were chosen to receive heavily subsidized consulting services (at a cost that was approximately 10 percent of the commercial rate). The mean number of employees was 14, and 72 percent were male-owned firms. The training impact was assessed with a single follow-up survey one to three months after the consulting. The authors find large point estimates for the impacts on sales and profits, which are sometimes significant depending on the measure used and the extent of trimming. However, the study faces many of the same challenges as the business training studies reviewed above. First, the firms in the sample are very heterogeneous, with a baseline coefficient of variation in sales of 3.7, and 2.4 even after trimming the top 1 percent. Second, even though all firms signed a statement of interest, only 80 of the 150 firms (53 percent) assigned to treatment participated in the consulting. Third, attrition rates were reasonably high, and there was additional item nonresponse on profits and sales even among those who were interviewed, so only 288 firms (66.7 percent) provided data on profits in the follow-up survey. These challenges are likely to face any similar government program offering subsidized consulting or business services to firms, such as the matching grant programs used in many World Bank private sector loans.

The final individualized consulting study is [Bloom et al. \(2013\)](#), who focus on a much smaller sample of 17 large textile firms in India. The typical firm in their sample has 270 employees, two plants, and sales of \$7.5 million per year. They provided 11 of these firms with five months of free intensive consulting from Accenture Consulting. The consultants averaged 781 hours per treated plant, working with the firms to implement 38 key management practices related to quality control, factory operations, inventory, human resource management, and sales and order management. They address the problem of small sample size by focusing on very homogeneous firms and collecting large amounts of data from them, including weekly data on quality, output, and inventories. They find that adopting these

management practices raised productivity by 17 percent in the first year through improved quality and efficiency and reduced inventory, and they find some evidence that within three years, adopting these practices led to the opening of more production plants. The results show that in large firms, at least, changing management practices can lead to substantial improvements in firm performance. However, the authors can only indirectly estimate the changes in profits from this effort.

What We Do Not Know

There are now a range of studies on a variety of business training programs that examine impacts on business practices, business outcomes, and (sometimes) outcomes for microfinance institutions. However, existing studies leave a number of open questions that are important in considerations of the case for policy action to support business training.

Who Does Training Help Most?

Our discussion above touches on heterogeneity in outcomes by the gender of the owner and, to some extent, across studies by firm size. Several studies have examined heterogeneity in other dimensions, such as the owner's education and baseline business skill levels, business sector, and interest in training. However, the low power of most studies to find average effects for the full sample indicates low power for examining the heterogeneity of effects.

As a result, the question of who benefits most from training—or which types of training are most suitable for which types of firms—remains unanswered. On one hand, poor subsistence firms whose owners run the business only because they cannot find a wage job may have very low business skills. Thus, it should be relatively easy for them to make improvements. However, the owners may be less interested or able to implement the practices taught, or these practices may only have an effect when businesses reach a larger scale. There is much talk of targeting gazelles—firms that grow rapidly—but even if the characteristics to identify such firms in advance can be defined, it is unclear whether these firms need the help or would grow rapidly anyway. Theoretically, it would be preferable to target firms in which skills are the binding constraint on growth, but there is little evidence to date to determine which firms these are, especially among the smallest firms.

How Does Training Help Firms, and Do Gains Come at the Expense of Other Firms?

Most studies have not explored the channels through which training affects business outcomes. In part, this omission reflects the lack of power in detecting an impact on profits in the first place. Does training enable firm owners to use the same inputs more efficiently—thereby reducing costs and wastage—or is the main impact due to increasing revenues at the same cost ratios by new marketing and sales efforts? The policy implications differ depending on the channels. In particular, one possibility is that gains for the treated firms are due to these firms taking customers from other firms.

Such spillovers have implications for both internal and external validity. If the increased sales are mainly due to taking business from the control group firms, then the stable unit treatment value assumption, which assumes that the outcomes of each firm are not affected by the treatment statuses of other firms, is violated. As a result, the experimental estimate no longer provides the average impact of training for the sample population. If the increased sales are mainly due to other firms not in the sample, the results of the experiment could be misleading with respect to the gain to society from scaling up the training program. It should be noted that spillovers might instead be positive if control or nonsample firms copy some of the techniques or new products introduced by firms that have participated in training. Indeed, this possibility is often given as one of the main justifications for public subsidies of matching grant programs that subsidize the purchase of business development services by SMEs. These issues are part of the broader question of how competition responds to newly trained firms. We do not know whether this deters some new firms from entering the industry, causes others to exit, or causes the incumbents who remain in business to make other changes to the way they run their businesses.

To investigate this issue, a much larger sample is needed. Experimental variation in the intensity of the treatment within different geographical areas could be used to test for and measure these spillovers. An example in the context of labor programs for youth is found in Crepón et al. (2011). A first attempt in this direction for business training is found in Calderon et al. (2012), who randomly assigned 17 villages into seven treatment villages and 10 control villages, with half the individuals in the treated villages assigned to training. Their preliminary analysis surprisingly finds little evidence for spillovers despite working in remote villages with 1,500 or fewer households and with firms that mostly make or sell goods for local consumption. However, it is unclear how much power is available to examine these spillovers given the relatively small number of villages included in the study.

Do Larger Impacts Emerge over Time?

Most of the studies take a single snapshot of the impact of training a relatively short time after training has ended. Two studies that have traced the trajectories of impacts suggest that effects may vary considerably over time. In [de Mel et al. \(2012\)](#), the impacts on business start-up fade over time as control firms catch up. [Bloom et al. \(2013\)](#) find that introducing management practices in larger firms shows immediate effects on quality and then slowly leads to changes in inventory levels, output, and productivity. The impacts begin to appear in terms of employment generation (through new plants opening) only after several years of using these practices. Given the interest of many governments in employment creation, studies that consider only a year or so after treatment may miss effects that take some time to be realized—or, conversely, we may find that effects that seem promising in the short term dissipate over time.

If Training Is so Helpful, Why Do Firms not Purchase It?

It is notable that all of the business training studies reviewed here offer the training for free, as do two out of the three consulting experiments, with the other offering a 90 percent subsidy. In part, this approach is used for research purposes, to ensure sufficiently high take-up and to provide evidence on how training influences a range of firms. Even with this approach, we have seen limits to demand, with some studies struggling to encourage people to undertake training even when it is offered without cost.

As a result, we know very little about what types of firms would choose to purchase training at market prices and the effects of training on this subgroup of firms. Public intervention is typically motivated by the belief that market failures prevent firms who would benefit substantially from training from purchasing this training at market prices or a belief that there are positive externalities from training that lead firms to underinvest relative to what is socially desirable.

Even if market failures exist, the first-best solution would be to fix these market failures rather than to give away training for free or highly subsidized rates. However, given the difficulty of alleviating some of these market failures in many developing countries, subsidizing training may be seen as a feasible second-best solution. Several potential constraints or market failures are discussed in the literature. The first, and the one for which there is the most support ([Karlan and Valdivia 2011](#); [Bloom et al. 2013](#)), is that of an information failure: entrepreneurs do not understand the value of business training. Those with the most to gain may understate the value the most because they do not realize how poorly their firms are run.

A second market failure is credit constraints. Firms may find it more difficult to borrow to finance training, an intangible asset, than to finance assets that could be

seized by a bank in the event of nonrepayment. There is strong evidence that many microenterprises are credit constrained (de Mel et al. 2009b), but there is much weaker evidence to support the view that this is the key constraint to purchasing business training services.

A third possibility is the failure of insurance markets. Firm owners may be reluctant to take training even if they think it has a high expected payoff because they are unable to insure against the possibility that it will not work. There is some recent evidence to support the view that risk is a constraint to start-up and investment in small businesses (Bianchi and Bobba Forthcoming), but no evidence of which we are aware shows that alleviating this constraint leads to more purchases of training.

A fourth possibility is supply-side constraints. Consulting or training services simply may not exist in the market. Thus, even if a firm wants to purchase these services, it is unable to. This is likely to be true in some countries and areas, but in many others, such services do exist.

Even with market failures, public financing is not justified if the gains to training are realized entirely by the firms being trained unless the financing is provided with the goals of either poverty alleviation (raising the incomes of these particular firm owners) or productivity enhancement. More typically, public spending is justified by claims of positive spillovers, whereby the public gains from training are believed to greatly exceed the private gains, causing firms to underinvest. Such externalities have yet to be demonstrated empirically.

The optimal policy response differs depending on which constraint binds, so making progress on the issue of why more firms do not purchase training or consulting is likely to have useful implications for policy efforts.

Practices or Personality?

Business training courses have traditionally focused on teaching particular practices that firm owners can implement in their firms. However, another school of thought is that the attitudes and personalities that business owners bring to the business are equally, if not more, important. Premand et al. (2012) report that one of the main objectives of the trainers in their study was to change the students' personalities to "turn them into entrepreneurs." They find that their intervention led to measurable and significant changes in several domains of personality.

There is also a range of training courses studied by psychologists that focus more on the personality of an entrepreneur than on specific skills (Glaub and Frese 2011). Glaub et al. (2012) find some evidence to support a positive impact of such training in Uganda. Although several studies have incorporated some aspect of aspirations or entrepreneurial attitudes into their content, to date, no research tests the relative contribution of each type of training.

Conclusions and Suggestions for Future Work

The last few years have seen rapid growth in the number of randomized trials evaluating business training programs, providing a growing body of evidence in an area with large policy interest. However, a number of challenges have hampered how much we can learn from these studies. Methodological concerns and heterogeneity in both training content and the characteristics of who is trained complicate comparisons across studies. Many of the key questions needed to justify large-scale policy interventions in this area remain unanswered. Researchers continue to learn more about how to better conduct firm experiments, suggesting that these difficulties are not insurmountable.

To learn from the next generation of studies, we believe that the following elements are needed.

1. Analyzing much larger samples or more homogeneous firms: Rather than more studies with 100 to 500 individuals in each treatment or control group, we need studies to move to samples of several thousand or more. This would increase the power of the studies and allow more consideration of the types of people for whom training is most effective. An alternative to large cross-sectional samples is to reduce the heterogeneity of the sample by focusing on firms within one industry and size category and collecting much more frequent time series data on these firms (McKenzie 2011, 2012).

2. Using better measurement of outcomes: Measuring firm profits and revenues has proved to be a challenge for many studies, and little evidence is available on how training changes a firm's production process. Further efforts to improve the measurement of financial information (and to ensure that there is not simply a measurement effect of training) are needed. Focus on a specific industry or sector may allow more detailed production-level monitoring of physical outputs and inputs.

3. Designing experiments to measure spillovers: These experiments could include greater use of global positioning system data to measure local spillovers (Gibson and McKenzie 2007) and randomization of the intensity of training at the local market level to determine whether effects differ when all firms competing in a local area are trained versus when only some of them are trained, building on the work of Calderon et al. (2012).

4. Measuring trajectories of outcomes over longer periods: The impacts of training may differ in the short and medium term. Measuring outcomes at multiple points in time would enable better understanding of whether effects take time to materialize or whether effects that emerge quickly persist.

5. Testing which elements of content matter: With larger samples, studies could build on the work of Drexler et al. (2012) and test different forms of training to determine which elements of business skills have the greatest impact and whether training should focus on entrepreneurial personality as well as processes.

However, researchers should avoid the temptation to perform this testing at the cost of insufficient power in each treatment arm.

6. Understanding market failures and building market-based solutions: Almost every study has given training away for free and experienced difficulties in take-up. There are many open questions concerning the development of a market for these business services and the types of policies that could overcome the market failures that prevent firms from using these markets.

Funding

This work was supported by the World Bank and the UN Foundation.

Acknowledgments

We thank the editor, three anonymous reviewers, Louise Fox, and Markus Goldstein for helpful comments and suggestions, and we thank the authors of the studies reviewed in this paper for providing us with additional details and clarifications on their studies.

Notes

David McKenzie is Lead Economist in the Finance and Private Sector Development group of the World Bank's Research Development; email address: dmckenzie@worldbank.org. Christopher Woodruff is Professor of Economics at the University of Warwick.

1. See <http://www.ilo.org/empent/areas/start-and-improve-your-business/lang-en/index.htm> [accessed September 6 2012].

2. All dollar amounts are U.S. dollars unless otherwise indicated.

3. Kaizen and 5S are Japanese systems for improving production efficiency based on a philosophy of continuous improvement and by improving workflow in a production process through standardized and efficient storage, set-up, and production.

4. See the working paper (McKenzie and Woodruff 2012) for a more technical discussion of the calculation of power in this table.

5. Note that Berge et al. (2011) take existing loan groups who meet on a given day in a given branch and randomly assign training to one of two days in each of the two branches. Thus, true randomization only involves choosing one of four possible allocations and has zero power according to permutation analysis. The authors claim that because loan groups are offered time on the basis of availability, this is as good as random, and so they proceed with analysis as if randomization was at the group level. Our table does the same, but this caveat should be noted.

6. A related concern is that people who take training may overreport profits or revenues after training to exaggerate how well their firms have benefited from training. The same robustness checks as described in the text can help to rule out this sort of behavior, as can detailed probing and observation from the surveyors.

7. This effect includes the impact of seed money given to the top placed business plans, but the authors argue via various checks that the impact is not driven by these grants.

8. The sales impact is insignificant when covariates are dropped or clustering is used to attempt to address the fact that randomization did not occur at the loan group level.

9. If we account for discount rates, opportunity costs, and risk aversion, the desired returns would have to be higher. However, a 25 percent increase in profits would still likely provide a very reasonable return to microenterprise training, even after accounting for these factors.

References

- Berge, L., I. Oppedal, K. Bjorvatn, and B. Tungodden. 2011. "Human and Financial Capital for Microenterprise Development: Evidence from a Field and Lab Experiment." NHH Discussion Paper Sam 1 2011. Norwegian School of Economics, Bergen, Norway.
- Betcherman, G., K. Olivas, and A. Dar. 2004. "Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries." World Bank Social Protection Discussion Paper no. 402. World Bank, Washington, DC.
- Bianchi, M., and M. Bobba. Forthcoming. "Liquidity, Risk, and Occupational Choices." *Review of Economic Studies*.
- Bjorvatn, K., and B. Tungodden. 2010. "Teaching Business in Tanzania: Evaluating Participating and Performance." *Journal of the European Economic Association* 8 (2–3): 561–70.
- Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics* 128 (1): 1–51.
- Bloom, N., and J. van Reenen. 2010. "Why Do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives* 24 (1): 203–24.
- Bruhn, M., D. Karlan, and A. 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. Yale University, New Haven, CT.
- . 2010. "What Capital is Missing in Developing Countries?" *American Economic Review Papers and Proceedings* 100 (2): 629–33.
- Bruhn, M., and B. Zia. 2012. "Stimulating Managerial Capital in Emerging Markets: The Impact of Business and Financial Literacy for Young Entrepreneurs." Mimeo. World Bank, Washington, DC.
- Calderon, G., J. Cunha, and G. de Giorgi. 2012. "Business Literacy and Development: Evidence from a Randomized Trial in Rural Mexico." Mimeo. Stanford University, Stanford, CA.
- Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora. 2011. "Do Labor Market Policies Have Displacement Effect? Evidence from a Cluster Randomized Experiment." Mimeo.
- Dar, A., and Z. Tzannatos. 1999. "Active Labor Market Programs: A Review of the Evidence from Evaluations." Social Protection Discussion Paper no. 9901. World Bank, Washington, DC.
- De Mel, S., D. McKenzie, and C. Woodruff. 2012. "Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka." Mimeo. World Bank, Washington, DC.
- . 2009a. "Measuring Microenterprise Profits: Must We Ask How the Sausage is Made?" *Journal of Development Economics* 88 (1): 19–31.
- . 2009b. "Are Women more Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns." *American Economic Journal: Applied Economics* 1 (3): 1–32.
- Drexler, A., G. Fischer, and A. Schoar. 2012. "Keeping it Simple: Financial Literacy and Rules of Thumb." Mimeo. London School of Economics.
- Duflo, E., R. Glennerster, and M. Kremer. 2008. "Using Randomization in Development Economics Research: A Toolkit." In T. P. Schultz and J. Strauss, eds., *Handbook of Development Economics*, Vol. 4, 3895–962. Amsterdam, The Netherlands: North Holland.

- Field, E., S. Jayachandran, and R. Pande. 2010. "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *American Economic Review Papers and Proceedings* 100 (2): 125–29.
- Gibson, J., and D. McKenzie. 2007. "Using Global Positioning Systems in Household Surveys for Better Economics and Better Policy." *World Bank Research Observer* 22 (2): 217–41.
- Giné, X., and G. Mansuri. 2011. "Money or Ideas? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan." Mimeo. World Bank, Washington, DC.
- Glaub, M., and M. Frese. 2011. "A Critical Review of the Effects of Entrepreneurship Training in Developing Countries." *Enterprise Development and Microfinance* 22 (4): 335–53.
- Glaub, M., M. Frese, S. Fischer, and M. Hoppe. 2012. "A Psychological Personal Initiative Training Enhances Business Success of African Business Owners." Mimeo. National University of Singapore Business School.
- Karlan, D., R. Knight, and C. Udry. 2012. "Hoping to Win, Expected to Lose: Theory and Lessons on Microenterprise Development." Mimeo. Yale University, New Haven, CT.
- Karlan, D., and M. Valdivia. 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *Review of Economics and Statistics* 93 (2): 510–27.
- King, E., and J. Behrman. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *World Bank Research Observer* 24 (1): 55–82.
- Klinger, B., and M. Schündeln. 2011. "Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America." *World Development* 39 (9): 1592–610.
- Mano, Y., A. Iddrisu, Y. Yoshino, and T. Sonobe. 2012. "How Can Micro and Small Enterprises in Sub-Saharan Africa Become More Productive? The Impacts of Experimental Basic Managerial Training." *World Development* 40 (3): 458–68.
- McKenzie, D. 2010. "Impact Assessments in Finance and Private Sector Development: What Have We Learned and What Should We Learn?" *World Bank Research Observer* 25 (2): 209–33.
- . 2011. "How Can We Learn Whether Firm Policies Are Working in Africa? Challenges (and Solutions?) for Experiments and Structural Models." *Journal of African Economics* 20 (4): 600–25.
- . 2012. "Beyond Baseline and Follow-Up: the Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210–21.
- McKenzie, D., and C. Woodruff. 2012. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" World Bank Policy Research Working Paper no. 6202. World Bank, Washington, DC.
- Premand, P., S. Brodmann, R. Almeida, R. Grun, and M. Barouni. 2012. "Entrepreneurship Training and Self-employment among University Graduates: Evidence from a Randomized Trial in Tunisia." Mimeo. World Bank, Washington, DC.
- Schulz, K., and D. Grimes. 2005. "Sample Size Calculations in Randomised Trials: Mandatory and Mystical." *Lancet* 365:1348–53.
- Sonobe, T., A. Suzuki, and K. Otsuka. 2011. "Kaizen for Managerial Skills Improvement in Small and Medium Enterprises: An Impact Evaluation Study." The World Bank, Washington, DC. <http://siteresources.worldbank.org/DEC/Resources/FinalVolumeIV.pdf>
- Valdivia, M. 2012. "Training or Technical Assistance for Female Entrepreneurship? Evidence from a Field Experiment in Peru." Mimeo. GRADE.