

Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries

Michael Grimm^{a, b, c} and Anna Luisa Paffhausen^a

^a University of Passau, Innstrasse 29, 94032 Passau, Germany; Ph.: ++49-851-5093310, F: ++49-851-5093312; Email: michael.grimm@uni-passau.de; anna.paffhausen@uni-passau.de.

^b Erasmus University Rotterdam, Kortenaerkade 12, 2518 AX The Hague, The Netherlands; Email: grimm@iss.nl.

^c Institute for the Study of Labor (IZA), Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany.

Corresponding author: Michael Grimm (michael.grimm@uni-passau).

Abstract – Worldwide, 600 million jobs are needed over the next 15 years to keep employment rates at their current level. Because most employment in low and middle income countries is in micro-, small and medium-sized enterprises, governments, non-governmental organizations and donors spend on targeted programs and broader policies to enhance employment creation in these firms. But despite these efforts, not much is known about which of these interventions are really effective. This systematic review synthesizes the existing evidence on the employment impact of these programs. The results show that the effects have so far been very modest. Even if many interventions were relatively successful in boosting self-employment, expanding employment in already existing firms is generally more difficult but eventually easier in somewhat larger firms compared to very small firms. This finding is also true in relative terms, but it is probably not fully independent from the contexts in which firms of different sizes have been observed. The effects of finance interventions have on average been weaker than the effects of entrepreneurship training or business development services. Our study also reveals that about a third of the interventions covered by this review are not primarily designed to create employment but rather strive for income stabilization and poverty reduction. A further striking finding is that the study design matters for the impacts found; randomized controlled trials find systematically smaller effects than quasi-experimental studies. A significant shortcoming of the literature is that almost nothing is known about long term effects and cost-effectiveness and many studies fail to provide a detailed analysis of why certain effects occurred or did not occur – making it hard to extrapolate lessons.

Key-words: Employment; Active Labor Market Policy; Firm Creation, Micro-, Small- and Medium-sized Firms; Impact Evaluations; Systematic Review.

JEL codes: D22, G21, J21, O10.

Acknowledgements: Financial support from the German Development Bank (KfW) is gratefully acknowledged.

1. Introduction

Creating new jobs, and in particular ‘good jobs,’ i.e. jobs in high-productivity sectors and offering decent working conditions, is one of the major challenges most low and middle income countries face. According to the 2013 World Development Report on jobs, 600 million jobs are needed worldwide over the next 15 years to keep employment rates at their current level (World Bank, 2012). Governments, non-governmental organizations and donors spend large amounts of money for targeted programs and broader policies to enhance employment creation and the creation of new firms. Because most employment in low and middle income countries is in micro-, small and medium-sized enterprises (MSMEs) (see e.g. Ayyagari et al., 2011), these firms are often targeted by such interventions.

This systematic review synthesizes the existing evidence on the employment impacts of these interventions and extracts the main lessons. We focus on the following five policy areas: (i) access to finance, (ii) entrepreneurship training, (iii) business development services, (iv) wage subsidies, and (v) improvements to the business environment (e.g. registration procedures). There are many other interventions and policies that may have employment effects such as improvements in energy supply, road construction or trade and exchange rate policies, but given that such policies are typically not targeted it is hard to establish causal evidence.

Our work builds on a few earlier reviews which however have not focused specifically on employment creation in MSMEs or they considered only a sub-set of the policies we focus on. McKenzie and Woodruff (2014) review the quality and findings from business training and entrepreneurship evaluations. They focus in particular on statistical power, measurement issues and attrition. Across the reviewed studies, they find only modest impacts on survivorship but stronger impacts on business creation. Bruhn and McKenzie (2013) review the rigorous evidence on entry regulation and formalization of micro-enterprises. They conclude that formalizing firms is generally difficult and the effects of formalization on firms are in most cases very modest if not insignificant. McKenzie (2010) reviews a small set of finance evaluations (all pre-2009) and concludes that many existing micro-finance programs fail to achieve their targets. Tripney et al. (2013) conducted a systematic review of post-basic technical and vocational education and training (TVET) interventions to improve employability and employment of TVET graduates in low and middle Income Countries (LMICS). While the authors are concerned with the same outcome as this systematic review, they do not explicitly focus on employment creation in MSMEs. Betcherman (2014) provides a more narrative review of studies that explore the effects of labor market regulations on employment and other outcomes. Finally, other authors focused on youth employment, but again, do not consider MSMEs specifically (Betcherman et al., 2007; Puerto, 2007).

A meta-analysis on entrepreneurship programs in developing countries conducted by Cho and Honorati (2014) is so far the most relevant synthesis of evidence with respect to the purpose of this

review. Cho and Honorati (2014) focus on various business outcomes and find that finance and training interventions to promote MSME development are more effective in changing intermediate outcomes like business knowledge and practice than in increasing a general set of labor market outcomes. For the latter, the combination of training and finance proves to be the most effective, though this depends also on the type of beneficiary that is being targeted. This review will substantially add to the work done by Cho and Honorati (2014) first, by updating and broadening the evidence base considerably, while taking into account also policy areas other than finance and training, and second, by choosing a strong focus on employment and business creation. Because of this difference in scope and the advancement in time, we cover 39 evaluation studies that have not been covered by Cho and Honorati (2014). Hence, the major originality of our review is the focus on the ‘job creation’ outcome. Considering a wide range of programs will give insight to which instruments under which conditions contribute the most effectively to this goal.

Our review shows that overall the impacts on employment are very modest. Although many interventions successfully affect intermediate outcomes such as management skills only very few interventions enhance job creation. The picture is a bit more optimistic if the focus is just on the set-up of micro-firms. The evaluated business training and business development programs show on average better results than the access to finance programs. Comparing successes and failures suggests that ‘intense treatments’ combined with the appropriate targeting are needed to have an impact. Interestingly, we also find that interventions evaluated with Randomized Controlled Trials (RCTs) show substantially less often significantly positive results than interventions evaluated with quasi-experimental designs. While this has certainly to do with the context and type of intervention RCTs have been used for, it may also imply that many quasi-experimental designs cannot fully get rid of selection effects. In addition, as we show, the very small sample sizes and hence the low power of many RCTs are responsible for their inability to detect significant effects if these are small.

The remainder of this review is organized as follows. In Section 2 we lay out our inclusion criteria and the search strategy. In Section 3 we propose a theory of change which will guide our analysis along the causal chain, linking program inputs and employment outcomes. In Sections 4 and 5 we present our search results and a narrative synthesis of the evidence. The results from a meta-regression analysis are presented and discussed in Section 6. In Section 7 we conclude.

2. Inclusion criteria and search strategy

2.1 Inclusion criteria

We include studies that explicitly focus on MSMEs in the formal as well as informal sector. We limit the analysis to urban as well as rural non-farm employment and firms, i.e. farms and employment on farms are not considered. Although there is no common definition to identify MSMEs, neither by

researchers nor by statistical offices, we use an employment criterion and set the threshold at 250 employees. We define micro-enterprises as firms with less than five workers. Small firms are firms with 5 to 19 workers and medium-sized firms are firms with 20 or more, but less than 250 workers. We limit our review to firms in low and middle income countries. We use the thresholds of the World Bank and consider countries as developing countries if they show a gross national income (GNI) below USD 12,476 per capita, calculated using the World Bank Atlas method. Specifically, the following income groups provided by the World Bank classification are included: low income (USD 1,025 or less per person and year), lower middle income (USD 1,026 - 4,035), and upper middle income (USD 4,036 - 12,475).

Regarding the outcome, we include studies if they measure the creation of jobs specifically for MSMEs. That is, we included any evaluation study that provides evidence for MSMEs separately from large firms. Whether the policy under study also affects large firms does not matter. We define employment creation as the emergence of new jobs in existing MSMEs (whether privately or publicly owned) and as jobs that arise through the creation of new MSMEs. The latter also includes self-employment. Whenever a certain intervention creates some and destroys other jobs simultaneously, we explicitly consider – if the data allows – both gross and net employment generation. We consider any form of employment under acceptable working conditions, conditional on the specific context studied. This includes paid employment, as well as paid and unpaid family employment.¹ Studies are not included if they focus exclusively on hours worked, labor intensity, wages or labor supply without considering employment per se. Interventions that target the youth or the labor force directly are only considered if they aim at the creation of new MSMEs or self-employment. Based on an initial screening we decided to classify all studies into five intervention categories: access to finance, entrepreneurship training, research and development, business development services, and private sector incentive schemes. In addition to these specific programmatic interventions, we also decided to include studies that analyze the impact of more general conditions, generally referred to as the business environment, on MSMEs.

Studies are included if they can establish a credible causal relationship between a programmatic or policy intervention and job creation in MSMEs. Hence, we include studies if they consist of an impact evaluation based on an experimental design such as an RCT, as well as on quasi-experimental designs including propensity-score matching, instrumental variables, regression discontinuity designs or difference-in-difference estimation. Since different methods have different strengths and weaknesses we discuss the problems that come with these techniques in Section 4.

Moreover, inclusion/exclusion is not based on publication status. If an identified study was still ongoing, the authors were contacted in order to check whether the results were already available for

¹ The latter usually implicitly paid.

inclusion into this review. Evaluations in any of the five following languages are included: English, Spanish, French, Portuguese, and German.

2.2 Search methods to identify relevant studies and data extraction

A range of different search methods were applied to ensure the identification of recent, ongoing, as well as unpublished studies. These include the screening of electronic databases, relevant websites, key journals as well as literature snowballing and contacting of researchers and key experts. Details on the search strategy can be found in Grimm and Paffhausen (2014).

From each included study we extracted systematically information about the author(s), the title, the year and the type of publication. Furthermore, the country or countries and the target population were characterized for each study, as well as the outcome(s) measured and the type(s) of intervention. With respect to the type(s) of intervention, we recorded whether the intervention was a stand-alone or joint intervention. If it was the latter, the different components of the intervention were documented. Whether employment or firm creation was a primary objective of the intervention was also recorded. Furthermore, included studies were characterized according to their study design, including their methodology, sampling procedures etc. Finally, information regarding the internal and external validity was extracted from the studies. Data extracted from included studies, especially those involving judgments by the coder, have been discussed extensively among the two researchers involved.

3. Theory of change

The considered interventions affect firms and prospective firms in many different ways. Some relax capital market constraints, others improve management skills and business practices, and again others reduce the cost of labor or ease formalization procedures. The figure below shows a simplified result chain for all interventions together. Next to the final outcome of interest these chains also show intermediate outcomes such as investment, productivity, output and profits. For all interventions it is obvious that employment does not necessarily have to respond. If capital market imperfections are relaxed and investment increases, employment effects will only occur if the investment is large enough and if labor is a complement to the investment and not a substitute. Labor-saving investments could even reduce employment. Credits that are used to increase inventories are also unlikely to increase employment. Interventions that increase productivity, such as training, will only have employment effects if output is increased following falling costs. Hence, the price elasticity of demand and the degree of competition matter. In other words growth at the intensive and extensive margin needs to be distinguished; only the latter goes hand in hand with more employment. Yet, if

employment is not increased, increasing productivity may secure survival of the firm and therefore prevent jobs from disappearing and thereby also make a contribution to employment.

[Figure 1 about here]

Employment effects also require that labor supply is sufficient, i.e. that workers queue for jobs and can be hired at any time. An assumption that is probably realistic for most urban areas in the developing world at least as long as unskilled workers are concerned. If skilled workers are concerned then this might not always be a good assumption. Some interventions may also imply negative externalities on non-beneficiaries. In a context in which economic growth is weak or completely absent, an increase in productivity in one firm may drive another firm out of the market and hence the net job effect might be zero, or at least reduced.

Temporary wage subsidies will have lasting impacts only if the subsidized job or ‘on the job training’ increases the worker’s productivity to such an extent that the firm continues hiring that person even if the subsidy expires. This would imply that the temporary subsidy allows reaching a higher growth trajectory that would not be reached without the subsidy. Obviously such programs may also work because they solve a behavioral bias, i.e. it may provide a hiring incentive to those that are very risk-averse and reluctant to hire, even if it would be beneficial for them.

A simplification of registration procedures will only have any effect on employment if the registration allows access to new markets or a reduction of costs of certain inputs such as public services or if it improves access to finance. But even then the above mentioned caveats apply.

Finally, it is important to note that these result chains have been developed from the perspective of the evaluator, based on the objectives of proponents of such interventions, which justify them on the basis that they eventually contribute to employment creation. They are, however, not necessarily the result chains of the implementers of such interventions and most certainly they are not the result chains of the beneficiaries. For instance, an NGO delivering a micro-credit program might do this with the aim of promoting female empowerment or increasing household income only. A researcher who evaluates this program then however might assess the effect on a broad range of outcomes, including employment in household enterprises. The person obtaining the micro-credit might be interested in merely smoothing consumption, keeping her business going or creating a new subsistence activity, but not necessarily employing more staff.

4. Search procedures and results

4.1. Search and identification

We organized our database search strategy around two alternative search variants. The first variant combined sets of search terms referring to the population, outcome and type of study and was conducted in February 2013. This search resulted in 1,200 hits. After removing duplicates, there were still 932 records left that needed to be screened. The second variant focused on the different intervention categories considered in this systematic review, was carried out in May/June 2013, and resulted in 2,446 hits.² After removing duplicates, 1,343 records remained for a careful screening.

We then merged the results of both database search variants, which together amounted to 2,275 hits. Again, duplicates had to be removed so that the final sample included 1,924 records. These records were screened successively, applying inclusion criteria, in a first step, to titles and abstracts only. This was done mainly by one researcher. However, in order to minimize bias, a second researcher randomly double-screened about a quarter of the studies that had been excluded by the first researcher. There were virtually no discrepancies in judgment for this sample of studies. Moreover, unclear cases were screened additionally by the second researcher, and a conclusion was reached through discussion, where necessary. Thereby, already about 85 percent of studies could be excluded. A common reason for early exclusion of studies was that they did not focus on low and middle income countries. In addition, although many were concerned with SMEs or smaller enterprise, these were no impact evaluations and typically did not document any changes in the outcomes relevant for this review. Obviously, a non-negligible part of the papers was also totally off topic.

The whole procedure left 139 studies for thorough full text screening. Out of these, four records could not be obtained in full text. However, further cross-checks suggested that they were, with a very high probability, irrelevant for the purpose of this review. There was one study that was judged relevant based on the abstract, but the study could not be obtained. The database search also identified 15 potentially relevant ongoing impact evaluations. We contacted the authors in order to see whether any preliminary results were already available. As of February 15, 2014 we had received information on the status of five ongoing impact evaluations. None of these however were in a stage in which any findings could already be included in our review. Since up to this point, we had not yet identified a single study for inclusion which had been conducted before 1990, we decided to definitively exclude studies that had been conducted before that date. For all remaining papers the final decision on inclusion or exclusion was always based on the full text of the study. Finally, 33 studies were included from the database search.

² Note that the second database search variant led to a contribution of only three additional studies to the final sample of included studies from the database search (a total of 33) that were not found by the first approach. This result gave us confidence in our search strategy and we decided to stop the database search here, not adding further database search variants since the expected benefits in terms of additional studies would not justify the corresponding effort of searching and screening.

The screening of websites of key donors and funders of MSME interventions as well as research institutions resulted in one further study that was included. In addition to the database search and website screening, a number of relevant academic journals were searched for studies to be included. This enabled us to identify three more studies for inclusion. Literature snowballing of the World Bank World Development Report 2013 and other reviews resulted in four more records that were included. The references of included studies were thoroughly screened, resulting in ten more studies to be included in the review. Furthermore, we contacted key researchers and provided them with a preliminary list of included studies as of September 9, 2013, asking whether they were aware of any further studies that met our inclusion criteria which we should include in the review. We received answers from three researchers as of February 15, 2014. This exercise resulted in no further study to be included in the review. Finally, two studies were included in the review that were already known to the authors but had not been identified through any of the sources mentioned above. Altogether, this comprehensive search process led to a final sample of 53 studies that have been included in this review. The entire process is illustrated in Figure 2.

[Figure 2 about here]

4.2 Short characterization of included studies

Population and context

The 53 studies included in this review cover a wide range of countries from all major world regions. The majority of studies focuses on Latin America. This can be explained by the fact that Latin American countries in particular have often experimented with active labor market policies over the past two decades and that they have started earlier than others to rigorously evaluate many of their policies. There are still only few studies that evaluate interventions implemented in Sub-Saharan Africa, although employment creation is a major preoccupation and considerable amounts of foreign aid are spent on private sector development in that region. Most studies have been conducted in upper middle income countries (32) and lower middle income countries (14). Only 7 evaluations were implemented in low income countries. The majority of the studies focus on micro-enterprises and self-employment (with up to five employees). A total of 36 studies fall into this group. Another 17 studies target larger firms, which could be categorized as small and medium-sized enterprises. These are generally already established and mostly registered (i.e. formal) enterprises.³

[Table 1 about here]

³The size classification is based on the thresholds described in section 2.1. Self-employment is classified as ‘*micro-enterprise*’. When a study was not clear on the definition that had been used to classify enterprises, but when it reported at least summary statistics regarding the number of employees, we considered the mean size of the enterprises. A number of studies do not in fact report the exact size of the firms analyzed, but have an obvious focus on MSMEs. In those cases we relied on the classification provided in the study.

Interventions

Most of the studies in our sample analyze interventions that aim to remove credit constraints of micro-, small and medium-sized firms (26). Some of the interventions focusing on access to finance are also combined with other interventions. In most cases these are entrepreneurship training interventions. For the sake of completeness such interventions are analyzed below in both groups - finance and entrepreneurship training - whenever justified, based on the specific features of the intervention evaluated. The sample of studies focusing on entrepreneurship training includes 20 studies. There are six studies in the area of business development services including for instance counseling, supplier development, product and process innovation, and the provision of working premises. The studies focusing on private sector incentive schemes all evaluate wage incentives. In this category we found three relevant studies, which all assess the impact on MSMEs. Finally, we identified five studies that assess interventions falling into the category of improvements to the business environment. These studies all deal with business registration.

Outcomes and impacts

As laid out before, we included studies that assessed the impact of interventions relevant to MSMEs on either changes in employment levels in these enterprises or the creation of new enterprises, including becoming self-employed. There are also studies that assess the impact on both outcomes which were then also both used. We always extracted the evaluators' most preferred estimate. Robustness checks were considered to assess the quality and reliability of the estimates, but are not retained for the meta-regression analysis below.

Some studies provide impacts at different points in time, such as short-term and long-term impacts. Others offer impacts for stand-alone and mixed interventions. In these cases we have more than one impact per study. In total, from the 53 included studies, we drew a total of 116 impact estimates. Most of these impacts (about 60 percent) relate to employment. The remaining impacts measure business creation and self-employment. We do not consider firm survival explicitly as only few studies provide results for this outcome. We express impacts in terms of their standardized mean difference (SMD), i.e. as the ratio between the change in the outcome due to the intervention divided by the standard deviation of the outcome in the control group (or at baseline). If the outcome is a binary outcome such as 'having set up a firm or not' we use the risk ratio. In those studies where such impact measurements were not directly provided (in fact the large majority) we computed them ourselves. However, some studies do not provide all the necessary information; in these cases we based the estimate just on the reported *t*-values of the impact and the sample sizes of treatment and control groups using the formulas given in Lipsey and Wilson (2001). Using different computation methods, on the one hand, reduces the comparability of the estimates. On the other hand it limits the number of studies for which no effect size can be computed.

Study design and methods applied

Of the 53 studies that have been included in this review, 27 studies employ quasi-experimental designs while 26 studies are based on RCTs. RCTs are especially relevant for impact evaluations in the area of entrepreneurship training and finance. Only one fourth of the studies that evaluate the impact of entrepreneurship training use a quasi-experimental design. Studies that assess finance interventions are divided relatively evenly across different study designs, with 13 studies employing an experimental design and another 13 an RCT. Within the group of RCTs, the studies focus exclusively on micro-enterprises and potential entrepreneurs. This is also the case for the entrepreneurship training interventions. RCTs are often considered the ‘gold-standard’ in the evaluation literature as they in principle avoid selection effects that can bias the results.

However, RCTs also come with significant drawbacks. The main problem we noticed is the weak power of many of these studies. Most of the included RCTs do not provide any power calculations. Hence, to assess their power to detect a 20 percent and 50 percent change in the outcome, we drew information on the sample size, on the control group mean, on the standard deviation (if available on stratification, on up-take, and on attrition from a sample of studies. For that purpose we concentrated on the 17 included RCTs that look at the number of jobs created as an outcome (ignoring the set-up of businesses).⁴ Based on the information collected, we computed the absolute effect size corresponding to a 20 percent and 50 percent change in the outcome. Similar to McKenzie and Woodruff (2014), we then compared these changes to the standard errors of the estimated impacts to make a judgment whether these changes could be detectable. Based on this assessment, we conclude that only about 5 out of the 17 studies are able to detect effect sizes of 20 percent and less. Although the picture looks much better for effect sizes of 50 percent and more, as 11 to 15 of the 17 studies would be able to detect such changes, it will be seen below that in 60 percent to 80 percent of the studies - depending on the type of intervention considered – the point estimate of the actual effect size is smaller than 20 percent. Hence the upper-bound effect size of 50 percent is hardly relevant. The second drawback with RCTs is their often low external validity, i.e. many of the programs evaluated through RCTs are targeted to specific groups and areas and run at a very small scale. Up-scaling those programs may not only be institutionally difficult and costly, but it may also come with a whole range of additional problems. We get back on this issue when discussing the studies below.

In contrast to the areas of finance and training, studies included in the areas of business development services, wage subsidies, and business environment almost exclusively rely on quasi-experimental designs such as difference-in-difference, propensity score matching or instrumental variables to deal with possible selection effects. These studies typically do not have problems with low power and perform much better with respect to external validity. However, they cannot always fully rule out

⁴ Since continuous outcomes require different formulas to make the power calculations than binary outcomes, we limited the exercise to the continuous outcome to have fully comparable results.

potential selection effects. This is particularly the case for those studies that just rely on cross-sectional matching (10 out of the 28 quasi-experimental studies), although studies where we had serious concerns about causality (see inclusion criteria above) were not included. Most included quasi-experimental designs rely on a difference-in-difference design (14 studies) where the remaining concern is typically only with unobservable time-variant confounders. Studies that just use either an instrumental variable for identification or a regression discontinuity design are rare (2 studies each).

5. Synthesis of the evidence

5.1 Access to finance interventions

Most of the 26 studies that deal with finance interventions examine the effectiveness of micro-credit schemes (20 studies), followed by conditional or unconditional cash- or in-kind grants (four studies) and two studies that just introduce changes to existing credit schemes. Field et al. (2011), for instance, evaluate the extension of the period until the first repayment is due. De Mel et al. (2013b) examine savings incentives. Not a single study could be identified that looks at the impact of micro-insurance on employment. All studies under consideration are listed and briefly presented in the Appendix. Some of the interventions analyzed here are stand-alone interventions, whereas others are combined with different forms of entrepreneurship training or even other interventions with a completely different scope. This is, for instance, the case for the study by Tarozzi et al. (2013), who look at the combination of micro-credit and family planning services. The credits and grants that are typically considered by the evaluations under review are in the range of USD 100 to USD 2,000. Loans provided through public credit lines targeted at small and medium-sized firms are on average larger.

With respect to employment creation most micro-credit schemes turned out to be rather unsuccessful; only 20 out of 54 impact estimates, which were drawn from the 26 studies on access to finance, show a statistically significant increase in employment or business creation. 32 out of the 54 treatment effects were not statistically significant. In two cases, a statistically significant negative effect was found. The sign and significance of program effects can be found in the Appendix. Table 2 provides the distribution of effect sizes. The high frequency of statistically insignificant effects also reflects the often very low power of these evaluations (see our discussion above). Positive effects on employment, if found at all, were only small, especially for already existing small and micro-enterprises. Major effects were achieved with respect to the creation of new (mostly micro-) enterprises and the expansion of already larger, well-established and profitable firms. But there is hardly any evidence for employment expansion in micro-enterprises, notable exceptions being a cash transfer program in Uganda (Blattmann et al., 2012) and a group lending scheme in the Philippines (Kondo et al., 2008).

The majority of insignificant effects (29 out of 32), as well as the two statistically significant negative employment effects are indeed found in evaluations that focus on micro-enterprises. More than 70

percent of the effects associated with micro-enterprises are statistically insignificant. In contrast for small firms, all treatment effects are positive and statistically significant. For medium-sized firms about half of the studies find positive significant effects. All this suggests that it is easier to create jobs in small and medium-sized enterprises than in micro-enterprises. However, three comments are in order. First, it is not straightforward to disentangle how, on the one hand, firm size and, on the other hand, the country context shape impacts. In low income countries most interventions target micro-enterprises whereas in middle and, in particular, upper-middle income countries interventions tend to target small to medium-sized firms. Second, very often the programs that target micro-enterprises as opposed to those that target small and medium-sized firms do not have employment creation as one of their primary objectives; they are in many cases primarily concerned with income stabilization and female empowerment. Third, the evaluations that deal with small and medium-sized firms predominantly use quasi-experimental evaluation designs which, as explained above, cannot always claim to fully get rid of potential selection effects. However, in contrast to experimental methods, evaluations based on quasi-experimental designs often rely on larger samples and less heterogeneous firms. Therefore their power and hence the probability of detecting significant effects is higher. These aspects will be further explored in the framework of a meta-regression analysis below.

[Table 2 about here]

Programs which explicitly target women also appear to be less successful in employment creation than programs without such a focus, suggesting that women in particular face additional constraints which need to be overcome in order to increase the return to finance. Women typically have lower levels of formal education, often have only limited access to other relevant services, are exposed to higher consumption demands from their family and relatives, often need to work from home and are often deprived of property rights – land rights in particular.

Analyzing impacts along the full causal chain, linking program inputs with final outcomes, can reveal whether such interventions have an impact on intermediate outcomes and under what conditions changes in intermediate outcomes also lead to changes in the employment outcome. We screened the 26 studies for information regarding the impact on investment, hours worked, productivity, output, sales, revenues and profits. The findings are selective, however, as not all studies explore the same intermediate outcomes.

First, it can be stated that most enterprises make use of the credit or cash grants provided, if they are directly offered, but the studies' findings suggest that the financial resources are primarily used as working capital, i.e. invested into inventories. Only on rare occasions do they result in fixed capital investments such as machines and buildings. De Mel et al. (2013b), for example, detect a significant effect of a savings scheme on investment in inventories, while there is no significant effect on fixed capital investment. Likewise Arraiz et al. (2012) find that credit had no impact on capital stocks and suggest that firms rather use credit to increase their working capital.

Studies that assess the impact on labor supply in terms of hours worked in the business find mixed results, with two studies detecting an increase in labor supply as a result of the intervention (see Banerjee et al., 2013; Blattman et al., 2012) but others finding no impact (see Crepón et al., 2011; De Mel et al., 2013b ; Karlan et al., 2012) or inconclusive evidence (see Augsburg et al., 2012).

A few studies assess the impact on productivity. The majority of these studies focuses on small and medium-sized enterprises and suggests that finance interventions do not affect productivity of established enterprises. It is however important to note, as De Negri et al. (2011) rightly point out, that this lack of impact might also be due to the fact that labor productivity is often proxied by real wages, which are more rigid than actual productivity.

Enterprise performance is typically measured in terms of output, sales, revenues, business expenditures and profits. Most studies considering the impact on sales do not find any impact, or only in the short-run, such as De Mel et al. (2013b). Of the studies focusing on revenues and business expenditures, most do not detect any effects either (see Banerjee et al., 2011; Karlan et al., 2012; Nelson, 2011) or find that both revenues and costs increase proportionally (Angelucci et al., 2013). Only very few studies assess effects on physical output, but those who do find an increase (see Arraíz et al., 2012, and Eslava et al., 2012). Finally, with respect to business profits, the majority of studies does not detect an impact (see Angelucci et al., 2013; Augsburg et al., 2012; Barnes, 2011; Crepón et al., 2011; De Mel et al., 2013b; Karlan and Zinman, 2011; Nelson, 2011). Karlan et al. (2012) find that capital grants even lowered profits. Banerjee et al. (2013) detected a positive impact, but only for existing, relatively large and the very profitable micro-businesses. Only Marcours et al. (2012) find that cash grants for investment resulted in higher profits from non-agricultural self-employment activities more generally. These effects were substantial even two years after the end of the intervention.

It could be that the limited evidence for an improvement in business profits is also due to rather short follow-up periods in the evaluation design. It might take some time for the entrepreneur to be able to use the capital treatment effectively and adjust accordingly in order to reap the benefits of increased investment. Otherwise it may simply imply that employment effects do not materialize because the finance interventions already fail in changing the necessary intermediate outcomes.

Apart from the general observation that many of these programs do not even intend to create jobs, reasons for low or no impacts seem to be that most loans are simply too small and their maturities too short to lead to large changes in the capital stock and the production technology which in turn could generate employment effects. For instance, a tailor who – thanks to a micro-credit – switches from a mechanical to an electric sewing machine may still not reach the profitability to immediately hire an additional worker although this tailor may well see an increase in performance as measured by revenues, profits and business investment. Put differently, if the studies find any growth at all, it is rather at the intensive than at the extensive margin. Field et al. (2011) show that the repayment period

also matters. Short grace periods seem to prevent poor entrepreneurs from investing since they fear not being able to repay on time. Moreover, in the cases where potential business starters or subsistence-type enterprises are targeted, a reason for failure may also be that very often there are competing needs. Hence, instead of investing, borrowers spend the credit on health care, education, housing improvements and so on. Based on such insights Karlan and Zinman (2011) suggested that micro-credit may need to be combined with detailed business planning and extraordinarily close monitoring in order to ensure that it leads to increased investment in the business. Yet all this also shows that capital constraints interact with other constraints and therefore, only addressing capital constraints is often not enough.

In summary, the programs under consideration are more effective in creating new firms. In contrast, employment expansion happens in small and medium-sized firms. The weak evidence for employment expansion in micro-enterprises has different causes. Most importantly, many of these programs do not even intend to create jobs and even if they do, these programs are often not well designed (i.e. loans are often too small and repayment periods too short) and they often lack a clear targeting strategy. Further flaws come from the evaluations themselves. Low power of many experimental studies is probably the most serious problem in this respect.

5.2 Entrepreneurship training

Overall we reviewed 20 studies that considered entrepreneurship training (see Appendix). These studies cover quite a heterogeneous set of policies: business skills training, business plan development, financial literacy training, technical and vocational training (in-class and on the job), and life skills training. Entrepreneurship training is sometimes quite general and sometimes specifically tailored to the businesses and difficulties of the beneficiaries. Most evaluations assess the impact of general entrepreneurship training in the areas of business management, accounting, financial literacy, or the development of vocational skills (15 cases). Bandiera et al. (2012), for instance, analyze the provision of vocational training in activities like hair-dressing, tailoring, computing etc. to adolescent girls. Other examples can be found in Bruhn and Zia (2011), Calderon et al. (2013) and de Mel et al. (2012), which assess interventions that provide general business and financial education trainings to micro-credit clients and individuals interested in starting a new business. A few interventions provide more specifically tailored assistance (4), which may come in the form of business plan development assistance (see Klinger and Schündeln, 2011; Jaramillo and Parodi, 2003) or management consulting services focusing on problem diagnosis and solving (see Bruhn et al., 2013). One study analyzes both an intervention that provides general business training to micro-entrepreneurs as well as the combination of this general training with individualized support (see Valdivia, 2011). The interventions analyzed here further vary substantially regarding their duration. While some training courses are delivered over a few days only (see for instance Bruhn and Zia, 2011), others are more substantial. The business training evaluated by Valdivia (2011) for example was delivered over twelve

weeks in three sessions per week that lasted three hours each. Vocational entrepreneurship training also tends to be more substantial in duration, varying further with the specific occupations for which the training is provided.

The majority of interventions targets micro-enterprises with up to five employees or aims to enhance self-employment in groups highly at risk of unemployment, such as youth. Often entrepreneurship training interventions, especially when provided for business start-up, are delivered jointly with start-up finance. The majority of studies included here are based on RCTs; only five employ a quasi-experimental design. All together, 36 treatment effects were analyzed.

Looking across all studies, 15 out of 36 treatment effects show significant positive employment effects (see Appendix). 21 treatment effects were not statistically significant. As discussed above and also shown for finance interventions, the low statistical power of many RCTs is probably partially causing the large number of null results. Interestingly, a few studies found negative employment effects. Drexler et al. (2013) found that training led to a reduction of employees for low-skilled business owners and both Valdivia (2011) and Calderon et al. (2013) found that micro-entrepreneurs who participated in general business training were more likely to close poorly performing businesses. This suggests that entrepreneurship training enhances the entrepreneurial spirit and forces (potential) entrepreneurs to think more carefully about their business model and its profitability. Entrepreneurship training can also prevent non-profitable business ideas from being started.

The majority of firms in this sample are micro-enterprises. Only two studies assess entrepreneurship training interventions for small-sized firms and there is only one study focusing on medium-sized firms (see Table 1). Of the 30 treatment effects we collected for micro-firms, 63 percent are insignificant with the remaining 37 percent being statistically significant and positive. Out of the five effect sizes for small firms, four are positive and statistically significant. The effect size for medium-sized firms is insignificant.

With respect to the impact of entrepreneurship training along the causal chain, information on a range of intermediate outcomes is available. Employment creation is typically not the primary aim of entrepreneurship training. Instead, those interventions rather aim at improvements in business and entrepreneurial skills, which are supposed to lead to improved business performance. Hence, most studies provide information on those types of intermediate outcomes. Most training interventions have, in fact, difficulties in changing actual business performance. Most programs produce significant improvements in business skills and behavioral skills, and sometimes even higher optimism and motivation (see Cho et al., 2012; Premand et al., 2012), although some of these changes might be due to a change in reporting behavior after the start of the intervention (Drexler et al., 2012). In addition, although business skills and practices are significantly higher after the entrepreneurship training interventions, this might not necessarily imply that skills are sufficiently developed to run or expand an enterprise, an issue also being raised by Cho et al. (2012). Of those studies that look at the impact

of training on investment, only one study finds a significant positive effect. Yet, in this particular intervention, investment was also a core concept emphasized in the training (see Bruhn and Zia, 2011). Only few studies assess the effect on labor supply, in terms of hours worked. The results from these studies are inconclusive. With very few exceptions, most evaluations do not find that entrepreneurship training led to increased sales or revenues. Finally, the training interventions analyzed here do not generally succeed in increasing profits. Bruhn and Zia (2011), for instance, do not find any treatment effects of business and financial education training on business performance, but they do find changes in business practices consistent with the topics taught in the course. Similarly Giné and Mansuri (2001) find that, especially for male entrepreneurs, business training increased business knowledge and practices, but had no impact on sales or profits. Furthermore, short-term positive effects often seem to vanish in the long run. Hence, in the case of training interventions, it appears that impacts can only be detected very early in the causal chain in the form of improved business skills and, in some cases, practices. But those effects are not sufficient enough to trigger some sort of virtuous cycle, which could lead to improvements in business performance and, ultimately, employment.

Several studies assess the influence of targeting. The evidence is mixed on whether the return to training is higher for those with initially lower skills and whether it is more helpful for male or female owned enterprises. The review suggests that training is more helpful for start-ups than for business expansion. However, since many interventions that aim at business start-up often also include some form of financial assistance, it is not always possible to isolate the effect of the training. De Mel et al. (2012), for instance, find the combination of a cash-grant with entrepreneurship training to be especially successful. The more tailor-made and substantial the training, the better, but the more complex programs are not necessarily the most successful. Drexler et al. (2013), for example, observe that general accounting training led to some improvement in objective reporting quality and business performance for higher-skilled participants, but had the opposite effect for less educated individuals, while a simpler, rule-of-thumb training was more effective for those participants with lower educational levels. Furthermore, from the studies it appears that training needs to address specific knowledge gaps and be ‘substantial’ in order to be effective. The consulting and mentoring services analyzed by Bruhn et al. (2013) eventually led to increases in the number of employees of treated businesses. These services were provided to beneficiary enterprises over a period of one year, in weekly four-hour consulting sessions. Likewise, Premand et al. (2012) find positive treatment effects of an intervention that is provided over a period of one academic semester.

To summarize, whereas many of these programs indeed are able to improve business skills, there is only weak evidence for substantial employment effects. However, low power of many of the evaluations is again part of the story, i.e. some of these programs may in fact have employment effects but they are too small to be detected by the studies. The combination of finance and training seems to work better than either of them in isolation. Moreover, training needs to be intensive: short-term interventions do not have a lasting impact.

5.3 Business development services and wage subsidies

In this sub-section we review nine studies, which are rather heterogeneous in the specific nature of the underlying interventions (see Appendix). Broadly, they fall under the heading of business development services and targeted subsidies. Four of the nine studies cover business development services in the narrow sense, i.e. they deal with supplier development, support for environmental audit, provision of working premises, etc. One of these studies covers conditional tax-breaks and fiscal incentives for technological innovations as well. Two studies measure the employment impact of grants for product and process innovations and three studies cover supply or demand side wage subsidies. The studies dealing with wage subsidies focus on Turkey and Sri Lanka, while the other studies cover Latin American countries almost exclusively. Overall, circumstances were less challenging in these countries than in those that hosted many of the micro-credit and entrepreneurship interventions analyzed before. Only one of these nine studies is based on an RCT design, while the others use a quasi-experimental approach, typically by exploiting variation in the policy across time and space in a difference-in-difference framework.

The studies show mostly positive and statistically significant employment effects (see Appendix). This holds regardless of the size of firms that are being evaluated. For medium-sized firms, about 62 percent of the thirteen treatment effects are statistically significant; for small sized firms, one out of two treatment effects is statistically significant and positive, while the other is insignificant; and for micro-enterprises three out of four treatment effects are statistically significant and positive. Overall it seems that business support services and targeted subsidies can contribute to employment generation if they are demand-driven, tailor-made and focused. Larger firms may need quite specific and sophisticated support, whereas small firms can be helped with rather simple improvements to their business. Wage subsidies as well as tax-breaks and fiscal incentives conditional on process and product innovations seem to be particularly effective. Interestingly, this has also been found by Kluge (2010) for a sample of European countries. His review showed that direct employment effects were the largest for wage subsidies and ‘services and sanctions’ conditional on certain productivity enhancing activities in comparison to other active labor market policies such as training or out-of-work income support.

However, the findings for low and middle income countries covered here are somewhat fragile, first, because the findings are based on a very small set of studies and, second, because almost all studies have to rely on a rather weak identification strategy and hence a bias through firms selecting themselves into such programs cannot be ruled out completely. It is also remarkable that nothing can be said about the impact of business support services in the East and South East Asian context, where at least in some countries such programs may have played an important role.

With respect to the impact of business development services along the causal chain, most studies consider effects on sales, productivity and sometimes output, exports and fixed capital. Generally, the

interventions are successful in increasing sales significantly, as well as exports, outputs and fixed capital. When productivity is considered, effects are also predominantly positive, which is additionally corroborated by studies assessing the impact on wages, since these studies find overall positive impacts. Tan (2009) is an excellent example of a study that assesses the effects of an intervention along the causal chain. He first assesses the impact of the services evaluated on a number of intermediate outcomes where effects were intended and finds that beneficiary enterprises were more likely to be involved in research and development, to have introduced new product lines and production processes, to have obtained internationally recognized quality certification, and to have provided their workers with training. These can be seen as underlying mechanisms, which explain why the programs led to increases in sales, labor productivity and finally employment.

Overall, the studies conclude that the interventions considered here were successful in raising the competitiveness of beneficiary firms, which enabled business expansion. The general positive employment effects of business development services can therefore be seen as a consequence of improvements in business performance which was a result of the intervention. Furthermore, there is also evidence that the time it takes for these effects to occur is substantial. For example, Lopez-Acevedo and Tinajero-Bravo (2010) assess the impact of different business support programs in Mexico over ten years; only after a four year period the effects become statistically significant, but then they even gain in magnitude as time elapses.

Two different programs that were evaluated in a similar context in Turkey allow for an interesting comparison of supply- and demand-oriented subsidies. The supply-oriented on-the-job training program targeted employees (Fretwell et al., 1999) whereas the more conventional wage subsidy programs targeted employers (Betcherman et al., 2010). Whereas the first was a failure, the second was a success, suggesting that employers may keep workers hired at a reduced rate when they are free to choose the workers they actually prefer. In contrast, voucher-based programs may leave little choice to firms and firms may interpret vouchers as a negative signal about the workers' productivity reducing their chance of being offered a long term contract. Again, these interpretations are based on only two studies and eventually, whether or not they are really valid can only be determined if several wage subsidy programs test and compare such specific design features.

Indeed the three studies on wage subsidies do not provide much evidence about the underlying mechanisms that are important to understand why or why not such employment effects are lasting beyond the period over which the subsidy is paid. Only de Mel et al. (2013b) systematically assess the effect of the wage subsidy on outcomes such as management practices, investment, sales, profits and personal income. Delivered as a stand-alone intervention, the wage subsidy has no effect, either on management practices or on investment. It neither led to increased sales nor profits. Only when combined with savings or training interventions, the wage subsidy significantly improved investments and sales. Profits, however, remained unchanged.

Finally, it should be noted that wage subsidies are in general a quite expensive intervention and the programs covered here are no exception. The pure wage subsidy program in Turkey entails costs per job-month created that correspond to roughly 94 percent of the total cost of employing a minimum wage worker. This may still seem acceptable if the jobs created are sustainable, but this is far from obvious (Betcherman et al., 2010). A major cost component is the dead weight loss produced by the fact that many workers that are hired under a subsidized rate would have been hired anyway. This is also confirmed by the experimental study in Sri Lanka (De Mel et al., 2010 and 2013b) which shows a strong correlation between pre-program hiring intentions and program uptake.

5.4 Interventions to promote formalization

In most low and middle income countries, the bulk of urban micro- and small enterprises is informal, i.e. most enterprises are not registered with the tax authority and operate outside most regulations. A key policy question is whether the performance of these firms can be improved and their size in terms of capital and staff be expanded through formalization. On the one hand, one may argue that formalization increases access to credit and other resources important for business success and expansion. On the other hand, formalizations could imply a significant increase in tax payments which have to be added to the bureaucratic act of formalization, which some think is typically already so costly that it alone prevents firms from becoming formal (De Soto, 1989).

As both costs and benefits of formalization are involved, the policy problem of formalization is two-fold: Which interventions are suited to enhance firms' formalization, and what are the effects of becoming formal? We identified five studies that can credibly establish a causal link between formalization and employment (see Appendix). Four of them concentrate on Brazil and Mexico, where significant reforms have been implemented to reduce the costs of formalization. The fifth focuses on Sri Lanka and is based on an RCT.

All studies show that it is difficult to get the average firm formalized, because the average firm is simply too small and not profitable enough to make use of the potential that formality offers. Programs that offer cheaper and easier formalization procedures seem to work for a relatively small group of entrepreneurs and for firms that show already a higher initial performance (De Mel et al., 2013a). Inspections seem to work somewhat better (Andrade et al., 2013⁵). It also seems easier to formalize firms while they are being set up than to formalize firms that already exist. This has been shown particularly in the case of Mexico (Bruhn, 2011).

Overall, based on the sample of studies analyzed here, we find that formalization – to the extent it works – yields positive employment effects. Out of the seven effect sizes we drew from the included studies only one is insignificant. For small and medium-sized firms, all effect sizes are statistically

⁵ This study is not included in our systematic review because it does not assess employment effects.

significant positive, while two of the three effect sizes for micro enterprises are statistically significant positive and the other insignificant.

According to Bruhn (2011) the SARE reform in Mexico increased the number of registered firms by five percent and wage employment by 2.2 percent. However, employment decreased in ineligible firms (high risk industries) and hence the net effect is less impressive. Kaplan et al. (2011) use a different identification strategy and a different data source to evaluate the same reform. They only observe formal firms that also have formally registered employees and must ignore formal self-employment and new informal jobs in newly formally registered firms. Similar to Bruhn (2011), Kaplan et al. (2011) confirm that the program enhanced formalization: the number of new monthly formal firms increased by 4.7 percent. It is however not possible to say whether these firms are new or whether they existed previously as informal firms. Moreover, according to the findings, the program caused a monthly increase in formal employment of 10 percent. Again, some of these workers may have worked before as informal workers. Overall Kaplan et al. (2011) judge these effects to be rather modest in terms of their magnitude and not necessarily lasting. Over two years, according to the authors, the program would imply 42,408 additional formal jobs, corresponding to just 0.2 percent of informal employment.

In Brazil, SIMPLES offered a simplified tax system and lower taxes for micro-enterprises. The program permitted an overall reduction of taxes of up to eight percent with the main objectives to increase formality and to enhance the competitiveness of micro-enterprises relative to larger firms. Fajnzylber et al. (2008) find that the policy led to an increase in employment of 0.4-0.5 employees per firm. When the authors account for selection effects by instrumenting formal status, they even find an effect double that size. The results show that the increase is particularly high for firms with at least one employee. Moreover, the increase is largely in paid employment, as the share of paid employees in employment increases by 10 to 40 percent. The authors attribute the positive employment effects of the SIMPLES program in Brazil to the reductions in social security payments for hired labor. They also correctly point out that this effect does not necessarily apply to any informal firm that would be formalized, but rather reveals what would happen if entrepreneurs register their firm at the time of start-up. Courseuil and Moura (2010) analyze also the SIMPLES reform, but focus on firms with at least five employees. They find an increase of 6 to 7.5 percent in the average employment level.

In contrast, de Mel et al. (2013a) focus on micro-enterprises in Sri Lanka and do not find any employment effect as a result of formalization.

Of the five studies considered here, only two studies present detailed evidence on causal mechanisms linking formalization and employment outcomes. For Brazil, Fajnzylber et al. (2011) find that, among those firms that do formalize, performance improves in terms of revenue and profits. Effects on fixed capital are inconclusive, since formalization appears to have no impact on access to credit or sales to larger firms, but rather on the probability that firms operate from a fixed location. In conjunction with

the positive employment effect mentioned earlier, the authors conclude that the mechanism through which formalization improves business performance is the creation of more permanent and larger scale operations (Fajnzylber et al., 2011). De Mel et al. (2013a) also do not find any evidence that formalization improves access to credit. They further cannot detect any impacts on sales, profitability or capital stock, though business practices, in particular the use of receipt books and advertising, improve as a result of formalization. Formalization appears to have a substantial positive effect on profits. But this effect is driven by very few firms in the upper tail of the distribution. Providing qualitative evidence on these firms, de Mel et al. (2013a) show that they indeed used their formal status to develop new forward and backward linkages, to receive official quality labels or to take a loan and invest in capital. Notwithstanding, these effects were not strong enough to have a sizeable effect on employment.

De Mel et al. (2013a) also consider potential impacts on attitudes of the entrepreneurs and find that trust in local governments they interacted with during the intervention improved and they tend to agree more often that it is important to pay taxes. However, there are no effects on actual tax payments.

In conclusion, it seems that programs that ‘force’ firms to formalize are unlikely to produce any significant employment effects as for many previously informal firms formality does not translate into extra profits but into additional costs. It seems the best incentive governments can provide for formalization is to offer useful public services in return. This, of course, does not imply that policies should not simplify administrative procedures, but that efforts need to go further. It is not the cost of registration but the expected benefits of formality that is pivotal for the decision to formalize, and only if such benefits exist it is likely that formality will also lead to additional jobs.

6. Meta-regression analysis

6.1 Method

To implement the meta-regression analysis we use two alternative impact measures: first, whether a given intervention had a positive significant impact on employment, firm ownership (start or continuation) or self-employment at the 10 percent significance level;⁶ and, second, the standardized effect size. How we derived the latter was explained above. Relying on standardized effect sizes ensures a certain comparability of impacts across studies. However, measured impacts still differ in terms of the temporal horizon they refer to and, of course, in the program costs that had to be incurred to produce a particular change in the outcome.

⁶ Since we had only four statistically significant negative impact estimates in the sample, we decided to lump together insignificant and statistically significant negative estimates.

For the sets of estimates where we just use a binary variable taking the value one if the effect of program i was significantly positive, we run a simple probit regression and explore the variation of that binary variable across large set of study characteristics X_i . Hence, the model reads:

$$\text{Probit}(y_i=1/x_i)=\Theta(X_i'\beta + \varepsilon_i), \quad (1)$$

where X_i includes the type and characteristics of the intervention, the term, ε_i , is the error term and Θ stands for the cumulative normal distribution. Since coefficients of a probit model cannot be directly interpreted, we compute and show marginal effects. We correct the standard errors for within-group correlations, since some studies contribute with more than one outcome. The application of weights ensures that each intervention counts only once in the sample.

We further conduct simple linear regressions, where the dependent variable is the standardized effect size. The regression model reads:

$$y_i = X_i'\alpha + \eta_i, \quad (2)$$

On the right hand side we use the same explanatory variables as in the probit model above. The term η_i captures the unexplained part of the variance in y , which also includes measurement error.

As explanatory factors we include the type of the intervention, whether the intervention was explicitly aimed at employment creation, whether it was provided in combination with other interventions, whether it targeted women, the types of enterprises targeted, the study design of the impact evaluation, as well as the country income category, and finally, the outcome measure, i.e. whether it is employment or firm ownership and the way the standardized effect size is computed. Table 3 below shows the summary statistics for the dependent and explanatory variables. We also considered testing whether the type of the implementing agency plays a role, for instance whether this was a public or private entity. However, this did not lead to any meaningful results. First many studies are not clear about the status of the implementing agency. In other cases the implementing agency might be private, but the intervention was developed and designed by a public agency, which at the end makes it hard to interpret the results. Apart from that, the sample size also puts a limit on the extent to which heterogeneity - in particular in terms of context and program characteristics - can be captured by the meta-regression analysis.

[Table 3 about here]

In total, we have 116 observations that can be used for the meta-regression analysis. 40 percent of these relate to business creation while the remaining relate to changes in employment in existing firms. Overall, 46 percent of impact estimates are positive and statistically significant at the 10 percent level. Around 70 percent of the estimates come from evaluations of interventions that are explicitly aimed at creating employment or new enterprises. Almost half of the impact estimates come from joint interventions, and more than half of the impact estimates have been obtained through RCTs.

Furthermore, the majority of evaluations focused on micro-enterprises and was conducted in upper middle income countries.

6.2 Results

Looking at the effectiveness across the different intervention categories, it can be seen that the included finance interventions had on average lower employment effects than the included training interventions (see Table 4). This is confirmed by both specifications, i.e. whether we use the dichotomous outcome or the standardized effect sizes. There are, however, no systematic and significant differences between training interventions and business development services, or wage incentive schemes, or interventions that improved the business environment.

The multivariate analysis confirms many of the findings discussed in the more narrative review above. Interventions of the type we consider seem to have more chances in establishing new firms than expanding existing firms.⁷ Those interventions that target small enterprises appear to be more successful in achieving significantly positive employment effects as compared to those that target micro-enterprises, implying that only a small share of micro-enterprises graduates, or that it is at least difficult to expand micro-enterprises in general. Interestingly, the impacts in medium-sized firms are not significantly different from the impacts measured for micro-enterprises, although descriptively we find positive significant impacts in medium-sized firms more often than in micro-enterprises. It is only if we control for the intervention category that the effect associated with medium-sized firms loses significance, though it remains positive. Separating intervention category effects from firm size effects is not straightforward as the distribution across different intervention categories is very different between micro- and medium-sized firms. Interventions targeted at micro-firms are mainly in the area of finance and training. Interventions targeted at medium-sized firms are mainly in the area of business development and wage subsidies and only to a lesser extent in the area of finance. If we look at finance interventions only, we see that 57 percent of the interventions targeted at medium-sized firms have a positive significant employment effect but only 28 percent of the interventions targeted at micro-enterprises firms. In conclusion, we think it is safe to say that it is less difficult to expand employment in medium-sized firms than in micro-enterprises.

[Table 4 about here]

Surprisingly, whether employment creation has been an explicit objective of the evaluated intervention does not correlate with larger employment effects. Combined interventions also did not systematically lead to larger employment effects, although we had seen above that the specific combination between finance and training often seems to work better than each of these two interventions in isolation.

⁷ However, as McKenzie and Woodruff (2014) pointed out, this finding may partly be due to the fact that power is generally higher for detecting binary outcomes (such as whether a new business started) than for detecting changes in rather continuous variables (such as employment).

Those interventions in our sample that targeted women specifically apparently had a lower chance of success, although this finding is only statistically significant in some specifications. This is consistent with a number of recent studies that find capital or business training to have no systematic effect on female enterprise performance (see for instance de Mel et al., 2008 and 2009; Fafchamps et al., 2011). The causal mechanisms underlying this finding were discussed above.

Finally, the multivariate analysis also confirms that programs that have been evaluated experimentally show fewer significant employment effects than programs that have been evaluated with quasi-experimental methods. Looking at finance interventions, for example, shows that 22 out of the 29 treatment effects that are based on an RCT show insignificant effects, while this is the case for only 10 out of the 25 treatment effects based on quasi-experimental methods. Hence, the failure rate is almost twice as high. To explore this further, we constructed further variables measuring the risk of bias, based on our own assessment for various dimensions of internal and external validity and following the criteria proposed in Waddington and Hombrados (2012). Although low statistical power of many RCTs is an important shortcoming, as we have explained above, controlling for this problem and other risk of biases (such as John Henry and Hawthorne effects, attrition, non-compliance, spill-overs and sample selection) could not substantially reduce the estimated coefficient associated with RCTs. However, we think that the fact that RCTs focus more prominently on small programs, very poor areas and very specific target groups (as compared to evaluations based on quasi-experimental designs) also explains why they show more frequently weak or null effects. In other words, RCTs are often applied in very specific cases and therefore their findings cannot always be generalized.

7. Conclusion

Overall the review shows that creating and enhancing employment is a very complex challenge. Many conditions have to be met before interventions in favor of individual enterprises do not only improve business performance but also lead to additional jobs. It is ‘a long way’ in the result chain from policy inputs to employment impacts, even more so if employment is intended to be sustainable and tied to acceptable and secure working conditions. Given the discrete nature of decisions to create a new business or to hire an additional employee, ‘intense treatments’ combined with the appropriate targeting are needed to have an impact. Surprisingly, the available evidence suggests that it is easier to enhance the creation of a micro- firm than expanding employment in a micro-enterprise, which would suggest that many of these micro-enterprises rapidly attain their optimal firm size. McKenzie and Woodruff (2014) found similar evidence for smaller set of business training and entrepreneurship evaluations. Based on our review, we can generalize this finding also to other types of interventions.

It seems that particularly (micro-) finance and training interventions achieve positive effects only very early in the result chain, improving management practices, skills and investments but without further or lasting results on business performance and, finally, employment. Many interventions seem to lead

to changes at the intensive margin, but fail to deliver productivity increases that go hand in hand with more jobs. This is coherent with the findings by Cho and Honorati (2014). Their meta-analysis of finance and training programs shows that vocational and business training programs in particular have positive effects on business outcomes such as improved knowledge and practice and sometimes income, but effects on a general set of labor market outcomes are quite weak and very often insignificant. They even find training alone to have a negative effect on labor market activity and to be effective only in combination with finance. This is something we cannot confirm. Our results show that even as stand-alone intervention training leads at least on average to positive employment outcomes, above all in terms of the creation of new businesses.

Business development services seem to generate somewhat better employment outcomes, but the evidence mainly comes from small and medium-sized enterprises in Latin American countries and therefore it is difficult to say whether these programs would be equally successful when targeted at micro-enterprises or at any enterprises in the African or Asian context. For interventions that intend to enhance formalization, we faced a similar problem. The evidence mainly comes from Latin America. The overall effects of formalization on employment have been positive in Brazil and Mexico; the magnitude of these effects is still debated. However, these studies also show that it is not easy to encourage firms to formalize in the first place.

Interestingly, our study also revealed that about a third of the interventions covered by this review are not primarily designed to create employment but rather strive for income stabilization and poverty reduction, a fact that has so far been ignored in the literature. Hence, one should not expect massive impacts on job creation if interventions were not even intended to deliver this result. We believe this is basically an issue of design and targeting. If business development in the form of capital accumulation and employment expansion is the objective, the type of entrepreneurs that needs to be targeted is different from those that are targeted through finance and entrepreneurship interventions that aim to reduce poverty and to empower women.

The purpose of such a review is also to highlight still existing knowledge gaps. We show that the available evidence is still sketchy particularly for large parts of Sub-Saharan Africa and Asia, regions in which the need for jobs will be the highest in the coming decades. Findings from Latin America, which dominate in this review, cannot necessarily be generalized and applied to other regions. Also only very few studies are able to assess the longer-term effects of their interventions and policies. Many studies also fail to provide a detailed analysis of why certain effects occurred or did not occur – making it hard to extrapolate lessons. Moreover, it is also hard to tell from the results whether new jobs were created or whether workers were just tracked away from other activities. Furthermore, almost none of the 53 studies provided a detailed cost effectiveness analysis, i.e. how much does it cost to create an additional job with a certain program compared to another? This gap should alert both implementers and researchers. Implementers should provide the necessary numbers and researchers

should go beyond the estimate of simple impacts, which is not really helpful for those who have to allocate resources across different interventions.

Finally, a striking finding of our review is that the study design matters for the impacts found. RCTs, which are typically seen as the ‘gold standard,’ find less frequently positive employment effects than other methods, controlling for the type of intervention, type of country and type of firm that is targeted. This may suggest that in many of the studies which are based on a weaker identification strategy, selection bias is still an issue. However, it can also not be ignored that many RCTs have low statistical power due to small sample sizes and that they are applied particularly to small programs, very poor areas and very specific target groups. Hence, RCTs seem to paint an overly pessimistic picture with respect to the potential of such policies and interventions to create jobs. So far this has been largely ignored. This contrast is not only important with respect to the implications of this study; it is also – we believe –an important insight that can enrich the debate about the advantages and disadvantages of different evaluation designs.

Appendix: Overview of included studies

Reference	Intervention	Job priority	Country	Size of enterprises targeted	Women targeted	Study design	Length of follow-up	Program effect(s)
Access to finance interventions								
Angelucci et al. (2013)	Micro-credit, joint-liability, stand-alone, targeted at women that have a business or self-employment activity or intend to start one.	yes	Mexico (urban, peri-urban, and rural)	Micro	yes	RCT	up to three years; on average 26 months	BC: insignificant E: insignificant
Arraíz et al. (2012)	Government-backed partial credit guarantees for Colombian MSMEs without enough collateral, stand-alone.	yes	Colombia (national)	Small and medium	no	Quasi-experimental: PSM + DID	up to two years	E – current yr: positive E – 1 yr after: positive E – 2 yrs after: positive
Attanasio et al. (2011)	Small loans; two different treatments: group-lending and individual loans, stand-alone. Target group: Relatively poor women in rural areas	yes	Mongolia (rural)	Micro	yes	RCT	1.5 years	BC – indiv. lending: insignificant BC – group lending: insignificant female BC – indiv. lending: insignificant female BC – group lending: positive
Augsburg et al. (2012)	Individual-liability micro-credit to ‘marginal’ borrowers, i.e. loan applicants who would otherwise be excluded from loans because of a lack of collateral, stand-alone.	yes	Bosnia and Herzegovina (national)	Micro	no	RCT	14 months	BC: positive
Banerjee et al. (2011)	Direct transfer of productive assets combined with provision of training (inoculation of savings habits and integration into micro-finance groups) to the ‘ultra poor’, particularly women.	no	India (rural West-Bengal)	Micro	yes	RCT	18 months	BC: insignificant
Banerjee et al. (2013)	Group-liability micro-credit loans ranging between \$200 at market exchange rates (or \$1,000 at PPP-adjusted exchange rates) \$400, stand-alone and targeted to women and the poor, but not the very poor.	no	India (Hyderabad)	Micro	yes	RCT	3 to 3.5 years	E –short term: insignificant E – long term: insignificant BC – short term: insignificant BC – long term: insignificant

Barnes (2001)	Group-liability micro-credit, accompanied by an orientation session that teaches sound business management practices, and loan officers provide management advice.	yes	Zimbabwe (urban)	Micro	no	Quasi-experimental: Matching	2 years	E: insignificant
Blattman et al. (2012)	Nearly unconditional, unsupervised group cash transfers to pay for vocational training, tools, and business start-up costs, stand-alone, targeted at the poor and underemployed youth.	no	Uganda (Northern Region)	Micro	no	RCT	2 years	E – males: significant E – females: negative
Bruhn and Love (2009)	Opening of Banco Azteca in pre-existing stores for electronics and household goods, offering a variety of financial services for low and middle-income customers, previously underserved by the traditional banking industry.	yes	Mexico (national)	Micro	no	Quasi-experimental: DID	up to 9 quarters	BC: positive E: insignificant
Crepón et al. (2011)	Micro-credit, joint-liability, as well as individual-liability targeted at existing enterprises.	no	Morocco (rural/semi-urban)	Micro	no	RCT	2 years	E: insignificant BC: insignificant
Da Silva et al. (2006)	Constitutional financing funds, stand-alone. Target group: firms in the northeastern region of Brazil, in particular SMEs.	yes	Brazil (North/North-East)	Small and medium	no	Quasi-experimental: PSM	3 years	E – program 1: positive E – program 2: positive
De Mel et al. (2013b)	Three different interventions: (1) a subsidized savings program, (2) temporary wage subsidies to incentivize hiring additional employees, and (3) a five-day training program based on the ILO's Improve Your Business (IYB) program, provided jointly (combination of two of these interventions) as well as stand-alone to male-owned enterprises with two or fewer paid employees.	yes	Sri Lanka (urban)	Micro	no	RCT	up to 2 years	E- F: insignificant E – F+T: insignificant E – F+WS: positive
DeNegri et al. (2011)	Public credit lines, stand-alone.	yes	Brazil (national)	Mostly small and medium	no	Quasi-experimental: DID	up to 5 years	E – short term: positive E – long term: positive
Eshetu et al. (2013)	Joint intervention: enabling legal framework and streamlining regulatory conditions as well as specific support services (financial and business development services including a credit and saving scheme trainings, technology transfer, counseling, provision of working premises etc.	yes	Ethiopia (urban Dire Dawa)	Micro	no	Quasi-experimental: PSM	up to 6 years	E: positive
Eslava et al. (2012)	Public credit lines, stand-alone.	yes	Colombia (national)	Small and medium	no	Quasi-experimental: DID and PSM	up to 4 years	E: positive
Field et al. (2011)	Introduction of a grace period of two months for the initiation of the repayment of micro-loans. Normally repayments start after two weeks.	no	India (Kolkata)	Micro	yes	RCT	3 years	BC: insignificant E: insignificant

Gubert and Roubaud (2011)	Individual-liability loans averaging €500 for urban micro-businesses and longer-term loans (from 24 to 36 months) to small and medium-sized enterprises (SMEs) averaging € 8,000 euros.	no	Madagascar (urban)	Micro	no	Quasi-experimental: Matching	up to 4 years	E: insignificant
Kaboski and Townsend (2005)	Village-level (micro-finance) institutions: (1) Production micro-credit groups; (2) Rice bank; (3) Women's group; (4) Buffalo banks.	no	Thailand (rural/semi-urban North-East and Central)	Micro	no	Quasi-experimental: IV	unclear	BC: insignificant
Karlan et al. (2012)	Two types of interventions: (1) specific management consulting services and (2) unconditional cash grant of approximately US \$133- Provided stand-alone as well as jointly to tailors and seamstresses.	no	Ghana (urban)	Micro	no	RCT	up to 11 months after the consulting stopped and 14 months after the capital drops	E – F: insignificant E – F+T: insignificant
Karlan and Zinman (2011)	Individual liability micro-credit, loans ranging from 5,000 to 25,000 pesos for existing enterprises, stand-alone.	yes	Philippines (two provinces and capital region)	Micro	no	RCT	11-22months	E:negative BC: negative
Kondo et al. (2008)	Group-liability micro-credit, stand-alone.	no	Philippines (national)	Micro	no	Quasi-experimental: DID	3 to 5 years	BC: positive E: positive
Macours et al. (2012)	CCT complemented either by vocational training or a productive investment grant targeted at rural households.	no	Nicaragua (rural)	Micro	no	RCT	2 years	BC: positive
Nelson (2011)	Large-scale, publicly-funded micro-finance initiative which helped to set-up and to support independent village banks.	yes	Thailand (rural/semi-urban North-East and Central)	Micro and small	no	Quasi-experimental: IV	up to 6 years	BC – low wealth: negative BC – middle wealth: positive BC – high wealth: insignificant
Resende (2012)	Constitutional Financing Funds: Loans at subsidized interest rates, stand-alone, targeted at micro- and small rural and industrial producers.	yes	Brazil (North/North-East)	Small and medium	no	Quasi-experimental: Matching	up to 6 years	E – 3yr period: positive E – 6yr period: positive
Tan (2009)	Seven different matching grants and credit programs and two other, open-ended, residual programs.	yes	Chile (national)	Small and medium	no	Quasi-experimental: DID + PSM	up to 10 years	E – techn. Assistance (BDS): insignificant E – cluster formation (BDS): insignificant E – technology dev. (BDS): insignificant E –any BDS: positive

Tarozzi et al. (2013)	Joint-liability micro-credit, combined with family planning program	yes	Ethiopia (rural)	Micro	no	RCT	up to 2 years	BC – Oromiya: insignificant BC – Amhara: insignificant
Entrepreneurship training								
Bah et al. (2011)	Financial and/or technical assistance analyzed jointly.	yes	Macedonia (national)	Mostly micro and small	no	Quasi-experimental: matching	up to 3 years	E – 1 st yr: positive E – 2 nd yr: positive E – 3 rd yr: positive
Bandiera et al. (2012)	Joint intervention that simultaneously provides: (1) vocational training to run/start small-scale enterprises; and (2) information on health and risky behaviors. Courses are supplemented by financial literacy courses and targeted at adolescent girls aged 14-20.	yes	Uganda (rural, urban and semi-urban)	Micro; target group not necessarily existing entrepreneurs	yes	RCT	2 years	BC: positive
Banerjee et al. (2011)	Direct transfer of productive assets combined with provision of training (inoculation of savings habits and integration into micro-finance groups) to the ‘ultra poor’, particularly women.	no	India (rural West-Bengal)	Micro; target group not necessarily existing entrepreneurs	yes	RCT	18 months	BC: insignificant
Bruhn and Zia (2011)	Three-day business and financial education training for micro-credit clients.	yes	Bosnia and Herzegovina (urban)	Micro	no	RCT	6 months	E: insignificant BC: insignificant
Bruhn et al. (2013)	Subsidized consulting and mentoring services for owners/managers of formal businesses. Consultants were asked to (1) diagnose the problems that prevented the enterprises from growing, (2) suggest solutions and (3) assist in implementing the solutions.	no	Mexico (Puebla)	Mostly micro and small	no	RCT	up to one year (short term) and between 1-3 years (long term)	E – short term: insignificant E – long term: positive
Calderon et al. (2013)	Stand-alone basic business training provided at no cost, focusing on the application of the concept discussed in class on the participants’ businesses. Target group: small, female headed firms in the retail and production sector.	no	Mexico (rural)	Micro	yes	RCT	up to 1 year (short term effects) and about 2.5 years (medium term effects)	E – below median profit: insignificant E – above median profit: insignificant
Cho et al. (2012)	Vocational training apprenticeship combined with entrepreneurial support and life skills training and, in some cases, start-up capital. Target group: vulnerable youth who are poor, orphaned, HIV/AIDS vulnerable, school dropouts.	yes	Malawi (national)	Micro; target group not necessarily existing entrepreneurs	no	RCT	4 months	BC: insignificant

De Mel (2013b)	Three different interventions: (1) a subsidized savings program, (2) temporary wage subsidies to incentivize hiring additional employees, and (3) a five-day training program based on the ILO's Improve Your Business (IYB) program, provided jointly (combination of two of these interventions) as well as stand-alone to male-owned enterprises with 2 or fewer paid employees.	yes	Sri Lanka (urban)	Micro	no	RCT	up to 2 years	E – T: insignificant E – T+F: insignificant E – T+WS: positive
De Mel et al. (2012)	ILO's 'Start-and-Improve Your Business program' provided to female current as well as potential business owners.	yes	Sri Lanka (urban)	Micro	yes	RCT	2 years	BC – T – short term: positive BC – T – long term: insignificant BC – T+F – short term: positive BC – T+F – long term: insignificant
Drexler et al. (2013)	Two different stand-alone training interventions: (1) Standard accounting training, and (2) rule-of-thumb training, which taught participants simple rules for financial decision making. Target group: micro-entrepreneurs interested in training.	no	Dominican Republic (urban)	Mostly micro	no	RCT	1 to 2 years	E: insignificant
Galasso et al. (2004)	Vouchers that entitled to hire an employee at a subsidized wage for 18 months. In a variant of that intervention employees received special skill training.	yes	Argentina (urban)	Micro	no	RCT	18 months	BC: positive
Giné and Mansuri (2011)	Hands-on business training based on the ILO's 'Know About Business' modules was added to micro-finance. In addition, one-on-one follow-up training sessions and beneficiaries had the opportunity to participate in a lottery for a loan up to seven times the average loan size.	no	Pakistan (rural)	Micro	no	RCT	18 months after training and 6 months after loan lottery	BC – beneficiary involved: insignificant BC – beneficiary not involved: insignificant
Jaramillo and Parodi (2003)	Two different programs providing training and finance to low-income/poor youth (18-25 years) either already owning a micro-enterprise or interested in establishing one. Focus was on business plan development.	yes	Peru (urban)	Micro; not necessarily existing entrepreneurs	no	Quasi-experimental: PSM	3 months (CID program) and 11 months ('JUMP' program)	E – CID: positive BC – JUMP: positive
Karlan and Valdivia (2011)	Training is added to microcredit program. The training included general business skills and strategy training, not client-specific problem solving. Target group: female micro-entrepreneurs who are micro-credit clients.	yes	Peru (regional)	Micro	yes	RCT	up to 2 years	E: insignificant BC: insignificant
Klinger and Schündeln (2011)	Multi-phased business plan competition	yes	El Salvador, Guatemala and Nicaragua (within-country distribution)	Micro and small; not necessarily existing entrepreneurs	no	Quasi-experimental: RDD	1 to 3 years	BC: positive

			unclear)					
Lopez-Acevedo and Tinajero-Bravo (2010)	Subsidies for SMEs to (1) hire independent instructors to design and deliver training, and (2) to reduce the costs of producing training materials, developing training programs, and assessing workers' skills based on labor competency standards.	yes	Mexico (national)	Mostly medium size	no	Quasi-experimental: PSM	up to 5 years	E: negative
Macours et al. (2012)	CCT complemented either by vocational training or a productive investment grant. Target group: Rural households	no	Nicaragua (rural)	Micro; not necessarily existing entrepreneurs	no	RCT	2 years	BC: positive
Premand et al. (2012)	Introduction of an innovative entrepreneurship track in the university curriculum, including entrepreneurship courses, external private sector coaching delivered by entrepreneurs, and business plan development. Target group: University students.	yes	Tunisia (national)	Micro; not necessarily existing entrepreneurs	no	RCT	up to 1 year	BC: positive
Steiner et al. (2010)	Stand-alone business training program promoting productive activities in the agricultural sector; and agro-industrial sector, as well as in services and industry and targeted at the unemployed youth (16-25) in rural and remote areas.	no	Colombia (rural)	Micro; not necessarily existing entrepreneurs	no	Quasi-experimental: PSM and DID	about 1 year	BC: significant
Valdivia (2011)	Stand-alone training (general and individualized). There were two different treatments: (1) Regular business training consisting of personal development, business development and management and productivity improvements; and (2) Additional individualized support in the form of technical assistance (TA) Target group: Female micro-entrepreneurs in Lima.	no	Peru (urban)	Micro	yes	RCT	10 months	BC – general training: insignificant BC – general + specific training: insignificant

Business development services (BDS)

Arráz et al. (2013)	Supplier development program providing public subsidies for projects aimed at strengthening the management of SMEs that supply large firms. Each project must include between 10 to 20 SMEs.	no	Chile (national)	SMEs	no	Quasi-experimental: PSM	up to three years	E: positive
Benavente (2007)	Technology development fund providing matching-grants for projects aimed at developing new products and improving production processes. Covers the development of prototypes and market testing.	yes	Chile (national)	SMEs	no	Quasi-experimental: PSM+DID	unclear	E: positive
Castillo et al. (2011)	Co-financing (up to 50%) for product innovation or process innovation.	no	Argentina (national)	SMEs	no	Quasi-experimental: PSM + DID	up to 8 years	E – product innovation: positive E – process innovation:

								positive
Eshetu et al. (2013)	Joint intervention: enabling legal framework and streamlining regulatory conditions as well as specific support services (financial and business development services including a credit and saving scheme trainings, technology transfer, counseling, provision of working premises etc.).	yes	Ethiopia (urban Dire Dawa)	Micro and small	no	Quasi-experimental: PSM	up to 6 years	E: positive
Lopez-Acevedo and Tinajero-Bravo (2010)	Four different programs including: (1) tax breaks, (2) finance audits and support for investments to reduce environmental risks, (3) fiscal incentives for technological innovation, and (4) a training of the industrial workforce.	yes	Mexico (national)	MSMEs	no	Quasi-experimental: PSM	up to 5 years	E – tax breaks: positive E – environment audit: negative E – fiscal incentives and techn. innovation: positive E – other state support: insignificant E – any program: positive
Tan (2009)	Seven different matching grants and credit programs, and two other, open-ended, residual programs.	yes	Chile (national)	Mostly medium-sized	no	Quasi-experimental: DID+PSM	up to 10 years	E – techn. assistance (BDS): insignificant E – cluster formation (BDS): insignificant E – technology dev. (BDS): insignificant E – any BDS: positive

Private sector incentive schemes (wage subsidies)

Betcherman et al. (2010)	Social security contribution and wage subsidies as well as land and electricity subsidies (the latter for particular cases).	yes	Turkey (national)	Small and medium-sized formal firms	no	Quasi-experimental: PSM + DID	1-2 years	E: positive
De Mel et al. (2013b)	Three different interventions: (1) a subsidized savings program, (2) temporary wage subsidies to incentivize hiring additional employees, and (3) a five-day training program based on the ILO's Improve Your Business (IYB) program, provided jointly (combination of two of these interventions) as well as stand-alone to male-owned enterprises with 2 or fewer paid employees.	yes	Sri Lanka (urban)	Micro	no	RCT	up to 2 years	E – WS: insignificant E – WS + F: positive E – WS + T: positive
Fretwell et al.	On-the-job training, averaging 4.5 months in length. Contracts with training agencies and enterprises were 'performance-based' with pre-	yes	Turkey (urban)	Small and medium-	no	Quasi-experimental:	1-2 years (unclear)	E: insignificant

(1999) negotiated job placement rates and trainees were provided with a token amount for living and travel expenses. sized Matching

Business environment: Interventions to promote formalization (BE)

Bruhn (2011)	Federal program targeted at municipalities that ‘one-stop’ firm registration offices allowing small firms to obtain a license to operate in two days or less and to postpone health and social security inspections for three months. The program reduced registration procedures from 30 to 2 days.	yes	Mexico (national)	MSMEs	no	Quasi-experimental: DID	up to 4 years	BC: positive
Courseuil and Moura (2011)	A tax incentives program that combines, simplifies and promotes the collection of federal taxes from micro-firms and small companies, with lower, though progressive, tax rates on the same base for calculation (gross revenue). The program combines reductions both in monetary and administrative costs of tax payment.	yes	Brazil (national)	Mostly medium-sized	no	Quasi-experimental: RDD	1 year	E: positive
de Mel et al. (2013a)	Information about procedures and costs and benefits of formalization, provided either alone or with capital.	yes	Sri Lanka (urban)	Mostly micro	no	RCT	up to three years	E: insignificant
Fajnzylber et al. (2011)	Simplified tax system for micro-firms, including also an overall reduction of taxes of up to 8%.	yes	Brazil (urban)	Micro	no	Quasi-experimental: RDD, IV and DID	up to 1 year	E – all micro-firms: positive E – firms with at least one employee: positive
Kaplan et al. (2011)	Federal program targeted at municipalities that ‘one-stop’ firm registration offices allowing small firms to obtain a license to operate in two days or less and to postpone health and social security inspections for three months. The program reduced registration procedures from 30 to 2 days.	yes	Mexico (national)	MSMEs	no	Quasi-experimental: DID	up to 4 years	E: positive BC: positive

Source: Own representation.

References

References of studies included in the systematic review

- Angelucci, M., D. Karlan and J. Zinman (2013), 'Win Some Lose Some? Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco', IZA Discussion Paper Series, No. 7439, IZA.
- Arráiz, I., F. Henríquez and R. Stucchi (2013), 'Supplier development programs and firm performance: evidence from Chile', *Small Business Economics*, Vol. 41, pages 277-293.
- Arráiz, I., M. Meléndez and R. Stucchi (2012), 'Partial Credit Guarantees and Firm Performance: Evidence from the Colombian National Guarantee Fund', OVE Working Paper, No. 02/2012, Office of Evaluation and Oversight (OVE), Inter-American Development Bank.
- Attanasio, O., B. Augsburg, R. De Haas, E. Fitzimons and H. Harmgart (2011), 'Group lending or individual lending? Evidence from a randomised field experiment in Mongolia', IFS Working Paper, No. 11/20, Institute for Fiscal Studies.
- Augsburg, B., R. De Haas, H. Heimgart and C. Meghir (2012), 'Microfinance at the margin: experimental evidence from Bosnia and Herzegovina', EBRD Working Paper, No. 146, European Bank of Reconstruction and Development.
- Bah, E. H., J. C. Brada and T. Yigit (2011), 'With a little help from our friends: the effect of USAID assistance on SME growth in a transition economy', *Journal of Comparative Economics*, Vol. 39, pages 205-220.
- Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul and M. Sulaiman (2012), 'Empowering adolescent girls: Evidence from a randomized control trial in Uganda', Mimeo, London School of Economics.
- Banerjee, A., E. Duflo, R. Chattopadhyay and J. Shapiro (2011), 'Targeting the Hard-Core Poor: An Impact Assessment', Mimeo, Poverty Action Lab, MIT.
- Banerjee, A., E. Duflo, R. Glennerster and C. Kinnan (2013), 'The miracle of microfinance? Evidence from a randomized evaluation', Mimeo, Poverty Action Lab, MIT.
- Barnes, C. (2001), 'Microfinance Program Clients and Impact: An assessment of Zambuko Trust, Zimbabwe', Assessing the Impact of Microenterprise Services, Report, United States Agency for International Development.
- Benavente, J.M., G. Crespi and A. Maffioli (2007), 'Public support to firm-level innovation: an evaluation of the FONTEC program', OVE Working Paper, No. 05/2007, Office of Evaluation and Oversight (OVE), Inter-American Development Bank.
- Betcherman, G., N. Meltem Daysal and C. Pagés (2010), 'Do employment subsidies work? Evidence from regionally targeted subsidies in Turkey', *Labor Economics*, Vol. 17, pages 710-722.
- Blattman, C., N. Fiala and S. Martinez (2012), 'Employment Generation in Rural Africa – Mid-Term Results from an Experimental Evaluation of the Youth Opportunities Program in Northern Uganda', DIW Discussion Papers, No. 1201, DIW.
- Bruhn, M. and L. Inessa (2009), 'The Economic Impact of Banking the Unbanked. Evidence from Mexico', World Bank Policy Research Working Paper, No. 4981, World Bank.
- Bruhn, M. and B. Zia (2011), 'Stimulating Managerial Capital in Emerging Markets. The Impact of Business and Financial Literacy for Young Entrepreneurs', World Bank Policy Research Working Paper, No. 5642, World Bank.
- Bruhn, M. (2011), 'License to sell: the effect of Business Registration Reform on Entrepreneurial Activity in Mexico', *Review of Economics and Statistics*, Vol. 93(1), pages 283-286.
- Bruhn, M., D. Karlan and A. Schoar (2013), 'The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico', World Bank Policy Research Working Paper, No. 6508, World Bank.

- Calderon, G., J. M. Cunha and G. de Giorgi (2013), 'Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico', Mimeo, Stanford University.
- Castillo, V., A. Maffioli, S. Rojo and R. Stucci (2011), 'Innovation Policy and Employment', IADB Working Paper, No. IDB-TN-341, Inter-American Development Bank.
- Cho, Y., D. Kalomba, M. Mobarak and V. Orozco (2012), 'The effects of apprenticeship training for vulnerable youth in Malawi', Mimeo, World Bank.
- Courseuil, C. H. and R. L. de Moura (2011), 'The Effects Of Tax Incentives For Small Firms On Employment Levels', FGV Working Paper series 'Ensaio Econômicos', No. 701, FGV/EPGE Escola Brasileira de Economia e Finanças, Getulio Vargas Foundation, Brazil.
- Crepón, B., F. Devoto, E. Duflo and W. Pariente (2011), 'Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation', MIT Working Paper, No. 6659, MIT.
- Da Silva, A. M. A., G. Mendes Resende, S. da Mota and R. Neto (2006), 'Avaliação Econômica dos fundos constituicionais de financiamento do Nordeste e do Norte (FNE e FNO)', Discussion Paper, No. 1207, Instituto de Pesquisa Economica Aplicada (IPEA), Brazil.
- De Mel, S., D. McKenzie and C. Woodruff (2012), 'Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka', World Bank Policy Research Working Paper, No. 6145, World Bank.
- De Mel, S., D. McKenzie and C. Woodruff (2013a), 'The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka', *American Economic Journal: Applied Economics*, Vol. 5 (2), pages 122-50.
- De Mel, S., D. McKenzie and C. Woodruff (2013b), 'What generates Growth in Microenterprises? Experimental Evidence on Capital, Labour and Training', Mimeo, World Bank.
- De Negri, J.A., A. Maffioli, C. M. Rodriguez and G. Vázquez (2011), 'The Impact of Public Credit Programs on Brazilian Firms', Inter-American Development Bank Working Paper, No. IDB-WP-293, Inter-American Development Bank.
- Drexler, A., G. Fischer and A. Schoar (2013): 'Keeping it Simple: Financial Literacy and Rules of Thumb', CEPR Discussion Paper, No. DP7994, CEPR .
- Eshetu, T., M. Ketema and H. Kassa (2013), 'Economic Impact of Support Service Program on Micro and Small Enterprises: The Case of Dire Dawa Administration, Ethiopia', *Agris on-line Papers in Economics and Informatics*, Vol. 5(1), pages 21-29.
- Eslava, M., A. Maffioli and M. Meléndez (2012), 'Second-tier Government Banks and Firm Performance. Micro-Evidence from Colombia', IADB Working Paper, No. IDB-WP-294, Inter-American Development Bank.
- Fajnzylber, P. W., F. Maloney and Gabriel V. Montes-Rojas (2011), 'Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program', *Journal of Development Economics*, Vol. 94, pages 262-276.
- Field, E., P. Rohini, J. Papp and N. Rigol (2011), 'Term Structure of Debt and Entrepreneurship: Experimental Evidence from Microfinance', Mimeo, Harvard University.
- Fretwell, D. H., J. Benus and C. J. O'Leary (1999), 'Evaluating the Impact of Active Labor Market Programs: Results of Cross Country Studies in Europe and Asia', Social Protection Discussion Paper Series, No. 9915, World Bank.
- Galasso, E., M. Ravallion and A. Salvia (2004), 'Assisting the transition from workfare to work: A randomized experiment', *Industrial and Labor Relations Review*, Vol. 58(1), pages 128-142.
- Giné, X. and G. Mansuri (2011), 'Money or ideas? A Field Experiment on Constraints to entrepreneurship in rural Pakistan', Mimeo, World Bank.

- Gubert, F. and F. Roubaud (2011), 'The impact of Microfinance Loans on Small Informal Enterprises in Madagascar. A panel data analysis', Social Protection Discussion Paper Series, No. 77931, World Bank.
- Jaramillo, M. and S. Parodi (2003), 'Jóvenes emprendedores', Instituto Apoyo, Peru.
- Kaboski, J. P. and R. M. Townsend (2005), 'Policies and impact: An analysis of village-level microfinance institutions', *Journal of the European Economic Association*, Vol. 3(1), pages 1–50.
- Kaplan, D. S., E. Pedra and E. Seira (2011), 'Entry regulation and business start-ups: Evidence from Mexico', *Journal of Public Economics*, Vol. 95, pages 1501-1515.
- Karlan, D., R. Knight and C. Udry (2012), 'Hoping to Win, Expected to Lose: Theory and Lessons on Microenterprise Development', CGD Working Paper, No. 312, Center for Global Development.
- Karlan, D. and M. Valdivia (2011), 'Teaching Entrepreneurship: Impact of Business Training on Microfinance clients and Institutions', *Review of Economics and Statistics*, Vol. 93(2), pages 510-527.
- Karlan, D. and J. Zinman (2011), 'Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation', *Science*, Vol. 332(1278).
- Klinger, B. and M. Schündeln (2011), 'Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America', *World Development*, Vol. 39(9), pages 1592-1610.
- Kondo, T., A. Orbeta, Jr., C. Dingcong and C. Infantado (2008), 'Impact of microfinance on rural households in the Philippines', PIDS Discussion Paper Series, No. 05/2008, Philippine Institute for Development Studies.
- Lopez-Acevedo, G. and M. Tinajero-Bravo (2010), 'Evaluating Enterprise support programs using panel firm data', World Bank Policy Research Working Paper, No. 5186, World Bank.
- Macours, K., P. Premand and R. Vakis (2012), 'Transfers, Diversification and Household Risk Strategies', World Bank Policy Research Working Paper, No. 6053, World Bank.
- Nelson, L. K. (2011), 'From Loans to Labor: Access to Credit, Entrepreneurship and Child Labor', Mimeo, University of California.
- 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', IZA Discussion Paper, No. 7079, IZA.
- Resende, G. M. (2012), 'Measuring Micro- and Macro-Impacts of Regional Development Policies: the case of the Northeast Regional Fund (FNE) Industrial Loans in Brazil, 2000-2006', *Regional Studies*, No. 01/2012, pages 1-19.
- Steiner, R., P. Acosta, N. Rojas, N. Millán, C. Castañeda and J. L. Gonzáles (2010), 'Evaluación de impacto del programa jóvenes rurales emprendedores del servicio nacional de aprendizaje – SENA', Unpublished Evaluation Report, FEDESARROLLO - Centro de Investigación Económica y Social, Colombia.
- Tan, H. (2009), 'Evaluating SME Support Programs in Chile Using Panel Firm Data', World Bank Policy Research Working Paper, No. 5082, World Bank.
- Tarozzi, A., J. Desai, K. Johnson and (2013): 'On the Impact of Microcredit: Evidence from a Randomized Intervention in Rural Ethiopia', BREAD Working Paper, No. 382, BREAD.
- Valdivia, M. (2011), 'Training or Technical Assistance? A Field Experiment to Learn What Works to Increase Managerial Capital for Female Microentrepreneurs', CAF Working Papers, No. 02/2011, CAF – Banco de Desarrollo de América Latina, Venezuela.

Other References

Andrade, G.H., M. Bruhn and D. McKenzie (2013), A Helping Hand of the Long Arm of the Law. Experimental Evidence on what Governments Can Do to Formalize Firms. World Bank Policy Research Working Paper, No. 6435, World Bank.

Ayyagari, M., A. Demirguc-Kunt and V. Maksimovic (2011), 'Small vs. Young Firms Across the World: Contribution to Employment, Job Creation, and Growth', World Bank Policy Research Working Paper, No. 5631, World Bank.

Betchermann, G. (2014), 'Labour Market Regulations: What do we know about their Impacts in Developing Countries?', forthcoming in *World Bank Research Observer*.

Betcherman, G., M. Godfrey, S. Puerto, F. Rother and A. Stavreska (2007), 'A review of interventions to support young workers: findings of the youth employment inventory', Social Protection Discussion Papers, No. 41412, World Bank.

Bruhn, M. and D. McKenzie (2013), Entry Regulation and Formalization of Microenterprises in Developing Countries, World Bank Policy Research Working Paper, No. 6507, World Bank.

Cho, Y. and M. Honorati (2014), 'Entrepreneurship programs around the Developing World: A Meta Analysis', forthcoming in *Labour Economics*.

De Mel, S., D. McKenzie and C. Woodruff (2008), 'Returns to capital in microenterprises: evidence from a field experiment', *The Quarterly Journal of Economics*, Vol. 123(4), pages 1329-1372.

De Mel, S., D. McKenzie and C. Woodruff (2009), 'Are women more credit constrained? Experimental evidence on gender and microenterprise returns', *American Economic Journal, Applied Economics*, Vol. 1(3), pages 1-32.

De Mel, S., D. McKenzie and C. Woodruff (2010), 'Wage subsidies for microenterprises', *American Economic Review, Papers and Proceedings*, Vol. 100, pages 614-619.

De Soto, H. (1989), 'The Other Path', Harper and Row, New York.

Fafchamps, M., D. McKenzie, S. R. Quinn and C. Woodruff (2011), 'When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana', NBER Working Papers, No. 17207, NBER.

Grimm, M. and A. L. Paffhausen (2014), 'Interventions for employment creation in micro, small and medium-sized enterprises – A systematic review', KfW-Research 'Studies and Proceedings', KfW Development Bank.

Kluge, J. (2010), 'The effectiveness of European active labor market programs'. *Labour Economics*, Vol. 17, pages 904-918.

Lipsey, M. W., D. B. Wilson (2001), 'Practical Meta-Analysis', Practical Meta Analysis. Applied Social Research Methods Series, Vol. 49.

McKenzie, D. (2010), Impact Assessments in Finance and Private Sector Development: What Have We Learned and What Should We Learn? *World Bank Research Observer*, Vol. 25 (2), pages 209-233-

McKenzie, D. and C. Woodruff (2014), What are we learning from business training and entrepreneurship evaluations around the developing world?, *World Bank Research Observer*, Vol. 29 (1), pages 48-82.

Puerto, O. S. (2007), 'Labor Market Impact on Youth: A Meta-Analysis of the Youth Employment Inventory', Mimeo, World Bank.

Tripney, J., J. Hombrados, M. Newman, K. Hovish, C. Brown, K. Steinka-Fry, E. Wilkey (2013), 'Post-Basic Technical and Vocational Education and Training (TVET) Interventions to Improve Employability and Employment of TVET Graduates in Low- and Middle-Income Countries: A Systematic Review', Campbell Collaboration.

Vaessen, J., F. L. Leeuw, S. Bonilla, A. Rivas, R. Lukach, J. Bastiaensen and N. Holvoet (2012), 'The effect of microcredit on women's control over household spending in developing countries', Protocol, Campbell Collaboration.

Waddington, H. and J. G. Hombrados (2012), 'Tool to assess risk of bias and internal validity of social experiments and quasi-experiments', Mimeo, 3ie.

World Bank (2012), 'World Development Report 2013: Jobs', World Bank.

Tables and Figures

Table 1: Regional distribution and basic characteristics of included studies

	Finance	Training	BDS/Wage	Formalization	Total
Region					
Latin America & Caribbean	9	11	5	4	29
Sub-Saharan Africa	6	2	1	0	9
South Asia	4	4	1	1	10
East Asia & Pacific	5	0	0	0	5
Europe & Central Asia	1	2	2	0	5
Middle East & North Africa	1	1	0	0	2
Firm size					
Micro	20	17	2	2	41
Small	3	2	2	2	9
Medium	3	1	5	1	10
Stand-alone or joint					
Stand-alone	18	6	1	3	28
Joint	6	12	7	2	27
Both	2	2	1	0	5
Empl. creat. primary objective					
Yes	16	11	7	5	39
No	10	9	2	0	21
Study design					
RCT	13	15	1	1	30
Quasi-experimental	13	5	8	4	30
Total	26	20	9	5	60

Notes: The total count does not add up to 53 because some studies are included in more than one category.

Source: Own computations based on information made available by the evaluations.

Table 2: Distribution of standardized effect sizes by intervention area

	Finance		Training		BDS/Wage		Formalization	
	Count	Share (%)	Count	Share (%)	Count	Share (%)	Count	Share (%)
Negative effect size (<0)	13	24.1	8	22.2	2	10.5	0	0
Small effect size (>0, <0.2)	33	61.1	16	44.4	12	57.9	5	71.4
Medium effect size (>0.2, <0.5)	7	13.0	5	13.9	2	10.5	0	0
Large effect size (>0.5, <1)	1	1.9	7	19.4	3	15.8	2	28.6
Total	54	100	36	100	19	100	7	100

Notes: Effect sizes are computed as the standardized mean difference (SMD), i.e. as the ratio between the change in the outcome due to the intervention divided by the standard deviation of the outcome in the control group (or at baseline). If the outcome is a binary outcome such as ‘having set up a firm or not’ the risk ratio is computed (-1). In those studies where such impact measurements were not directly provided they were computed based on the available information. However, some studies do not provide all the necessary information, in these cases we based the estimate just on the reported *t*-values of the impact and the sample sizes of treatment and control groups using the formulas given in Lipsey and Wilson (2001). This implies that effect sizes are not fully comparable across studies and hence can only roughly reflect the order of magnitude of program impacts.

Source: Own computations based on information made available by the evaluations.

Table 3: Description of sample of impacts used for the meta-regression analysis

Variable	Observations	Mean	Std. Dev.	Min	Max
Positive significant program effect	116	0.457	0.500	0	1
Effect size ^{a)}	115	0.145	0.283	-0.891	1.5
Training	116	0.310	0.465	0	1
Finance	116	0.466	0.501	0	1
BDS	116	0.121	0.327	0	1
Private sector incentives and business environment	116	0.103	0.306	0	1
Empl. creat. primary objective	116	0.690	0.465	0	1
Joint intervention	116	0.491	0.502	0	1
Intervention targets women	116	0.198	0.400	0	1
Evaluation based on an RCT	116	0.534	0.501	0	1
Micro-enterprises	116	0.664	0.474	0	1
Small enterprises	116	0.147	0.355	0	1
Medium enterprises	116	0.190	0.394	0	1
LIC	116	0.086	0.282	0	1
LMIC	116	0.345	0.477	0	1
UMIC	116	0.569	0.497	0	1
Measured outcome is business creation	116	0.371	0.485	0	1

Notes: ^{a)} Regarding the computation of effect sizes, see Note to Table 2.

Source: Own representation.

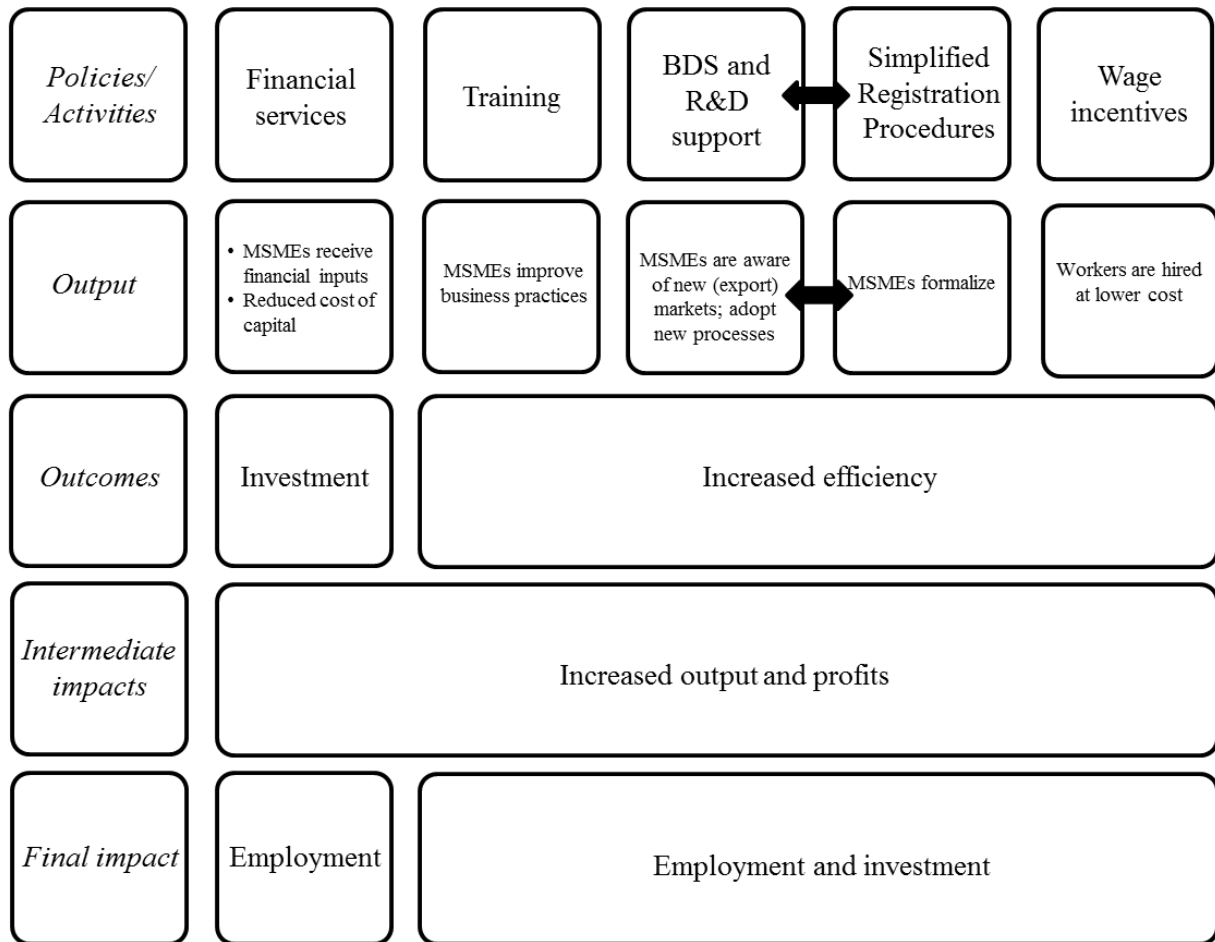
Table 4: Results from the meta-regression analysis

	(1) Positive significance, unweighted	(2) Positive significance, unweighted	(3) Positive significance, unweighted	(4) Positive significance, unweighted	(5) Positive significance, weighted	(6) Effect size, unweighted	(7) Effect size, weighted
Program type							
Training	Ref.			Ref.	Ref.	Ref.	Ref.
Finance	-0.048 (0.123)			-0.227 (0.147)	-0.294** (0.136)	-0.217** (0.102)	-0.245* (0.130)
Business development services	0.225 (0.156)			-0.012 (0.186)	-0.007 (0.197)	-0.0723 (0.123)	-0.082 (0.134)
Private sector incentives and business environment	0.332** (0.138)			0.154 (0.205)	0.060 (0.225)	-0.105 (0.112)	-0.106 (0.144)
Firm size							
Micro-enterprises		Ref.		Ref.	Ref.	Ref.	Ref.
Small enterprises		0.482*** (0.104)		0.410** (0.173)	0.371** (0.186)	0.126 (0.136)	0.0457 (0.154)
Medium-sized enterprises		0.176 (0.170)		0.107 (0.250)	0.0712 (0.243)	-0.102 (0.141)	-0.182 (0.174)
Empl. creat. primary objective		0.0856 (0.133)		0.050 (0.137)	0.024 (0.152)	0.022 (0.083)	0.040 (0.098)
Joint intervention		0.0001 (0.128)		-0.110 (0.170)	-0.209 (0.148)	0.0004 (0.0739)	-0.032 (0.0871)
Intervention targets women		-0.244* (0.128)		-0.216 (0.140)	-0.272* (0.143)	-0.027 (0.081)	-0.072 (0.101)
Evaluation based on an RCT				-0.334* (0.176)	-0.365** (0.172)	-0.107 (0.140)	-0.087 (0.156)
Outcome is business creation			-0.135 (0.097)	0.130 (0.132)	0.240* (0.138)	0.014 (0.051)	0.011 (0.055)
Country income category							
LIC				Ref.	Ref.	Ref.	Ref.
LMIC				-0.053 (0.214)	-0.035 (0.235)	-0.102 (0.114)	-0.111 (0.125)
UMIC				-0.186 (0.233)	-0.146 (0.247)	-0.106 (0.150)	-0.072 (0.180)
Effect size measure is SMD						0.138 (0.098)	0.154 (0.121)
Square root of sample size				-0.0001 (0.0002)	-0.0001 (0.0002)	-0.0002 (0.0001)	-0.0002 (0.0001)
Intercept						0.317 (0.215)	0.340 (0.243)
Observations	116	116	116	116	116	115	115
Pseudo R2	0.051	0.156	0.013	0.210	0.205	0.220	0.197

Notes: Robust standard errors clustered at the level of studies in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SMD means 'Standardized Mean Difference'.

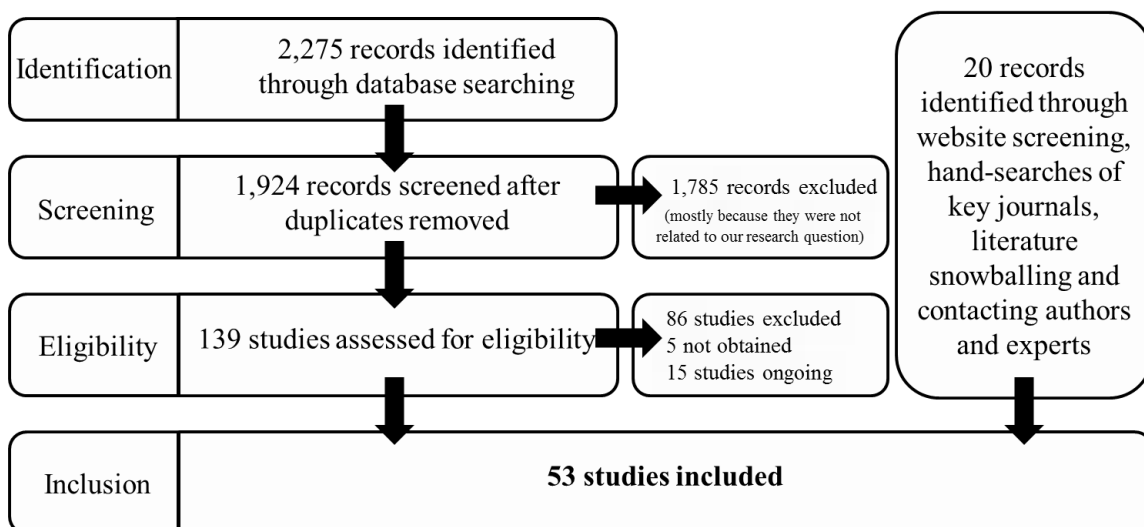
Source: Own data.

Figure 1: Simplified results chain linking interventions and employment outcomes



Source: Own representation.

Figure 2: Selection of studies



Source: Own representation.