MATHEMATICA
Policy Research

REPORT

Evaluation Design for the Georgia II Industry-Led Skills and Workforce Development Project

August 15, 2017

Evan Borkum

Ira Nichols-Barrer

Irina Cheban

CONTENTS

TABLES

FIGURES

1. Introduction

Recognizing that workforce education gaps and skills shortages significantly hinder Georgian economic growth, the Government of Georgia and the Millennium Challenge Corporation (MCC) developed a $140 million compact to improve the quality of education in science, technology, engineering, and math, and in turn foster a more skilled Georgian labor force. This work builds on other Government of Georgia efforts, including a 2013 Vocational Education and Training Development Strategy, with the goal of making vocational programs that meet the economy's labor skills needs more available and flexible (Ministry of Education and Science [MES] 2013). The five-year compact, which entered into force in July 2014, includes three projects that focus on general education, workforce development, and higher education. The Industry-led Skills and Workforce Development (ISWD) project, with a total investment of $16 million, aims to increase the supply of Georgians with technical skills relevant to the local economy through investments in Technical and Vocational Education and Training (TVET). MCC recently contracted with Mathematica Policy Research to evaluate the implementation and potential effects of the ISWD project.

This report describes Mathematica's design for the evaluation of the ISWD project. We begin by presenting an overview of the ISWD program logic, and briefly reviewing the existing literature on the impacts of vocational training programs in other countries. Next, we discuss the key evaluation questions and our methodological approach to address them, data sources and outcomes, and our analysis approach. We then discuss some of the key limitations and challenges that we expect the evaluation to face. Finally, we describe the administrative details related to implementing the evaluation.

2. Overview of the Industry-led Skills and Workforce Development project

The ISWD project aims to improve the alignment between the skills of Georgian TVET graduates and the skills demanded by the labor market. The Millennium Challenge Account-Georgia (MCA-Georgia) is managing the implementation of the project and has subcontracted the implementation to a consortium led by PEM GmbH. The project comprises the following four activities:

- **Activity 1, Program Improvement Competitive Grants (PICG),** will fund Georgian TVET providers on a competitive basis to establish new or improved training courses that reflect industry demand for skills. The 10 institutions selected to receive grants will establish approximately 29 new courses and improve 12 existing courses. These include courses in areas such as information technology, agriculture, pisciculture, maritime operations, mountain guiding, railways, and aviation. Most of these courses are at TVET levels 4 and 5, which are training courses for upper secondary school graduates (Appendix Table A.1 provides a complete list of the funded courses). This activity accounts for the bulk of the project funding—$12 million of the total $16 million—with private industry making an additional contribution of about $7 million to the new and improved courses.

- **Activity 2, Strengthening TVET Provider Practice (STPP),** will provide small grants on a competitive basis to develop innovative tools for formal and informal TVET providers, and will identify and promote the uptake of best practices across the TVET sector. The grants are available to TVET providers and other institutions, including educational

establishments, public or private companies, and professional and nongovernmental organizations. Seven grants totaling about $69,000 were awarded in the first round in April 2016; an additional 10 grants totaling about $172,000 were awarded in the second (and likely final) round in April 2017 (Appendix Table A.2 provides a complete list of the STPP grants).

- **Activity 3, Strengthening TVET Sector Policy,** will provide technical assistance to the MES related to TVET sector policy. To reflect the latest priorities of the MES, the efforts under this activity have been consolidated into three main areas: (1) promoting increased business engagement in TVET; (2) improving and promoting the quality and attractiveness of TVET; and (3) supporting the enhancement of learning and qualifications opportunities for adults. The first area includes support for public-private partnerships for TVET provision, the establishment of sector skills councils that identify sector-level skills gaps and develop plans to address them, and increased engagement between employers and TVET providers at the local level. The second area includes support for enhanced marketing of TVET, improved career education and guidance at schools and TVET providers, development of a quality assurance framework for TVET courses, and implementation of a unified national qualifications framework to ensure that TVET qualifications are recognized. Finally, the third area focuses on supporting the validation of non-formal and informal learning, which enables adults with work experience but no formal qualifications to obtain formal recognition of their expertise.

- **Activity 4, Annual TVET Conference,** will serve as a forum for dialogue and information sharing among TVET stakeholders, and the dissemination of best practices. The first conference took place in July 2016, and the second is scheduled for October 2017. The conference will be complemented by other public relations and outreach events to promote the projects' objectives and Georgian TVET more generally, such as awards ceremonies for project grants and a multimedia communications strategy to publicize the project.

In Figure 2.1, we provide a logic model of the ISWD project, which is a modified version of one originally developed by MCA-Georgia and MCC. (We modified the original logic model for the evaluation to more clearly highlight the key pathways through which the project activities are expected to influence the ultimate outcomes of interest and that the evaluation will examine.) The logic model indicates that, in the short term, Activity 1 is expected to lead to an increase in the availability of industry demand-driven TVET courses (the PICG-funded courses). These courses—as well as Georgian TVET courses more generally—are expected to benefit further from improved quality and closer alignment with industry needs through the adoption of best practices disseminated by Activity 2, as well as the implementation of policy changes supported by Activity 3.[1] By encouraging interaction among stakeholders, Activity 4 is expected to support the other activities—for example, by facilitating dissemination of best practices (Activity 2) and greater industry engagement in TVET (Activity 3).

In the medium term, the combination of project activities is expected to increase the availability of graduates with higher-level skills in areas of industry demand. This is expected to

---

[1] The logic model does not explicitly include the enhancement of learning and qualifications opportunities for adults, one of the components of Activity 3, because it is not closely related to the other project activities. Therefore, we do not expect to focus on it as part of the evaluation.

result in greater industry satisfaction with local TVET programs, which will lead to greater industry co-investment in the sector. In turn, this increase in investment should feed back into an even larger increase in the availability of graduates with industry-demanded skills.

Finally, in the long term, the logic model suggests that industry will engage fully in the TVET sector. The close alignment of graduates' skills with market needs will lead to increased incomes through higher employment rates, which reflects higher demand for their skills, and higher wages for those who are employed, which reflects their higher productivity. Ultimately, these outcomes are expected to contribute to increased economic growth and reduced poverty in Georgia (the Georgia Compact's overarching goal, not shown in the logic model).

Figure 2.1 The ISWD logic model
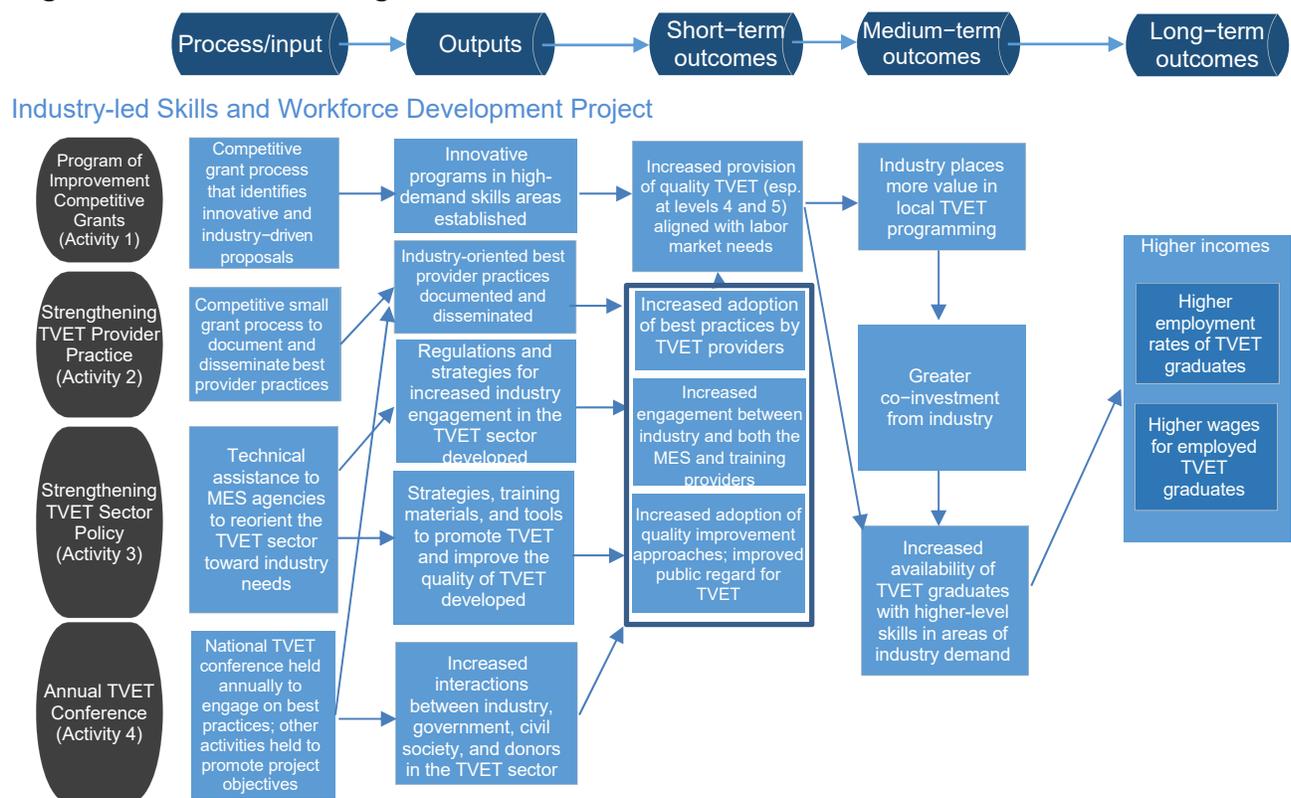
Note: MES = Ministry of Education and Science; TVET = Technical and Vocational Education and Training.

3.   Literature review

In this section, we review the literature on the labor market impacts of vocational training programs in low- and middle-income countries. The program logic for the ISWD project anticipates that the project will eventually result in improved labor market outcomes for Georgian TVET graduates—specifically, increased employment and higher wages. Outside of Georgia, a number of high-quality impact studies have examined the relationship between vocational training programs and these outcomes. These prior studies provide useful context and could help to indicate whether the ISWD program logic represents a plausible hypothesis about how the project activities could affect the ultimate outcomes of interest.

McKenzie (2017) recently reviewed 12 such impact studies that used an experimental design, which provides the highest standard of evidence (Table 3.1).[2] Only three of the nine studies that measured employment as an outcome found a statistically significant impact, and the mean impact was only 2.3 percentage points.[3] However, there is some evidence of larger impacts on formal employment, with a mean impact across studies of 3.6 percentage points—suggesting that training might shift workers from the informal to the formal sector.[4] Only two of nine studies that examined earnings as an outcome found a statistically significant impact, although most estimates were positive, with a mean of a 17 percent increase and median of an 11 percent increase. McKenzie concluded that the impacts of vocational training on employment and earnings are modest in most studies, although they are positive in some cases. He also suggested that few of these programs are likely to pass a simple cost-benefit test given the high costs of training and uncertainty about the sustainability of labor market impacts over time.

The findings from Mathematica's recent impact evaluation of MCC-funded scholarships for vocational training in Namibia (Borkum et al. 2017),[5] which also used an experimental design, are consistent with the modest impacts described above. The trainee scholarships were provided by issuing competitive grants to training providers and aimed to fund training in high-priority skill areas. Although the evaluation found that receiving a scholarship offer had large impacts on the probability of enrolling in and completing vocational training, especially among women, there was no evidence of positive impacts on employment and wages. A complementary qualitative study suggested that the process TVET providers used to assess market demand was not fully developed when the grants were made, which could partially explain the project's limited labor market impacts.

Additional evidence on the implementation and effects of vocational training programs in developing countries can be drawn from performance evaluations of specific programs. These evaluations often use mixed qualitative and quantitative methods and—in contrast to impact evaluations—are characterized by the lack of a rigorously defined comparison group. A review of the literature on youth workforce development over the past decade (U.S. Agency for International Development 2013) identified approximately 15 performance evaluations of vocational training programs in developing countries. The findings on the success of these programs in terms of employment and earnings, workforce readiness, and skills development

---

[2] The literature also includes several relevant quasi-experimental impact evaluations. However, a review by Tripney et al. (2013) found that the quality of these studies is highly variable, making it difficult to interpret the similarly variable findings on labor market impacts. In addition, other studies have found that evaluations of the same training program using different quasi-experimental methodologies can yield very different results (Ibarrarán and Rosas Shady, 2009; Delajara et al., 2006). Therefore, we focus our review on the higher-quality experimental studies summarized in Table 3.1, which have largely superseded these older quasi-experimental studies.

[3] McKenzie's review focused on the impacts of the offer of training; the impacts on those who actually received training was 20 to 40 percent larger, depending on the take-up rate of the offer of training in a particular study.

[4] The definition of formal sector employment varies across studies. It is typically based on employment in a job that includes legally mandated benefits in a given country context, such as health insurance, injury compensation, or social security contributions.

[5] The final evaluation report is currently under review at MCC; we expect it to be publicly available in fall 2017.

Table 3.1. Experimental evaluations of vocational training programs in low- and middle-income countries

Country	Study	Population	Follow-up period relative to end of training	Impact of the offer of training				Cost per trainee (USD)
				Employment (percentage points)	Formal employment (percentage points)	Earnings (percent)	Formal earnings (percent)	
Argentina	Alzúa et al. (2016)	Low-income youth	18 months	n.r.	**8.0**	n.r.	**64.9**	$1,722
		Low-income youth	33 months	n.r.	4.3	n.r.	23.1	
Colombia	Attanasio et al. (2011)	Low-income youth	14 months	**4.5**	**6.4**	**11.6**	**27.1**	$750
	Attanasio et al. (2015)	Low-income youth	up to 10 years	n.r.	**4.2**	n.r.	**13.6**	
Dominican Republic	Card et al. (2011)	Low-income youth	12 months	0.7	2.2	10.8	n.r.	$330
	Ibarrarán et al. (2014)	Low-income youth	18 to 24 months	-1.3	1.8	6.5	n.r.	$700
	Ibarrarán et al. (2015)	Low-income youth	6 years	-1.4	2.6	-1.9	n.r.	$700
	Acevedo et al. (2017)	Low-income youth	3 years	0.7	n.r.	n.r.	n.r.	n.r.
India	Maitra and Mani (2017)	Low-income women	18 months	**8.1**	n.r.	**95.7**	n.r.	$13
Kenya	Honorati (2015)	Low-income youth	14 months	**5.6**	n.r	29.7	n.r.	$1,150
Malawi	Cho et al. (2013)	Low-income youth	4 months	n.r.	n.r.	-19.6	n.r.	n.r.
Peru	Diaz and Rosas (2016)	Low-income youth	36 months	1.6	**3.8**	13.4	n.r.	$420
		Low-income youth	36 months	n.r.	4.5	n.r.	n.r.	
Turkey	Hirshleifer et al. (2016)	Unemployed	1 year	2.0	2.0	5.8	8.6	$1,700
		Unemployed	2.5 years	n.r.	-0.1	n.r.	-0.8	

Source: McKenzie (2017).

Notes: Impacts that are statistically significant at the 5 percent level are in bold.

n.r. = not reported.

were generally mixed, and depended on the features and context of the particular program. (For examples of specific performance evaluations, see Asian Development Bank [2013] and Kelly et al. [1998]).

Overall, the literature suggests that, although the effects of many vocational training programs in developing countries are modest, these programs can be successful in specific contexts. The success of any given program likely depends on factors such as social, economic, and labor market conditions; existing skill levels of targeted groups; and training program characteristics. To the best of our knowledge, no large-scale, rigorous evaluations of vocational training programs in Georgia or other countries in the Caucasus region have taken place, so the likely range of effects in the Georgian context are unclear.

In addition, there are some important differences between many of the training programs studied in the literature and the PICG-supported courses. First, the PICG-supported courses are substantially longer in duration and involve training at higher technical levels than many of these programs. The PICG-supported courses are expected typically to be between 9 and 24 months in duration (a median of about 18 months) and are offered at levels 4 and 5, which are only available to secondary school graduates. In contrast, many of the other training programs in the literature are only a few months in duration and target a less educated group of trainees. Second, PICG-supported courses have a much stronger market-related focus and a higher degree of industry involvement than most of the other training programs in the literature. Employers have been closely involved in developing the PICG-supported courses—for example, through co-funding arrangements (including monetary and in-kind contributions) and by advising providers on curriculum development—which could lead to stronger market alignment of these courses relative to other programs. Third, the ISWD project was explicitly designed to integrate market-relevant training with complementary activities and broader sector reforms aimed at improving the quality and market relevance of vocational training, which was not the case for most other training programs. The expected changes in the TVET sector resulting from these complementary activities—for example, adoption of best practices by TVET providers, increased employer and engagement, and improved public perceptions of TVET—could support broader improvements in outcomes for graduates of Georgian TVET programs. It is possible that these features of the ISWD project could make it more successful than typical vocational training programs.

4. Evaluation design

In this section, we describe our design for the evaluation of the ISWD project, including the evaluation's key research questions, analytical methods, study sample, and the time frame for the data that we plan to collect.

4.1. Evaluation type

Evaluation studies generally fall into one of two categories: performance evaluations, which measure key outcomes and assess the contribution of a program to these outcomes without using a counterfactual, and impact evaluations, which use a counterfactual to rigorously estimate a program's causal effects on key outcomes. To evaluate the ISWD project, we propose a mixed-methods performance evaluation, which will include two studies: (1) a quantitative outcomes

study of Activity 1;[6] and (2) a qualitative study assessing all project activities (Activities 1–4). To evaluate the possible effects of Activity 1, the outcomes study will measure the training and labor market outcomes of trainees in PICG-supported courses and, to the extent possible, compare those outcomes with those of a relevant sample of trainees who attended non-supported courses. The qualitative study will explore implementation of all the project activities, the potential mechanisms driving the results observed in the Activity 1 outcomes study, and the likelihood of sustainability across all ISWD initiatives after the compact ends.

4.2. Evaluation questions

Table 4.1 presents the key evaluation questions and the approaches we will use to answer them.

Table 4.1. Evaluation questions and approaches to answering them

Evaluation questions	Approaches to answering them
Activity 1 – Program Improvement Competitive Grants	
1. How did the implemented PICG courses compare with the original grant proposals, and what were the reasons for any deviations?	*Outcomes study* • Analyze trainee survey data to document course content, perceived quality, and training approaches (teaching materials, practices, and delivery methods; use of laboratories; availability of internships; receipt of career guidance; etc.) *Qualitative study* • Analyze documents from grantees to describe course offerings and practices, and compare with original grant proposals • Interview teachers of PICG-supported courses to understand changes in curriculum, training, and teaching methods associated with PICG-supported courses • Conduct trainee focus groups to explore perceived strengths and weaknesses of course content, training quality, and training approaches • Interview grantees, PEM, MCA-Georgia, the National Centre for Education Quality Enhancement (NCEQE), and employers and other partners who co-invested or supported course development to understand the course development and authorization process, and reasons for deviations from original plans
2. Did trainees enroll in PICG-supported courses and graduate from them at targeted levels? a. To what extent did women or members of socially disadvantaged groups (defined by language, region of origin, or other socio-demographic characteristics) enroll and graduate? b. Did these patterns differ across training sectors and grantees?	*Outcomes study* • Analyze trainee survey data to measure enrollment and graduation rates, overall and by subgroup *Qualitative study* • Interview grantees and trainee focus groups to better understand enrollment and graduation patterns

---

[6] As we discuss in Section 4.3, a complementary impact evaluation of Activity 1 *might* become feasible later in the evaluation, under certain conditions. However, at this stage we only plan to conduct a descriptive performance evaluation.

Evaluation questions	Approaches to answering them
3. What were the labor market outcomes (employment and wages) for graduates from PICG-supported courses? a. How did the outcomes of these graduates compare to those of graduates from other, non-supported courses? b. Did these results differ by gender or other socio-demographic measures? c. Did the results differ across training sectors and grantees?	*Outcomes study* • Analyze trainee survey data to measure labor market outcomes among trainees in PICG-supported courses, overall and by subgroup • Compare labor market outcomes for trainees in PICG-supported courses to those of trainees in a sample of non-supported courses (benchmarking approach) • Compare labor market outcomes for trainees in improved PICG-supported courses to those of trainees in the same courses before they were improved (pre-post design) *Qualitative study* • Interview employers to understand the job-search, hiring, and wage-setting processes • Interview graduates of PICG-supported certificate courses to understand effects on their labor market experiences
4. What were employer perceptions of the graduates from the PICG-supported courses, and how did the availability of these graduates affect their hiring and training plans? a. Do employer perceptions of graduates from PICG-supported courses differ according to gender or socio-demographic categories?	*Qualitative study* • Interview employers to understand the reasons underlying the perceptions of graduates from PICG-supported courses, as well as changes in hiring and training processes
5. Will PICG-supported courses be sustained after the compact? a. What are the main challenges to sustaining these courses, and how can they be overcome? b. How has the level of engagement between employers and grantees changed after the compact?	*Qualitative study* • Interview grantees to document plans to sustain new courses and assess the main barriers to doing so • Complement with an analysis of administrative data from grantees to assess enrollment levels for cohorts who enrolled in PICG-supported courses after the compact • Analyze grantee financial data on the revenues and costs of PICG courses to assess financial viability • Interview grantees, employers, and industry groups to investigate engagement between employers and providers and how this has changed over time
Activities 2–4 (Strengthening TVET Provider Practice; Strengthening TVET Sector Policy; Annual Conference)	
6. What are TVET providers' perceptions of the best practices identified and disseminated by the project, to what extent have they adopted them, and what are the main barriers to doing so? a. How were best practices identified and disseminated in practice? b. How has the adoption of best practices affected TVET providers, including the grantees and other providers? c. Is the adoption of best practices sustainable, and is the extent of adoption likely to increase in the future?	*Qualitative study* • Analyze documents from STPP grantees to describe the best practices identified and the dissemination process • Interview STPP grantees, PEM, and MCA-Georgia to understand the identification and dissemination process for best practices • Interview STPP grantees and other providers to explore perceptions and adoption of practices, their effects, and their perceived sustainability

Evaluation questions	Approaches to answering them
7. To what extent have the MES and its agencies adopted the policy reforms supported by the project, (for example, those related to industry engagement, marketing of TVET, and quality improvement) and what have been the main challenges in doing so? a. How has the adoption of these reforms affected or expected to affect the TVET sector, and in what time frame? b. Are the policy reforms supported by the project sustainable, and how are these policies expected to evolve? c. Is there any evidence of a broader shift toward higher-level, industry-driven courses in the Georgian TVET sector? If so, what was the role of the project, and if not, why not?	*Qualitative study* • Interview the MES to explore the adoption of reforms and remaining challenges to doing so • Interview employers, industry groups, and providers to understand the effects of the reforms on key stakeholders • Interview the MES and the Georgian Association of Private Colleges, as well as review administrative data on TVET course offerings, to understand broader changes in these offerings and the underlying reasons for these changes
8. How and to what extent has the annual TVET conference influenced providers, employers, the MES, and other TVET sector stakeholders? a. Who attended and financed the conference, and what were its main areas of focus? b. Is the conference likely to be sustained in the future?	*Qualitative study* • Conduct in-depth interviews with providers, employers, the MES, and other stakeholders to explore participation in and the effects of the conference and its perceived sustainability • Analyze administrative data on TVET conference attendance and financing • Directly observe conference activities and record descriptive information about conference implementation and engagement between participants

4.3. Methods

This section describes the outcomes study and qualitative study in further detail. We also briefly discuss the possibility of an additional study component—an impact evaluation of Activity 1 using a randomized controlled trial—which might be feasible later in the evaluation under certain conditions.

a.   Outcomes study

The Activity 1 outcomes study will describe trainee outcomes that are linked to evaluation questions 1–3. This includes an analysis of trainees' training experiences, graduation rates, and key post-graduation labor market outcomes—in particular, employment rates and wages.[7] We will measure trainees' outcomes using data from a trainee tracer survey that follows TVET enrollees after graduation and into the labor market.

We also plan to compare the labor market outcomes of trainees in the PICG-supported courses to those of trainees in a set of purposefully selected non-supported courses (related to evaluation question 3). Because it was not feasible to use a rigorous approach to identify these

---

[7] We will also examine whether graduates of PICG-supported courses enroll in further training and whether they are "productively engaged" (defined as being employed or enrolled in training). However, the program logic does not posit substantial enrollment in further training soon after graduating from the PICG-supported courses, which are already at an advanced level. Therefore, we focus on employment and wages as our primary labor market outcomes.

non-supported courses,[8] we recognize that these comparisons are unlikely to yield causal estimates of the impacts of Activity 1. Nevertheless, by carefully selecting the non-supported courses and assessing baseline differences between the types of trainees enrolling in PICG-supported courses and non-supported courses, these comparisons will enable us to place the estimated trainee-level outcomes into the existing vocational training context in Georgia. These descriptive results will also help to assess whether the linkages between project activities and trainee outcomes assumed in the project logic and economic rate of return (ERR) model are likely to have occurred in practice.

We will use two complementary approaches to select the set of non-supported courses for these comparisons:

- **A benchmarking approach**, which will select non-supported courses that trainees in PICG-supported courses considered or applied to when they applied to the PICG-supported courses. To identify which courses trainees considered, we will gather information from trainees in PICG-supported courses as part of the baseline tracer survey that we plan to conduct soon after they enroll. For each PICG-supported course, the most commonly identified non-supported courses that the trainees considered or applied to will serve as the benchmark for comparison. In this way, we expect to identify non-supported courses that are most relevant to the types of trainees who enrolled in the PICG-supported courses. If this approach proves not to be feasible (for example, if the survey responses about relevant courses are very diffuse), we will instead select a broad cross section of non-supported public and private Georgian TVET courses using course databases from the MES and the Georgian Association of Private Colleges.[9]

Regardless of the approach we use to identify the benchmarking courses, the benchmarking results will be purely descriptive and cannot be interpreted as the impacts of the PICG grants. In particular, the types of trainees who enroll in PICG-supported courses might differ from the types of trainees who enroll in the benchmarking courses, and these differences could drive much of the differences in their outcomes. Our analysis will attempt to address this issue using statistical matching techniques, such as covariate matching or propensity score matching, which are designed to balance the characteristics of trainees in the PICG-supported and benchmarking course samples. If it is possible to identify a sample of trainees in benchmarking courses with very similar baseline characteristics to trainees in PICG-supported courses (as measured by the trainee characteristics collected in the baseline survey), we will be able to rule out that differences in these characteristics are driving the

---

[8] In particular, because the PICG-supported courses are unique in the Georgian context, few comparable non-supported courses are available to conduct course-level matching and implement an impact evaluation using a matched comparison group design. The selection mechanism for the PICG grants also precluded an impact evaluation using a course-level random assignment or regression discontinuity design. An impact evaluation using an individual-level random assignment design *might* be feasible at a later stage of the project, as we describe later in this section; however, this would complement rather than replace the approaches described here.

[9] As we note in Section 4.5, we will have to delay the timing of the baseline tracer survey for trainees in benchmarking courses to provide sufficient time to identify these courses (based on the baseline survey responses of trainees in PICG-supported courses). However, we expect the benchmarking courses to be several months in duration (like the PICG-supported courses), so we expect to be able to conduct the baseline survey for trainees in the equivalent academic cohort in these courses while they are still enrolled.

differences in outcomes. However, given that the set of baseline characteristics will be limited, we will still not be able to rule out that other, unobserved differences in trainees characteristics (for example, differences in trainee motivation) are contributing to the differences in outcomes.

- **A pre-post design**, which focuses on the 15 PICG-supported courses that were improved (rather than introduced as new courses), and compares the outcomes of trainees in these courses with those of earlier cohorts in the same courses before they were improved. In particular, we will compare the outcomes of the final cohort of trainees in each pre-improvement course, captured one year after they graduate, to the outcomes of the first cohort of trainees in the improved course, captured one year after *they* graduate. These pre-post estimates cannot be interpreted as the impacts of the grant-funded improvements, because several confounding factors could be driving pre-post changes in outcomes in addition to these improvements. For example, the typical profile of trainees might change after the improvements, and there could be different labor market conditions in the years in which trainees in pre-improvement and improved courses enter the labor market. Because the differences in the types of trainees in the pre-improvement and improved courses might be substantial (for example, due to changes in admissions requirements) and the sample sizes in many of the pre-improvement courses are relatively small, we will likely not be able to use statistical matching techniques to account for differences in trainee profiles. Instead, we will attempt to control for the influence of observed characteristics on labor market outcomes in a simple pre-post regression model.

It will be valuable to conduct both of these descriptive analyses, because each analysis has different strengths and weaknesses. An important strength of the benchmarking analysis is that it will enable us to benchmark the outcome estimates for all of the PICG-supported courses, whereas the pre-post design will only focus on the subset of these courses that existed before the project. In addition, statistical matching approaches (if feasible) could help ensure that differences in trainee profiles are not driving the benchmarking analysis estimates, whereas in the pre-post design we would only be able to partly control for these differences using a simple regression model (matching is not likely to be possible in the pre-post design because sample of comparison trainees will be small, and because the pre-improvement courses are likely to enroll trainees with different demographic profiles and academic credentials than the improved versions of those courses). The benchmarking analysis will also compare trainees who enter the labor market contemporaneously and hence experience similar labor market conditions, which is not the case for the pre-post design. In contrast, an important strength of the pre-post design is that it compares the outcomes of trainees enrolled in the same providers, who have selected the same subject areas for their vocational education. Therefore, the pre-post design accounts for potentially important unobserved differences in providers and trainees that might partly drive the benchmarking analysis estimates even after accounting for differences in observed characteristics. Overall, if the descriptive results from both analyses are broadly consistent, this would provide strong suggestive evidence that a positive program impact is likely to have occurred (or not to have occurred). If the results are inconsistent, we would seek to conduct additional analyses to understand why (for example, restricting the benchmarking analysis sample to the improved courses to investigate whether differences in providers are driving the differences in results).

b.   Qualitative study

The qualitative study will draw primarily on interviews and focus groups with key stakeholders, complemented by contextual information from grantee documents, administrative data, and grantee financial records. As we describe in Section 6.3, we will systematically organize and synthesize the key themes that emerge from each data source and triangulate the findings across these sources to answer the evaluation questions. The qualitative study will focus on the following five areas:

- **Implementation of PICG-supported courses and their sustainability after the compact** (relevant to evaluation questions 1, 2 and 5). We will seek to understand how the PICG grantees developed courses, how they calibrated curricula and instruction to industry needs, how they managed the authorization process for these courses, and how and why implementation differed from the original proposals. To do this, we will analyze data from interviews with all PICG grantees and entities involved in implementation (MCA-Georgia, MCC, and PEM), teachers of PICG-supported courses, employers and other partners who co-invested in these courses or supported course development, and the National Centre for Education Quality Enhancement (NCEQE, the MES entity responsible for course authorization), together with information from grantee documents. The analysis of interview data will also explore the perceived sustainability of the PICG-supported courses, which we will also investigate using grantee administrative data on trainee enrollment and financial records.

- **Trainees' and employers' perceptions about the potential benefits of PICG-supported courses** (relevant to evaluation questions 3 and 4). To capture trainees' perspectives about their training and labor market experiences, we will analyze transcripts from focus group discussions with trainees enrolled in PICG-supported courses, and draw on descriptive information from the trainee tracer survey conducted for the outcomes study. To capture employers' perspectives regarding the quality of the PICG-supported courses and implications for their hiring and training practices, we will analyze data from both qualitative interviews and focus groups with employers in sectors relevant to the PICG-supported courses and with the main industry groups in Georgia. This qualitative data collection from employers will include semi-structured interviews with employers who have hired PICG course graduates and focus groups with a broader cross-section of employers in relevant sectors who did not hire PICG course graduates. The analysis of qualitative data from trainees and employers will also examine differences by trainee gender and other socio-demographic characteristics. For example, we will examine the extent to which curricula, pedagogical approaches, and employer engagement (such as through provision of internships or hiring of PICG graduates) were inclusive with respect to gender, language group, and disability status. Finally, although our evaluation focuses on the fully accredited versions of these PICG courses—which underlie the project logic and ERR estimates—we will also conduct interviews with graduates of shorter (unaccredited) certificate courses established by the PICG grantees. These interviews will enable us to explore the motivation of trainees for enrolling in these courses, their perspectives about their training, and how their certification has affected them on the labor market.

- **Implementation of best-practice grants and adoption of best practices** (relevant to evaluation question 6). To document how best-practice grants were implemented, we will

review administrative data on the scope of each STPP grant and reports from grantees on funded activities. We will also analyze data from interviews with the entities involved in implementation, focusing on the grant selection process and the successes and challenges of implementation. To evaluate the success of the grants in developing and disseminating best practices, we will analyze data from interviews with STPP grantees to document how practices were shared, and analyze qualitative data from a broader sample of TVET providers to assess whether these practices were adopted more widely among non-grantees.

- **Implementation and potential effects of national changes in TVET policy** (relevant to evaluation question 7). We will analyze qualitative interview data from MES staff and the entities involved in implementation to understand progress made in each of the policy areas supported by Activity 3, and to assess the likely sustainability of new policy initiatives after the compact ends. Interviews with other major donors in the TVET sector will provide additional information on changes in these policy areas, as well as broader changes in the sector. We will also analyze interview data from employers, industry groups, and TVET providers to assess the level of industry engagement and coordination with TVET providers and how this changed over time. Analyzing interview data from TVET providers will also enable us to further explore the implementation and effects of efforts related to TVET marketing and quality improvement (for example, marketing strategies, career education and guidance, and quality assessments). Finally, we will analyze data from interviews with the MES and Georgian Association of Private Colleges, together with administrative data on course offerings, to assess whether there is a broader shift to industry-demand driven TVET courses and the role of the project in facilitating this shift (if any).

- **Implementation and potential effects of the compact's annual TVET conferences** (relevant to evaluation question 8). To assess the effectiveness of compact-sponsored TVET conferences, we will review conference documentation recording levels of participation among TVET providers, employers, and MES policymakers, as well as the sources of financial support for the conference (in particular, the contributions of industry). We will also analyze qualitative interview data gathered from each of these types of conference attendees to explore the perceptions of stakeholders regarding the effects of conference attendance. Finally, members of the evaluation team will directly observe conference activities and record descriptive information about conference implementation and engagement between participants.

c.   Possible additional study: trainee-level randomized controlled trial

As the evaluation progresses, we will also assess the feasibility of conducting an impact evaluation using an individual-level randomized controlled trial, which would provide more rigorous evidence to answer evaluation question 3. Under this approach, providers would first screen applicants for each course using whatever criteria they see fit. In courses in which the number of eligible applicants exceeded the number of available enrollment slots, we would then work with providers to randomly select the trainees who are offered admission to the course. Comparing the outcomes of the trainees who were randomly selected for admission (the treatment group) to those who were not (the control group) would provide rigorous evidence of the impacts of the offer of PICG-supported training on trainees.

For this approach to be feasible, several conditions must hold. First, demand for PICG-supported courses (after screening for eligibility) would have to exceed the number of available slots in a sufficient number of courses and by a sufficient amount to provide the necessary statistical power. Second, we would have to secure buy-in from key stakeholders, especially grantees, for a random assignment mechanism to select those offered admission from among eligible applicants. Third, we would need to develop an appropriate data collection strategy for locating and surveying all of the program applicants, including those who declined an admission offer or did not receive an admission offer. We will revisit these conditions after we have more information about the demand for the PICG-supported courses. In particular, if administrative and qualitative data from grantees suggest there is sufficient excess demand for the initial cohorts in these courses, we will explore with MCC and other stakeholders the possibility of conducting this impact evaluation for subsequent cohorts. These discussions would include an assessment of whether such a design is feasible within the time frame and budget of our current evaluation contract with MCC, given that any potential randomized controlled trial would involve analyzing data from a later set of trainee cohorts.

4.4. Study sample and power calculations

The evaluation design for the ISWD project requires us to collect primary data from two different samples: (1) a sample of trainees in PICG-supported courses, benchmarking courses, and pre-improvement courses (for the trainee outcomes analysis); and (2) a sample of key stakeholders for qualitative data collection (for the qualitative study). Below we describe each of these samples in further detail.

a.   Trainee sample

 The sample for the trainee outcomes analysis will include trainees in all PICG-supported courses, as well as trainees in the benchmarking and pre-improvement courses that we identify. We will select each of these trainee samples as follows:

- **PICG-supported courses**. We had initially planned to survey the universe of trainees in the first cohort in each PICG-supported course that experienced the fully established version of the course. However, discussions with stakeholders suggested that, because these courses are so new, most of them are likely to evolve substantially after the first cohort. For example, there could be substantial changes in terms of recruitment efforts, types of trainees enrolled, course content and delivery, and partnerships with employers for internships or job placements. Therefore, in many cases, the training and labor market experiences of the first cohort might not reflect those of subsequent cohorts enrolled in a more developed version of these courses. (The courses would continue to evolve over time, but the changes after the first cohort would likely be the most substantial.)

    To address this concern, we will follow several steps to select the sample of trainees in each PICG-supported course. First, we will collect baseline information for all trainees in the first cohort enrolled in the fully established version of the course, as initially planned. Collecting this information is relatively low-cost and will provide us with maximum flexibility in selecting the sample. Second, before trainees in the first cohort are due for follow-up, we will discuss with the relevant grantee and entities involved in implementation how and to what extent the course is likely to differ for the second cohort. Third, based on this

information, we will determine on a case-by-case basis whether to conduct the follow-up with trainees in the first cohort and/or whether to collect baseline information from trainees in the second cohort so that we can conduct a follow up with them.

- **Benchmarking courses.** For each cohort of trainees that we include in the PICG-supported course sample, we will include in the sample trainees in the equivalent cohort in the relevant benchmarking courses. Using equivalent cohorts will ensure that trainees in PICG-supported and benchmarking courses experience similar labor market conditions. In addition, it is possible that the set of relevant benchmarking courses may differ for different cohorts; this could occur if changes in recruitment in PICG-supported courses change the types of trainees who choose to enroll. For example, if a course's second cohort began to enroll students who were more inclined to consider higher-level TVET courses, we would consider adding a higher-level benchmark-course to the study sample in response.

- **Pre-improvement courses.** The sample for pre-improvement courses consists of the last cohort of trainees in each course before it was improved. In contrast to the other trainee samples, this comparison sample is fixed: regardless of which PICG cohort(s) are ultimately included in the study, this pre-improvement sample will remain the same.

To maximize the statistical precision of our estimates, we would ideally include the universe of trainees in the selected cohort(s) in PICG-supported, benchmarking, and pre-improvement courses. This is especially important for subgroup analyses by provider or by trainee characteristics. For example, it would maximize the precision of the estimated employment and wage rates for graduates from specific PICG-supported providers or courses, which may be of interest to MCC and other stakeholders (although these estimates might still be relatively imprecise for providers or courses with few enrolled trainees).

At this stage, however, our expected sample size of trainees is uncertain because enrollment levels are still uncertain. For the purposes of the statistical power calculations below, we assume a total sample of about 1,550 trainees in PICG supported courses. This corresponds to the total number of trainees expected to enroll in the first cohort in these courses, based on information provided by PEM in early August 2017. The final available sample will depend on actual enrollment numbers and our decision about the number of cohorts to include. If enrollment levels are high and we include more than one cohort in the sample, it might be necessary to draw a random sample of trainees in some courses rather than surveying all of them (in that case we would draw a sample in courses with the highest enrollment levels). For now we assume a similar total sample size of 1,550 trainees in benchmarking courses, although we will only know the number of trainees in these courses once we identify them. In contrast to the uncertainty for PICG-supported and benchmarking courses, we know that there were 400 trainees in the relevant cohort in pre-improvement courses, and we plan to survey all of them. Combined, we currently assume a total sample of about 3,500 trainees across all types of courses.

Table 4.2 presents our calculations of minimum detectable differences (MDDs) for the trainee outcomes analysis using the sample described above, focusing on the key outcome of employment. The power calculations assume an 80 percent survey response rate and a comparison employment rate of 76 percent, following the assumption in MCC's ex-ante ERR

model.[10] We present the MDD calculations separately for the descriptive benchmarking approach and pre-post design.[11]

Table 4.2. MDD calculations for trainee outcomes analysis

	Benchmarking approach			Pre-post design		
	Full sample	50 percent subgroup	25 percent subgroup	Full sample	50 percent subgroup	25 percent subgroup
Number of PICG courses	43	43	43	15	15	15
Number of comparison courses	43	43	43	15	15	15
PICG trainee sample size	1,550	775	388	960	480	240
Comparison trainee sample size	1,550	775	388	400	200	100
MDD for employment (percentage points)	4.6	6.4	9.1	8.3	11.7	16.6

Note: MDD calculations assume a two-tailed test with a 5 percent significance level and 80 percent power. Expected sample sizes for PICG-supported and pre-improvement comparison courses are based on communications with the PICG grantees; we assume that the sample size for the benchmarking comparison courses will be the same as for PICG-supported courses. We also assume that the response rate to the follow-up trainee tracer survey will be 80 percent and that mean employment in the comparison group will be 76 percent, based on the assumption in the ex-ante ERR model. Finally, we assume an R^2 value of 0.1 due to statistical matching in the benchmarking analysis and the inclusion of covariates in the pre-post analysis model.

We estimate that, for the full sample, we will be able to detect an MDD of 4.6 percentage points for employment using the benchmarking approach, and 8.3 percentage points using the pre-post design. The MDDs are larger for subgroups, but for the benchmarking approach we will still be able to detect a difference of 9.1 percentage points even with a relatively small subgroup comprising one-quarter of the full sample. However, the equivalent subgroup MDDs for the pre-post design are substantially larger, suggesting that it will be more difficult to detect changes for small subgroups using this design.

The MDDs for employment in Table 4.2 are higher than many estimates of the impacts of vocational training programs reported in the literature. However, as we discussed in Section 3, the estimates in the literature are highly context-specific, and it is possible that, given close industry involvement in PICG course development, the PICG courses could align more with

---

[10] We do not present MDDs for the percentage change in wages, the other key labor market outcome, because we do not have information about the standard deviation of wages in the study's comparison groups (one of the primary inputs needed to carry out the calculations).

[11] The MDD calculations do not account for dependence (clustering) between the outcomes of trainees in the same course. The reason is that we view the estimated differences in outcomes as being between fixed populations of trainees, rather than a sample at a given point in time. For the benchmarking approach these populations are the first cohort of trainees in the PICG-supported courses and the equivalent cohort in benchmarking courses; for the pre-post design they are the first cohort of trainees in improved PICG-supported courses and the last cohort in the pre-improvement courses. Therefore, as long as we do not attempt to generalize the results to a broader population of courses or cohorts, we do not need to account for clustering. These generalizations would not be appropriate given that there is a fixed set of PICG-supported courses and that the outcomes of future cohorts in these courses might differ as the courses evolve.

industry needs than other, more typical vocational training programs. Therefore, it is also possible the increases in employment and wages associated with PICG training could be higher in this context than the effects observed elsewhere. We also believe changes of this magnitude are relevant because the estimated MDDs for the full sample in both the benchmarking approach and pre-post design are within the change of 9 percentage points expected in MCC's ex-ante ERR model. For the benchmarking approach, analyses for even relatively small subgroups have an MDD close to this expected change. Although the evaluation's descriptive design might not be able to estimate causal impacts of the program, observing whether differences of this magnitude occur will be a useful indicator of whether projected impacts in the ERR are plausible.

We also examined the sensitivity of our MDD calculations to the assumed sample size in case the sample size is lower than expected because PICG-supported courses do not meet their enrollment targets. For example, if PICG-supported courses only enroll half the number of trainees that we assumed above—which we view as a worst case scenario—we calculated that the MDDs for employment would be 6.4 percentage points for the benchmarking approach and 9.9 percentage points for the pre-post design (not shown).[12] Therefore, even in this worst case scenario, the MDD for the benchmarking approach is within the expected change in the ex-ante ERR model and the MDD for the pre-post design is only slightly higher than the expected change (although it would become more challenging to detect differences for subgroups). In addition, as discussed above, it is possible that we will include trainees in more than one cohort for at least some PICG-supported courses, which would enable us to increase the sample size even if enrollment in the first cohort is below expectations. Therefore, lower than expected enrollment in PICG-supported courses is unlikely to be a major challenge for the proposed design.

b. Qualitative sample

Table 4.3 shows the planned respondents, sample sizes, and sampling approach for qualitative data collection. Our plan recognizes the high degree of diversity among the PICG grantees and the STPP grantees, both in provider characteristics and the specific activities funded by the grants. This diversity suggests that implementation experiences, project effects, and sustainability might differ substantially across providers. Therefore, to fully address the evaluation questions, the planned sample sizes of provider-specific respondents (in particular, providers, trainees, and employers) are relatively large and seek to maximize the coverage of grantees with the available resources. For example, in the final round of data collection, we propose interviewing four employers of graduates from each PICG grantee, a total of 40 employers. This will enable us to obtain some grantee-specific findings about the extent to which graduates' skills meet employers' needs, and to contrast these findings across PICG grantees. (We propose interviewing a smaller sample of 10 employers in the interim round because PICG trainees will not yet have entered the labor market and we will therefore draw the sample from the smaller pool of employers who co-invested in the PICG-supported courses.)

---

[12] The MDD for the benchmarking approach in this scenario is identical to that for the 50 percent subgroup shown in Table 4.2. The equivalent MDD for the pre-post design is lower than that for a 50 percent subgroup in Table 4.2 because the sample size in pre-improvement courses would still be fixed at 400 trainees.

Table 4.3. Respondents and sample sizes for qualitative data collection

Respondent	Interim round, 2018		Final round, 2021	
	Sample size	Sampling approach	Sample size	Sampling approach
Interviews and focus groups to be conducted by local data collection partner				
STPP grantees and related TVET providers	8 interviews	8 STPP grantees, selected as those whose practices have the best potential for wider adoption (based on discussions with PEM)[a]	18 interviews	8 STPP grantees (same as interim round); 10 non-grantees, identified as potential best-practice adopters by the selected STPP grantees
Trainees	12 trainee focus groups	One mixed-gender focus group per PICG grantee, and two female-only cross-grantee focus groups (each covering five grantees); each focus group will include 8–12 trainees, with participants selected to be diverse in terms of socio-demographic characteristics		
Graduates of PICG-supported certificate courses			16 interviews	Up to 4 graduates per certificate course
Teachers	10 individual or small-group interviews	One teacher or small group of teachers per PICG grantee, selected from those teaching PICG-supported courses		
Employers	10 interviews	One employer per PICG grantee, selected from those who co-invested in the PICG-supported course or supported course development	40 interviews[b]	Four employers per PICG grantee, selected from the most common employers of PICG gradates (according to the follow-up trainee tracer survey)
			2 focus groups[c]	Employers in sectors covered by PICG who did not hire PICG graduates, with 6–8 employers per focus group selected to be diverse in terms of sector and size
			2 focus groups[c]	Employers in sectors outside of those covered by PICG but that could potentially benefit from Activities 2–4, with 6–8 employers per focus group selected to be diverse in terms of sector and size[d]
Interviews to be conducted by Mathematica				
PICG grantees	10 interviews	All 10 PICG grantees	10 interviews	All 10 PICG grantees

Respondent	Interim round, 2018		Final round, 2021	
	Sample size	Sampling approach	Sample size	Sampling approach
International PICG partners	5 interviews	Up to one international partner per PICG grantee, selected from those who supported course development		
Industry groups	2 interviews	Georgia Chamber of Commerce; Georgia Employers' Association	2 interviews	Georgia Chamber of Commerce and Georgia Employers' Association
MES and NCEQE staff	2 interviews	Head of TVET department; NCEQE staff	2 interviews	Head of TVET department; NCEQE staff
Georgian Association of Private Colleges			1 interview	Staff familiar with courses offered by association members
PEM	4 interviews	Team leader and key program staff		
MCA-Georgia and local MCC staff	5 interviews	Key program staff		
Other donors in the TVET sector	2 interviews	European Union delegation; another major international donor	2 interviews	European Union delegation; another major international donor

Note: Blank cells indicate that we will not collect qualitative data from a particular respondent in the relevant round.

[a] We will review the final presentations compiled by all 17 STPP grantees to gain a full understanding of all the best practices that were developed, even for those grantees that we do not interview.

[b] These qualitative interviews will also gather basic descriptive quantitative information about the employers.

[c] If arranging focus groups with employers proves to be too logistically difficult, we will interview them instead.

[d] We will use the interim round of data collection to identify the non-PICG sectors and employers most likely to be affected by Activities 2–4, if any (for example, through interviews with MES staff and an analysis of attendance at the annual TVET conference).

There is substantial overlap in the respondents across the two rounds of qualitative data collection, which will enable us to explore the evolution and sustainability of the project activities after the compact ends. For example, repeating interviews with grantees will provide an opportunity to uncover important insights about whether supported programs and practices can be sustained after financial support ends (as was assumed in the project logic and ERR), and in what form. However, there are also some important differences in the respondents across the two rounds. First, the interim round will include entities involved in implementation—MCA-Georgia, local MCC staff, and PEM—to capture their responses toward the end of the compact, after which their involvement will end. Second, the final round will include a broader set of providers and employers than in the interim round to provide a more comprehensive assessment of the long-term effects and sustainability of the project activities.

4.5. Time frame

Table 4.4 summarizes the planned timing of data collection for the various primary data sources we will draw on for the evaluation.

Table 4.4. Timing of data collection for the ISWD project evaluation

Type of data	Respondents	Timing
Trainee tracer surveys	Trainees in PICG-supported courses and non-supported comparison courses (benchmarking or pre-improvement courses)	Baseline while enrolled in training (in mid-2017 for pre-improvement courses, and between late 2017 and early 2019 for other courses); follow-up one year after training (between late 2018 and early 2021)[a]
Qualitative data: key informant interviews and focus groups	TVET providers (PICG grantees, STPP grantees, and others); trainees from PICG-supported courses; employers; industry groups; MES staff; Georgian Association of Private Colleges; PEM staff; MCA-Georgia and MCC; other donors	Interim round in late 2018 (toward the end of the compact); final round in 2021 (post-compact)[b]

[a]Survey dates will vary by course, depending on the course start and end dates.

[b]Some respondents vary across the two rounds of qualitative data collection (Table 4.3).

For the trainee tracer survey, we plan to survey trainees in PICG-supported courses and comparison non-supported courses while they are still enrolled (at baseline) and again one year after they graduate (at follow-up). This one-year follow-up period is typical in the literature discussed in Section 3 (the typical period is 12 to 18 months, although some studies do have a longer follow-up), and was the same period used in our recent evaluation of MCC's vocational training investments in Namibia. The exact timing of the baseline and follow-up trainee tracer survey for a particular course will depend on when the course starts and the duration of training. Therefore, both the baseline and follow-up rounds of data collection will take place over several months.

According to the latest (but still preliminary) course timeline gathered from grantees by our local consultant, most of the PICG-supported courses are expected to start between the third quarter of 2017 and the second quarter of 2018 (Appendix Table A.1). Therefore, we expect the baseline round of the trainee tracer survey for the first cohort of trainees in PICG-supported courses to take place over this same period.[13] The baseline for the benchmarking courses is expected to take place over a similar period, but with some delay to enable us to identify those courses. Given this delay and the possible inclusion of a second cohort of trainees for some PICG-supported courses, the baseline data collection could potentially extend until early 2019. In May and June 2017, our local consultant collected baseline data from trainees in pre-improvement courses that will be part of the pre-post design because it was necessary to do so while trainees were still enrolled in those courses.

---

[13] Courses conducted by the Georgia Mountain Guide Association started earlier, in late 2016, well before the other PICG-supported courses. To ensure that we are able to include these courses in the baseline data collection to be procured by MCA-Georgia in summer 2017 and that we are measuring the outcomes of the fully established version of these courses, we plan to focus on the second cohort enrolled in these courses (which will enroll starting in late 2017). Therefore, the baseline survey timeline for these courses will be similar to the other PICG-supported courses.

According to the course timeline, the follow-up round will start in the third quarter of 2018, one year after the first graduates of the final cohorts in the pre-improvement courses enter the labor market. This round would continue until April 2021 as additional graduates from pre-improvement courses, PICG-supported courses, and benchmarking courses enter the labor market. This cutoff date, about nine months before the end of the evaluation contract in January 2022, would give us sufficient time to process and analyze the data for the final report. Under the current projected timeline for PICG courses, the evaluation's projected end-date would exclude one of the 43 PICG-supported courses from the follow-up survey round (the signalization, centralization and blocking mechanic course conducted by the Georgian Railways grantee). Extending the evaluation contract to the third quarter of 2022 would enable us to include this course. It will also likely be necessary to extend the evaluation contract if we include an additional cohort of trainees in some of the PICG-supported courses in our sample (as discussed above). We will discuss the optimal timing further with MCC and MCA-Georgia after we have finalized the course timelines with the grantees and have assessed the extent to which it is necessary to survey an additional cohort of trainees.

The timing for the study's qualitative data collection activities is less complex. We will conduct two rounds of qualitative data collection: an interim round in late 2018, toward the end of the compact, and a final round in 2021, toward the scheduled end of the evaluation contract. The interim round will focus on implementation issues and preliminary results at a point when implementation is still active or relatively recent and key implementation-related stakeholders (such as PEM and MCA-Georgia) are still available in country. The final round will explore the longer-term effects of the project activities and their sustainability in the post-compact period.

5. Data sources and outcomes

In this section, we describe each of the primary and secondary data sources that we will draw on for the evaluation in further detail, including the key outcomes that we will capture through each source.

5.1. Data from the trainee tracer survey

The baseline trainee tracer survey will be a short survey of trainees enrolled in PICG-supported, benchmarking, and pre-improvement courses. It will capture detailed trainee contact information, background characteristics, information about other training courses considered, and trainees' expectations for future employment and wages. This information will enable us to contact trainees for the follow-up survey, describe trainee background characteristics, identify relevant benchmarking courses, identify relevant employers for qualitative interviews and focus groups, and eventually compare trainees' initial expectations to their actual outcomes. The baseline survey will be self-completed by trainees in the classroom while they are still enrolled in training, and will largely be conducted by an MCA-procured local data collection firm.[14]

The follow-up survey will be a longer survey that captures detailed information about trainees' training experience, as well as labor market outcomes such as employment and wages.

---

[14] As noted earlier, our local consultant collected baseline data for trainees in pre-improvement courses in May and June 2017 because it was necessary to mobilize quickly to collect these data while the trainees were still enrolled. The MCA-procured data collector will collect baseline data for all PICG-supported and benchmarking courses.

A local data collection firm will conduct the survey one year after trainees graduate from each course; because the staggered timeline of follow-up data collection will extend beyond the compact period, MCA-Georgia will procure some of these surveys (those conducted during the compact period) and Mathematica will procure the post-compact data collection. To maximize the follow-up survey response rate, we will use a multi-pronged strategy to locate respondents and have them complete the survey, which we will develop in close coordination with MCA-Georgia and the local data collection firm. This strategy will likely include a combination of emails (with a link to a self-completed web-based survey), phone contacts (a web-based survey completed over the phone by the survey team), and additional contact information from the baseline survey (social media contact information and phone numbers of relatives or friends) to obtain updated phone contact information for nonrespondents. We will also consider visits to nonrespondents' permanent home addresses, the work sites of major employers of graduates, or both to conduct face-to-face surveys.

The MES also conducts follow-up tracer surveys of graduates from all public TVET courses, six months after they graduate. However, it will be important to conduct a separate tracer survey for our study for several reasons. First, we are interested in measuring course completion rates and the labor market outcomes of those who enrolled but did not graduate; this will require us to collect information from all enrollees rather than just graduates, whereas the MES tracer survey focuses on graduates only. Second, we are interested in measuring longer-term labor market outcomes measured one year rather than six months after graduation. Third, the response rate to the MES survey has been about 50 percent—we are seeking a higher response rate (closer to 80 percent) for our study, to limit possible non-response bias. This might be realistic given that we are surveying a smaller set of courses than MES and can devote additional resources to locating respondents using the multi-pronged strategy described above. Finally, we plan to use a more comprehensive survey instrument that will capture additional information relative to the MES survey—for example, detailed information on training experiences. Although conducting a separate tracer survey implies that some trainees will be interviewed by both MES and our local data collector, these interviews would occur six months apart and would therefore not be unduly burdensome for respondents (in addition, given the expected differences in response rates, a large fraction of trainees would only be surveyed by our local data collector).

We summarize the planned contents of the baseline and follow-up survey in more detail in Table 5.1 (the full baseline survey instrument is included as Appendix B; the follow-up instrument is still to be developed and the contents in Table 5.1 are therefore preliminary).

Table 5.1. Preliminary contents of the baseline and follow-up trainee tracer surveys

Domain	Survey contents
Baseline survey	
Contact information	Primary and secondary phone number; email address; Facebook contact information; name, phone number, and email of relatives or friends
Training information	Provider, name, and level of other training courses to which trainee applied; initial perceptions of current training
Expectations for the future	Expected main activity one year after graduation; expected wage; main employers of interest
Demographics and background information	Gender; age; marital status; number of children; disability status; home language; region of origin; education level; parental education level; score in vocational training entry exam; prior training completed and other concurrent training; work experience and current employment status
Follow-up survey	
Training receipt	**Completion of training**; highest level of training completed; perceptions of training quality; course content and pedagogical approaches (use of laboratories, practical component, teaching practices, etc.); receipt of and duration of internships; career guidance received; job placement assistance received; enrollment in further training after graduation
Employment	**Paid employment**; productive engagement (employed *or* engaged in further training); formal sector employment; self-employment; hours per week and days per month worked; time to find a job; job satisfaction; availability for work and job-seeking activities (for those not working)
Wages	**Monthly wages from employment (or profits from self-employment)**

Note: Key outcomes for the follow-up analysis are in bold.

5.2. Qualitative data

We will develop a qualitative data collection protocol for each type of respondent in each of the two rounds of qualitative data collection. Although we will tailor the protocols for each respondent type, they will all cover similar topics related to the research questions (Table 5.2 presents illustrative areas of focus for each type of respondent). We anticipate that a combination of Mathematica staff and an MCA-procured local data collection firm will collect the qualitative data in the interim round; for the final round, Mathematica will hire a local data collection firm directly.

Table 5.2. Illustrative areas of focus for qualitative data collection

	Illustrative areas of focus	
Respondent	**Interim round, 2018**	**Final round, 2021**
PICG grantees, STPP grantees, and related TVET providers	- Successes and challenges of implementation - Nature of and reasons for deviations from original implementation plans, including course development and industry collaboration - Level and patterns of demand for training in PICG-supported courses - Perceived sustainability of PICG-supported courses and risks to achieving long-term outcomes - Dissemination activities and potential for broader best-practice adoption (STPP grantees) - Perceived effects of policy changes supported by the project - Involvement in and perceived effects of annual TVET conference	- Changes to PICG-supported courses since the end of the compact - Level and patterns of demand for training in PICG-supported courses since the end of the compact - Sustainability of PICG-supported courses and barriers to continued sustainability - Perceptions of best practices, extent of adoption, and effects of adoption (STPP grantees and non-grantees) - Perceived effects of policy changes supported by the project
Trainees	- Motivation for enrollment in PICG-supported courses and types of trainees who enroll - Perceived strengths and weaknesses of course content, training quality, and training approaches - Plans and expectations for further training and employment	
Graduates of PICG-supported certificate courses		- Background information (quantitative), for example: gender, age, language, education, employment status, and work experience - Motivation for enrollment in PICG-supported certificate courses - Perceived strengths and weaknesses of course content, training quality, and training approaches - Effects of certification on labor market experiences (for example, employment, position, wages, and job mobility)
Teachers	- Changes in curriculum and teaching methods (for example, teaching materials, practices, and delivery methods) associated with the PICG-supported courses - Training and professional development associated with the PICG-supported courses, and additional training needs - Perceptions of trainees in PICG-supported courses (for example, their ability, interest, and motivation) - Main challenges of teaching PICG-supported courses, and how these might be addressed	
Employers and other	Motivation for involvement with PICG-supported courses	Background information (quantitative), for example: nature and location of business,

Respondent	Illustrative areas of focus	
	Interim round, 2018	**Final round, 2021**
PICG partners	• Nature and extent of involvement during the course development and implementation phases, satisfaction with the process, and key challenges faced • Nature of and reasons for deviations from original course development plans • Plans for continued involvement with PICG-supported courses after the compact, if any • Extent to which initial expectations about these courses have been or are likely to be met • Perceived sustainability of PICG-supported courses and risks to achieving long-term outcomes • Involvement in and perceived effects of annual TVET conference	annual revenues, total number of employees, breakdown of employment by primary occupation and other employee characteristics (for example, tenure, gender, and nationality) • Perceived skills of graduates from Georgian TVET programs before the PICG-supported courses were introduced • Awareness of PICG-supported courses and perceptions of these courses • Current hiring, training, and wage-setting practices, how these have changed over time, and reasons for these changes • Key challenges faced in hiring employees with the right skills, how these have changed over time, and reasons for these changes • Satisfaction with skills of graduates from PICG-supported courses and graduates from other courses, including key skills gaps • Differences in perceptions of graduates from PICG-supported courses by gender and other socio-demographic characteristics • Extent and nature of engagement with TVET providers, and future plans • Extent and nature of engagement with the MES, and future plans • Perceived sustainability of PICG-supported courses • Involvement in and perceived effects of annual TVET conference
Industry groups	• Extent and nature of engagement between industry, TVET providers, and the MES • Perceived skills of graduates from Georgian TVET programs • Awareness and perceptions of PICG-supported courses	• Changes in the extent and nature of engagement between industry, TVET providers, and the MES, and reasons for these changes • Perceptions of graduates from PICG-supported courses • Perceptions of broader trends in the skills of TVET graduates
MES and NCEQE staff	• The authorization process for PICG-supported courses, implications for course design, and related challenges • Status of various policy changes supported by the ISWD project, expectations for further changes, and related challenges • Perceived effects of policy changes supported by the project • Other policy changes in the TVET sector • Involvement in and perceived effects of annual TVET conference	• Long-term status of various policy changes supported by the ISWD project • Perceived long-term effects of policy changes supported by the project • Other policy changes in the TVET sector • Extent of a broader shift to industry demand-driven TVET courses, role of the ISWD project, and remaining barriers
PEM, MCA-Georgia, and	• Successes and challenges of implementation	

	Illustrative areas of focus	
Respondent	**Interim round, 2018**	**Final round, 2021**
local MCC staff	• Nature of and reasons for deviations from original implementation plans • Perceived sustainability and risks to achieving long-term outcomes	
Georgian Association of Private Colleges		• Awareness of best practices by members, extent of adoption, and effects of adoption • Perceived long-term effects of policy changes supported by the project • Involvement in and perceived effects of annual TVET conference • Extent of a broader shift to industry demand-driven TVET courses, role of the ISWD project, and remaining barriers
Other donors in the TVET sector	• Nature and scope of current and planned investments in the TVET sector • Interaction and coordination with ISWD project • Perceived sustainability of ISWD project and risks to achieving long-term outcomes	• Nature and scope of current and planned investments in the TVET sector • Perceptions of whether the ISWD project had a significant long-term impact • Perceived sustainability of ISWD project

Note: Blank cells indicate that we will not collect qualitative data from a particular respondent in the relevant round.

5.3. Project documents and administrative data

To complement the information gathered through the qualitative data collection effort, we will draw on several types of project documents and administrative data. First, we will examine the original PICG proposals and related documents (such as due diligence reports and cost-benefit analyses) in detail. This will enable us to investigate any deviations from the original plans by collecting grant-specific qualitative data and examining documents the grantees produced during the grant period (for example, curricula, teacher training materials, strategy plans). We will also be able to compare the estimated employment and wage rates from the trainee outcomes analysis to the assumed rates that grantees used to justify their PICG proposals. Second, administrative data on enrollment trends in PICG-supported courses will help us explore the long-term sustainability of these courses. Third, financial data from PICG grantees on their revenues and the costs of training will be useful to assess the long-term financial viability of the PICG-supported courses.[15] Fourth, administrative data on TVET course offerings from the MES (public courses) and the Georgian Association of Private Colleges (private courses) will help us assess broader changes in these offerings over time. Fifth, we will analyze documents from STPP grantees (including the final PowerPoint presentations compiled by each grantee) to describe the best practices identified and the dissemination process. Finally, information on TVET annual conference attendance and financing will contribute to our analysis of the effects of the conference.

---

[15] Public colleges' revenues are determined by a government formula. However, we could still assess how that formula is determined and whether it covers training costs, in addition to examining the revenues and costs for private colleges.

6. Analysis plan

In this section, we describe the main types of quantitative and qualitative analyses that we plan to conduct for the evaluation.

6.1. Quantitative analysis

We will conduct the following quantitative analyses using data from the trainee tracer survey:

- **A descriptive outcomes analysis of tracer survey data for trainees in PICG-supported courses.** For each trainee-level outcome that we examine, we will present both graphical and numerical descriptions of the averages for the full sample of trainees in PICG-supported courses. We will also provide similar descriptions of background trainee characteristics such as gender, education level, marital status, language group, region of origin, parental education, and disability status. To complement the overall averages, we will also present averages for specific providers or courses, and for other subgroups of interest to MCC (for example, those defined by gender or language group). However, the estimated averages will be less precise for these subgroup analyses and will depend on the sample sizes for each subgroup.

- **A descriptive benchmarking outcomes analysis using data from the trainee tracer survey.** For this analysis, we will estimate the average characteristics and outcomes of trainees in benchmarking courses and contrast them with the respective averages for trainees in PICG-supported courses, using numerical and graphical comparisons. The simplest analysis approach would be to conduct these comparisons using the overall averages for trainees in the PICG-supported courses and benchmarking courses. This approach would be appropriate if we use a broad cross-section of non-supported courses as the benchmark, rather than identifying specific benchmarking courses for each PICG-supported course using data from the baseline trainee tracer survey. If we are able to identify specific benchmarking courses, we will reweight the benchmarking course sample to reflect the distribution of trainees in the PICG courses before estimating the averages for the benchmarking course trainees. For example, if a certain course with a large number of trainees is serving as a benchmark for a PICG course with a small number of trainees, we would reweight that benchmarking course to ensure that it makes a relatively small contribution to the overall benchmarking average. With appropriate reweighting, we could also conduct benchmarking analyses for specific PICG providers and courses, although provider- and course-specific estimates might be imprecise if the sample size of trainees is low.

  We will also attempt to implement statistical matching techniques to balance the observed characteristics of trainees in the PICG-supported and benchmarking courses. A covariate matching approach would involve grouping trainees into cells defined by combinations of discrete trainee characteristics (for example, gender, education categories, and age categories). Then we would estimate the differences in outcomes for trainees in PICG-supported and benchmarking courses within each cell (for example, male trainees with a grade 12 education who are ages 18 to 24) and combine these estimates using the number of trainees in PICG-supported courses in each cell as weights. An alternative approach, propensity score matching, would involve estimating a single metric for each trainee based on his or her characteristics (known as the propensity score), and comparing trainees in

PICG-supported courses to trainees with similar propensity scores in benchmarking courses. If we identify specific benchmarking courses for each PICG-supported course, we could implement these matching approaches separately for each PICG-supported course and its benchmarking courses, and then combine the estimates. For these matching approaches to be feasible, it is necessary that the characteristics of trainees in PICG-supported and benchmarking courses are not too dissimilar; otherwise, we will not be able to find appropriate matches for a large fraction of trainees in PICG-supported courses. If statistical matching is feasible, the benchmarking analysis would provide more credible evidence about the program's potential effects, although it would still fall short of a rigorous impact analysis given the potential for unobserved differences between trainees in PICG-supported and benchmarking courses.

- **A pre-post analysis of tracer survey data for trainees in improved courses.** This analysis will compare the outcomes of trainees in improved PICG-supported courses to those of trainees who enrolled in the pre-improvement versions of the same courses. To conduct this comparison, we will use the following regression model:

$$Y_{ijt} = \alpha + \beta POST_t + \gamma X_{ij} + \delta_j + \varepsilon_{ijt} \ ,$$

where Y_{ijt} is the outcome for trainee i enrolled in course j at time t (where t is before or after the improvements); $POST_t$ is a binary indicator for the trainee being enrolled after the PICG-supported improvements; X_{ij} is a set of pre-determined trainee characteristics such as gender and education level; δ_j is a set of binary indicators, one for each course, which ensure that the pre-post comparisons are made within each course and then averaged over the full sample; and ε_{ijt} is a random error term. The estimated value of the coefficient β represents the pre-post change in the outcome of interest for the average trainee; with appropriate reweighting we can also estimate the change for the trainee in the average improved course or average provider in the sample.[16]

6.2. Cost-benefit analysis

According to documentation MCC provided to Mathematica, MCC produced an ex-ante cost benefit analysis model with an estimated ERR of 14 percent for the ISWD project. The ERR is a summary statistic that is used by MCC to determine the cost-effectiveness of its investments. Conceptually, it is the discount rate at which the cumulative benefits of a program over time are exactly equal to its costs; a higher (positive) ERR represents higher benefits and lower costs.

MCC's ex-ante ERR model focuses on Activity 1, which accounts for most of the ISWD project funding and has the most clearly defined benefits. The costs in the ERR model include the total PICG investment amount from both MCC and the private sector, as well as tuition costs. The main benefits are higher earnings of PICG trainees compared to their earnings had they taken existing courses. These higher earnings are driven by: (1) a higher expected employment rate (an increase of 9 percentage points); and (2) higher expected wages for those employed (an

---

[16] This is relevant because a large fraction of trainees in the improved course sample are receiving training from one provider; therefore, the pre-post estimate for the average trainee will largely be driven by the pre-post change for this provider. Although this estimate is of interest because most of the improved course beneficiaries are receiving training from this one provider, it might mask interesting differences in pre-post changes across courses and providers.

increase of about 24 percent from a base of 319 GEL per month, in 2010 currency). MCC plans to update the ex-ante ERR model in summer 2017 to reflect updated information on investment costs, the number of expected trainee beneficiaries, the timing of the PICG courses, and so on.

Because the evaluation design for the ISWD project does not currently include an impact evaluation, we will not be able to produce a comparable ex-post ERR estimate that uses rigorous evidence of the employment rates and wages of trainees in PICG-supported courses relative to the counterfactual in which the PICG training did not exist. However, our outcomes analyses can still shed light on whether the assumed improvements in employment and wages, as well as the estimated base wage rate (a crucial parameter in the model), are reasonable. For example, if employment rates are similar for graduates from PICG-supported and benchmarking courses, then large increases in employment might not be plausible. Similarly, if the wage rate for trainees in benchmarking courses is very different from the assumed baseline wage level, it would be important to consider revising this parameter.

6.3. Qualitative analysis

To analyze qualitative data, Mathematica will use qualitative transcript-coding software to organize and synthesize the key themes that emerge from document reviews, in-depth interviews, and focus groups. More specifically, we will follow four steps to analyze the data (Creswell 2009):

1. **Raw data management.** Raw data management is the process of organizing such data into meaningful units of analysis (that is, from audio files to transcripts). During this step, we will review all data and eliminate any that are incomplete or not useful to our analysis.

2. **"Chunking" and initial coding.** Often referred to as *data reduction*, this step will allow us to read the transcripts several times and obtain a holistic sense of the data. We will develop a detailed initial coding scheme—a set of themes we might encounter in the interview and focus group transcripts, which are mapped to the research questions and logic model (for example, initial themes might include "implementation challenges," "employer/provider engagement," and "adoption of new practices"). We will also begin developing internal memos to accompany the broader coding themes.

3. **Detailed coding.** This step will involve refining the coding scheme and recoding data as we look at the data in greater depth. We will use Atlas.ti or NVivo software to review and code the transcripts based on the initial codes developed during the "chunking" process. We will expand and refine these codes during the coding exercise and subsequent analysis of the coded transcripts iteratively as additional themes emerge.

4. **Data interpretation and writing.** Analyzing the coded transcripts will involve triangulating the findings across stakeholders to highlight mechanisms, context, and similarities and differences in perspectives. For example, for Activity 1, the analysis of interviews with employers of graduates from PICG-supported courses might identify differences in employers' perceptions of and satisfaction with graduates across PICG grantees. By comparing these to grantee-specific implementation challenges reported by implementers and grantees, as well as to differences in training experiences and perceptions reported by PICG trainees, we will seek to identify why some PICG-supported courses were more successful in meeting employers' needs than others.

For Activity 2, we will compare the dissemination steps reported by STPP grantees and implementers to non-grantees' reported awareness, perceptions, and adoption of best practices. This should enable the study to determine the extent to which practices were likely to be widely known and adopted, and which dissemination mechanisms are likely to have been most successful.

For Activity 3, we will compare the progress on specific MES-reported policy reforms to the effects observed by key stakeholders in the TVET sector. For example, the analysis will cross-check MES-reported progress on national regulations for public-private partnerships and sector skills councils with the perceptions of and involvement in these mechanisms by TVET providers and employers. This will enable us to assess how and to what extent the implemented reforms have affected or are likely to affect the TVET sector.

Finally, for Activity 4, we will analyze qualitative interview data from participants in the annual TVET conferences and observe conference activities directly, to assess whether these events are likely to have encouraged deeper levels of collaboration and engagement among TVET stakeholders.

6.4. Analysis of project documents and administrative data

Our analysis of quantitative administrative data will be descriptive in nature. For example, we will describe trends in enrollment in PICG-supported courses, the range and averages of revenues and costs for these courses, the training levels and areas of broader public and private TVET course offerings, and the numbers and types of attendees at the annual TVET conference.

To analyze project documents such as the original PICG grant proposals and related documents, we will conduct a structured document review. This review will be conceptually similar to the coding approach described above for qualitative data analysis and will enable us to summarize the key content of these document in an organized manner. For example, we could summarize the planned activities for each PICG grant in areas such as curricula and training modules, training materials, teacher recruitment and training, laboratories and other infrastructure development, and pedagogical approaches. From this review, we will be able to both develop grant-specific areas of focus for the qualitative data collection (for example, asking about specific types of infrastructure that were planned) and identify additional administrative data or documents that might support the analysis of implementation (for example, numbers of teachers trained or teacher training materials).

7. Limitations and challenges

Our evaluation of the ISWD project faces some important challenges and limitations that we will attempt to address to the extent possible:

- **Absence of a rigorous counterfactual.** Our evaluation design is a descriptive performance evaluation. The study will not identify a rigorous counterfactual for trainees in PICG-supported courses—that is, we cannot confidently determine what the labor market outcomes of these trainees would have been in the absence of the project. Our benchmarking and pre-post analyses will provide information about outcomes in non-PICG courses, but this should be viewed as suggestive and not rigorous evidence about the counterfactual

situation. After PICG courses begin, the study team will assess whether it might be possible to provide more rigorous impact estimates through a future randomized controlled trial.

- **Potential for low response rates in the trainee tracer survey.** Our ability to provide quantitative evidence on trainee outcomes that is generalizable to the full group of PICG beneficiaries depends on achieving high response rates to the tracer survey. Otherwise, there may be a concern that only certain types of trainees—for example, those with better outcomes—are appearing in the data and driving the findings. We will use several strategies to ensure high response rates, including collecting detailed contact information at baseline and using a multi-pronged approach to locate trainees for the follow-up survey (as discussed in Section 5). In the analysis, we will also be able to get a sense of a lower bound on some outcomes by assuming a "worst case" scenario—for example, by assuming that all those who do not respond are unemployed.

- **Potential for heterogeneous effects across PICG grantees.** There is substantial variation across PICG-supported courses in terms of course content, course design, and delivery approaches. In addition, grantees could differ in terms of their capacity to effectively implement the new courses and could face different challenges related to implementation and sustainability. Therefore, it is likely that the effects of new and improved courses on labor market outcomes will vary substantially across grantees. Our ability to explore this variation in effects quantitatively through the trainee outcomes analysis will be limited, because small provider-level sample sizes will result in imprecise estimates of labor market outcomes for specific grantees. However, the qualitative study will enable us to focus in some detail on the experiences of specific grantees. In particular, we plan to conduct a detailed structured document review focused on each of the grantees, together with two rounds of in-depth qualitative interviews with each grantee to fully understand their specific experiences with implementation and sustainability. In addition, by conducting qualitative interviews with a relatively large sample of employers in the final round of data collection, we will be able to explore differences in employer perceptions across different types of grantees, including whether there is variation in the extent to which their PICG-supported courses met employers' needs.

8. Administrative details

8.1. Institutional review board requirements and clearances

Mathematica will prepare and submit institutional review board (IRB) applications for approval of the research and data collection plans for the baseline trainee tracer survey, follow-up trainee tracer survey, and each round of qualitative data collection. The application materials include three sets of documents: (1) a research protocol, which will draw heavily on the present design report and adds more information about plans for protecting study participants' confidentiality and human rights; (2) copies of all data collection instruments, including statements of informed consent; and (3) a completed IRB questionnaire that summarizes the key elements of the research protocol, plans for protecting participants' human rights, and possible threats to participants if their confidentiality were compromised. Based on experience, we expect that the study will qualify for expedited review because it presents minimal risk to participants. If so, the IRB can typically review the application within one week of its submission.

IRB approval is valid for one year from the date of approval and must be renewed annually. We expect that the annual renewals will require minimal updates to the core application materials. In addition, if data collection instruments change substantially from those that the IRB approved, then we must reapply for approval. Small changes to the instruments (such as rewording or reordering of questions or editing changes) do not require reapplication, but the finalized instruments must be submitted to the IRB for documentation.

After Mathematica drafts the IRB research protocol, we will coordinate with MCA-Georgia to ensure the data collector and local stakeholders agree on the data collection protocol. The data collector's contract with MCA-Georgia (for data collection that occurs during the compact) or Mathematica (for data collection that occurs after the compact) must specify that it will abide by the IRB's recommendations. The data collector and Mathematica must also sign an IRB authorization agreement stating that the data collector will adhere to the IRB-approved data collection procedures and protocols.

8.2. Data access, privacy, and documentation

After producing each of the interim and final reports, we will prepare corresponding de-identified data files and codebooks that MCC can make available to the public. We will de-identify these data files, user manuals, and codebooks according to the most recent guidelines set forth by MCC. The public-use data files will be free of personal or geographic identifiers that would allow users to directly identify individual respondents or their households, and we will remove or adjust variables that could introduce reasonable risks of deductively disclosing the identity of individual participants. Mathematica will remove all individual identifiers, including names, addresses, telephone numbers, and any other similar variables. We will also remove unique and rare data using local suppression, replacing these observations with missing values instead. If necessary, we will also use top and bottom coding, setting upper and lower bounds to remove outliers and collapse any variables that make an individual highly visible depending on geographic or other factors (such as home language) into less easily identifiable categories. Our manner of data perturbation will not significantly degrade the data.

8.3. Dissemination plan

Mathematica will present interim and final evaluation findings in person to MCC and to stakeholders in Georgia. The interim analysis will occur after completing the baseline trainee tracer survey and interim round of qualitative data collection. This analysis will produce an interim report that summarizes the characteristics of the sample from the trainee tracer survey and the qualitative findings related to implementation and early results from all activities, which we expect to submit in March 2019. Following the follow-up trainee tracer survey, employer survey, and final round of qualitative data collection, we will produce a final evaluation report for the ISWD project, which we expect to submit in September 2021.[17] Table 8.1 shows how the

---

[17] This date is based on the current evaluation contract, and allows for several months of revisions before the end of the contract in January 2022. If the timeline is adjusted so that we can include additional PICG courses in the follow-up survey, include additional trainee cohorts in the sample, or conduct an impact evaluation (as discussed in Section 4), we would adjust the timing of data collection and the final report accordingly.

timing of the interim and final reports aligns with the planned data collection activities for the evaluation.

We will work with MCC to increase the visibility of the study's findings, particularly among education policymakers and development practitioners. We will collaborate with MCC and stakeholders to identify a variety of forums—including conferences, workshops, and publications—during which to share results and encourage donors, implementers, and policymakers to integrate the findings into future programming. For example, in addition to the project's final report, we will develop issue briefs summarizing and visualizing key findings for a broader audience of readers and stakeholders. Potential conferences for presenting evaluation findings will include forums hosted by the Comparative International Education Society, the American Evaluation Association, or the Association for Public Policy Analysis and Management. We will also seek to publish a peer-reviewed article disseminating the study's results in academic or sector-specific journals focused on vocational education systems in developing countries.

Table 8.1. Evaluation timeline and reporting schedule

Year	2017			2018				2019				2020				2021				2022	
Quarter	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2
Data collection								*End of compact*												*End of evaluation*	
Baseline trainee tracer survey[a]	░		░	░			░														
Follow-up trainee tracer survey						░	░	░				░	░	░							
Qualitative data						░	░										░	░			
Reporting																					
Interim and final evaluation reports							░	░										░	░		

[a] Baseline data collection in Q2, 2017, is for trainees in pre-improvement courses. Baseline data collection for PICG-supported and benchmarking courses will begin in Q4, 2017.

8.4. Evaluation team roles and responsibilities

Mathematica's project team has extensive experience conducting mixed-methods, multicomponent, large-scale evaluations in the fields of general and vocational education. Mr. Ira Nichols-Barrer will serve as the program manager, acting as the primary point of contact for MCC. He will manage the relationships with government agencies and other local entities and contractors, while supervising the evaluation design and implementation process and ensuring high data quality. Dr. Evan Borkum is the principal investigator for this evaluation, providing methodological and technical oversight and serving as a senior analyst overseeing the study's quantitative data collection and analysis process. Dr. Candace Miller will serve as a senior analyst overseeing the study's qualitative data collection and analysis process. Ms. Irina Cheban will serve as the project analyst, supporting data collection, analysis, and reporting efforts on the evaluation. Dr. Natia Gorgadze will serve as the project's in-country consultant, providing substantive knowledge of Georgia's education system and assisting with the study's data collection and other local evaluation management tasks.

8.5. Budget

At this time, Mathematica does not anticipate that the ISWD evaluation design and data analysis plans described in this report will require changes to the total evaluation budget figure presented in the study's original proposal. Mathematica will work closely with MCC and MCA-Georgia to ensure data collection is feasible within the compact's budget parameters.

References

Acevedo, Paloma, Guillermo Cruces, Paul Gertler, and Sebastián Martínez. "Living Up to Expectations: How Job Training Made Women Better Off and Men Worse Off." Cambridge, MA: National Bureau of Economic Research, March 2017.

Alzúa, María Laura, Guillermo Cruces, and Carolina Lopez. "Long-Run Effects of Youth Training Programs: Experimental Evidence from Argentina." *Economic Inquiry,* vol. 54, no. 4, 2016, pp. 1839–1859.

Asian Development Bank. "Viet Nam: Vocational and Technical Education Project." Manila, Philippines: Asian Development Bank, 2013.

Attanasio, Orazio, Adriana Kugler, and Costas Meghir. "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics,* vol. 3, no. 3, 2011, pp. 188–220.

Attanasio, Orazio, Arlen Guarín, Carlos Medina, and Costas Meghir. "Long Term Impacts of Vouchers for Vocational Training: Experimental Evidence for Colombia." Cambridge, MA: National Bureau of Economic Research, 2015.

Borkum, Evan, Arif Mamun, and Malik Khan Mubeen. "Evaluation of the Vocational Training Grant Fund in Namibia: Final Report." Draft report submitted to the Millennium Challenge Corporation. Washington, DC: Mathematica Policy Research, March 2017.

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas-Shady, and Yuri Soares. "The Labor Market Impacts of Youth Training in the Dominican Republic." *Journal of Labor Economics,* vol. 29, no. 2, 2011, pp. 267–300.

Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, and Victor Orozco. "Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-out Behavior." Policy Research Working Paper No. WPS 6545; Impact Evaluation Series No. IE 99. Washington, DC: World Bank Group, 2013.

Creswell, John. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* Los Angeles: Sage, 2009.

Delajara, Marcelo, Samuel Freije, and Isidro Soloaga. "An Evaluation of Training for the Unemployed in Mexico." Working paper 0906. New York: Inter-American Development Bank Office of Evaluation and Oversight. 2006.

Diaz, Juan Jose, and David Rosas. "Impact Evaluation of the Job Youth Training Program ProJoven." IADB Working Paper No. 693. Washington, DC: Inter-American Development Bank, 2016.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Cano. "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey." *Economic Journal,* vol. 126, no. 597, 2016, pp. 2115–2146.

Honorati, Maddalena. "The Impact of Private Sector Internship and Training on Urban Youth in Kenya." Policy Research Working Paper No. WPS 7404; Impact Evaluation Series. Washington, DC: World Bank Group, 2015.

Ibarrarán, Pablo, Jochen Kluve, Laura Ripani, and David Rosas. "Experimental Evidence on the Long-Term Impacts of a Youth Training Program." IDB Working Paper No. 657. Washington, DC: Inter-American Development Bank, 2015.

Ibarrarán, Pablo, Laura Ripani, Bibiana Taboada, Juan Miguel Villa, and Brigida Garcia. "Lifeskills, Employability, and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design." *IZA Journal of Labor and Development,* vol. 3, no. 10, 2014.

Ibarrarán, Pablo, and David Rosas Shady. "Evaluating the Impact of Job Training Programmes in Latin America: Evidence from IDB Funded Operations." *Journal of Development Effectiveness,* vol. 1, no. 2, 2009, pp. 195–216.

Kelly, Alan, Richard Willing, and Marivic Puzon. "The Philippines-Australia Technical and Vocational Education Project (PATVEP)." AusAID Evaluation No. 4. Canberra, Australia: AusAID, 1998.

Maitra, Pushkar, and Subha Mani. "Learning and Earning: Evidence from a Randomized Evaluation in India." *Labour Economics*, vol.45, 2017, pp.116–130.

McKenzie, David. "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence" *IZA Discussion Papers*, No. 10655, 2017.

Ministry of Education and Science, Georgia. "Vocational Education and Training Development Strategy for 2013-2020." Tbilisi, Georgia: Ministry of Education and Science, 2013.

Tripney, Janice, Jorge Garcia Hombrados, Mark Newman, Kimberly Hovish, Chris Brown, Katarzyna T. Steinka-Fry, and Eric Wilkey. "Post-Basic Technical and Vocational Education and Training (TVET) Interventions to Improve Employability and Employment of TVET Graduates in Low- and Middle-Income Countries." *Campbell Systematic Reviews,* vol. 9, no. 9, 2013.

U.S. Agency for International Development. "State of the Field Report: Examining the Evidence in Youth Workforce Development." Washington, DC: U.S. Agency for International Development, 2013.

APPENDIX A:

LIST OF PICG-SUPPORTED COURSES AND STPP GRANTS

Table A.1. Program Improvement Competitive Grants (PICG)-supported courses, by provider

Provider name	Course name	Level	Course type	Expected number of trainees	Expected start date	Expected end date
1. Agricultural University of Georgia	Farmer	4	New	80	Oct 2017	Nov 2019
	Veterinary service specialist	5	New	50	Oct 2017	Aug 2019
	Viticulturist-oenologist	5	New	70	Oct 2017	Apr 2019
2. Batumi State Maritime Academy	Crane operator	3	New	15	Oct 2017	Sep 2018
	Crane operator	4	New	15	Oct 2017	Sep 2018
	Fishing vessel navigator	5	New	60	Oct 2017	Dec 2018
	Logistics and port management	4	New	10	Oct 2017	June 2018
	Logistics and port management	5	New	10	Oct 2017	June 2018
	Welder	3	New	30	Oct 2017	Sep 2018
	Welder	4	New	24	Oct 2017	Sep 2018
3. Community College Spectri	Air-conditioning technician	4	Existing	65	Oct 2017	Oct 2018
	Electrician	4	Existing	65	Oct 2017	Oct 2018
	Electrician	4	Existing	40	Oct 2017	Oct 2018
	Water supply systems exploitation	4	New	70	Oct 2017	Oct 2018
	Water sewage systems exploitation	5	New	70	Oct 2017	Oct 2018
	Welder	4	Existing	40	Oct 2017	Oct 2018
	Welder	5	Existing	65	Oct 2017	Oct 2018
4. Georgia Railway Transport College	Electrical mechanic	4	Existing	15	Apr 2018	Apr 2020
	Locomotive driver	4	New	15	Apr 2018	Apr 2020
	Rail car mechanic	4	Existing	15	Apr 2018	Apr 2020
	Rail carrier	4	Existing	20	Apr 2018	Apr 2020
	Rolling stock mechanic	4	Existing	15	Apr 2018	Apr 2020
	Signalization, centralization and blocking mechanic	5	Existing	15	Apr 2018	Oct 2020
	Testing inspector	4	New	15	Apr 2018	Apr 2019
	Track repairer	3	Existing	15	Apr 2018	Oct 2019
5. Georgian Aviation University	Aircraft maintenance technician: mechanical (airplane turbine)	5	New	45	Oct 2017	Oct 2019
	Aircraft maintenance technician: avionics	5	New	40	Oct 2017	Oct 2019
	Helicopter pilot	5	New	20	Oct 2017	Oct 2019
6. Georgian Institute of Public Affairs	Occupational health, safety and environmental specialist/manager	5	New	40	Oct 2017	Aug 2019
	Occupational health, safety and environmental specialist/manager[a]	5	New	40	Mar 2018	Apr 2020

Provider name	Course name	Level	Course type	Expected number of trainees	Expected start date	Expected end date
7. Georgian Mountain Guide Association	Mountain guide	4	New	15	Oct 2016	Oct 2018
	Trekking guide	3	New	15	Oct 2016	Oct 2017
	Ski teacher (levels 1-3)	1-3	New	30	Oct 2016	Oct 2017
8. Georgian Technical University	Electrical technician, high voltage	4	New	24	Nov 2017	Jan 2019
	Industrial automation technician	5	New	24	Nov 2017	Nov 2019
	Manufacturing supervisor	5	New	24	Nov 2017	May 2018
	Mechanical engineering technician	5	Existing	24	Nov 2017	Nov 2019
	Mechatronics technician	5	Existing	24	Nov 2017	Nov 2019
9. Vocational College Phasizi	Fish breeding technician	4	New	55	Oct 2017	Jan 2020
	Fish lab technician	4	New	30	Oct 2017	Jan 2020
	Fish processing specialist	4	New	15	Oct 2017	Jan 2020
10. Vocational College Tetnuldi	IT support technician	3	Existing	110	Oct 2017	Oct 2018
	Computer network administrator	5	Existing	30	Apr 2018	Apr 2020
	Network and systems technician	4	New	40	Apr 2018	Oct 2019
Total				**1,549**		

Source: Personal correspondence with PEM.

[a] This is the same course as in the row immediately above, but enrollment will be staggered across two different cohorts.

Table A.2. Strengthening TVET Provider Practice (STPP) grants

Grantee	Project name
Round 1: $68,994	
1. Business Academy of Georgia (SBA)	Development of Assessment Tools for the Entrepreneurship and Introductory Practice Modules
2. Community College AISI	Teacher's Professional Development Practice
3. Georgian Employers' Association	Non-formal Educational Program in the Work-Based Learning Format
4. Georgian Technical University	Development of E-learning Course in IT
5. Kutaisi Public School #33	Social Enterprise in Public Schools
6. Mindstream Ltd.	Career Planning & Employer Communication Strategy
7. Vocational College Icarosi	Employers Forum for Industry Engagement in TVET Sector
Round 2: $172,186	
1. Akaki Tsereteli State University	New Professional Personnel for the Use of Solar Energy
2. Community College Akhhali Talga	Supporting individual learning paths of TVET students
3. EasySoft LTD	Learning platform of innovation technology for professional education
4. EMIS	Introducing Informal Education Recognition Methodology
5. Georgian Adult Education Network (GAEN)	Strengthening entrepreneurial training in non-formal education
6. GeoTuran Ltd.	Euro Master 2017
7. ISET Policy Institute	Strengthening entrepreneurial training in formal TVET system
8. Mindworks Ltd.	Flipped Classroom Deployment in BLACKSEA and ERKVANI
9. The Georgian Patriarchate Community College of Decorative Gardening	Promoting TVET related to decorative gardening professions among general school students
10. Vocational College Modus	Vocational training through distance learning, based on theory, practice and visual media

Source: The ISWD project website: http://www.iswd.ge/.

APPENDIX B:

BASELINE TRAINEE TRACER SURVEY INSTRUMENT

Georgia Technical Vocational Education and Training Baseline Survey[18]

Course information sheet

This course information sheet should be completed by the survey team with training provider staff before starting the survey. Only one course information sheet is required for each group of trainees surveyed.

X1.   Code of vocational training program (please assign a unique code; this should match the code on the consent statement, Y1)

|__|__|

X2.   What is the name of this vocational training provider?

_____

X3.   What is the name of this vocational training program?

_____

X4.   What is the level of this vocational training program? [Select one].
1 ❑ LEVEL I
2 ❑ LEVEL II
3 ❑ LEVEL III
4 ❑ LEVEL IV
5 ❑ LEVEL V
6 ❑ DOES NOT HAVE A LEVEL

X5. What month and year did this vocational training program start?

Month   Year
|__|__| |__|__| __| __|

X6. What month and year will this vocational training program end?

Month   Year
|__|__| |__|__| __| __|

---

[18] This version of the baseline survey reflects lessons from administering the survey to trainees in pre-improvement courses in May and June 2017.

Y.1 Code of vocational training program (prefill) |__|__|

Georgia Technical Vocational Education and Training Baseline Survey

<u>Consent statement:</u>

Mathematica Policy Research, a U.S. based evaluation firm, is conducting a study of vocational training in Georgia. We plan to gather information about employment outcomes of trainees who graduated from vocational training courses in in order to study improvements in the vocational education system. Our study is funded by the Millennium Challenge Corporation, an agency that provides assistance to other countries' development projects, and is being carried out with the support of Millennium Challenge Account - Georgia. If you agree to participate in this survey, we will gather information about your experience with existing vocational training courses and anticipated future employment. We will also collect your contact information to follow up with you one year after your course ends to talk about employment outcomes. In addition, we will gather information on your national identification number and NAEC test scores (if applicable) directly from the administrative records held by your vocational training provider.

The survey is expected to take 20 minutes. Any information you provide that can identify you will be kept confidential by the parties conducting this study, including MCC employees, employees of the survey firm, and researchers, to the maximum extent permitted by the laws of the United States of America and the laws of the Republic of Georgia. These users will use data for statistical purposes only.

Your participation is voluntary and you may choose not to answer any or all questions for any reason. In other words, you have the alternative to not participate and there will be no consequences for nonparticipation. You may contact Dr. Natia Gorgadze at *[Local Phone Number]*, if you have questions, concerns or complaints about the study or your rights as a participant. If you have any questions for us, please feel free to ask at any time.

Y2. Please indicate your decision whether to participate in the study by checking one of the boxes below. If you agree to participate in the study, please provide your first name, last name, and sign to confirm your participation.

❑ I AGREE TO PARTICIPATE IN THIS STUDY

First name: _____

Last name: _____

Signature: _____

❑ I DO NOT WISH TO PARTICIPATE IN THIS STUDY

SECTION A: TRAINING INFORMATION AND EXPECTATIONS FOR THE FUTURE

A1. What other training programs were you interested in when you applied to your current program? Please give the provider name, program name, and level for up to five training programs you were interested in at the time, ranked in order of interest (starting with the program in which you were most interested). Also, please indicate which of these programs you applied to, if any.

Provider name	Program name	Level	Did you apply?
1._____	_____	❑ I ❑ II❑ III ❑ IV❑ V❑ NONE	❑ YES ❑ NO
2. _____	_____	❑ I ❑ II❑ III ❑ IV❑ V❑ NONE	❑ YES ❑ NO
3._____	_____	❑ I ❑ II❑ III ❑ IV❑ V❑ NONE	❑ YES ❑ NO
4._____	_____	❑ I ❑ II❑ III ❑ IV❑ V❑ NONE	❑ YES ❑ NO
5._____	_____	❑ I ❑ II❑ III ❑ IV❑ V❑ NONE	❑ YES ❑ NO

A2. What do you expect to be your main activity one year after you graduate from your current training program? [Select one].

        1 ❑ EMPLOYED IN A FULL-TIME JOB
        2 ❑ EMPLOYED IN A PART-TIME JOB
        3 ❑ SELF-EMPLOYED
        4 ❑ ENGAGED IN FURTHER TRAINING OR STUDY
        5 ❑ UNEMPLOYED AND SEARCHING FOR A JOB
        6 ❑ OTHER (SPECIFY _____)

If you did not select options 1, 2, or 3 in A2, please skip question A3 and proceed to question A4.

A3. Please provide the monthly salary or profits you expect to receive from your job or self-employment one year after graduation, in GEL (before taxes).

    Expected monthly salary or profits one year after graduation
    |__|__|__|__|__|__| GEL

A4. Please give the names and locations of up to three employers you would be interested in working for one year after graduation, ranked in order of interest (starting with the employer in which you are most interested).

Employer name   Location

1._____   _____

2. _____   _____

3._____   _____

SECTION B: BACKGROUND INFORMATION

B1. Are you male or female?
       1 ❑ MALE
       2 ❑ FEMALE

B2. How old were you when you started the vocational training program in which you are currently enrolled?
       |__|__| Years

B3. What is the main language spoken in your home? [Select one].
       1 ❑ GEORGIAN
       2 ❑ ABKHAZ
       3 ❑ ARMENIAN
       4 ❑ AZERBAIJANI
       5 ❑ RUSSIAN
       6 ❑ OTHER (SPECIFY _____)

B4. What is your region of origin? [Select one].

       1 ❑ ABKHAZIA
       2 ❑ SAMEGRELO-ZEMO SVANETI
       3 ❑ GURIA
       4 ❑ ADJARA
       5 ❑ RACHA-LECHKHUMI AND KVEMO SVANETI
       6 ❑ IMERETI
       7 ❑ SAMTSKHE-JAVAKHETI
       8 ❑ SHIDA KARTLI
       9 ❑ MTSKHETA-MTIANETI
       10 ❑ KVEMO KARTLI
       11 ❑ KAKHETI
       12 ❑ TBILISI
       13 ❑ ANOTHER COUNTRY (SPECIFY _____)

B5. Are you currently married?
       1 ❑ YES
       2 ❑ NO

B6. How many children do you have? Please do not leave blank; if you do not have any children, enter zero.
       |__| Children

B7. Do you have a physical disability, such as a serious hearing or vision problem that cannot be corrected or a condition that substantially limits basic physical activities such as walking or climbing stairs?
       1 ❑ YES
       2 ❑ NO

B8. Do you have an emotional or mental condition that makes it difficult to learn or fully participate in education and training?

     1 ❑ YES
     2 ❑ NO

B9. What is the highest level of education that your father completed? [Select one].

     1 ❑ LESS THAN GRADE 9
     2 ❑ GRADE 9
     3 ❑ GRADE 9 AND ADDITIONAL EDUCATION OR TRAINING
     4 ❑ GRADE 12
     5 ❑ GRADE 12 AND ADDITIONAL EDUCATION OR TRAINING

B10. What is the highest level of education that your mother completed? [Select one].

     1 ❑ LESS THAN GRADE 9
     2 ❑ GRADE 9
     3 ❑ GRADE 9 AND ADDITIONAL EDUCATION OR TRAINING
     4 ❑ GRADE 12
     5 ❑ GRADE 12 AND ADDITIONAL EDUCATION OR TRAINING

B11. What is the highest grade of general education that you have completed? [Select one].

     1 ❑ LESS THAN GRADE 9
     2 ❑ GRADE 9
     3 ❑ GRADE 10
     4 ❑ GRADE 11
     5 ❑ GRADE 12

B12. Before enrolling in your current vocational training program, did you <u>complete</u> any other education or training program beyond general education? [Select one; if you have completed more than one type of additional training, please select the highest level completed].

     1 ❑ NO
     2 ❑ YES, VOCATIONAL TRAINING
     3 ❑ YES, UNIVERSITY EDUCATION
     4 ❑ YES, OTHER EDUCATION OR TRAINING

<u>If you did not select options 2, 3 or 4 in B12, please skip question B13 and proceed to question B14</u>

B13. Please give the provider name, program name, level, program duration, and completion date for the highest level training program that you <u>completed</u> before enrolling in the current program.

Provider name   Program name   Level

_____ _____ ❑ I ❑ II❑ III ❑ IV❑ V❑ NONE❑ OTHER (SPECIFY_____)

Duration of program   Month and year you completed the program
|__|__| Months |__|__| |__|__| __| __|

B14. Are you <u>currently enrolled</u> in any training program besides the program where we are interviewing you?

    1 ❑ YES
    2 ❑ NO

<u>If you did not select option 1 in B14, please skip question B15 and proceed to question B16</u>

B15. Please give the provider name, program name, level, program duration, and expected completion date for the other program in which you are currently enrolled.

Provider name   Program name   Level

 _____   _____   ❑ I ❑ II❑ III ❑ IV❑ V❑ NONE❑ OTHER (SPECIFY_____)

Duration of program   Month and year you expect to complete the program
|__|__| Months |__|__| |__|__| __| __|

B16. Are you currently employed (select one)?
    1 ❑ YES, EMPLOYED IN A FULL-TIME JOB
    2 ❑ YES, EMPLOYED IN A PART-TIME JOB
    3 ❑ YES, SELF-EMPLOYED
    4 ❑ NO

<u>If you did not select option 1, 2, or 3 in B16, please skip question B17 and proceed to question B18</u>

B17. Is your current employment related to your current field of study?
    1 ❑ YES
    2 ❑ NO

B18. How many total years of work experience do you have? Please do not leave blank; if you have never worked, enter zero.

    |__| Total years of work experience

B19. How many years of work experience do you have that are <u>related to your current field of study</u>? Please do not leave blank; if you have never worked in this field, enter zero.

    |__| Years of work experience related to current field of study

SECTION C: PERCEPTIONS OF CURRENT VOCATIONAL TRAINING PROGRAM

C1. In your opinion, what is the quality of the instructors in the vocational training program that you are currently attending? [Select one].
    1 ❑ EXCELLENT
    2 ❑ GOOD
    3 ❑ AVERAGE
    4 ❑ POOR

C2. In your opinion, what is the quality of the written materials in the vocational training program that you are currently attending? [Select one].

      1 ❑ EXCELLENT
      2 ❑ GOOD
      3 ❑ AVERAGE
      4 ❑ POOR

C3. In your opinion, what is the quality of the tools and equipment in the vocational training program that you are currently attending? [Select one].

      1 ❑ EXCELLENT
      2 ❑ GOOD
      3 ❑ AVERAGE
      4 ❑ POOR

C4. In your opinion, what is the overall quality of the vocational training program that you are currently attending? [Select one].

      1 ❑ EXCELLENT
      2 ❑ GOOD
      3 ❑ AVERAGE
      4 ❑ POOR

C5. Overall, has the vocational training program that you are currently attending met your expectations?

      1 ❑ YES
      2 ❑ NO

SECTION D: CONTACT INFORMATION

As we mentioned earlier, we would like to conduct another round of this survey in the future to learn about your experiences after graduation. For this reason, we would like to collect your contact information.

D1. What is your mobile phone number?

      |__|__|__|__|__|__|__|__|__|

D2. If we cannot reach you on that number, is there a second number we can use?

      |__|__|__|__|__|__|__|__|__|

D3. What is your e-mail address?

      _____

D4. What name do you use on Facebook? Please give the name in English or Georgian, exactly as used on Facebook.

      _____

D5. What is the first name and surname of a family member or a close friend who would be able to locate you if your contact information changes?

      _____

D6. What is this person's relationship to you?
    1 ❏ MOTHER
    2 ❏ FATHER
    3 ❏ SIBLING
    4 ❏ SPOUSE
    5 ❏ OTHER RELATIVE
    6 ❏ FRIEND

D7. What is the best contact phone number for this person?
    |__|__|__|__|__|__|__|__|

D8. What is this person's e-mail address, if you know it?

    _____

 D9. What is the first name and surname of an additional family member or a close friend who would be able to locate you if your contact information changes?

    _____

D10. What is this person's relationship to you?
    1 ❏ MOTHER
    2 ❏ FATHER
    3 ❏ SIBLING
    4 ❏ SPOUSE
    5 ❏ OTHER RELATIVE
    6 ❏ FRIEND

D11. What is the best contact phone number for this person?
    |__|__|__|__|__|__|__|__|

D12. What is this person's e-mail address, if you know it?

    _____

D13. What is your permanent home address?

    _____

    _____

    _____

Thank you again for your time! If you have any questions about this survey, please contact Dr. Natia Gorgadze at *[Local Phone Number]*.

SECTION E: ADMINISTRATIVE DATA (to be completed by the survey team from training provider records)

E1. National ID: |__|__|__|__|__|__|__|__|__|

E2. NAEC exam score (if available): |__|__| __|

Improving public well-being by conducting high quality, objective research and data collection

MATHEMATICA
Policy Research