MATHEMATICA Policy Research

REPORT

Evaluation of the Vocational Training Grant Fund in Namibia: Baseline Report

August 28, 2015

Evan Borkum
Arif Mamun
Linus Marco
Malik Khan Mubeen

Submitted to:
Millennium Challenge Corporation
875 15th Street, NW
Washington, DC 20005
Project Officers: Algerlynn Gill and Emily Travis
Contract Number: MCC-10-0114-CON-20 (MCC-13-TO-0001)

Submitted by:
Mathematica Policy Research
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Arif Mamun
Reference Number: 40233.241

This page has been left blank for double-sided copying.

ACKNOWLEDGMENTS

We greatly appreciate the support of the many people whose efforts contributed to this report. We especially thank our monitoring and evaluation colleagues for their input and support: Algerlynn Gill and Emily Travis of the Millennium Challenge Corporation (MCC), and Kandi Shejavali and Kofi Owusu-Tieku of the Millennium Challenge Account-Moldova (MCA-Namibia). We also appreciate the insights of Martin Wilkinson at MCA-Namibia, and Mao Tjiroze at the MCC Resident County Mission in Namibia.

NORC at the University of Chicago was responsible for managing the initial data collection effort; we thank Sam Haddaway, Kareem Kysia, and Aparna Ramakrishnan at NORC for their hard work and cooperation. In Namibia, we are grateful to Survey Warehouse (SW), who collected the baseline data under contract to NORC. We especially thank Franko Venter, Lizl Stoman, and Piet Stoman at SW for leading the data collection effort on the ground. We also thank the SW staff who conducted the interviews for their perseverance and patience throughout the process.

Our colleagues at Mathematica also provided valuable support for the baseline data collection and analysis. We are especially grateful to Kristen Velyvis, who provided input on and oversight of data collection activities. Matthew Andersen and Elena Moroz also supported the data collection effort, and Anne Hower, Divya Parthasarathy, and Kathy Shepperson worked diligently to program and test the baseline survey instrument. Anu Rangarajan carefully reviewed the draft report and provided excellent comments that greatly improved it. We also thank Felita Buckner for formatting the report, and Molly and Jim Cameron for their careful editorial work.

This page has been left blank for double-sided copying.

ACRONYMS

COSDEC	Community Skills and Development Center
ITT	Intent to treat
LCDRS	Levy collection, distribution, and reporting system
MCC	Millennium Challenge Corporation
MCA-N	Millennium Challenge Account Namibia
MDI	Minimum detectable impact
MRC	Multidisciplinary Research Center
NHIES	Namibia Household Income and Expenditure Survey
NQA	Namibia Qualifications Authority
NTA	Namibia Training Authority
NTF	National Training Fund
RPL	Recognition of prior learning
SLA	Service-level agreement
ToT	Treatment-on-the-treated
TP	Training provider
VTGF	Vocational Training Grant Fund

This page has been left blank for double-sided copying.

CONTENTS

TABLES

FIGURES

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

To promote economic growth and reduce poverty in Namibia, the Millennium Challenge Corporation (MCC) signed a $304.5 million compact with the Government of the Republic of Namibia in 2009. The compact, which was formally completed in September 2014, included three projects: tourism, agriculture, and education. The education project sought to address the shortage of skilled workers in Namibia and limitations in the education system's capacity to create a skilled workforce. The shortage of skilled workers in Namibia manifests in a high national unemployment rate, especially among women and the youth. The overall unemployment rate in Namibia was estimated at 34 percent in the 2009–2010 Namibia Household Income and Expenditure Survey (NHIES), while it was 39 percent for women and 54 percent for 20 to 24 year olds (Namibia Statistics Agency, 2012). One of the key activities under the education project was the vocational training activity, which focused on expanding the availability, quality, and relevance of vocational education and skills training in Namibia.

The vocational training activity consisted of three subactivities: (1) grants for high-priority vocational skills programs offered by public and private training providers through the Vocational Training Grant Fund (VTGF); (2) technical assistance to establish a National Training Fund (NTF), intended to provide a sustainable source of funding for vocational training programs in Namibia; and (3) improvement and expansion of Namibia's network of Community Skills and Development Centers (COSDECs), which provide vocational training targeting marginalized populations—primarily out-of-school youth but also including low-skilled adults. MCC contracted with Mathematica Policy Research to conduct an evaluation of the vocational training activity covering all three subactivities.

This report focuses on the evaluation of the main component of the VTGF subactivity, which involved awarding grants to training providers to conduct trainings in specific high-priority skills areas. Training providers that received these grants used them to award scholarships to eligible applicants. The first VTGF grants were awarded in the fourth quarter of 2010; the last grant was awarded in the third quarter of 2014. Other components of the VTGF subactivity, not covered in this report, included additional capacity-building grants for training providers and pilots of two initiatives that will be fully implemented under the NTF.

The evaluation design for the VTGF subactivity involves a rigorous impact evaluation complemented by a qualitative implementation analysis. The impact evaluation uses a random assignment design to determine the effects of the VTGF-funded scholarships on recipients' training and labor market outcomes. This report presents the findings from quantitative baseline data collected to support the impact evaluation.

A. Research questions

The VTGF subactivity evaluation will address 11 key research questions (Table ES.1); the implementation analysis will address some of these questions, and the impact analysis will address others. The analysis of baseline quantitative data in this report primarily seeks to support the impact analysis.

Table ES.1. Research questions addressed by the VTGF evaluation

Analysis type	Research questions
Implementation analysis	1. Was the VTGF subactivity implemented as planned? 2. How were the VTGF grants managed? 3. What were beneficiaries' perceptions of the VTGF grants? 4. How did employers hire VTGF graduates, and what were their perceptions of the graduates? 5. Were the recognition of prior learning (RPL) and employer-provided training pilots implemented as planned? How did employers' perceptions of and attitudes toward their RPL-certified employees change after they became certified? How did the perceptions of employees about their job security and mobility change?
Impact analysis	6. To what extent did those offered the opportunity of training through the VTGF receive more training relative to nonfunded qualified applicants? 7. To what extent did the VTGF improve employment outcomes for VTGF-funded trainees relative to nonfunded qualified applicants? 8. To what extent did VTGF-funded trainees have higher earnings and income relative to nonfunded qualified applicants? 9. To what extent did increased earnings result from increased wages while employed versus increased employment? 10. Did the effects of the VTGF-funded training vary by trainee characteristics? 11 What key characteristics or practices of training providers were associated with stronger impacts on employment and earnings?

VTGF = Vocational Training Grant Fund, RPL = Recognition of Prior Learning (piloted under the VTGF).

B. Impact evaluation design and role of the baseline survey

The impact evaluation uses a random assignment design to answer the relevant research questions (Table ES.1). Under this design, eligible applicants to each VTGF-funded training in which the number of applications exceeded the number of available slots were randomly assigned by the training provider either to a group that was offered a VTGF scholarship (treatment group) or one that was not (control group). Training providers conducted random assignment of applicants separately for each VTGF-funded training after that training was funded. Random assignment therefore occurred on a rolling basis between the fourth quarter of 2010 and the third quarter of 2014, as MCA-N and then the Namibia Training Authority (NTA) funded additional trainings.

The treatment and control groups for each training are expected to be equivalent, on average, except for the offer of VTGF funding. Therefore, differences in the outcomes of the treatment and control groups at some time after the training period (which, in this evaluation, will be one year after the end of training) can be attributed to the impact of the VTGF funding. Our analysis approach will enable us to combine the estimates from all the trainings included in the evaluation to estimate the overall impacts of the VTGF funding on trainee outcomes. The impact evaluation will be accompanied by an implementation analysis, which will use largely qualitative data to explore the VTGF's implementation (including all subactivity components) and help interpret the quantitative impact estimates.

In a typical random assignment evaluation, the analysis of baseline data serves two main purposes: (1) to estimate baseline levels of key characteristics and indicators for the analysis sample, providing important context for the evaluation and a benchmark against which the ultimate impacts can be assessed; and (2) to enable us to assess the degree of similarity between

the treatment and control groups on these baseline characteristics, thus confirming the validity of our random assignment design.

However, an important challenge in conducting the VTGF baseline analysis was a delay in the timing of the baseline survey relative to the start of training. Ideally, the baseline survey would have been conducted close to the random assignment date and before the start of training; in practice, it almost always was conducted after training had started (on average, about 3.4 months after training had started). Several factors led to these delays, including a delay in launching the baseline survey for the first few trainings, the typically short time between random assignment and the start of training, the grouping of trainings with different start dates in the same survey release, and the need for multiple attempts over an extended period of time to contact some respondents. The delay in the baseline survey until after the start of training could affect the interpretation of findings from the baseline data because indicators related to applicants' training, employment, earning, and income may have been affected by the VTGF training. Thus, the estimated baseline levels and treatment-control differences for these indicators may not represent the true baseline situation, limiting our ability to provide baseline context and to assess baseline similarity between the treatment and control group members.

Nevertheless, the baseline survey data provide information on several demographic characteristics that are unlikely to have been affected by the timing of baseline survey. We were able to use the baseline data on these applicant characteristics to provide useful context for the evaluation and assess the similarity between the treatment and control group members.

C. Data collection

The Millennium Challenge Account Namibia (MCA-N), and NORC at the University of Chicago collected the baseline data analyzed in this report through a survey of applicants to each VTGF training who had been randomly assigned to a treatment or control group. Applicants were surveyed through a computer-assisted telephone interview system; they provided information on their demographic characteristics, vocational training history, employment status, earnings, household income, and related indicators.

The baseline data were collected between December 2011 and July 2014. The long fielding period for the baseline survey reflects the fact that MCA-N and then the NTA awarded the VTGF grants (and training providers conducted random assignment) at several points throughout the compact period, and the baseline survey for applicants to each training was expected to be conducted soon after random assignment. MCA-N collected baseline data for the initial cohorts between December 2011 and August 2012. NORC (in partnership with Survey Warehouse, a local data collection firm) took over the data collection for subsequent cohorts in February 2013, and continued to collect baseline data after Mathematica joined the evaluation in mid-2013.

The baseline analysis sample consisted of 1,406 applicants in the treatment and control groups from 28 VTGF trainings, conducted by 11 training providers. The trainings ranged from one to 21 months in duration, with a median duration of 8 months. The analysis sample reflects an overall baseline survey response rate of 74 percent (78 percent in the treatment group and 71 percent in the control group).

D. Summary of findings

Our analysis of the demographic characteristics of the sample suggested that the baseline treatment and control samples were very similar (Table ES.2). Specifically, the differences between the treatment and control groups in gender, age, marital status, level of education, household size, and parental education were small and statistically insignificant. These results increase our confidence that there were no underlying differences between the two groups.

In contrast, there were large and statistically significant differences in the baseline indicators most closely related to the follow-up outcomes of interest in all of the domains we will analyze at follow-up (Table ES.3). In particular, the treatment group was significantly more likely to have been enrolled in vocational training in the five years before the baseline survey, significantly less likely to be employed at the time of the survey, and had significantly lower average earnings.

However, these differences likely do not reflect underlying treatment-control differences that could threaten the validity of the design, but rather reflect a delay in the timing of the baseline survey relative to training. Specifically, because in most cases the survey was conducted several months after the start of VTGF training, many baseline indicators in the treatment group would have been affected by enrollment in VTGF training. For example, because most treatment applicants would have already started training by the time they were surveyed, indicators of training (for example, whether they have ever received vocational training) are likely to be higher than in the control group. Similarly, if individuals in the treatment group gave up a job to

Table ES.2. Demographic characteristics of the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Female	741	665	64.3	63.7	0.5	0.827
Age, mean (years)	741	664	26.8	26.7	0.1	0.674
Unmarried	739	664	93.1	92.7	0.4	0.802
Respondent's education						0.795[a]
Less than grade 10	740	664	3.6	4.7	-1.0	
Completed grade 10	740	664	29.5	29.7	-0.2	
Completed grade 12	740	664	62.9	61.1	1.8	
Higher	740	664	4.0	4.5	-0.5	
Household size, mean (number)	739	665	5.5	5.8	-0.3	0.193
Mother's education						0.818[a]
Less than grade 10	552	544	62.7	61.0	1.8	
Completed grade 10	552	544	18.3	20.6	-2.3	
Completed grade 12	552	544	13.7	13.9	-0.2	
Higher than grade 12	552	544	5.3	4.6	0.7	
Father's education						0.873[a]
Less than grade 10	449	473	53.2	51.4	1.7	
Completed grade 10	449	473	17.4	17.9	-0.5	
Completed grade 12	449	473	21.1	23.1	-2.1	
Higher	449	473	8.3	7.5	0.8	

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.

[a] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

Table ES.3. Key indicators of the analysis sample, by domain (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Training receipt						
Any vocational training in previous 5 years	740	665	50.1	28.0	22.1	0.000***
Completed vocational training in previous 5 years	653	643	17.7	15.6	2.0	0.278
Enrolled in vocational training at the time of the survey	740	664	41.0	13.8	27.3	0.000***
Employment and productive engagement						
Employed in the past 7 days [b]	740	665	36.8	46.1	-9.2	0.002***
Engaged in any productive activity [c]	740	665	63.8	53.4	10.5	0.000***
Earnings and income						
Monthly gross earnings from wages, mean ($N) [d]	706	643	649	894	-245	0.001***
Monthly gross household income, mean ($N) [d]	548	530	2,302	2,310	-8	0.938

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.
[a] p-value from a test of joint significance across all categories using seemingly unrelated regressions.
[b] Worked at least one hour for pay, profit, or family gain in the previous seven days, or had a job to which to return.
[c] Employed in past seven days or enrolled in vocational training at the time of the survey.
[d] Computed using the midpoint of each category in the survey (which are more detailed than the categories presented here) and the lower bound for the highest category.

enroll in training, their employment and earnings would be lower than those in the control group. Thus, the VTGF baseline survey may not have captured a true baseline measure of these indicators.

Overall, the implications of the observed treatment-control differences are that the impact evaluation design still is likely to be valid. Baseline demographic characteristics, which were unaffected by VTGF trainings at baseline, were very similar in the treatment and control groups, increasing our confidence that random assignment was successful. Differences between the treatment and control groups in baseline indicators related to outcomes were large and statistically significant, but likely capture the effects of training at the time of the baseline interview (conducted after the start of the VTGF-training for most cases) and so do not threaten the validity of the design. These findings also suggest that we will have to be cautious about which baseline indicators to use as control variables in the follow-up impact analysis. These controls will primarily include demographic characteristics; controlling for the other baseline indicators could result in biased impact estimates.

E. Summary of internal validity risks

The main risks to the interval validity of our random assignment design are (1) underlying differences between the treatment and control groups that could affect follow-up outcomes, and (2) noncompliance with randomly assigned treatment status. In this setting, noncompliance is

most likely to result from treatment applicants who do not take up the offer of VTGF-funded training. The available information from VTGF implementation suggests that this situation may not be widespread; in any case, we will account for noncompliance by producing treatment-on-treated (ToT) estimates. As to underlying treatment-control differences, our baseline analysis does not suggest any such differences in demographic characteristics. There are many large and statistically significant differences in key baseline indicators; however, as described above, we believe that these differences are related to the fact that in most cases the baseline survey was conducted after training had started. Therefore, we do not believe that they threaten the validity of the design.

F. Plans for future data collection and reporting

The follow-up data collection for the VTGF impact evaluation is currently underway. For the follow-up survey, we will survey applicants to each training approximately one year after the end of training; we will therefore release trainings for follow-up based on the training end date (trainings with similar end dates will be released together). Between March and August 2014, NORC managed the collection of follow-up data for the applicants to the first set of trainings that were due for follow-up and worked with Survey Warehouse to conduct the survey. Mathematica took over management of the follow-up data collection starting in February 2015, when the next cohort was due for follow-up. Mathematica also is working with Survey Warehouse to implement the follow-up survey and will manage the data collection for all remaining sample members. We expect to complete the follow-up survey data collection in November 2015. We will analyze the follow-up data to produce a final VTGF evaluation report, a draft of which is expected to be submitted to MCC by the second quarter of 2016 and finalized by the end of 2016. The report will draw on both the quantitative follow-up data and the qualitative data related to implementation to answer the key research questions related to the VTGF subactivity comprehensively.

I. INTRODUCTION

To promote economic growth and reduce poverty in Namibia, the Millennium Challenge Corporation (MCC) signed a $304.5 million compact with the Government of the Republic of Namibia in 2009. The compact, which was formally completed in September 2014, included three projects: tourism, agriculture, and education. The education project sought to address the shortage of skilled workers in Namibia and limitations in the education system's capacity to create a skilled workforce, cited as some of the most serious constraints to Namibia's economic diversification and achievement of broad-based economic growth (U.S. Agency for International Development 2003; World Bank 2013). The shortage of skilled workers in Namibia manifests in a high national unemployment rate, especially among women and the youth. The overall unemployment rate in Namibia was estimated at 34 percent in the 2009–2010 Namibia Household Income and Expenditure Survey (NHIES), while it was 39 percent for women and 54 percent for 20 to 24 year olds (Namibia Statistics Agency, 2012). The education project consisted of several activities that aimed to improve the quality of Namibia's workforce by enhancing the equity and effectiveness of basic, vocational, and tertiary education.

The vocational training activity was one of the key activities under the education project. It focused on expanding the availability, quality, and relevance of vocational education and skills training in Namibia, and consisted of three subactivities: (1) grants for high-priority vocational skills programs offered by public and private training providers through the Vocational Training Grant Fund (VTGF); (2) technical assistance to establish a National Training Fund (NTF), intended to provide a sustainable source of funding for vocational training programs in Namibia; and (3) improvement and expansion of Namibia's network of Community Skills and Development Centers (COSDECs), which provide vocational training targeting marginalized populations—primarily out-of-school youth but also including low-skilled adults. MCC contracted with Mathematica Policy Research to conduct an evaluation of the vocational training activity covering all three subactivities.

The evaluation design for the main component of the VTGF subactivity—awarding funding to training providers to provide scholarships for vocational training—involves a rigorous impact evaluation complemented by a qualitative implementation analysis. The impact evaluation uses a random assignment design to determine the effects of the VTGF-funded scholarships on recipients' training and labor market outcomes. The impact evaluation will be accompanied by an implementation analysis, which will use largely qualitative data to explore the VTGF subactivity's implementation and help interpret the quantitative impact estimates. This report presents the findings from quantitative baseline data collected to support the impact evaluation.

In the rest of this chapter, we begin with a review of the findings from previous research on the impacts of vocational training programs in developing countries to situate the VTGF evaluation in Namibia in the research literature. We then describe the VTGF subactivity and its program logic in further detail, and provide a roadmap for the rest of the report.

A. Literature review

Although a large body of literature documents rigorous evidence on the impacts of vocational training programs in developed countries,[1] the evidence for developing countries is much more limited. A recent review of impact evaluations of vocational training programs targeted at youth in lower- and middle-income countries (Tripney et al. 2013) identified 26 studies that used an experimental or quasi-experimental design to estimate impacts on labor market outcomes (very few of the cited studies used an experimental design, which provides the highest standard of evidence). On average, these studies found positive impacts on outcomes such as paid employment and earnings. However, there was substantial variation in impacts across studies, and the average impact on paid employment was much lower when the review considered only higher quality studies. Given the variation in the quality of the quasi-experimental studies and the estimated impacts across studies, the authors caution that it is difficult to draw strong inferences about the impacts on vocational training programs more generally from the available literature.[2]

The handful of experimental evaluations of vocational training programs in developing countries conducted to date have found mixed results. These include the following:[3]

- Card et al. (2011) conducted an experimental evaluation of a subsidized training program for low-income, out-of-school youth in the Dominican Republic. The authors found no statistically significant impacts on employment approximately a year after graduation, but marginally significant and positive impacts of about 10 percent on wages among those employed.
- In contrast, Attanasio et al. (2011) found more positive results from an experimental evaluation of a similar training program aimed at disadvantaged youth in Colombia, with positive impacts of about 7 percent on employment and almost 20 percent on wages for female trainees approximately a year after the end of the program, but no significant impacts for men.
- Hirshleifer et al. (forthcoming) conducted an experimental evaluation of a large-scale vocational training program in Turkey, which provided unemployed individuals with three months of training offered through a range of private and public providers. The evaluation

[1] See Card et al. (2010) for a meta-analysis of training programs and other active labor market programs in the United States and Europe. Specific examples of large randomized evaluations of vocational training programs in the United States include the Job Training Partnership Act study (Bloom et al. 1997) and Job Corps (Schochet et al. 2008).

[2] Differences in methods and data across studies have sometimes even led to widely varying results for the same program. For example, Ibarrarán and Rosas Shady (2009) noted that seven evaluations of the same training program in Peru using data from different cohorts produced a wide range of estimated impacts. Similarly, Delajara et al. (2006) reported a wide range of estimated program impacts for a training program in Mexico, which they attributed to differences in the evaluation methodology.

[3] Additional experimental evaluations that have not yet released estimates of impacts on labor market outcomes include Hicks et al. (2011) on a voucher program for out-of-school youth in Kenya, and Cho et al. (2013) on a three-month apprenticeship program for vulnerable youth in Malawi. We will continue to monitor the evolving impact evaluation literature on vocational training programs so we can benchmark our final evaluation results against those from other studies.

found no statistically significant impacts on employment or labor income one year after training; even impacts on outcomes that seemed positive and significant after one year (such as measures of employment quality) had dissipated after three years, based on administrative data.

- Maitra and Mani (2014) conducted an experimental evaluation of a 6-month vocational training program in stitching and tailoring for unemployed women in India. Six months after training, program participants were significantly more likely to be employed (6 percentage points), work additional hours (2.5 hours per week), and earn more (150 percent). These short-run impact estimates were all sustained in a second follow-up conducted 18 months after training.
- Blattman et al. (2014) evaluated the impact of providing cash grants to groups of poor unemployed youth in rural Uganda to help them become self-employed artisans. Grant recipients invested in both vocational training provided by local artisans or small local training institutes and tools and materials to start their own businesses. After four years, compared to youth in the control group, the grant recipient youth were twice as likely to be engaged in a skilled trade as nonrecipients, and had substantially higher earnings (38 percent) and work hours (17 percent).
- Alcid (2014) experimentally evaluated a program that provided youth in rural Rwanda with training related to broad work readiness skills and more specialized technical skills (including vocational training), as well as internship opportunities. Six months after the program was completed, youth in the treatment group had significantly higher work readiness skills and were 12 percentage points more likely to be employed than those in the control group.

Overall, the existing literature on evaluations of vocational training programs in developing countries has important gaps, especially with regard to rigorous evaluations. Few high quality impact evaluations of these programs exist, and experimental evidence is especially limited, particularly in settings that are similar to the Namibian setting. Given the limited available evidence and substantial variation in impacts found in available studies across developing regions and countries (possibly due to differences in social, economic, and labor market conditions, existing skill levels of targeted groups, and training program characteristics), any additional rigorous evidence would be extremely valuable. In addition, few of the existing studies have integrated impact evaluation findings with a well-designed implementation analysis to understand the mechanisms behind estimated impacts; this is particularly relevant, given the substantial variation in impacts in the literature.

The impact evaluation of the VTGF subactivity in Namibia will make an important contribution to addressing some of these gaps and provide useful information for Namibian policymakers. In particular, it will provide rigorous evidence on the impact of funding for vocational training in Namibia through an experimental design, complemented by a strong implementation analysis to help interpret the findings. The evaluation findings will be especially valuable for policymakers in Namibia and Sub-Saharan Africa, given the lack of rigorous evidence on vocational training impacts in the region.

B. The VTGF subactivity

The VTGF subactivity was designed to provide funding for vocational skills programs in high-priority areas while the NTF was being set up. It was also intended to serve as a pilot for future vocational training funding under the NTF, which will involve a broader system-wide reform of the vocational training sector but has many features similar to those of the VTGF. The VTGF subactivity included several components; the impact evaluation discussed in this report focuses on the key component of awarding grants to training providers. Under this component, the VTGF solicited grant applications for conducting trainings in specific high-priority skills areas. Training providers receiving those VTGF grants used them to award scholarships to eligible disadvantaged applicants.[4] The scholarships, which covered tuition and included a subsistence allowance, were intended to increase access to training for these applicants. Providers that were awarded VTGF grants could also apply for an additional capacity-building grant, which they could use for a variety of purposes related to increasing their capacity (such as purchasing new tools and equipment or improving or expanding their infrastructure).

The first grants were awarded in the fourth quarter of 2010, and the last grant was awarded in the third quarter of 2014. A total of 14 training providers received VTGF grants, with some of these receiving more than one grant (for different intakes of trainees).[5] The NTA—the same government body that will oversee the NTF—managed most of the grants. (The MCA-N managed the initial grants.)

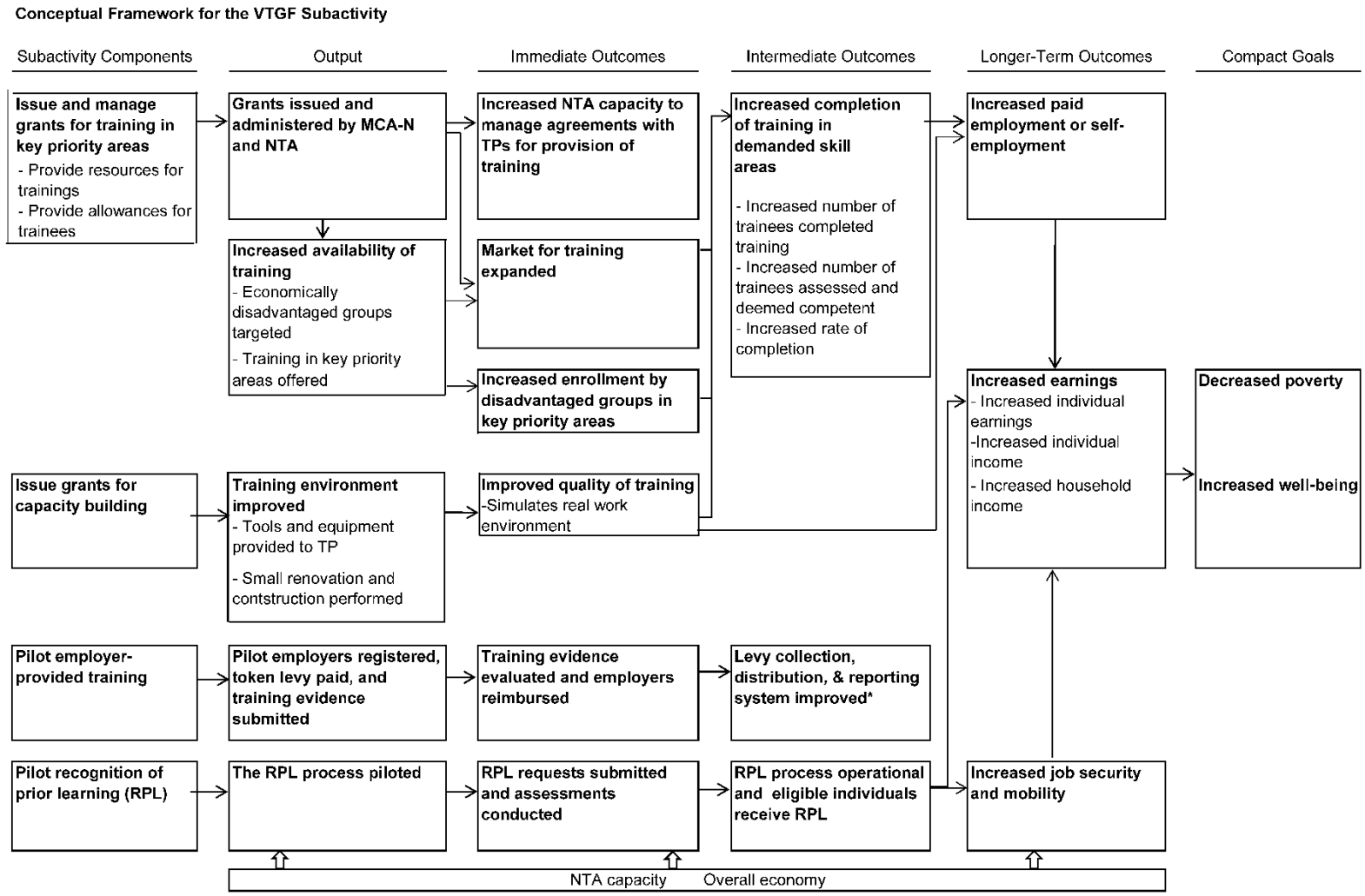
The remaining components of the VTGF were pilots of two other initiatives that will be fully implemented under the NTF. The first was piloting the reimbursement of employers for the costs of employer-sponsored training under the NTF's levy collection, distribution, and reporting system (LCDRS), in which employers register, pay a (token) levy, and submit training evidence to get reimbursed. The second was the recognition of prior learning (RPL) program, which helps people with experience in a certain vocational skills area but without formal training to compile a portfolio of evidence of their work experience and have their skills formally assessed and certified. We are evaluating these two pilots through the largely qualitative implementation analysis and do not cover them in this report; a recent evaluation report presents findings from the analysis of the first round of qualitative data for these two pilots (Mamun et al. 2015).

In Figure I.1, we provide a logic model that illustrates how the components of the VTGF are expected to contribute to the ultimate compact goals of decreased poverty and increased economic well-being. The left-hand column lists the components of the subactivity and the second column shows the direct outputs of these components. The key outputs included the NTA's administration of grants and increased availability of training for the disadvantaged (resulting from the VTGF grants), improved equipment and infrastructure (resulting from the capacity-building grants), and implementation of the RPL and employer-provided training pilots.

[4] "Disadvantaged" was defined as having an annual household income of less than N$250,000 (about US$27,000 at the average exchange rate over the baseline survey period) after subtracting training costs for other household members who might be participating in training at the time.

[5] As mentioned earlier, not all of these grants are included in the evaluation; in Chapter II, we list those included.

Figure I.1. VTGF logic model



Assumptions:

A1. Training providers are on the path to NTA registration and NQA accreditation.
A2. RPL certificate is the same and valued in the same way as a traditional vocational training certificate.
A3. Training is of sufficient quality.
* Detailed levy development and processes are reflected in the NTF conceptual model (Mamun et al. 2014).
NQA = Namibia Qualifications Authority; NTA = National Training Authority; SLA = service-level agreement; TP = training provider.

The third, fourth, and fifth columns in the logic model show the immediate, intermediate, and longer-term outcomes, respectively, of the investments under the VTGF subactivity. In the immediate term, the NTA's capacity to manage service-level agreements, under which training providers commit to certain milestones and disburse funds based on achieving these milestones, was expected to increase through their experience in managing the VTGF grants. The grants themselves were expected to increase the quality of training through the investments in tools and infrastructure using the capacity-building grant for those training providers who received one, increase enrollment of disadvantaged groups targeted by the grants, and expand the market for training through the competitive bidding process for grant funds. The employer-provided training and RPL pilots also were expected to be conducted (culminating in reimbursement of employers and assessment of candidates, respectively), and the lessons learned synthesized. The sixth and final column in the logic model shows the compact goals to which the subactivity was ultimately expected to contribute, namely decreased poverty and increased wellbeing.

To examine the VTGF subactivity's implementation and gauge early perceptions of its likely effects, we analyzed qualitative data collected from VTGF stakeholders toward the end of the compact (Mamun et al. 2015). These stakeholders included MCA-N, NTA, training providers, VTGF-funded trainees and control group members, and potential employers of VTGF-funded trainees. Focusing on the training grant component—the one covered in this report—we found the grants largely were implemented as planned. Through the VTGF experience, the NTA gained valuable experience in managing grants. Specifically, it was able to improve its understanding of the costs of training, how to compare costs across providers and skill areas, and how to manage grants by setting and monitoring the progress of training providers toward concrete milestones. However, the process for awarding the grants was less competitive than envisaged, due to the limited number of training providers with the capacity to offer courses in the fields and levels targeted for training or expand their capacity to accommodate additional courses or trainees. Thus, rather than selecting grant recipients from a large number of applicants, the implementers had to actively seek out providers for participation. In addition, the process for determining market demand was not as scientific as planned, which posed a possible risk to the employment prospects of trainees if the skill areas in which they trained were not in high demand.

Nevertheless, stakeholder perceptions of the likely impacts of VTGF-funded training generally were positive. Focus group discussions with VTGF trainees who still were engaged in training revealed that they felt the quality of the training they were receiving was high and they would be able to translate that experience into a positive labor market experience. They also said that a board and lodging allowance for all trainees, introduced during implementation, was important in increasing attendance and reducing dropouts—thus increasing the likelihood that trainees would graduate successfully from their training programs. The employers we interviewed had a very positive view of the training providers that received VTGF grants, which ultimately could be reflected in these employers hiring VTGF-funded trainees. Employers and training providers noted existing relationships through which they partner for job attachments or even direct hire of trainees, although it is unlikely that all VTGF-funded trainees will be accommodated through these relationships. Although the successful implementation of the VTGF training grants and positive perceptions of the potential labor market prospects of trainees are both encouraging, the VTGF impact evaluation will enable us to provide more rigorous evidence of whether the expected impacts on trainees materialized.

C. Roadmap for the report

The balance of this report presents the baseline analysis for the VTGF impact evaluation and is structured as follows. In Chapter II, we review the key research questions for the evaluation and evaluation design the baseline data were intended to inform, and describe the baseline data, sample, and analysis approach. In Chapter III, we present the findings from the baseline analysis. We conclude with Chapter IV, in which we summarize the implications of the baseline findings for the evaluation; assess the challenges to internal and external validity; and describe our future plans for data collection, analysis, and reporting.

This page has been left blank for double-sided copying.

II. EVALUATION DESIGN AND BASELINE ANALYSIS

In this chapter, we review the evaluation design for the VTGF evaluation and describe the baseline analysis we conducted. We begin by listing the key research questions for the evaluation and providing an overview of the type of evaluation we are implementing. We then describe the baseline data analyzed in this report—including the data collection process, key indicators, and sample—and discuss our baseline analysis approach.

A. Research questions

The evaluation of the VTGF subactivity will address 11 key research questions, which we have grouped into the following two sets (see Mamun et al. 2014, for the full list of subquestions associated with these research questions):

1. Implementation analysis

1. Was the VTGF subactivity implemented as planned?
2. How were the VTGF grants managed?
3. What were beneficiaries' perceptions of the VTGF grants?
4. How did employers hire VTGF graduates, and what were their perceptions of the graduates?
5. Were the RPL and employer-provided training pilots implemented as planned? How did employers' perceptions of and attitudes toward their RPL-certified employees change after they became certified? How did the perceptions of employees about their job security and mobility change?

2. Impact analysis

1. To what extent did those offered the opportunity of training through the VTGF receive more training relative to nonfunded qualified applicants?
2. To what extent did the VTGF improve employment outcomes for VTGF-funded trainees relative to nonfunded qualified applicants?
3. To what extent did VTGF-funded trainees have higher earnings and income relative to nonfunded qualified applicants?
4. To what extent did increased earnings result from increased wages while employed versus increased employment?
5. Did the effects of the VTGF-funded training vary by trainee characteristics?
6. What key characteristics or practices of training providers were associated with stronger impacts on employment and earnings?

To answer the key research questions, we are using an integrated mixed-methods approach that includes a qualitative implementation analysis and a quantitative impact evaluation (the evaluation design is described in further detail below). The qualitative implementation analysis will use insights obtained from trainees, control group members, MCA-N staff, NTA staff, training providers, and other stakeholders, enabling us to address questions 1 through 5. The

impact evaluation will complement the implementation analysis and enable us to provide quantitative estimates of impacts on the key outcomes of training, employment, and earnings and income (questions 6 through 9), and explore the variation in impacts (questions 10 and 11). The implementation analysis will provide important context for interpreting the impact estimates.

B. Evaluation design

As mentioned above, the VTGF subactivity evaluation includes a qualitative implementation analysis and a quantitative impact evaluation. This report focuses on the quantitative baseline data designed to inform the impact evaluation. However, because the implementation analysis will also ultimately inform the impact evaluation, we also briefly describe that component below. Further details on the evaluation design are available in the evaluation design report (Mamun et al. 2014).

1. Implementation analysis

To inform the implementation analysis, we collected qualitative data between October and November 2014 on the implementation of the VTGF subactivity. The Multidisciplinary Research Center (MRC) at the University of Namibia collected the data, and Mathematica provided oversight. Focus groups and interviews conducted with a range of stakeholders (Table II.1) gathered in-depth information on how the VTGF subactivity implementation occurred in practice, the experiences of VTGF trainees and training providers, the availability of non-VTGF training opportunities, the experiences of RPL participants, and other relevant topics. Mathematica prepared a report presenting findings from the analysis of the first round of qualitative data (Mamun et al. 2015). A second round of qualitative data, to be collected in October 2015, will focus primarily on other subactivities but may include some interviews with a new cohort of RPL participants to inform the research question related to RPL (question 5).

Table II.1. Qualitative data collected for the VTGF implementation analysis in October–November 2014

Data source	Type of data collection
VTGF trainees	Focus groups
VTGF control group members	Interviews
MCA-N staff	Interviews
NTA staff	Interviews
GOPA Consulting Group consultants[a]	Interviews
Training providers	Interviews
Employers	Interviews
RPL participants	Interviews

[a] Provided technical assistance to the NTA related to management of VTGF grants and the RPL program.

2. Impact evaluation

To estimate the impact of trainings funded by the VTGF subactivity on key trainee outcomes, we are using a random assignment design. Under this design, eligible applicants for each VTGF-funded training in which the number of applicants exceeded the number of available slots were randomly assigned by the training provider either to a group that was offered VTGF

funding (the treatment group) or one that was not (the control group).[6] Because the offer of funding was randomly assigned, the treatment and control groups for each training should be similar in all respects, on average, except that the treatment group received the offer of funding. Thus, the control group can be used to approximate the counterfactual: the average experience of the treatment group in the absence of the offer of funding. Any differences in outcomes that arise between the treatment and control group after random assignment thus can be attributed to the cumulative impact of the offer of VTGF funding up to that point. By combining the data for all trainings included in the evaluation, we will be able to estimate the overall impact of the offer of VTGF funding.

The random assignment procedure for the VTGF evaluation was conducted as follows. Applications for a particular VTGF-funded training were solicited, typically through advertisements in a national newspaper. Applicants had to satisfy the income criterion mentioned earlier (annual household income of under N$250,000 after subtracting training costs for other household members participating in training at the time), as well as other criteria specified by the training provider. The training provider then screened the applicants based on additional provider-specific criteria (for example, using grades in the final school-leaving examinations) to identify a final pool of applicants for random assignment. The training provider generally conducted random assignment using the random number generator in Microsoft Excel, with MCA-N and/or the NTA (for NTA-administered grants) providing oversight. Training providers conducted random assignment separately for each VTGF-funded training after that training had been funded and they had solicited applications from potential trainees. Random assignment therefore occurred on a rolling basis between the fourth quarter of 2010 and the third quarter of 2014, as MCA-N and then the Namibia Training Authority (NTA) funded additional trainings.

The impact evaluation will rely primarily on follow-up data on training and labor market outcomes collected from treatment and control group members approximately one year after the end of the relevant training. Because the timing of the trainings varies in terms of their start date and duration, the timing of the follow-up data collection (one year after the end of training) will also vary across trainings. Despite these differences in timing, our analysis approach will enable us to combine the data for all trainings and obtain a valid estimate for the impact of the *offer* of training (intent-to-treat, or ITT, estimates) by comparing the outcomes of the treatment and control groups. We will also be able to adjust the ITT estimates to determine the impact of training on those who actually enrolled in or completed it. These treatment-on-treated (ToT) estimates could differ from the ITT estimates if, for example, not all of those in the treatment group take up the offer of training.

C. Baseline survey data and analysis

In this section, we describe the baseline data and analysis we present in this report. We begin by describing the purpose of the VTGF baseline survey, and its timing and content. We then describe our baseline analysis, including the indicators on which we will focus, the sample sizes and response rates, and the analysis approach.

[6] Throughout this report, we use the term "training" to refer to one of these distinct random assignment groups (a single training provider thus may encompass multiple VTGF "trainings").

1. Purpose of the baseline survey

In a typical random assignment evaluation, the analysis of baseline data serves two main purposes: (1) to estimate baseline levels of key characteristics and indicators for the analysis sample, providing important context for the evaluation and a benchmark against which the ultimate impacts can be assessed; and (2) to enable us to assess the degree of similarity between the treatment and control groups on these baseline characteristics, thus confirming the validity of our random assignment design.

However, as we discuss in further detail below, an important challenge in using the VTGF baseline data for these purposes was that the baseline survey was conducted after the start of training for most applicants. This could affect the interpretation of findings from the baseline data because available baseline indicators related to applicants' training, employment, and earning and income—the key outcomes for the evaluation—may have been affected by the VTGF training. Thus, the estimated baseline levels and treatment-control differences for these indicators may not represent true baseline situation, limiting our ability to provide baseline context and to assess baseline similarity between the treatment and control group members. Further, this suggests that controlling for these baseline indicators in the impact analysis—which is typically done to control for baseline differences that may have arisen by chance and to improve the precision of the impact estimates—could result in biased impact estimates.

In view of this challenge, we refocus the baseline analysis on the following purposes: (1) estimating baseline levels of key characteristics and indicators for applicants in the control sample, which we argue is less likely to have been affected by the timing of the survey than the treatment sample and is therefore more representative of the true baseline situation, (2) verifying that demographic characteristics, which are unlikely to have been affected by the timing of the baseline survey, were similar in the treatment and control groups at baseline, and (3) for completeness, documenting the extent of baseline differences between the treatment and control groups in indicators related to follow-up outcomes, while recognizing that this is not a test of the validity of the random assignment design because of the delayed timing of the baseline survey.

2. Timing of baseline survey

The data for the VTGF baseline analysis are drawn from a survey of eligible applicants who were randomly assigned to VTGF trainings. MCA-N and NORC surveyed eligible applicants between December 2011 and July 2014 using a computer-assisted telephone interview system. This period corresponds roughly to that during which random assignment was conducted for various VTGF-funded trainings. Ideally, applicants for a given training would be surveyed after random assignment (at which point their telephone contact information was made available for survey purposes) but before the start of training, so that the baseline data would provide information on VTGF applicants before training.

However, in practice, the baseline survey almost always was conducted after training had started—often several months later (Table II.2). Specifically, 35 percent of the sample was surveyed more than 3 months after training had started, and about 20 percent was surveyed more than 6 months later (the mean was 3.4 months after the start of training, and the median was 2 months after the start of training). Less than 1 percent of the sample was surveyed on or before

the training start date. In some cases (about 8 percent of the treatment sample, not shown in the table), these delays meant that the baseline was conducted after training had been completed.

Several factors led to these delays in conducting the baseline survey relative to the start of training. First, a delay occurred in starting the baseline survey, so MCA-N conducted surveys for the first few trainings several months after the start of training (three trainings started between October 2010 and July 2011, but the baseline surveys only began in December 2011). Second, because of the typically short time between random assignment and the start of training, it would have been challenging to complete the baseline survey before the start of training. Third, for logistical reasons, applicants to different trainings were released for the baseline survey in batches. Some of the trainings grouped together in the same batch might have had earlier start dates, so there was a longer delay in conducting the survey for these trainings. Finally, if applicants could not be contacted initially, further attempts were made to contact them periodically as the baseline survey progressed. In some cases, successful contacts were made several months after the sample was released.

As mentioned above, the delay in the baseline survey until after the start of training could affect the interpretation of findings from the baseline data because available baseline indicators related to applicants' training, employment, and earning and income may have been affected by the VTGF training. For example, because most treatment applicants would have already started training by the time they were surveyed, indicators of training (for example, whether they have ever received vocational training) are likely to differ from the control group even if no difference in training existed between the two groups before random assignment. Similarly, if individuals in the treatment group gave up a job to enroll in training, their employment and earnings would be lower than those in the control group. Thus, observed differences might simply reflect the early effects of the training rather than underlying differences that could threaten the validity of the design. We will therefore consider the implications of the delayed baseline when we interpret the treatment and control differences we observe.

Table II.2. Timing of the VTGF baseline survey relative to the start of training for the VTGF baseline analysis sample (percentages, unless otherwise noted)

	Full sample
On or before training start date	0.9
Within 1 month of training start date	13.2
1 to 3 months after training start date	50.9
4 to 6 months after training start date	15.1
7 to 12 months after training start date	11.9
More than 12 months after training start date	7.9
Mean number of months between training start date and survey	3.4
Median number of months between training start date and survey	2.0
Sample size	**1,406**

3. VTGF baseline survey content

The VTGF baseline survey contained several sections (Table II.3). It collected data on basic demographic characteristics of the applicants, together with a range of outcome measures relevant to the research questions. These outcomes focused on the applicants' vocational training history, employment status, and earnings and income. It also gathered extensive contact information for applicants to facilitate their being contacted for the follow-up survey.

Some changes were made to the baseline data collection over time. MCA-N collected baseline data for the initial cohorts between December 2011 and August 2012. NORC (in partnership with Survey Warehouse, a local data collection firm) took over the data collection for subsequent cohorts in February 2013, making a handful of changes to the survey instrument. When Mathematica joined the evaluation in August 2013, we made a small number of further changes to the instrument for the remaining cohorts. The changes made to the instrument over time were relatively minor, and involved adjusting the wording of some questions, adding or removing some questions, and making some changes in question order and skip patterns. Despite these changes, the basic survey instrument and methodology remained similar over time, enabling us to combine data from different periods for the analysis.

4. Baseline indicators and relationship to follow-up outcomes

Typically, the baseline analysis would focus on demographic characteristics and other baseline indicators that have a strong relationship with the primary outcomes of interest for the follow-up impact analysis. If these characteristics and baseline indicators are similar in the treatment and control groups at baseline, we would be more confident that there are no underlying differences between the two groups that could bias the results of the impact analysis. These characteristics and indicators would also be the most useful control variables for the impact analysis.

We see two main challenges in identifying baseline indicators that are strongly correlated with follow-up trainee outcomes in the key domains for the evaluation—training receipt, employment and productive engagement, and earnings and income. First, as mentioned above, the baseline survey typically was conducted after training had started. Thus, some of these indicators were very likely to have been affected by the VTGF trainings, even at baseline, in turn affecting estimated treatment-control differences. Second, the baseline survey was designed before the evaluation design was finalized. Although the broad outlines of the design were known at the time, important details such as the research questions and key outcomes for the impact analysis were not. Because most the baseline data collection had been completed when Mathematica joined the evaluation and fully developed the evaluation design, there was little value in making major changes to the baseline survey at that time. Thus, the baseline indicators most ideally related to the ultimate outcomes of interest are not always available in the baseline data.

Table II.3. VTGF baseline survey sections

Section	Key topics covered
Identifying and contact information	Name; identification number; date of birth; region of residence; town and region of origin; telephone contact numbers; email address; postal address; contact information of friend or relative
Demographic information	Age; gender; marital status; nationality
Education (excluding vocational training)	Highest level of education; whether moved for education; desire for further education; challenges to further education
Vocational training	Enrollment in vocational training (in previous five years and as of survey date); total months of vocational training; sectors and skill areas of vocational training; job attachments; perceived quality of vocational training; dropout from vocational training; completion of vocational training (including sectors, skill areas, and institutions); accessibility of vocational training (*)
Employment and earnings	*Employment status:* whether currently employed; availability for employment; whether actively seeking employment *Among those employed:* number of jobs currently held; hours and days worked; type of employment (part-time, full-time, or self-employed); help received in finding employment; relevance of employment to training; whether employment is paid; job tenure; job satisfaction; size and sector (formal or informal) of workplace; source of information about job; occupation and sector of employment; monthly income from employment; number of dependents on earnings *Among those unemployed:* duration of unemployment; reasons for unemployment; whether previously employed (including satisfaction and reason for leaving); willingness to consider vocational training in the future (*)
Household demographics and income	Household size; ownership status of dwelling (*); monthly household income; main sources of household income; relationship of respondent to head of household; parental education; orphan status

(*) = Removed from the survey when Mathematica joined the evaluation in August 2013.

In Table II.4, we present the primary outcomes of interest for the follow-up impact analysis by domain (Mamun et al. 2014) and identify the ideal baseline indicators related to these outcomes. We then list the most closely related indicators available in the baseline survey and assess the implications of using these indicators rather than the ideal ones in the baseline analysis. The key findings from this exercise are as follows:

- In the **training receipt** domain, the primary follow-up outcomes for the impact analysis are enrollment in and completion of vocational training since the random assignment date. For the baseline analysis, enrollment in and completion of training *before* the random assignment date would be the ideal and most closely related indicators. However, these indicators are not available in the baseline survey. Instead, respondents were asked about their enrollment in and completion of vocational training in the previous five years. For those in the treatment group who had already started it, the VTGF training would be included in the five-year reference period, thus affecting reported baseline measures of training enrollment (the same applies to months of vocational training, an alternative measure of prior training). However, most of these respondents did not have time to *complete* the VTGF-funded training by the time they were surveyed at baseline; only about

8 percent of the treatment group members could have completed the training. For this reason, completion is likely to be a valid baseline measure for most sample members.

- In the **employment and productive engagement** domain, the impact analysis will focus primarily on whether the respondent was employed in the 12-month post-training period. Another primary outcome is whether the respondent was employed or engaged in further vocational training over this period. (We include vocational training in this measure of "productive engagement" because VTGF training might prompt recipients to undertake further training rather than enter employment.) Ideally, we would have similar measures for the pre-training period, such as whether a respondent was employed before random assignment (either over a 12-month period before random assignment, at any time before random assignment, or as of the random assignment date). However, the baseline survey focuses primarily on measuring the employment status of the respondent as of the survey date. In the treatment group, this status is likely to have been affected by having started VTGF training, because individuals participating in training would have been less available for employment.
- In the **earnings and income** domain, the primary follow-up outcome is average monthly earnings in a 12-month post-training period. Closely related secondary outcomes (not shown in Table II.4) are individual and household income in the month before the follow-up survey. The ideal baseline indicators would thus be measures of earnings and income over a given period before random assignment. However, in the baseline survey, respondents were asked only about their average current monthly earnings and current monthly household income. Again, both of these measures could plausibly be affected by treatment group members starting training—for example, if an individual gave up a job to enroll in training, his or her earnings and household income might have decreased.
- Almost all of the key baseline indicators are therefore likely to have been affected by the delay in the timing of the baseline, and may not represent a true baseline condition. We will therefore interpret any baseline treatment-control differences in these indicators with caution and will not use them as control variables in the follow-up impact analysis.
- In contrast, the demographic characteristics available in the VTGF baseline survey—such as respondent gender, age, marital status, level of education, household size, and parental education—are mostly independent of the intervention and likely to be correlated with follow-up trainee outcomes. Assessing baseline treatment-control differences in these characteristics is therefore a valid check of whether random assignment is successful, and we will also be able to use these directly as control variables in the impact analysis.

In Table II.5, we present the full set of characteristics and indicators from the baseline survey that we analyze in this report, arranged by domain (the available baseline indicators most closely related to the follow-up outcomes listed in the fourth column of Table II.4 are highlighted in bold). Although the baseline survey includes additional indicators besides those listed here, many of these are not directly relevant to the finalized evaluation design, which was not available when the survey was developed. For this reason, to ensure that our analysis is relevant for the evaluation, we have focused on the more concise set of indicators in Table II.5.

Table II.4. Follow-up outcomes and related baseline indicators, by domain

Domain	Primary follow-up outcomes for impact analysis	Ideal baseline indicators in domain	Available baseline indicators in domain	Implications
Training receipt	Enrollment in any vocational training since the random assignment date	Enrollment in any vocational training before the random assignment date	Enrollment in any vocational training in the 5 years before the survey date Months of vocational training in the 5 years before the survey date	Enrollment in any vocational training and months of training before the survey date could all be affected by VTGF training
	Completion of any vocational training since the random assignment date	Completion of any vocational training before the random assignment date	Completion of any vocational training in the 5 years before the survey date	Completion of training before the survey date could all be affected by VTGF training; however, completion of vocational training before the survey date was not affected by VTGF training for most cases.
Employment and productive engagement	Any paid job held in the 12-month post-training period	Any paid job held before the random assignment date (in a 12-month period, at any time, or at the random assignment date)	Employed as of the survey date	Employment as of the survey date could be affected by VTGF training
	Any paid job held or engaged in further education and training in the 12-month post-training period (productive engagement)	Any paid job held or engaged in vocational training before the random assignment date (productive engagement)	Employed or engaged in vocational training as of the survey date (productive engagement)	Employment and vocational training as of the survey date could be affected by VTGF training
Earnings and income	Average monthly earnings in the 12-month post-training period	Average monthly earnings before the random assignment date (in a 12-month period or at the random assignment date)	Average current monthly earnings	Current monthly earnings could be affected by VTGF training

Table II.5. Baseline characteristics and indicators for the VTGF baseline analysis, by domain

Domain	Baseline characteristics or indicators
Demographic characteristics	Gender; age; marital status; respondent's education; household size; parental education
Vocational training	Vocational training in the previous 5 years: **any training**, **months of training**, any job attachment, **any completed training**, any certificate received
	Enrollment in vocational training at survey date
Employment and productive engagement	Employment at time of survey: **employed**, **productively engaged**, employment status (employed/unemployed/not in labor force), type of employment, formal employment, number of jobs, hours worked per week
	Employment features for those employed: size of workplace, training relevance to job, job tenure, job satisfaction
	Unemployment features for those unemployed: ever employed, unemployment duration
Earnings and income	**Average monthly gross earnings**
	Sources of household income; average monthly gross household income

Note: Indicators most closely related to the follow-up outcomes (fourth column of Table II.4) are highlighted in bold.

5. Baseline analysis sample and response rates

The targeted sample for the VTGF evaluation consists of applicants to the 28 trainings listed in Table II.6. Eleven different providers conducted these trainings, some of which provided multiple trainings; the trainings ranged from one to 21 months in duration (the median duration was 8 months). The list of trainings in Table II.6 does not cover the full set of trainings funded by the VTGF activity. Specifically, it excludes 26 trainings for which there was no control group (typically because there were sufficient slots to accommodate all applicants),[7] 21 trainings for which the follow-up survey date (one year after the end of training) would fall outside of the evaluation period,[8] and 9 trainings for which there were severe violations of random assignment

[7] Random assignment was conducted for a handful of these trainings. However, after accounting for applicants who applied to multiple trainings (as described below, each applicant was linked to the first included training to which they applied for the purposes of the evaluation), these trainings were left with no control group.

[8] Of the 21 trainings that fall outside of the original evaluation period, 11 would be due for follow-up in late 2016, and 10 would be due for follow-up in mid-late 2017. Because MCC is now planning to extend the period of performance for the evaluation contract to November 2017, it might be possible to include the 11 trainings due in late 2016 in the evaluation (3 of these trainings have small control groups and might have to be dropped if there are no control responses, so there might only be 8 additional trainings). However, the improvement in minimum detectable impacts (MDIs) by including the trainings with a late-2016 follow-up would be relatively small—likely less than 10 percent of the current MDIs. For example, the MDIs for having completed vocational training or being employed would decrease by less than 1 percentage point. Further, adding this handful of trainings would not make the study sample substantively more representative of all VTGF trainings, and would also result in a delay of more than a year in the submission of the draft final evaluation report. We therefore recommend against including these trainings in the evaluation.

(the first three intakes of COSDEC Benguela). These 56 excluded trainings comprise about 45 percent of the total number of VTGF-funded trainees.

Table II.6. VTGF trainings included in the evaluation

Training provider	Course	Intake	Start date	Expected course duration (months)	Number of treatment applicants	Number of control applicants
NATH	Tour Guiding	1	4-Oct-10	21	50	33
Wolwedans	Hospitality & Tourism	1	11-Jan-11	5	31	3
Wolwedans	Hospitality & Tourism	2	11-Jul-11	7	35	25
Wolwedans	Hospitality & Tourism	3	7-Feb-12	9	39	11
ABTCC	Food & Beverage/ Housekeeping	1	4-Sep-12	1	15	16
ILSA	Reception Management & Office Administration	1	1-Oct-12	13	118	27
IUM[a]	Hospitality & Tourism	1	5-Jan-13	16	59	29
IUM[a]	Hospitality & Tourism	1	5-Jan-13	16	243	142
VVTC	Front Office	1	3-Jun-13	14	12	6
VVTC	Food Production	1	3-Jun-13	14	10	21
VVTC	Housekeeping & Food Preparation	1	3-Jun-13	14	13	7
VVTC	Food & Beverage Services	1	3-Jun-13	14	12	6
OVTC	Hospitality & Tourism	1	4-Mar-13	17	35	22
ZVTC	Plumbing	1	8-Jul-13	13	20	68
ZVTC	Hospitality & Tourism	1	8-Jul-13	13	20	168
ZVTC	Office Administration & Computing	1	8-Jul-13	13	16	212
ZVTC	Bricklaying	1	8-Jul-13	13	20	24
KAYEC	Carpentry	1	1-Oct-13	4	15	18
KAYEC	Shuttering	1	1-Oct-13	4	15	4
KAYEC	Concrete Work	1	1-Oct-13	4	15	16
KAYEC	Concrete Work	2	17-Mar-14	3	9	1
COSDEC Benguela	Office Administration & Computing	4	14-Apr-14	4	30	16
NamWater	Grader	2	19-May-14	3	10	4
NamWater	Bulldozer	2	19-May-14	3	10	2
NamWater	Forklift	2	19-May-14	3	20	5
KAYEC	Shuttering	3	25-Jun-14	2	30	25
KAYEC	Carpentry	3	25-Jun-14	2	30	22
KAYEC	Concrete Work	3	25-Jun-14	2	23	4
Total	--	--	--	--	**955**	**937**

Notes: Table excludes 26 trainings with no control group (2 NATH trainings, 2 ZVTC trainings, 5 KAYEC trainings, 10 RVTC trainings, 1 NamWater, 4 NAMCOL, and 2 COSDEC Benguela trainings); 21 trainings not covered by the evaluation period (4 NAMCOL trainings, 14 NIMT trainings, and 3 NamWater trainings); and 9 trainings with severe violations of random assignment (9 COSDEC Benguela trainings).

Number of treatment and control applicants corrects for multiple applications; applicants are linked to the first included training to which they applied.

[a] IUM hospitality and tourism trainings were conducted at two separate sites with separate random assignment; these are treated as separate trainings for evaluation purposes.

Strictly, the baseline analysis and ultimate impact estimates apply only to the applicants to the included trainings, not to all VTGF applicants. If the characteristics of applicants to the excluded trainings or the nature of these trainings differed from the included trainings, the evaluation findings might not generalize to the broader VTGF sample. For some of the excluded trainings, the difference in the nature of the training may have been manifested by the absence of a control group, which may indicate weaker levels of interest among potential applicants for these trainings. The difference in the nature of some other excluded trainings may be captured by the longer durations of the trainings as they extend beyond the evaluation period. However, most of the excluded trainings were conducted by training providers that had other VTGF trainings included in the evaluation (NIMT being an exception), so these excluded trainings may be similar to those that were included in terms of applicant characteristics and content. Taken together, we may want to be cautious about generalizing the impact estimates to the full VTGF sample.

There were 1,892 unique applicants to the included trainings, including 955 assigned to the treatment group and 937 assigned to the control group (Table II.7). Of these applicants, 55 (3 percent) applied to multiple trainings;[9] these applicants were linked to the first included training for purposes of the evaluation. Of the 1,892 unique applicants, 1,406 completed a baseline survey, resulting in a response rate of 74 percent (78 percent in the treatment group and 71 percent in the control group). These 1,406 baseline respondents constitute the analytic sample used for the VTGF baseline analysis.

Table II.7. VTGF baseline survey sample sizes and response rates

	Full sample	Treatment	Control
Number of unique applicants	1,892	955	937
Number of applicants who applied to multiple trainings[a]	55	18	37
Number of completed surveys	1,406	741	665
Response rate (percent) [b]	74.3	77.6	71.0

[a] All multiple applicants applied to two included trainings, except for one member of the control group, who applied to three trainings.

[b] Number of completed surveys divided by the number of unique applicants.

6. Baseline analysis approach

In addition to describing the characteristics of the sample, the baseline analysis seeks to assess the similarity between the treatment and control groups in characteristics and key indicators at baseline (although, as discussed above, we expect to see differences in key indicators because of the timing of the baseline survey and will interpret them with caution). Given the use of random assignment, the basic method for doing so is simply to compare mean outcomes between the treatment and control groups. However, in this setting we must take into account an important feature of the design, namely that random assignment was conducted separately by each training, with different proportions of treatment and control applicants.

[9] This reflects applications to the included trainings only. Some applicants may also have applied to excluded trainings, but we did not capture all of these applications in our applicant database.

Therefore, simple comparisons of treatment and control means might be biased because treatment status is correlated with specific trainings (Duflo et al. 2008).

To address this, we conducted the baseline analysis using the following ordinary least squares regression framework, which includes training fixed effects:

(1) $Y_{ij} = \alpha + \beta T_{ij} + \lambda_j + \varepsilon_{ij}$

In this regression, Y_{ij} is an outcome for individual *i* who applied to training *j*; T_{ij} is a binary indicator for applicants who were assigned to the treatment group; λ_j is a vector of training fixed effects, each of which is equal to one if individual *i* was assigned as part of training *j* and zero otherwise; and ε_{ij} is an error term. The vector of training fixed effects, λ_j, accounts for differences in treatment assignment proportions across trainings—effectively, the treatment-control differences are estimated separately for each training and then combined. As mentioned above, about 3 percent of applicants applied to multiple trainings. For the purposes of the evaluation, these applicants are linked to the first included training to which they applied, so each individual is associated with only one of the training fixed effects. The coefficient β in regression (1) gives the treatment-control difference adjusted for training fixed effects. In the results that we present in Chapter III, we use the coefficients from regression (1) to adjust the treatment and control means using the estimated fixed effects.

This page has been left blank for double-sided copying.

III. BASELINE FINDINGS

In this chapter, we examine the demographic characteristics of our sample and baseline indicators in the three main outcome domains for the ultimate impact analysis: (1) vocational training, (2) employment and productive engagement, and (3) earnings and income. However, as mentioned in Chapter II, delays in the timing of the baseline survey pose a challenge to conducting the baseline analysis for indicators in the main outcome domains, because these indicators are likely to have been affected by the start of the VTGF trainings. This suggests that they may not be true baseline indicators, and that the estimated levels and treatment-control differences should be interpreted with caution. In contrast, the demographic characteristics of the sample are unlikely to have been affected by the timing of the survey. Therefore, the analysis of these characteristics should provide a valid description of the sample, and a valid assessment of the baseline similarity between the treatment and control groups.

The baseline characteristics and indicators we will examine were summarized in Table II.5; in the text below, we analyze them by domain. To structure the discussion in each domain, we first describe the characteristics and indicators for the control group, and then assess their differences with the treatment group and the implications of these differences for the evaluation.

We focus on describing the indicators for the control group because we believe that this is the best available approximation to the true baseline situation, and therefore provides useful context for the evaluation. Although the control group members might have been affected by the delay in timing of the baseline because they would have had time to seek out alternative training and employment opportunities, the average delay in the baseline was only a few months relative to the start of VTGF training (a median of two months and a mean of about three months). Given the limited training opportunities and high youth unemployment rate in Namibia, it is unlikely that many training or employment opportunities were immediately available in such a short timeframe to the control group members. Further, the estimated treatment-control differences in key indicators that we present below did not vary much if the sample was restricted based on the length of the delay in the baseline survey. This suggests that the control group was largely not taking advantage of the VTGF training period to take up other opportunities. Therefore, while we cannot be certain that the levels of key indicators in the control group indicators were not affected by the timing of the baseline survey, they are likely to provide a reasonable proxy for the pre-VTGF training context. In contrast, the baseline indicators for the treatment group are likely to have been strongly affected by enrollment in VTGF training.

A. Demographic characteristics

The typical respondent in the control group was an unmarried female in her mid-20s who had completed grade 12 (Table III.1). About 64 percent of the control group was female, the average age was about 27, and 93 percent was unmarried. These respondents tended to live in relatively large households, with an average household size of 5.8, compared to the estimated average Namibian household size of 4.7 (Namibia Statistics Agency 2012). Almost all respondents (95 percent) in the control group had completed at least junior secondary school (grade 10), and about 66 percent had completed at least grade 12 (senior secondary). Parental education, a characteristic that could be correlated with respondents' outcomes, such as

Table III.1. Demographic characteristics of the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Female	741	665	64.3	63.7	0.5	0.83
Age						0.54[a]
Younger than 20	741	664	6.1	4.5	1.6	
20–24	741	664	39.9	38.8	1.0	
25–29	741	664	27.9	30.9	-2.9	
30–34	741	664	13.5	14.4	-0.9	
35 or older	741	664	12.6	11.4	1.2	
Mean (years)	741	664	26.8	26.7	0.1	0.67
Unmarried	739	664	93.1	92.7	0.4	0.80
Respondent's education						0.80[a]
Less than grade 10	740	664	3.6	4.7	-1.0	
Completed grade 10	740	664	29.5	29.7	-0.2	
Completed grade 12	740	664	62.9	61.1	1.8	
Higher	740	664	4.0	4.5	-0.5	
Household size						0.48[a]
1	739	665	3.7	3.7	0.0	
2–5	739	665	56.2	52.6	3.6	
More than 5	739	665	40.1	43.7	-3.6	
Mean (number)	739	665	5.5	5.8	-0.3	0.19
Mother's education						0.82[a]
Less than grade 10	552	544	62.7	61.0	1.8	
Completed grade 10	552	544	18.3	20.6	-2.3	
Completed grade 12	552	544	13.7	13.9	-0.2	
Higher than grade 12	552	544	5.3	4.6	0.7	
Father's education						0.87[a]
Less than grade 10	449	473	53.2	51.4	1.7	
Completed grade 10	449	473	17.4	17.9	-0.5	
Completed grade 12	449	473	21.1	23.1	-2.1	
Higher	449	473	8.3	7.5	0.8	

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.

[a] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

household income, typically is lower than respondents' own level of education. For example, only about 19 percent of mothers and 31 percent of fathers of control group respondents had completed at least grade 12, compared with 66 percent of the respondents themselves.

Comparisons of the demographic characteristics among treatment and control group members suggest that random assignment was successful. All of the differences between the two groups examined were small in magnitude, and none was statistically significant. This finding suggests that random assignment was successful in creating treatment and control groups that are similar along these dimensions. Because these characteristics are unlikely to be affected (or to have been affected in the past) by the VTGF funding, we intend to use them as control variables to improve the precision of the follow-up impact analysis.

B. Vocational training

The VTGF baseline survey captured information on respondents' participation in vocational training in the five years before the survey and their enrollment status as of the survey date. About 28 percent of the control group members reported having enrolled in any vocational

training in the previous five years, although only 16 percent reported having completed a training and 13 percent having completed a training with a certificate (Table III.2). This gap between training enrollment and completion could reflect both high drop-out rates from training (about 17 percent of the control group who had been enrolled reported that they had dropped out of a training, not shown in the table), as well as trainings that were ongoing at the time of the baseline survey (14 percent of the control group reported being engaged in vocational training at the time of the survey). On average, the control sample had experienced a total of 2.6 months of vocational training in the five years before the survey date.[10]

The pattern of vocational training is very different among the treatment group members, who on average were significantly more likely to have been enrolled in training. About 50 percent of treatment group respondents had ever enrolled in vocational training (22 percentage points higher than the control group), and 41 percent were enrolled in training at the time of the survey (27 percentage points higher than the control group). The higher reported levels of training enrollment in the treatment group are likely to reflect participation in VTGF training (although we cannot be certain from the information available in the survey), which would be captured in both enrollment as of the survey date and in the enrollment during the five-year reference period.[11] Average total months of training also was significantly higher among the treatment group members (3.9 months, compared to 2.6 months), as was the percentage reporting that they underwent a job attachment in the previous five years (21 percent versus 13 percent).[12] Again, these treatment-control differences likely reflect that many respondents in the treatment group already had started their VTGF-funded training at the time of the survey; in some cases, they may even have started their job attachments. In contrast, the percentage reporting that they had completed vocational training (or received a certificate) in the previous five years is statistically similar in the treatment and control groups. Because most VTGF trainees would not have had time to complete their course by the time the baseline was conducted, the completed trainings reported are more likely reflect pre-VTGF trainings.

C. Employment and productive engagement

A primary outcome of interest for the VTGF evaluation is employment (defined as either wage employment or self-employment). In addition, because VTGF funding might encourage trainees to undertake further training rather than employment, the evaluation will examine

[10] A total of 46 respondents reported that they had received vocational training in the previous 5 years, but had completed zero months of training. This could be reporting error, or reflect respondents who had completed less than one full month of training by the time of the survey. Because we could not be certain that these zero responses were misreports, we did not correct them—the estimated mean months of training would not be affected in any case.

[11] Because participation in VTGF-funded training is likely to be reflected in the measures of enrollment in vocational training and almost all of these trainings had started by the time applicants were interviewed, we find it puzzling that the fraction of treatment group members who reported to be enrolled in training (in the preceding five years, or at the time of the survey) is not closer to 100 percent. Possible explanations include: (1) dropout rates for these trainings are very high, which we will learn more about in the follow-up survey; and/or (2) our data on training start dates and/or respondent reports on current training enrollment status are erroneous, but there is no way for us assess whether this is true. We will have to wait for follow-up survey data to assess VTGF-training enrollment and completion rates.

[12] A "job attachment" is a short period of practical work experience gained during or at the end of some trainings; it entails trainees working for an employer in their field of training.

impacts on "productive engagement," defined as an individual being employed in the seven days preceding the survey or engaged in vocational training at the time of the survey. For this reason, we examined indicators related to employment and productive engagement at baseline (Table III.3).

Table III.2. Vocational training indicators for the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Vocational training in the previous 5 years:						
Any training	740	665	50.1	28.0	22.1	0.00***
Total months of training						0.00***[a]
None	735	662	54.7	74.2	-19.4	
Less than 6 months	735	662	20.2	7.6	12.6	
6–11 months	735	662	11.8	7.2	4.6	
12 months or more	735	662	13.3	11.1	2.3	
Mean (months)	735	662	3.9	2.6	1.3	0.00***
Had any job attachment or internship	740	664	20.6	12.7	7.9	0.00***
Completed training	653	643	17.7	15.6	2.0	0.28
Completed training with certificate	653	643	14.2	12.5	1.7	0.37
Enrolled in vocational training at the time of the survey	740	664	41.0	13.8	27.3	0.00***

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.

[a] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

About 46 percent of the control group members reported that they worked for pay, profit, or family gain in the seven days before the survey.[13] When we looked at engagement in any productive activity, which accounts for participation in vocational training as well as employment, we found that 53 percent of the control sample was productively engaged when the survey was conducted. We also examined a more detailed breakdown of employment status, in which individuals not employed are defined as unemployed if they were available to work if offered a job in the previous seven days, and out of the labor force if they were not available to

[13] This definition of paid employment is aligned with the official Namibia Statistics Agency definition of employment (Namibia Statistics Agency 2012).

Table III.3. Employment and productive engagement indicators for the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Employed in the past 7 days [a]	740	665	36.8	46.1	-9.2	0.00***
Engaged in any productive activity[b]	740	665	63.8	53.4	10.5	0.00***
Current employment status:						0.00***[d]
Employed [a]	708	662	38.3	46.5	-8.2	
Unemployed [c]	708	662	50.4	47.3	3.1	
Not in the labor force	708	662	11.3	6.2	5.1	

Source: VTGF baseline survey.
Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.
*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.
[a] Worked at least one hour for pay, profit, or family gain in the previous seven days, or had a job to which to return.
[b] Employed in past seven days or enrolled in vocational training at the time of the survey.
[c] Broad definition: available to work if offered job (does not include job search).
[d] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

work (for example, if they were studying, ill, or homemakers).[14] This analysis suggests that 47 percent of the control group was employed, another 47 percent was unemployed, and 6 percent was out of the labor force.

There were large and statistically significant differences between the treatment and control groups in employment and engagement in productive activity. In particular, employment was significantly higher in the control group (46 percent, compared to 37 percent in the treatment group), and engagement in productive activity was significantly lower in the control group (53 percent, compared to 64 percent in the treatment group). Again, this difference likely reflects the start of VTGF training by some members of the treatment group, who may have switched from employment (or unemployment) to training. This would have resulted in lower employment rates and higher productive engagement rates, the latter driven by increased training enrollment.[15] Because these differences likely reflect the timing of the baseline survey relative to the start of VTGF training, we again do not believe they threaten the validity of the evaluation design and we will not use these measures as control variables in the regression-adjusted impact analysis.

[14] This is the "broad" definition of unemployment used by the Namibia Statistics Agency. The "strict" definition classifies those who were not actively seeking work as out of the labor force and does not count them as unemployed. The broad definition may be especially relevant in the Namibian labor market context, in which the concept of seeking work may not be appropriate (Namibia Statistics Agency 2012).

[15] Enrollment in vocational training and employment are not mutually exclusive; about 12 percent of the control sample and 14 percent of the treatment sample report being engaged in both (not shown in the table).

To provide further descriptive information about the nature of employment and the experiences of the unemployed at baseline, we also examined several other indicators for the control group members who were employed or unemployed (Table III.4). Because these indicators are conditional on employment and there are statistically significant differences in employment rates between the treatment and control groups, treatment-control differences in these indicators could reflect differential selection into employment. For this reason, we focus on describing these indicators for the control group rather than analyzing treatment-control differences.

The most common type of employment for the employed control group sample was part-time employment (57 percent), followed by permanent employment (33 percent), and self-employment (10 percent). About 60 percent of the employed control group members was employed in the formal sector (that is, employed with an employer who is registered for tax purposes). Almost all of those who were employed (95 percent) had a single job. About 78 percent of the employed control group members worked a full 40-hour week, with the mean employed respondent in the control sample working 48 hours per week. For the employed, there was substantial variation in the size of the workplace, with 46 percent of the control group members working for a small employer (up to 10 workers), and 37 percent for a large employer (more than 20 workers). The mean employed respondent in the control sample had been in her or his job for about 28 months, 46 percent expressed being satisfied or very satisfied with their job, and 40 percent felt that their training was relevant to their job (increasing the relevance of training is one of the overall goals of the vocational training activity).

Focusing on the unemployed, only about 48 percent of those unemployed in the control group had ever held a job—suggesting that long-term unemployment is an important challenge for this group. About 68 percent of the unemployed in the control group had been unemployed for more than two years, with a mean unemployment duration of 47 months.

Table III.4. Features of employment and unemployment among the control group members (percentages, unless otherwise indicated)

	Control sample size	Control mean
Among respondents who are employed: [a]		
Type of current employment:		
Self-employed	284	9.9
Full-time	284	33.5
Part-time	284	56.7
Currently employed in formal sector	248	60.1
Currently holds only 1 job	288	95.1
Hours worked per week, all jobs		
1–19	282	6.7
20–29	282	6.0
30–39	282	8.9
40 or more	282	78.4
Mean (hours)	282	47.6
Number of workers at workplace		
1	278	9.7
2–10	278	36.3
11–20	278	16.5
21 or more	278	37.4
Job tenure:		
Less than 6 months	246	32.1
6–11 months	246	9.8
12–23 months	246	19.9
More than 23 months	246	38.2
Mean (months)	246	28.4
Job is relevant to education and training	287	40.4
Very satisfied or satisfied with job [b]	278	45.7
Among respondents who are unemployed:		
Ever employed	264	48.1
Duration of unemployment: [c]		
Less than 6 months	225	8.9
6–11 months	225	4.0
12–23 months	225	18.7
More than 23 months	225	68.4
Mean (months)	225	47.1

Source: VTGF baseline survey.
Note: Sample sizes within each panel vary because of item nonresponse.
[a] Based on main job if respondent holds multiple jobs.
[b] Options were very satisfied, satisfied, dissatisfied, very dissatisfied, or no opinion.
[c] For respondents who never were employed, this is the time since they left school.

D. Earnings and income

Another important goal of the VTGF subactivity is to improve the earnings (wage or self-employment income) of trainees. We thus examined baseline monthly earnings, setting these at zero for those not employed at the time of the survey (Table III.5). About 57 percent of the control group members had zero baseline earnings by this measure, reflecting the low levels of employment discussed above. Overall, mean monthly earnings for the control group was about

N$894 (US$97).[16] Only 7 percent of the control group members with nonzero earnings reported that anyone else depended on their earnings.

Table III.5. Monthly earnings for the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Monthly gross earnings from wages:						0.03**[a]
None	706	643	66.6	57.2	9.3	
N$800 or less	706	643	7.6	9.1	-1.5	
N$801–N$1,999	706	643	12.7	15.5	-2.8	
N$2,000–N$2,999	706	643	6.7	8.5	-1.7	
N$3,000 or more	706	643	6.4	9.7	-3.3	
Mean (N$)[b]	706	643	649	894	-245	0.00***
Any dependents on earnings among those with nonzero earnings	261	285	11.0	6.8	4.1	0.01***

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.

[a] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

[b] Computed using the midpoint of each category in the survey (which are more detailed than the categories presented here) and the lower bound for the highest category.

Mean earnings were significantly lower in the treatment group (N$649), again reflecting the lower levels of employment in this group. This difference suggests that baseline earnings may have been affected by the start of VTGF trainings and is unlikely to be a true baseline measure. For this reason, they are unlikely to reflect true underlying treatment-control differences; we would not consider this difference as a threat to the validity of random assignment, and we would not use this measure as a control variable in the regression-adjusted impact analysis.

We also examined monthly household income, a broader measure of well-being than earnings (Table III.6). In the control group, wages (of all household members, not just the respondent) are the most common source of household income (80 percent of households), followed by business activities (33 percent), and remittances (28 percent). This pattern was very similar in the treatment group.

Mean monthly household income was N$2,310 in the control sample (US$250).[17] Despite the significant treatment-control differences in respondent earnings noted above, mean household income is similar in the treatment and control samples. Other sources of household

[16] Earnings were reported in predetermined categories (not the same as the categories in Table III.5, which have been collapsed for ease of presentation). We used the midpoint of each category in the survey to convert these into monetary amounts. For the highest category in the survey ($N5,000 or more), we used the lower limit (N$5,000); thus, the estimated mean income is likely to be a lower bound of the true mean. To convert into U.S. dollars, we used the average exchange rate of US$1 = $N9.23 over the baseline survey period (December 2011 to July 2014), obtained from www.oanda.com.

[17] We converted reported categories for income into a continuous measure using category midpoints, similar to our approach for earnings.

income (including earnings of other household members) may simply dominate the contribution of respondent earnings, leading to similar overall household income. These figures also suggest that the average applicant's annual household income (N$27,720 in the control sample and $27,624 in the treatment sample) fell well below the cutoff for VTGF eligibility of N$250,000. Although it is reassuring that we found mean household income to be similar in the treatment and control groups at baseline, respondents' own earnings at baseline may constitute a sizeable component of household income. Because respondents' own earnings are likely to have been affected by the VTGF trainings, household income at baseline is also likely to have been affected by the VTGF trainings. Consequently, we do not intend to use baseline household income as a control variable in the regression-adjusted impact analysis.

Table III.6. Monthly household income for the analysis sample (percentages, unless otherwise indicated)

	Treatment sample size	Control sample size	Treatment mean	Control mean	Difference	*p*-value
Income sources:[a]						
Wages or pay from job	720	648	78.1	79.9	-1.8	0.49
Agricultural sales	719	646	19.3	17.4	1.8	0.45
Business activities	719	646	35.2	32.7	2.5	0.40
Pensions	718	646	20.1	19.4	0.7	0.78
Remittances	719	645	26.6	28.2	-1.6	0.56
Other	719	643	1.5	1.9	-0.3	0.70
Monthly gross household income:						0.52[b]
N$500 or less	548	530	8.5	9.5	-0.9	
N$501–N$800	548	530	14.2	16.2	-2.0	
N$801–N$1,999	548	530	29.9	28.1	1.8	
N$2,000–N$2,999	548	530	18.7	14.1	4.6	
N$3,000–N$3,999	548	530	6.4	9.2	-2.8	
N$4,000–N$4,999	548	530	6.6	6.2	0.3	
More than N$5,000	548	530	15.6	16.7	-1.0	
Mean (N$)[c]	548	530	2,302	2,310	-8	0.94

Source: VTGF baseline survey.

Note: All means and differences are regression adjusted for training fixed effects. Sample sizes vary because of item nonresponse.

*/**/*** Treatment-control difference is statistically significant at the 10/5/1 percent level of significance using a two-tailed test.

[a] Categories can add up to more than 100 percent because respondents could select multiple options.

[b] p-value from a test of joint significance across all categories using seemingly unrelated regressions.

[c] Computed using the midpoint of each category in the survey (which are more detailed than the categories presented here) and the lower bound for the highest category.

This page has been left blank for double-sided copying.

IV. CONCLUSION

In this chapter, we summarize the key findings from our analysis of the VTGF baseline data, focusing on the implications for the validity of our random assignment design. We also recalculate the minimum detectable impacts (MDIs)—the smallest impacts on key outcomes that our design will be able to distinguish statistically from zero—based on updated parameters from the baseline data. We then discuss the main challenges to the internal validity of the evaluation and assess its likely external validity to a broader population of potential trainees in Namibia. Finally, we outline our future plans for data collection and reporting related to the evaluation.

A. Summary of findings

Our analysis of the demographic characteristics of the sample suggested that the baseline treatment and control samples were very similar. Specifically, the differences between the treatment and control groups in gender, age, marital status, level of education, household size, and parental education were small and statistically insignificant. These results increase our confidence that there were no underlying differences between the two groups.

In contrast, there were large and statistically significant differences in the baseline indicators most closely related to the follow-up outcomes of interest in all of the domains we will analyze at follow-up. In particular, the treatment group was significantly more likely to have been enrolled in vocational training in the five years before the baseline survey, significantly less likely to be employed at the time of the survey, and had significantly lower average earnings. However, we determined that these differences likely do not reflect underlying treatment-control differences, but rather reflect a delay in the timing of the baseline survey relative to training. Because of this delay, many of these indicators would have been affected by the start of training; thus, the VTGF baseline survey may not have captured a true baseline measure of these indicators.

Overall, the implications of the observed treatment-control differences are that the impact evaluation design still is likely to be valid. Baseline demographic characteristics, which were unaffected by VTGF trainings at baseline, were very similar in the treatment and control groups, increasing our confidence that random assignment was successful. Differences between the treatment and control groups in baseline indicators related to outcomes were large and statistically significant, but likely capture the effects of training at the time of the baseline interview (conducted after the start of the VTGF-training for most cases) and so do not threaten the validity of the design. These findings also suggest that we will have to be cautious about which baseline indicators to use as control variables in the follow-up impact analysis. These controls will primarily include demographic characteristics; controlling for the other baseline indicators could result in biased impact estimates.

B. Updated MDIs

In our original design report (Mamun et al. 2014), we computed MDIs for the random assignment design based on our best estimates of sample sizes and other parameters available at the time. Using the baseline data, we were able to revise the MDI calculations based on more up-to-date information (Table IV.1). Specifically, we used the baseline control group means and

standard deviations for the primary outcomes of interest, whereas our original MDI calculations were based on estimates of these parameters from the 2009-2010 NHIES data. Because we expect the control group to have been largely unaffected by VTGF trainings at baseline, it is appropriate to use the control group parameters. We also updated the sample size for the trainings that will be included in the evaluation, which is smaller than the sample size that we originally assumed. The revised MDI calculations assume a response rate of 70 percent in the treatment group and 60 percent in the control group at follow-up.[18] This assumption reflects the difference in response rates between the treatment and control groups at baseline, and the likelihood that some baseline respondents will not respond to the follow-up survey.[19]As in the design report, we present MDIs for both ITT impact estimates (which will give the impact of the offer of VTGF training) and ToT impact estimates (which will give the impact of enrolling in or completing VTGF training).

Under the updated assumptions, the evaluation will be powered to detect ITT impacts of about 5.5 percentage points for training completion (a relative increase of 35 percent compared to the estimated control group mean of 16 percent), the primary outcome in the vocational training domain. In the other domains, we will be able to detect ITT impacts of 7.6 percentage points for employment (a relative increase of 16 percent compared to the estimated control group mean of 46 percent), 7.6 percentage points for productive engagement (a relative increase of 14 percent compared to the estimated control group mean of 53 percent), and N$209 for monthly earnings (a relative increase of 22 percent compared to the estimated control group mean of N$894).

We also are interested in separately analyzing impacts for certain subgroups, such as those defined by gender, baseline household income, or language group. For women, who comprise approximately 60 percent the full sample, the MDIs are about 30 percent larger than for the full sample. For applicants with a baseline household income of no more than $800, who comprise approximately 25 percent of the full sample, the MDIs are twice as large as those for the full sample. This finding suggests that our ability to estimate precise subgroup-specific impacts may be limited, especially for small subgroups.

The MDIs for the ToT estimates are larger than those for the ITT estimates because they account for noncompliance. Further, because not all of those who enroll in training will complete it, the MDIs for the ToT estimates of training completion (the impact of VTGF funding on those who complete training) are larger than those for training enrollment (the impact of VTGF funding on those who enroll in training). These MDIs for the ToT estimates of the impacts of

[18] The total sample size used in the current MDI calculations is smaller by 208 than that used for the calculations presented in the design report, reflecting updated information on the number of applicants in each training and the trainings that can be included in the evaluation. Also, the follow-up survey response rates used in the current calculations are lower by 5 percentage points for both treatment and control groups than those used for the calculations presented in the design report. We use a more conservative follow-up survey response rate assumption because of the somewhat lower than expected response rates we have observed in the follow-up survey from the early batches of the sample.

[19] The MDIs are not very sensitive to lowering these response rates; for example, they increase by 9 percent (not percentage points) if the response rates are 60 percent in the treatment group and 50 percent in the control group.

training enrollment and completion are about 15 percent and 28 percent larger than those for the ITT estimates, respectively.

Table IV.1. Updated minimum detectable impacts for the VTGF impact evaluation

	Completed vocational training [percentage]	Employed [percentage]	Productively engaged [percentage]	Monthly earnings [Namibian dollars]
Estimated baseline mean	15.6	46.1	53.3	894
Standard deviation	36.3	49.8	49.9	1,372
ITT Impacts[a]				
Full sample	5.5 (35%)	7.6 (16%)	7.6 (14%)	209 (23%)
Subgroup (60 percent)[b]	7.1 (46%)	9.8 (21%)	9.8 (18%)	269 (30%)
Subgroup (25 percent)[c]	11.0 (71%)	15.2 (33%)	15.2 (28%)	417 (47%)
ToT Impacts[a]				
Full sample (impact of enrollment)	6.3 (41%)	8.7 (19%)	8.7 (16%)	240 (27%)
Full sample (impact of completion)	7.1 (45%)	9.7 (21%)	9.7 (18%)	267 (30%)

Sources: Authors' calculations, using data from the VTGF baseline survey.

Note: MDI calculations assume a two-tailed test with a 95 percent confidence level and 80 percent power. The assumed baseline means and standard deviations are based on the control group means and standard deviations in the VTGF baseline survey. Using the total eligible sample sizes of 955 treatment cases and 937 control cases, we then assume a follow-up survey response rate of 70 percent in the treatment group and 60 percent in the control group. The assumed response rate reflects a lower response rate in the control group relative to the treatment group at baseline and additional expected nonresponse at follow-up (baseline response rates were 78 percent in the treatment group and 71 percent in the control group). We also assume that baseline covariates explain 10 percent of the variation in outcomes. For the ToT estimates, we assume that 90 percent of those awarded a VTGF grant (VTGF treatment group) will enroll in VTGF-funded training, and that 90 percent of those who enroll will complete training. We also assume that 3 percent of those not initially awarded a VTGF grant (VTGF control group) eventually will enroll in VTGF-funded training by applying to another training.

ITT = intent-to-treat; ToT = treatment-on-treated.

[a] MDIs expressed in percentage terms are relative to the baseline control group mean.

[b] For example, women (64 percent of the baseline sample).

[c] For example, household income of no more than N$800 per month (24 percent of the baseline sample).

C. Summary of internal validity risks

The main risks to the interval validity of our random assignment design are (1) underlying differences between the treatment and control groups that could affect follow-up outcomes, and (2) noncompliance with randomly assigned treatment status. In this setting, noncompliance is most likely to result from treatment applicants who do not take up the offer of VTGF-funded training. The available information from VTGF implementation suggests that this situation may not be widespread; in any case, we will account for noncompliance through our ToT estimates. As to underlying treatment-control differences, our baseline analysis does not suggest any such differences in demographic characteristics. There are many large and statistically significant differences in key baseline indicators; however, as described above, we believe that these

differences are related to the fact that in most cases the baseline survey was conducted after training had started. Therefore, we do not believe that they threaten the validity of the design.

Nevertheless, we still have a concern that there may be underlying differences between the treatment and control groups in our eventual impact analysis sample because of follow-up survey attrition (survey nonresponse). Attrition potentially could lead to biased impact estimates for two reasons. First, if the attrition rates between the treatment and control groups are very different at follow-up, different types of individuals might be attriting in the two groups, leading to a lack of comparability in the analysis sample. Despite a relatively large difference in attrition between the treatment and control groups at baseline (7 percentage points), our baseline analysis suggests that the two groups were broadly comparable along baseline demographic characteristics that would not have been affected by trainings. The difference in baseline attrition rates thus does not appear to have led to a treatment-control nonequivalence at baseline. However, if this gap widens at follow-up, concerns about a lack of comparability would increase. Second, if the overall rate of attrition is very high, different (likely unobservable) factors could be affecting attrition in each group, even if the attrition rates are similar. Again, this could lead to a lack of comparability between the treatment and control analysis samples that could bias the impact estimates.

Because of the concern about sample attrition, we will attempt to maximize response rates in the follow-up survey through several strategies. For example, we will make multiple attempts to contact respondents at different times if they do not respond initially. We also will use contact information from several sources, including the original random assignment information and alternate contact information from the baseline survey (which includes a secondary contact number and a contact number for a friend or relative), to maximize our chances of contacting each sample member. Based on promising results from a pilot with early batches of the follow-up sample, we will also mail up to two postcards to initial non-responders, asking them to contact the data collection team by telephone or text message and offering them an incentive of N$10 (about US$1) in cell phone airtime if they complete the survey (to ensure fairness, all those who complete the follow-up survey will receive this incentive, regardless of whether they completed it in response to the postcards or not). Finally, when we conduct the impact analysis, we will re-examine the baseline comparability of the treatment and control groups based on the follow-up analysis sample, which may differ from the baseline sample. Comparability on unobservable characteristics would not be guaranteed if the differential or overall follow-up attrition rates are high. However, if the follow-up treatment and control samples are comparable, confidence in the validity of our impact estimates would increase.

D. External validity

The VTGF impact evaluation was not designed to ensure external validity. The impact estimates from the VTGF evaluation will apply to a specific group: those who met MCC and VTGF training providers' requirements and were interested in applying for VTGF funding. They do not necessarily apply to a broader population, such as unemployed youth in Namibia, who could be drawn into vocational training in Namibia in the future, once the NTF is fully operational. Nevertheless, we tried to assess the possible extent of applicability of the impact estimates to the broader population of potential trainees by comparing the characteristics of the VTGF baseline analysis sample to those of the 2009–2010 Namibia Household Income and Expenditure Survey (NHIES) sample. The NHIES is a nationally representative household survey that gathers information on demographic characteristics, income, consumption, and

expenditure. It includes several demographic measures available in the VTGF baseline survey (such as age, gender, and education) and several of the baseline VTGF indicators (such as employment, work hours, and monthly household income).

Ideally, we would compare the characteristics and baseline indicators of the VTGF sample to the group of potential trainees in the NHIES sample. If these characteristics and indicators are similar, we would be more confident that our impact estimates largely are applicable to this broader population. However, the group of potential trainees is not readily identifiable in the NHIES sample. For this reason, our approach is to impose various assumptions on the characteristics of this group. We begin by assuming that it simply includes any individuals over the age of 15 (the lowest age category in the NHIES that includes young adults) who currently are not attending school. We then impose progressively stricter assumptions—namely, that the group of potential trainees has (1) a similar age distribution to the VTGF sample; (2) a similar age and gender distribution; and (3) a similar age, gender, and education distribution. In each case, we reweight the NHIES sample accordingly and compare the characteristics of the reweighted sample to those of the VTGF sample.[20] For example, under assumption (2), we reweight the NHIES data to match the proportion of the VTGF sample that falls in each age-gender category (such as females ages 20–24). We then assess the likely degree of external validity by comparing the VTGF and NHIES samples on characteristics for which we did not enforce similarity. For example, under assumption (2), age and gender will be mechanically similar in the two samples, but we still can compare education, employment, and household income.

There are several important caveats to this analysis. First, there are a limited number of characteristics and indicators that are comparable across the NHIES and VTGF data. Thus, even if the samples appear similar across these limited available measures, the VTGF sample still could be different in other (unobserved) ways to the broader NHIES sample. Second, there are some differences in measurement across the two surveys, particularly in household income. Specifically, the NHIES recorded household income in much greater detail by asking separately about various categories of income—it might thus have captured additional income sources and be more accurate relative to the VTGF survey. Third, the 2009–2010 NHIES data (the latest publicly available round) are less recent than the VTGF baseline data, so the comparison could partly reflect time trends or annual fluctuations (for example, in employment rates). Finally, even if this exercise suggests that the characteristics of the two samples are similar, the VTGF impact estimates still might not be externally valid if the features of training, the broader vocational training landscape, or the overall economic situation in Namibia changes substantially in the future. For these reasons, the results should be viewed only as suggestive regarding the possible external validity of the findings.

[20] Because the VTGF treatment sample might already have been affected by VTGF trainings at baseline, as discussed above, we focus on comparing the VTGF control sample to the NHIES sample. Although we cannot be certain that the control sample was unaffected by the delay in the baseline survey, we argued in Chapter III that these effects are likely to be small given the expected limited training and employment opportunities available to the control group between random assignment and the baseline survey.

Table IV.2. Comparison of the VTGF sample to the NHIES 2009–2010 sample

	VTGF control sample mean	NHIES sample mean, full sample[a]	NHIES sample mean, reweighted by age	NHIES sample mean, reweighted by age and gender	NHIES sample mean, reweighted by age, gender, and education
Demographic characteristics					
Age					
Younger than 20	4.5	7.2	4.5	4.5	4.4
20–24	38.8	17.2	38.8	38.8	37.6
25–29	30.9	16.3	30.9	30.8	30.0
30–34	14.4	14.3	14.4	14.5	14.0
35–39[b]	7.4	11.3	7.4	7.4	7.2
40 or older	4.0	33.6	4.0	4.1	6.9
Mean (age)[b]	26.7	36.1	27.1	27.1	27.8
Female	63.7	54.5	54.4	63.7	64.7
Education					
Less than grade 10	4.7	32.1	24.3	23.6	4.5
Grade 10	29.7	39.9	45.5	46.0	29.0
Grade 12	61.1	20.5	24.6	24.8	62.2
Higher than grade 12	4.5	7.5	5.6	5.6	4.4
Baseline indicators					
Employed	46.1	57.7	54.9	53.9	53.9
Weekly hours worked:					
Zero	55.6	48.5	50.9	52.0	49.8
1–39	10.6	11.2	10.9	11.0	9.5
40 or more	33.8	40.2	38.2	37.1	40.7
Mean (hours)	21.1	24.2	23.5	22.9	24.1
Monthly household income					
N$500 or less	9.5	31.3	30.7	31.1	24.8
N$501–N$1,999	44.3	20.3	20.7	20.6	18.0
N$2,000–N$2,999	14.1	6.5	6.7	6.6	6.2
N$3,000–N$3,999	9.2	4.9	5.1	5.1	5.5
N$4,000–N$4,999	6.2	3.8	4.0	3.9	4.4
N$5,000 or more	16.2	33.2	32.8	32.6	41.1

Source: VTGF baseline survey and 2009–2010 NHIES survey.

Notes: N = 643–665 in the VTGF baseline survey control sample and N = 19,328-18,137 for the NHIES sample.

[a] Includes all out-of-school respondents 15 years or older with nonmissing information for age, gender, and education level.

[b] In the NHIES survey, age was reported in categories. To obtain a continuous age measure, we used the midpoint of each category and the lower bound for the highest category.

This analysis suggests that, after imposing the simple restriction to out-of-school individuals over 15 years old, the NHIES sample differs from the VTGF control sample in several ways (Table IV.2). Specifically, the NHIES sample is older (mean age of 36 versus 27), comprises a lower percentage of females (55 percent versus 64 percent), and is less educated (72 percent did not complete grade 12, versus 34 percent).[21] Looking at the baseline indicators examined in Chapter III, the NHIES sample is more likely to be employed (58 percent versus 46 percent) and works slightly more weekly hours on average (24 versus 21). The distribution of monthly household income is more unequal in the NHIES sample compared to the VTGF sample, with a substantially higher percentage in the lowest income category of $N500 or less (31 percent versus 10 percent) and the highest income category of $N5,000 or more (33 percent versus 16 percent).[22] The difference in the highest income category could partly reflect the VTGF household income restriction for eligibility, which required applicants to have an annual household income of less than N$250,000 (about 12 percent of the NHIES sample have an annual household income above this cutoff). However, despite the differences in the distribution of monthly household income, the median category is the same in both samples (N$501–N$1,999).

We reweighted the NHIES sample to enforce similarity with the VTGF sample across several dimensions, as described above. Reweighting the NHIES sample by the VTGF age distribution would lead to greater similarity by age (mechanically), but only slightly reduce the discrepancies along other dimensions. The same is true for the reweighting by age and gender (except that, mechanically, gender would become more similar), and the reweighting by age, gender, and education (except that, mechanically, education would become more similar).

Overall, an important finding from this analysis is that even if those drawn into vocational training in the future are similar to VTGF applicants in age and/or gender, they may be substantially less educated on average. Thus, if vocational training is available to these individuals (for example, through COSDECs, which target their training to secondary school dropouts), our impact estimates may not generalize to them. The broader sample of potential trainees also has a higher average level of employment than the VTGF sample, but this difference is relatively small and unlikely to be associated with large differences in impacts. However, there is a difference in the distribution of household income, even though median income is similar. If future trainings draw in individuals from the lowest (or highest) parts of the income distribution, the impacts might differ from those estimated in the VTGF evaluation.

In sum, it is difficult to draw strong conclusions about the external validity of the VTGF impact estimates, given the caveats described above. Our analysis suggests that caution may be necessary in generalizing these estimates to vocational training in Namibia more broadly,

[21] Because the NHIES sample size is very large, almost all differences (even very small ones) with the VTGF sample are statistically significant. Thus, we do not present tests for significant differences in characteristics and indicators between the two samples; instead, we focus our discussion on the magnitude of the differences.

[22] Income was reported in categories in the VTGF baseline data and as a continuous measure in the NHIES data; it thus was not possible to compare mean income meaningfully because reported VTGF income was capped at the highest category.

especially if future trainings appeal to a broader group than the VTGF trainings, such as those with lower levels of education or household income.

E. Plans for future data collection and reporting

The follow-up data collection for the VTGF impact evaluation is currently underway. For the follow-up survey, we will survey applicants to each training approximately one year after the end of training; we will therefore release trainings for follow-up based on the training end date (trainings with similar end dates will be released together). Between March and August 2014, NORC managed the collection of follow-up data for the first set of trainings that was due for follow-up and worked with Survey Warehouse to conduct the survey. Mathematica took over management of the follow-up data collection starting in February 2015, when the next batch of sample members was due for follow-up. Mathematica also is working with Survey Warehouse to implement the follow-up survey and will manage the data collection for all remaining samples. We expect to complete the follow-up survey data collection in November 2015.

We will analyze the follow-up data to estimate the impacts of the VTGF funding on trainees' training and labor market outcomes. The impact analysis will involve comparing outcomes in the treatment and control samples for each training one year after the end of training. Using a regression framework similar to equation (1) in Chapter II, we will be able to combine the data for all trainings despite differences in the timing of the follow-up survey, and estimate the overall impact of the VTGF funding. As mentioned earlier, we intend to control for key demographic characteristics in this regression analysis to improve the precision of the estimates, although we will investigate the robustness of our results to the omission of these controls. We will also conduct additional analyses to answer the key research questions, such as analyses of impacts by trainee gender and by the duration or type of training.

We will use the results from the impact analysis to produce a final VTGF evaluation report, a draft of which is expected to be submitted to MCC by the second quarter of 2016 and finalized by the end of 2016. The report will draw on both the quantitative follow-up data and the qualitative data related to implementation to answer the key research questions related to the VTGF subactivity comprehensively.

REFERENCES

Alcid, Annie. "A Randomized Controlled Trial of Akazi Kanoze Youth in Rural Rwanda." Report submitted to USAID. Waltham, MA: Education Development Center, October, 2014.

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.

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *Quarterly Journal of Economics*, vol. 129, no. 2, May 2014.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, and Johannes M. Bos. "The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study." *Journal of Human Resources,* 1997, pp. 549–576.

Card, David, Jochen Kluve, and Andrea Weber. "Active Labour Market Policy Evaluations: A Meta-Analysis." *Economic Journal,* vol. 120, no. 548, 2010, pp. F452–F477.

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." IZA Discussion Paper No. 7408, May 2013.

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.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. "Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics,* vol. 4, edited by T. Schultz and John Strauss. Amsterdam and New York: North Holland, 2008.

Hicks, Joan H., Michael Kremer, Isaac Mbiti, and Edward Miguel. "Vocational Education Voucher Delivery and Labor Market Returns: A Randomized Evaluation Among Kenyan Youth." Report for Spanish Impact Evaluation Fund Phase II. Washington, DC: World Bank, 2011.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Cano. "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey." *Economic Journal*, forthcoming.

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.

Maitra, Pushkar, and Subha Mani. "Learning and Earning: Evidence from a Randomized Evaluation in India." GCC Working Paper Series 14-05, University of Pennsylvania, Population Studies Center, February 2014.

Mamun, Arif, Luke Heinkel, Evan Borkum, Kristen Velyvis. "Evaluation of MCC's Vocational Education and Training Activity in Namibia: Findings from Analysis of the first round of Qualitative Data." Draft report submitted to the Millennium Challenge Corporation. Washington, DC: Mathematica Policy Research, March 31, 2015.

Mamun, Arif, Evan Borkum, Kristen Velyvis, Luke Heinkel, and John Middleton. "The Millennium Challenge Corporation's Vocational Training Activity in Namibia: Evaluation Design Report." Report submitted to MCC. Princeton, NJ: Mathematica Policy Research, May 13, 2014.

Namibia Statistics Agency. "Namibia Household and Expenditure Survey 2009/10." Windhoek, Namibia: 2012.

Schochet, Peter Z., John Burghardt, and Sheena McConnell. "Does Job Corps Work? Impact Findings from the National Job Corps Study." *American Economic Review,* vol. 98, no. 5, 2008, pp. 1864–1886.

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 2013*, vol. 9, 2013.

U.S. Agency for International Development. "USAID/Namibia Country Strategic Plan FY 2004–2010." Washington, DC: USAID, 2003.

World Bank. "Country Partnership Strategy for the Republic of Namibia for the period FY2014–FY2017." June 26, 2013.

www.mathematica-mpr.com

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

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research