

**The Millennium Challenge
Corporation's Vocational Training
Activity in Namibia: Evaluation
Design Report**

May 13, 2014

Arif Mamun
Evan Borkum
Kristen Velyvis
Luke Heinkel
John Middleton



MATHEMATICA
Policy Research

This page intentionally blank

Contract Number:
MCC-10-0114-CON-20 (MCC-13-TO-
0001)

Mathematica Reference Number:
40233.112

Submitted to:
Millennium Challenge Corporation
875 15th Street, NW
Washington, DC 20005
Telephone: (202) 521-3636
Project Officer: Algerlynn Gill

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

**The Millennium Challenge
Corporation's Vocational
Training Activity in Namibia:
Evaluation Design Report**

May 13, 2013

Arif Mamun
Evan Borkum
Kristen Velyvis
Luke Heinkel
John Middleton

MATHEMATICA
Policy Research

This page intentionally blank

ACRONYMS

CCR	Compact Completion Report
COSDEC	Community Skills and Development Center
COSDEF	Community Skills and Development Fund
ERR	Economic Rate of Return
GRN	Government of the Republic of Namibia
IRB	Institutional Review Board
ISC	Industrial Skills Committee
ITT	Intent-to-Treat
IV	Instrumental Variables
LCDRS	Levy Collection, Distribution, and Reporting System
MCA-N	Millennium Challenge Account Namibia
MCC	Millennium Challenge Corporation
MDI	Minimum Detectable Impact
MSE	Micro and Small Enterprise
NAMCOL	Namibian College of Open Learning
NQA	Namibia Qualifications Authority
NTA	Namibia Training Authority
NTF	National Training Fund
OLS	Ordinary Least Squares
RPL	Recognition of Prior Learning
SLA	Service-level Agreement
SME	Small and Medium Enterprise
ToT	Treatment on the Treated
TP	Training Provider
VET	Vocational Education and Training
VTC	Vocational Training Center
VIGF	Vocational Training Grant Fund

This page intentionally blank.

CONTENTS

ACRONYMS	1
I INTRODUCTION.....	1
A. Background.....	1
B. Overview of the Vocational Training Subactivities	1
1. VTGF Subactivity	1
2. NTF Subactivity	5
3. COSDEC Subactivity.....	7
C. Roadmap for the Report	10
II LITERATURE REVIEW AND POLICY RELEVANCE OF THE EVALUATION	11
A. Summary of Existing Evidence	11
B. Gaps in the Literature	12
C. Policy Relevance of the Evaluation.....	13
III EVALUATION DESIGN.....	15
A. VTGF Subactivity: Key Research Questions	15
1. Implementation Analysis.....	15
2. Impacts on Key Outcomes	17
3. Variation in Impacts.....	17
B. VTGF Subactivity: Methodological Overview	18
C. VTGF Subactivity: Approach to the Impact Analysis.....	19
1. Impact Evaluation Design.....	19
2. Random Assignment Procedure.....	21
3. Analysis Plan.....	21
4. Statistical Power	27
5. Adherence to Randomly Assigned Control Status.....	28
D. VTGF Subactivity: Approach to the Implementation Analysis	29
E. VTGF Subactivity: Limitations and Challenges	30
F. NTF Subactivity: Key Research Questions	32

G.	NTF Subactivity: Overview of the Performance Evaluation and Key Challenges.....	34
H.	COSDEC Subactivity: Key Research Questions.....	35
I.	COSDEC Subactivity: Overview of the Performance Evaluation	37
J.	COSDEC Subactivity: Analytic Approach to the Performance Evaluation	39
	1. Outcomes Analysis.....	39
	2. Qualitative Implementation Analysis.....	40
K.	COSDEC Subactivity: Limitations and Challenges	41
L.	Cost Analysis	41
IV	DATA SOURCES AND OUTCOMES.....	43
	A. Quantitative Data Sources and Outcomes.....	43
	1. VTGF Data Source	43
	2. COSDEC Data Sources	44
	3. Outcomes from VTGF and COSDEC Tracer Surveys.....	44
	B. Qualitative Data Sources	47
V	EVALUATION ADMINISTRATION AND MANAGEMENT.....	51
	A. Summary of Institutional Review Board Requirements and Clearances.....	51
	B. Preparing Data Files	51
	C. Dissemination Plan	52
	D. Evaluation Team: Roles and Responsibilities	52
	E. Evaluation Timeline and Reporting Schedule	53
	REFERENCES.....	R.55

TABLES

III.1	Overview of the Evaluation Approach	15
III.2	VTGF Research Questions and Evaluation Approach	20
III.3	Types of Impact Estimates for the VTGF Evaluation.....	24
III.4	Minimum Detectable Impacts for the VTGF Impact Evaluation.....	27
III.5	COSDEC Research Questions and Evaluation Approach.....	38
IV.1	Primary and Key Secondary Outcomes for the VTGF and COSDEC Evaluations.....	46
IV.2	Status of Subactivities at the Time of Qualitative Data Collection	47
IV.3	Plans for Qualitative Data Collection.....	50

This page intentionally blank.

FIGURES

I.1	Logic Model for the VTGF Subactivity	4
I.2	Logic Model for the NTF Subactivity.....	6
I.3	Logic Model for the COSDEC Subactivity	9
V.1	Timeline for the Evaluation.....	54

This page intentionally blank.

I. INTRODUCTION

A. Background

Despite well-functioning physical infrastructure, rich natural resources, and relatively strong public administration, Namibia still suffers from the social and economic inequalities it inherited upon independence from South Africa in 1990. Although a per-capita income of US\$5,425 in 2012 suggests that Namibia is an upper middle income country, its income distribution remains among the most unequal in the world (World Bank 2013; Namibia Statistics Agency 2012a). The legacies of apartheid and colonialism include an economy that relies heavily on the extraction and export of resources, with little value added and limited potential for promoting economic growth; a low-skilled workforce; and vast disparities in quality of life between the economically dominant minority and an impoverished majority. The shortage of skilled workers and limitations in the education system's capacity to create a skilled workforce are 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).

To address some of these challenges, the Millennium Challenge Corporation (MCC) signed a \$304.5 million Compact with the Government of the Republic of Namibia (GRN), with the primary goal of reducing poverty through economic growth. The Compact, which entered into force in September 2009, includes three projects: a tourism project, an agriculture project, and an education project. The education project consists of several activities that aim to improve the quality of Namibia's workforce by enhancing the equity and effectiveness of basic, vocational, and tertiary education.

The vocational training activity is one of the key activities under the education project. It focuses on expanding the availability, quality, and relevance of vocational education and skills training in Namibia, and consists of three subactivities: (1) competitive grants for high-priority vocational skills programs offered by public and private vocational 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 targeted to marginalized populations, including primarily out-of-school youth, but also low-skilled adults.

MCC has selected Mathematica Policy Research (Mathematica) to conduct an evaluation of the vocational training activity, covering all three subactivities. This report describes our design for the evaluation. Next, we provide a more detailed description of each of the three subactivities and the associated program logic; we then provide a roadmap for the rest of this report.

B. Overview of the Vocational Training Subactivities

1. 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, and has several components. Our evaluation is focusing on the key component of the awarding of grants to training providers through a

competitive bidding process. The VTGF solicits grant applications for conducting trainings in specific high-priority skills areas. Initially, it was intended that numerical targets for the number of trainees required in specific skills areas would be identified by industry representative bodies known as industrial skills committees (ISCs), which would rely on labor demand studies, industry knowledge, and national-level planning documents (such as the fourth National Development Plan and the National Human Resource Plan 2010–2025) to identify the key priority areas for skill development. In practice, our understanding is that the ISCs were not fully operational during the VTGF implementation period, so that they identified high-priority areas more broadly. Training providers who receive VTGF grants use them to award scholarships to eligible disadvantaged applicants.¹ The scholarships, which cover tuition and include some subsistence allowance, are intended to increase access to training for these applicants. Providers who are awarded these grants can also apply for an additional capacity-building grant, which they can 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 grant mechanism is intended in part to serve as a pilot for future funding under the NTF, and it has many similar features: (1) grants are awarded to providers on a competitive basis; (2) providers must sign service-level agreements (SLAs) committing themselves to certain milestones to receive each tranche of grant funds; and (3) most of the grants are managed by the Namibia Training Authority (NTA), the same government body that will oversee the NTF (the initial grants were managed by the Millennium Challenge Account Namibia, MCA-N). The awarding of VTGF grants is ongoing; the first grants were awarded in the fourth quarter of 2010, and the last grant is scheduled to be awarded in the second quarter of 2014. A total of 16 training providers will receive VTGF grants, with some of these receiving more than one grant (for different intakes of trainees).

The remaining components of the VTGF are pilots of two other initiatives that will be fully implemented under the NTF. The first is reimbursement of employers for the costs of employer-provided training under the NTF's levy collection, distribution, and reporting system (LCDRS), in which employers register with the NTA, pay a (token) levy, and submit training evidence for reimbursement (we describe this initiative in more detail under the NTF subactivity). The second is the recognition of prior learning (RPL) program, which helps people who are experienced in a certain vocational skills area but do not have formal training to compile a portfolio of evidence of their work experience and have their skills formally assessed and certified. Unlike the VTGF grant funds, the employer-provided training and RPL pilots are not restricted to disadvantaged applicants. Both of these pilots are currently close to completion.

As noted above, the VTGF is similar in some respects to the envisaged fully functional NTF; however, there are also some potentially important differences: (1) there is a difference in timing: during the period of the Compact, the VTGF will be funding trainings (including through scholarships, employer-based training pilot, and the RPL pilot), while the NTF funding mechanism will still largely be in the process of being set up, and will only be fully functional after the end of the Compact; (2) some key processes used to award and manage training funds might differ under the

¹ “Disadvantaged” is defined as having an annual household income of less than N\$250,000, or about US\$23,500 at current exchange rates, after subtracting training costs for household members.

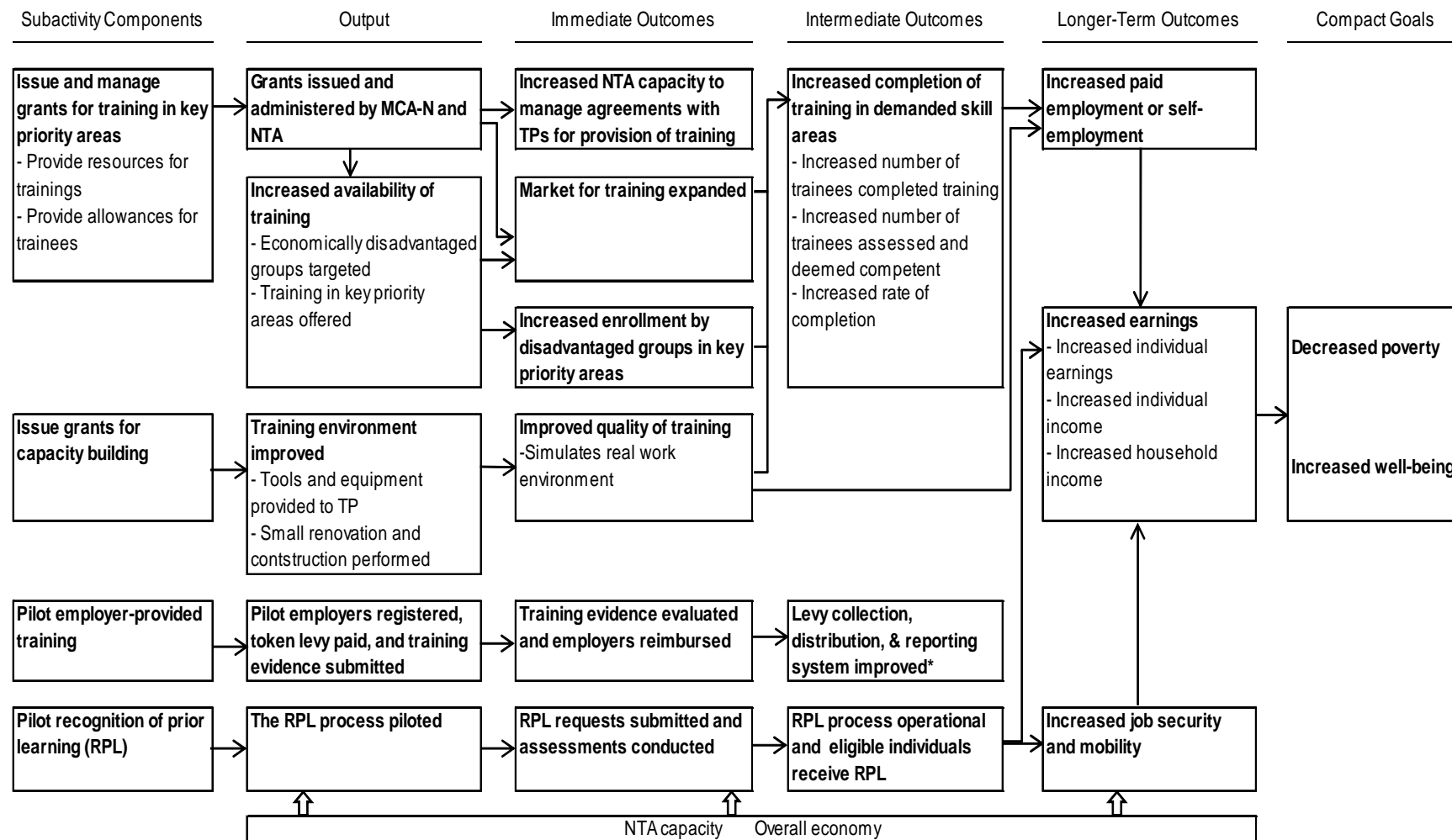
NTF, especially if lessons from the VTGF are used to refine these process under the NTF (for example, in terms of the features and management of SLAs); (3) there will be a difference in how market-relevant trainings are identified, with ISCs playing a much larger role under the NTF than under the VTGF; (4) the VTGF grants focused explicitly on disadvantaged applicants, while the NTF grants will not;² (5) the RPL and employer-provided pilots were a relatively small component of the VTGF, but will feature much more prominently in the NTF—specifically, 50 percent of the funds collected by the NTF will be allocated to employer-provided training and 35 percent will be allocated for training in key-priority areas (GOPA Consultants, 2013); (6) to the extent that NTA registration and NQA accreditation is associated with training quality, quality could be higher under the NTF because the VTGF only required training providers to be in the *process* of being registered and accredited (although some already were registered and/or accredited), while the NTF will require them to be fully registered and accredited; and (7) the NTF is a longer-term intervention that is expected to be retained post-Compact and to result in broad changes to the VET system as a whole (as we discuss below); these system-wide changes are not expected with the VTGF due to its smaller scale and short-term nature.

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 next column shows the direct output of these components. As described earlier, the outputs include administration of grants by the NTA 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.

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 capacity of the NTA to manage SLAs and to disburse funds to providers based on achievement of milestones is expected to increase through their experience managing the VTGF grants. The grants themselves are 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 (in particular, the new tools will more closely simulate those found in the modern work environment), improve enrollment of disadvantaged groups targeted by the grants, and expand the market for training through the competitive bidding process for grant funds (specifically, one new provider—the Namibian College of Open Learning [NAMCOL]—entered the vocational training market). The employer-provided training and RPL pilots are also expected to be conducted (culminating in reimbursement of employers and assessment of candidates, respectively), and the lessons learned synthesized. In the intermediate term, this is expected to result in increased

² However, the relatively high household income cutoff used for the VTGF grants suggests that it is unlikely that VTGF population is substantially more disadvantaged than the population that will apply for vocational training in the future. Specifically, the median household income in Namibia reported in the 2009-2010 Namibia Household Income and Expenditure Survey (NHIES) was about N\$35,000, and the 90th percentile of household income was about N\$100,000 (Namibia Statistics Agency, 2012b). This is well below the threshold of N\$250,000 used for VTGF, even after allowing for income growth since 2009-2010, suggesting that the VTGF income cutoff would not have been a binding constraint for most applicants.

Figure I.1. Logic Model for the VTGF Subactivity



Assumptions:

- A1. Training providers on the road 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.
- NQA = Namibia Qualifications Authority; NTA = National Training Authority; SLA = service-level agreement; TP = training provider.

completion of training (mechanically, through increased enrollment, and by reduced dropout due to the subsistence allowance for board and lodging) and improvements in the capacity of the NTA to support RPL and employer-provided training programs. In the long term, this is expected to improve the labor market outcomes of trainees in terms of further training, employment, and earnings (and, for RPL candidates specifically, to increased job security and mobility) and to lead to the ultimate goals. Key contextual factors for these pathways to be achieved include the NTA having capacity to manage training funds, training providers offering a sufficient quality of training, and a favorable overall economic condition in the country.

2. NTF Subactivity

The NTF is a public fund created by the Vocational Education and Training (VET) Act of 2008 to provide a sustainable source of funding for vocational training in Namibia, with a focus on high-priority skills development. The NTA will manage the NTF, which will be funded through a payroll levy on participating firms.³ The Compact is funding a technical adviser (GOPA consultants) to support establishment of the NTF and piloting of the LCDRS.

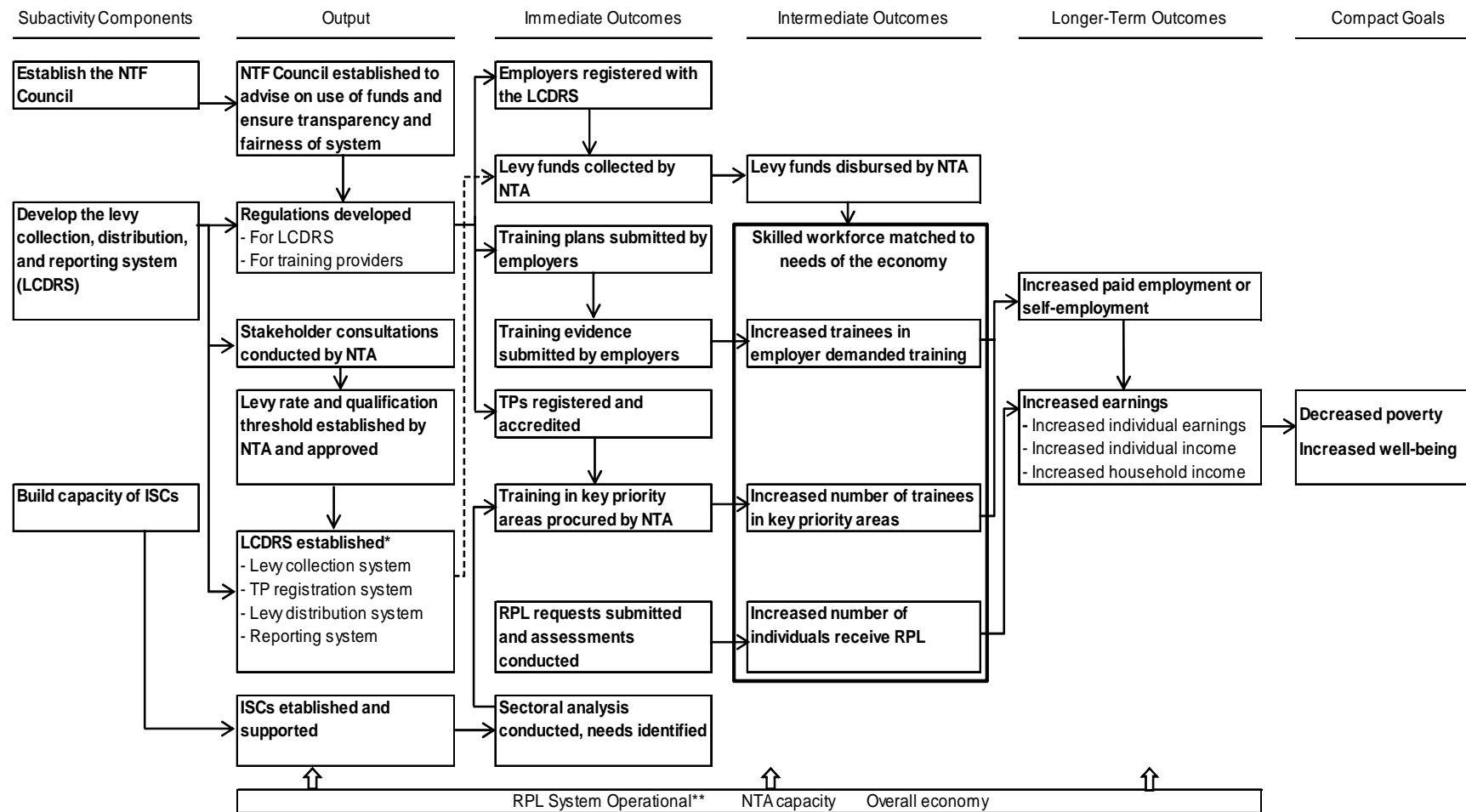
The LCDRS will require all employers in Namibia with a payroll above a certain threshold (currently determined to be over N\$1million) to register with the NTF and pay an annual payroll levy (currently determined to be 1 percent). The levy rate and qualification threshold were determined after consultation with partners from industry, and have been formally approved by the government. The funds raised through the LCDRS are intended to be disbursed as (1) competitive grants for training providers and RPL programs in key priority areas, as determined primarily by ISCs (piloted in part under the VTGF); and (2) reimbursement for employer-provided training, which will require employers to submit evidence of training (also piloted under the VTGF).⁴ To date, implementation has focused on establishing the NTF framework and related regulations, and piloting the training procurement system through the VTGF. Implementation of the full levy collection system is expected to begin in April 2014, and full implementation of the disbursement system will follow approximately a year later.

As the logic model in Figure I.2 illustrates, the direct outputs of the NTF subactivity include establishment of the NTF council, development of regulations, and piloting of all aspects of the LCDRS framework and the system itself. In addition, capacity-building support will be provided to the ISCs, which have a critical role to play in identifying key priority areas for funding under the LCDRS. Together, these outputs are intended to result in a fully functioning LCDRS in the immediate term, in which employers are fully interacting with the system by registering, paying the levy, and applying for reimbursement for training; key priority areas are identified by ISCs and

³ The NTF will also receive a direct allocation from the Ministry of Education. This will make up approximately half of its funding, but is specifically intended to support public training institutions. The main policy change is that the NTF will now be managing these funds.

⁴ As mentioned earlier, of the funds raised through the LCDRS, approximately 35 percent are intended for grants to training providers and 50 percent for employer reimbursements. The remaining 15 percent are intended for administrative costs (GOPA Consultants, 2013).

Figure I.2. Logic Model for the NTF Subactivity



Assumptions:

- A1. NTA receives employer database from Inland Revenue Services, integrates it with the NTA database and updates it regularly.
- A2. Sufficient compliance in levy payment occurs after the LCDRS is in place.
- A3. ISC identifies key priority areas using information from the national plans, market surveys, and stakeholder input.
- A4. RPL certificate is the same and valued in the same way as a traditional vocational training certificate.
- A5. Individuals eligible for RPL, that is, have work experience but lack certifications, are aware of the RPL system.
- A6. Training is of sufficient quality and may be improved to the extent lower-quality providers are screened out of the system.
- A7. Registration and accreditation processes are fair, transparent, and effective.

Risk: Insufficient employer compliance and lack of levy enforcement.

* Procurement of training piloted under VTGF.

**See VTGF model for details of RPL pilot and operationalization

ISC = Industrial Skills Committee; NTA = Namibia Training Authority; NTF = National Training Fund; RPL = recognition of prior learning; TP = training provider.

training in these areas is procured from registered and accredited providers;⁵ and RPL candidates are assessed and certified. In the intermediate term, funds will be continually disbursed to employers (based on approved evidence of trainings) and training providers (based on meeting milestones specified in their SLAs), and more people will be trained through these mechanisms and receive RPL certificates. Combined, this will result in a skilled workforce well matched to the needs of the economy; in the long term, this will lead to improvements in employment, further training, and income, contributing to the attainment of the Compact’s ultimate goals. The key contextual factors include participation and compliance of employers, the ability of ISCs to accurately identify key priority areas, and training being of sufficient quality.⁶

3. COSDEC Subactivity

COSDECs are community-based institutions that provide basic levels of vocational training to clients from disadvantaged backgrounds—particularly out-of-school youth who lack access to the formal vocational training system—to improve their employment prospects. COSDECs offer two main types of programs: (1) national programs, which typically last between 7 and 10 months and target out-of-school youth (for example, bricklaying, plumbing, and carpentry);⁷ and (2) local programs, which typically are shorter and can be taught as center-based programs at the COSDEC itself or as outreach programs delivered in the community on an ad-hoc basis (for example, beadwork, jam-making, and basic computer literacy). The courses offered are based on annual needs assessments conducted in the entire catchment region by each COSDEC through interviews and focus groups with employers, out-of-school youth, community leaders, and other stakeholders.

The Compact is funding the construction or renovation of seven COSDECs^{8,9} and the provision of new tools and equipment in the COSDECs, as well as the construction of an arts and crafts center and a bulk store.¹⁰ Four of the COSDECs will also include a small- and medium-

⁵ All training providers must be registered and accredited to be eligible for NTF funding. The registration process is managed by the NTA, and it includes such infrastructure requirements as sufficient physical space and appropriate tools. Accreditation is managed by the Namibia Qualifications Authority (NQA) and includes similar (but not identical) requirements for registration, as well as additional requirements, such as adequate management systems and trainer qualifications. The NQA also accredits specific courses offered by accredited training providers, which must include defined competencies or “unit standards.” The requirement for training providers to be registered and accredited to receive NTF funds is intended to encourage more providers to meet these standards.

⁶ Although the quality of training is not directly targeted by the NTF, the requirement for providers to be registered and accredited to receive funds may lead to an overall quality improvement.

⁷ The national programs are equivalent to levels 1 and 2, the most basic levels of training (the NQA classifies trainings into levels from 1 to 10, representing different levels of difficulty in learning and the application of knowledge and skills). However, most of the COSDEC national programs are not yet formally accredited (although they do follow the required unit standards); accreditation of the centers is one intended outcome of the subactivity.

⁸ Technically, three of the COSDECs are new, and four are renovated. However, all seven COSDEC sites already had a COSDEC; the only substantive difference between new construction and renovation is that the newly constructed sites will be in a different physical location in the same community (in most cases, the original site will be retained to further increase the physical capacity of the COSDEC).

⁹ The only COSDEC in Namibia not included in the subactivity is the Benguela (Luderitz) COSDEC, although it is included as one of the training providers receiving funding under the VTGF subactivity.

¹⁰ The arts and crafts center in Swakopmund is conceptually distinct from a typical COSDEC. The bulk store was included through COSDEC funding but was simply intended to allow the Namibian College of Open Learning to free

enterprise (SME) unit that will provide a physical workspace, subsidized materials, and other supports to enable graduates to start their own small enterprises.¹¹ To complement these physical improvements, the Compact is funding a consultant (Transtec) to provide technical support to the Community Skills and Development Fund (COSDEF), the umbrella body that supports the COSDECs, as well as to management of the new COSDECs. The technical support for COSDEC management includes support for the development of strategic plans to enable COSDECs to become NTA-registered and NQA-accredited institutions (and offer accredited courses), improvements in financial management, improvements in the needs assessments used to determine the courses offered, and development of strategies to market the COSDECs in their catchment areas. It also includes pedagogical training for COSDEC trainers, many of whom have vocational skills and industry experience but no formal pedagogical training. To date, construction has been completed in all the COSDECs, and they are beginning to receive their new equipment (a limited number of trainings have begun using existing equipment); it is expected that all the COSDECs will be fully operational during the first half of 2014.

As the logic model in Figure I.3 illustrates, in the immediate term, the technical support to the COSDECs is expected to result in improved management practices, more relevant trainings (through improved needs assessments), increased awareness of COSDECs in the catchment area (through marketing initiatives), and adoption of NQA-accredited unit standards. In addition, the physical improvements and new equipment, together with the improved pedagogical skills of instructors, are expected to result in improved quality of trainings. The physical expansion of the COSDECs will increase access to trainings and enable COSDECs to offer additional types of training, increasing overall enrollment. In the intermediate term, the new infrastructure and tools, as well as management improvements, will enable COSDECs to be formally registered and accredited and to offer courses officially accredited by the NQA. More trainees are expected to complete training through the COSDECs and to use the SME units to help start their own enterprises. In the long term, this will improve the labor market outcomes of trainees through increased training,¹² employment, and earnings—particularly among the disadvantaged—and contribute to the ultimate Compact goals. A potential risk of the intervention is that the improved COSDEC infrastructure and the quality of the trainings offered may attract applicants with higher educational attainment, and consequently, more disadvantaged applicants may get crowded out. Other contextual factors that may affect the achievement of the outcomes include the reforms being implemented at the NTA (such as the piloting of the funding for training and procurement of training), and the overall economic conditions in the country.

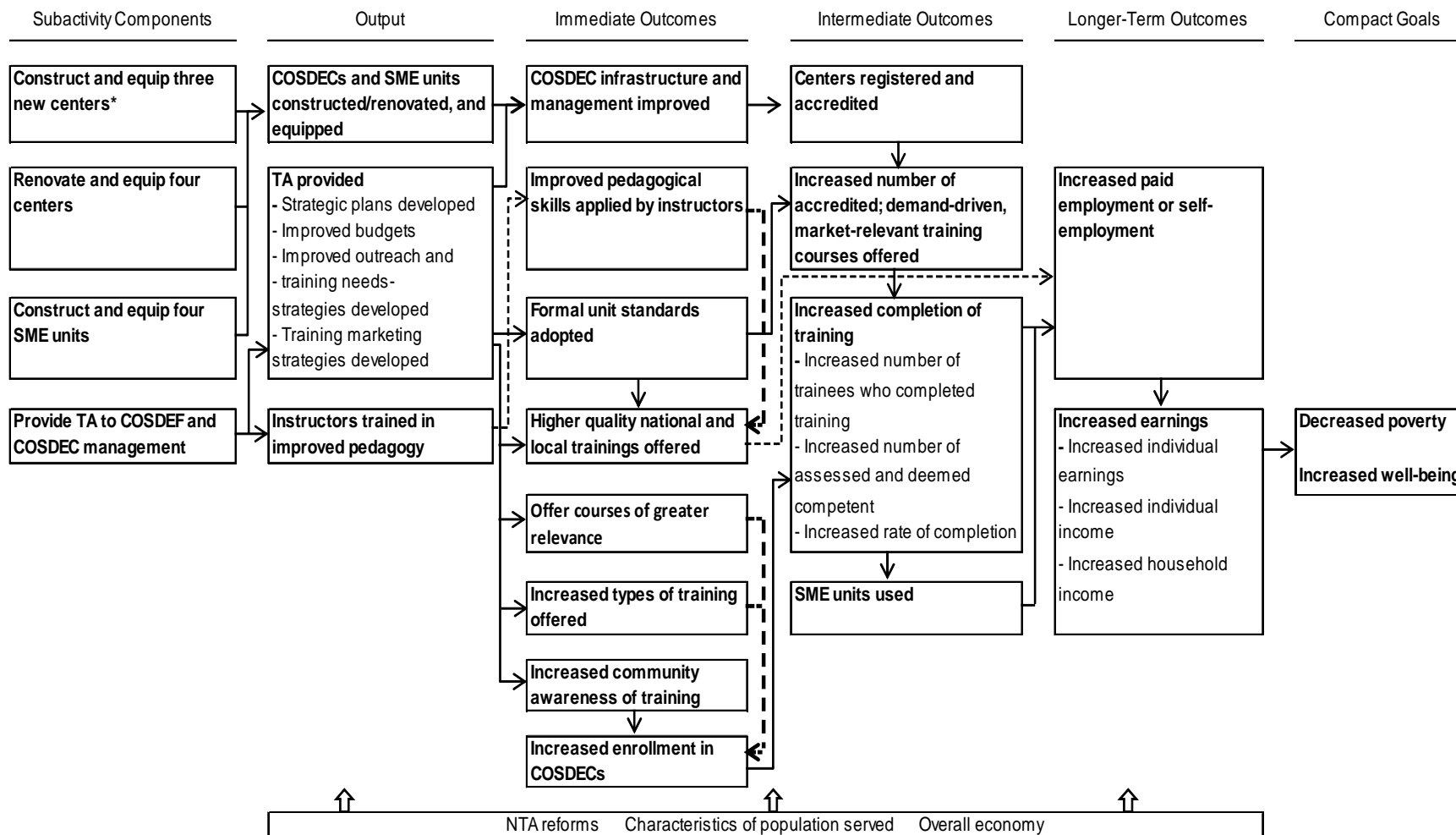
(continued)

up space for vocational training. Our evaluation will focus on the COSDECs themselves, not on these ancillary components of the subactivity; for this reason, they are not included in the COSDEC logic model.

¹¹ Although these are called SME units, their focus is on enabling trainees to use their skills to start small and micro enterprises, not medium enterprises. Therefore, a better term for them might be “micro- and small-enterprise” (MSE) units.

¹² The accreditation of COSDECs is key to enabling trainees to undertake further training beyond the very basic levels 1 and 2 offered in the COSDECs. Currently, COSDEC qualifications are not recognized by other training institutions, especially public vocational training centers (VTCs), so most COSDEC graduates cannot gain admission to higher levels of training at these other institutions.

Figure I.3. Logic Model for the COSDEC Subactivity



Assumptions:

A1. Limited availability of qualified trainers will not constrain improvements in training quality and relevance.

Risk: Applicants with higher educational attainment crowd out more disadvantaged applicants.

* A new arts and crafts center and a bulk store is also being built under the COSDEC subactivity.

NTA = Namibia Training Authority; SME = small and medium enterprise; TA = technical assistance; TP = training provider.

C. Roadmap for the Report

The rest of this report describes our planned design for the evaluation of the vocational training activity in detail, and is structured as follows:

- In Chapter II, we provide an overview of the literature on relevant vocational training programs to provide context for the evaluation of this activity and to identify gaps in the literature that the evaluation might address. We also describe the policy relevance of the evaluation in the Namibian and broader developing country context.
- In Chapter III, we provide an overview of the evaluation and detail the design for each subactivity—including the key research questions (informed by the logic models described above) and methodology.
- Chapter IV describes the quantitative and qualitative data sources and key outcomes we will use to inform our evaluation design.
- Chapter V discusses a number of issues related to the administration and management of the evaluation. These include our plans to obtain institutional review board (IRB) approval for data collection, prepare data files for broader use by policymakers and the research community, and disseminate the results of the evaluation to maximize their impact on policy. It also describes the roles and responsibilities of the members of the Mathematica evaluation team and lays out the evaluation timeline (including data collection, analysis, and reporting) in detail.¹³

¹³ We will also provide MCC with an estimated budget for the planned data collection activities one week after submission of this draft report; however, this will not be included in the report.

II. LITERATURE REVIEW AND POLICY RELEVANCE OF THE EVALUATION

A. Summary of Existing Evidence¹⁴

Although a large body of literature documents rigorous evidence on the impacts of vocational training programs in developed countries,¹⁵ the evidence in developing countries is much more limited—especially from experimental evaluations (see Tripney et al. 2013 for a recent review). The few impact evaluations in developing countries have primarily been conducted in Latin America (see Betcherman et al. 2004 and Ibarra and Rosas Shady 2009). These studies find that the impacts of vocational training programs that target youth on the key labor market outcomes of employment and wages are, on average, larger than estimates for programs in developed countries, although large variation exists both within and across countries.¹⁶ However, nearly all these studies used non-experimental methods (such as propensity score matching) to construct a comparison group, which might have led to biased estimates if individuals were selected into training based on unobserved characteristics. An additional concern is that differences in methods and data across studies could have led to widely varying results, even for the same program.¹⁷

A handful of published experimental evaluations—two in Latin America and one in sub-Saharan Africa—found mixed results. In the Dominican Republic, Card et al. (2011) conducted an experimental evaluation of a subsidized training program for low-income, out-of-school youth; they 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. The third experimental study (Blattman et al. 2014) evaluated the impact of providing cash grants to groups of poor unemployed youths in rural Uganda to help them become self-employed artisans. Grant recipients invested in vocational training (provided by local artisans or small local training institutes) and tools and

¹⁴ This section draws on a literature review we conducted at the start of the evaluation, which includes more detail and has been shared with MCC and MCA-N as a separate memo (Borkum et al. 2013).

¹⁵ 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).

¹⁶ Many of the programs for vocational training that target youth in Latin America are broadly similar in structure and are based on the “Joven” program introduced in Chile in 1990 (Ibarra and Rosas Shady 2009). They typically last approximately six months and include a classroom training component (focusing on occupational skills in demand in the labor market, as well as general skills, such as communication) followed by a work experience component at a firm. The trainings also typically provide a subsidy to participants for maintenance and transport. Many of these program features are similar to those of the VTGF subactivity in our evaluation.

¹⁷ For example, Ibarra 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.

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 non-recipients, and had substantially higher earnings (38 percent) and work hours (17 percent). There are also several additional experimental evaluations that are still underway, and have not published their final results (see, for example, Hicks et al. 2011 on a voucher program for out-of-school youth in Kenya; Maitra and Mani 2012 on a vocational training program in stitching and tailoring for unemployed women in India; 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 that we can benchmark our final evaluation results against those from other studies (although impacts are likely to vary based on program characteristics, location, and subgroups examined).

Additional evidence on implementation of vocational training programs is drawn from performance evaluations of specific programs. These evaluations often use mixed quantitative and qualitative methods and—in contrast to impact evaluations—are characterized by the lack of a rigorously defined comparison group. A recent review of the literature on youth workforce development over the past decade (U.S. Agency for International Development 2013) identified approximately 15 performance evaluations of vocational training programs in developing countries. In general, these evaluations found the programs to be at least partially successful, although the specific findings and recommendations depended on the features and context of the particular program (for examples of specific performance evaluations, see Asian Development Bank 2013 and Kelly et al. 1998).

Another relevant strand of the literature focuses on cross-country policy studies and thematic evidence reviews, and includes several studies in Sub-Saharan Africa. The most recent of these studies (Johanson and Van Adams 2004) found that reforms to the vocational training sector in the region were proceeding unevenly and skills development continued to be plagued by high costs, poor quality, and lack of connection to the labor market in many countries. Johanson (2009) conducted a cross-country review that focused on vocational training funds, many of which are funded by levies on employers and are similar to the NTF being established under the Compact. The study isolated several key factors for the success of these funds, including sufficient administrative capacity and autonomy, strong employer buy-in and participation (even in fund allocation decisions), a sufficient industrial base to generate fund revenue, and encouraging competition among training providers for funding. Two of the main challenges cited were ensuring that funds assist small enterprises (which may have limited capacity to conduct trainings of their staff), and reducing bureaucratic barriers to participation in the fund by employers.

B. Gaps in the Literature

Overall, the existing literature on evaluations of vocational training programs in developing countries—and in Sub-Saharan Africa in particular—has several important gaps. These gaps are especially apparent with regard to rigorous impact evaluations of vocational training programs. First, there have been few impact evaluations of these programs, and rigorous experimental evidence is especially limited. Given the likely variation in impacts even across developing regions and countries 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. Second, limited evidence exists on the cost-effectiveness of these programs. Only some of the existing impact evaluations included cost-benefit analyses, and the large variation in estimated

benefits across studies—even for the same program—casts some doubt on the conclusions from these analyses. Third, few of the existing evaluations collected information beyond one year after training, so that little is known about longer-term effects of these training programs. Fourth, few of the existing studies have integrated impact evaluation findings with an implementation analysis to understand the mechanisms behind estimated impacts and the reasons for any variation in impacts (which is substantial in the existing literature). Our evaluation of MCC’s vocational training activity in Namibia will make an important contribution to addressing some of these gaps and will provide valuable information for policymakers in Namibia and elsewhere that would otherwise not be available.

C. Policy Relevance of the Evaluation

The evaluation is expected to make an important contribution to policymaking in Namibia in the context of the country’s evolving vocational training sector.

As we describe in Chapter III, the impact evaluation of the VTGF will provide rigorous estimates of the effects of vocational training on the outcomes of trainees. These estimates will provide some evidence regarding the likely benefits of the similar funding mechanism that will be implemented to expand vocational training under the NTF although, as we note in Chapter III, Section E, some caution is necessary in generalizing the VTGF estimates to the NTF.

The performance evaluation of the NTF subactivity (and, to some extent, the implementation analysis conducted as part of the impact evaluation for the related VTGF subactivity) will provide valuable information to guide the improvement of the functioning and efficiency of the NTF, which will become the primary funding source for vocational training in Namibia. The VTGF and NTF evaluations will also be valuable for policymakers in Sub-Saharan Africa and developing countries more broadly, given the current paucity of rigorous evidence on the impact of vocational training programs (as discussed earlier) and the increased use of national training funds globally (Johanson 2009).

Finally, the COSDEC evaluation will yield data on the current situation in the COSDECs and inform efforts to strengthen the COSDEC network, including integrating them into the wider vocational training sector. Because COSDECs typically target disadvantaged youth, this has important implications for the government’s ability to alleviate disadvantage and marginalization among this group in access to, and completion of, vocational training.

This page intentionally blank

III. EVALUATION DESIGN

The evaluation of the vocational training activity will involve an integrated, mixed-methods approach that will include quantitative and qualitative methods. Although the three subactivities have similar ultimate goals and, together, are intended to result in comprehensive and meaningful improvements in the vocational training sector in Namibia, the specific target population, time frame, and mechanisms differ. Therefore, each subactivity will be evaluated separately (Table III.1 summarizes the evaluation approach for each subactivity). Nevertheless, we will also explore interactions between the subactivities where relevant—particularly between the VTGF and NTF subactivities, because the VTGF is intended to inform the development of the NTF. In the rest of this chapter, we describe the evaluation approach for each subactivity in detail, including the key research questions, the analytic approach, and key limitations and challenges.

Table III.1. Overview of the Evaluation Approach

Subactivity	Evaluation Approach
VTGF Subactivity	Impact evaluation with random assignment of applicants, including a quantitative impact analysis and qualitative implementation analysis
NTF Subactivity	Performance evaluation
COSDEC Subactivity	Performance evaluation, including an outcomes analysis for trainees

A. VTGF Subactivity: Key Research Questions

The evaluation of the VTGF subactivity will address eight key research questions, and related subquestions, which we have grouped into the following categories:

1. Implementation Analysis

1. Was the VTGF subactivity implemented as planned?
 - a) How did actual implementation compare to planned implementation, and what were the reasons for any deviations from plans?
 - b) What were the main challenges to implementation, and how were these addressed?
 - c) How were training providers selected for VTGF grants?
 - i. How was market demand determined?
 - ii. How did training providers determine which trainings to bid for, and what were their experiences with the bidding process?
 - iii. How, if at all, did training providers adapt (for example, in the features of their trainings or placement of graduates) to win a VTGF grant? To what extent did VTGF competition move providers towards registration and accreditation?
 - iv. How did training providers apply for capacity-building grants, and how were these grants allocated?
 - d) How did training providers identify qualified trainees?

- i. How was the income threshold for VTGF applicants determined, and how was it enforced in practice?
 - ii. To what extent did the selection process target and reach intended beneficiaries?
 - iii. What were the providers' perceptions of the trainee selection process?
 - e) To what extent did the VTGF create additional training slots for participating providers, during the subactivity, and in the future?¹⁸
2. How were the VTGF grants managed?
 - a) How well did processes and procedures for disbursing funds and reporting on milestones work?
 - b) What kind of process changes have been made since the NTA started managing VTGF grants? Has the NTA transferred any of the VTGF processes to management of the NTF funding?
 - c) To what extent has the VTGF experience sufficiently prepared the NTA to take on full responsibility for managing NTF grants?
3. What were beneficiaries' perceptions of the VTGF grants?
 - a) How, and to what extent, did the VTGF expand training opportunities available to potential applicants?
 - b) What were VTGF trainees' perceptions regarding the board and lodging allowance? Did they use it for its intended purpose? To what extent did it reduce dropouts from or increase attendance in training?
 - c) Did VTGF graduates find jobs for which they were trained? If not, why not?
4. How did employers hire VTGF graduates, and what were their perceptions of the graduates?
 - a) How do employers typically search for new workers and make hiring decisions? How did they come to hire VTGF graduates?
 - b) Do the trade-specific technical skills, and the non-technical skills of trainees meet employers' needs? What are the key skills gaps? Have employers had to retrain VTGF graduates or upgrade their skills after employment?
 - c) Overall, how satisfied are employers with the work performance of VTGF graduates?
5. Were the RPL and employer-provided training pilots implemented as planned? How did employers' perceptions of and attitudes towards their RPL-certified employees change after they became certified? How did the perceptions of employees about their job security and mobility change?

¹⁸ The creation of additional training slots that would not have existed in the absence of the VTGF subactivity—for example, through the development of new training programs, the addition of slots to existing trainings, or increasing the number of rounds of existing trainings offered—is referred to as “additionality”.

2. Impacts on Key Outcomes

6. To what extent did those offered the opportunity of training through the VTGF receive more training relative to nonfunded, qualified applicants?
 - a) To what extent did the probability of training enrollment and completion increase?
 - b) To what extent did the probability of enrollment in and completion of different types of training increase?
 - c) To what extent did skills and knowledge increase?
 - d) Did the VTGF meet its targets in terms of the number of trainees who completed training? If not, why not?
7. To what extent did the VTGF improve employment outcomes for VTGF-funded trainees relative to nonfunded, qualified applicants?
 - a) To what extent did it increase the likelihood of employment?
 - b) To what extent did it increase engagement in productive activities—including employment and further education and training?
 - c) What was the pattern of employment for VTGF trainees?
 - i. How long did it take for trainees to find employment?
 - ii. What percentage of the trainees found employment in their field of training?
 - iii. What was the pattern of job tenure?
 - iv. What was the pattern of job mobility?
 - v. Did the VTGF meet its targets in terms of the percentage of trainees finding employment? If not, why not?
8. To what extent did VTGF-funded trainees have higher earnings and income relative to nonfunded, qualified applicants?
 - a) To what extent did they have higher individual earnings?
 - b) To what extent did they have higher individual incomes?
 - c) To what extent did they have higher household incomes?
 - d) Did the VTGF meet its targets in terms of trainee income? If not, why not?
9. To what extent did increased earnings result from increased wages while employed, versus increased employment?

3. Variation in Impacts

10. Did the effects of the training funded by the VTGF vary by trainee characteristics?
 - a) Did they vary by gender?

- b) Did they vary by pretraining income?
- c) Did they vary by language group?

11. What key characteristics or practices of training providers were associated with stronger impacts on employment and earnings?

The first five research questions will analyze the implementation of the VTGF subactivity from different perspectives. In addition to providing important context for the analysis of the other research questions (particularly those on the impacts of VTGF grants on trainees), the implementation analysis will provide valuable evidence to inform the implementation of similar training mechanisms envisaged under the NTF. It will analyze the processes used by MCA-N and the NTA to identify market demand and select training providers for VTGF grants, the experiences of providers in bidding for grants, the process used by providers to identify qualified applicants (including the extent to which the program adequately targeted the economically disadvantaged), and the management of the VTGF grants by MCA-N and the NTA. The implementation analysis will also explore the VTGF training grants from the perspective of beneficiaries (in terms of the training itself and their subsequent labor market experience) and employers (in terms of whether VTGF graduates' skills meet their labor market needs). In addition to the VTGF grants for training providers, the implementation analysis will also focus on the RPL pilot and the employer-provided training pilot, which are being implemented under VTGF funding to test the systems of the NTF (question 5).

The next four research questions (questions 6 through 9) are designed to measure the effects of VTGF funding on immediate (receipt of training) and longer-term (employment and income) trainee outcomes. This will test a crucial component of the VTGF program logic: that VTGF funding led to increased enrollment in, and completion of, training and improved the labor market outcomes of trainees. This analysis will explore impacts on several measures of employment and income (see Chapter IV for details). It will also include a detailed analysis of the pattern of employment for VTGF-funded trainees, to yield a better understanding of labor market experiences post-training. We will also explore the extent to which any impacts on earnings are driven by increases in productivity (so that trainees earn more when employed) or simply by increased employment (so that fewer trainees have zero earnings); this is important in understanding the mechanisms underlying any changes in earnings that are observed.

The final two questions will explore variation in the impacts of the VTGF subactivity across different types of trainees and training providers. This will include an analysis of differences in impacts by key trainee characteristics, focusing primarily (but not exclusively) on three subgroups of particular interest to MCC and relevant in the Namibian context: gender, pre-training household income, and language group. We will also explore associations between key characteristics of training providers and provider-specific impacts to identify the characteristics most strongly related to effective trainings.

B. VTGF Subactivity: Methodological Overview

To answer the key research questions, we will use an integrated mixed-methods approach that will include a rigorous impact evaluation through a random assignment design. The quantitative impact analysis will be complemented by a qualitative implementation analysis. The impact analysis will enable us to provide quantitative estimates of impacts on the key outcomes of training, employment, and income (questions 3 through 6), and to explore the variation in impacts (questions

7 and 8). The implementation analysis will provide important context to interpret the impact estimates, and will enable us to answer questions not amenable to an impact analysis (questions 1 and 2). Table III.2 summarizes the approach taken for each research question.

C. VTGF Subactivity: Approach to the Impact Analysis

1. Impact Evaluation Design

To estimate the impact of VTGF funding on key outcomes, we will use a random assignment design. Under this design, eligible applicants for each VTGF-funded training intake¹⁹ in which the number of applicants exceeds the number of available slots are randomly assigned into two groups: those who are offered VTGF funding (the treatment group) and those who are not (the control group). Because the offer of funding is 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. Therefore, the control group can be used to estimate 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 can thus be attributed to the cumulative impact of the offer of VTGF funding up to that point. As mentioned earlier, the awarding of the VTGF grants is ongoing, and so is the random assignment process. The first random assignment for a VTGF training was conducted in late 2010, and the last will take place in early 2014.

The counterfactual in the VTGF evaluation design is driven by the choices of the control group. In the absence of VTGF funding, members of the control group can choose to: (1) not participate in any training; (2) obtain training at a VTGF-funded provider by reapplying to later VTGF trainings, or through another funding source (some, but not all, funded providers have additional non-VTGF slots available); or (3) obtain training at a non-VTGF funded provider. We do not have experimental control over this choice; the evaluation was designed to compare the treatment group to the full control group, which consists of some combination of choices (1), (2), and (3). However, because of random assignment, we would expect the treatment group to make similar choices in the absence of VTGF funding—therefore this combined counterfactual is valid. The impact estimates using this counterfactual can be interpreted as the combined impact of the VTGF training grants and capacity building grants relative to a situation in which the VTGF did not exist: the treatment group benefits from both of these, while the control group largely does not.²⁰

¹⁹ Some training providers receive VTGF funding for multiple training intakes and/or for multiple training courses for the same intake. Each group of applicants—to a specific training provider, intake, and course—is typically randomized separately. Throughout this report, we use the term “training” to refer to one of these distinct random assignment groups (a single training provider may therefore encompass multiple VTGF “trainings”).

²⁰ The exceptions are control group members who obtain training at VTGF-funded providers by reapplying to later VTGF trainings, or through other funding sources (choice (2)). These individuals would benefit directly from VTGF scholarships and/or the capacity building grant, which would not have existed absent the intervention. However, based on information on VTGF random assignment conducted to date, we believe that relatively few controls are in this group, so this is not likely to substantially change our interpretation.

Table III.2. VTGF Research Questions and Evaluation Approach

Research Question	Approach
1. Was the VTGF subactivity implemented as planned?	Qualitative implementation analysis
2. How were the VTGF grants managed?	Qualitative implementation analysis
3. What were beneficiaries' perceptions of the VTGF grants?	Qualitative implementation analysis
4. How did employers hire VTGF graduates, and what were their perceptions of the graduates?	Qualitative implementation analysis
5. Were the RPL and employer-provided training pilots implemented as planned?	Qualitative implementation analysis
6. To what extent did those offered the opportunity of training through the VTGF receive more training relative to nonfunded, qualified applicants?	
a. To what extent did the probability of training enrollment and completion increase?	Impact analysis (enrollment and completion)
b. To what extent did the probability of enrollment in and completion of different types of training increase?	Impact analysis (enrollment in and completion of different types of training)
c. To what extent did skills and knowledge increase?	Impact analysis (completion with external assessment and/or certificate)
d. Did the VTGF meet its targets in terms of the number of trainees who completed training? If not, why not?	Comparison of number of completers to program targets; qualitative implementation analysis
7. To what extent did the VTGF improve employment outcomes for VTGF-funded trainees relative to nonfunded, qualified applicants?	
a. To what extent did it increase the likelihood of employment?	Impact analysis (employment)
b. To what extent did it increase engagement in productive activities—including employment and further education and training?	Impact analysis (productive activities)
c. What was the pattern of employment for VTGF trainees?	Quantitative descriptive analysis
d. Did the VTGF meet its targets in terms of the percentage of trainees finding employment? If not, why not?	Comparison of employment to program targets; qualitative implementation analysis
8. To what extent did VTGF-funded trainees have higher earnings and income relative to nonfunded, qualified applicants?	Impact analysis (earnings and income)
	Comparison of employment to program targets; qualitative implementation analysis
9. To what extent did increased earnings result from increased wages while employed, versus increased employment?	Bound impacts on wages for those employed
10. Did the effects of the training funded by the VTGF vary by trainee characteristics?	Impact analysis disaggregated by trainee characteristics
11. What key characteristics or practices of training providers were associated with stronger impacts on employment and earnings?	Correlational analysis of training provider-specific impacts and provider characteristics

Note: For conciseness, research questions are only broken down by subquestion where the approach varies by subquestion. The implementation analysis will inform all research questions, even those addressed by other approaches.

2. Random Assignment Procedure

The random assignment procedure for the VTGF evaluation is conducted as follows. Applications for a particular VTGF-funded training are solicited, typically through advertisements in a national newspaper. Applicants must satisfy the income criterion mentioned earlier (annual household income of under N\$250,000, after subtracting training costs for other household members), as well as other criteria specified by the training provider (for example, some training providers use VTGF funding to offer scholarships for level 2 trainings, and therefore require applicants to hold a level 1 certificate). Applicants are then screened by the training provider, based on additional provider-specific criteria (for example, using symbols in the final grade 12 examinations) to identify a final pool of applicants for random assignment. Random assignment was generally conducted in Microsoft Excel by the training provider. Specifically, a random number generator is used to randomly order applicants for a particular training, and applicants are selected into the treatment group starting at the top of the list until all available training slots are filled (the number and ratio of treatment and control applicants therefore varies by training, depending on the number of applicants and the number of available slots; on average, the control group is about twice as large as the treatment group). Oversight of this process is provided by MCA-N, the NTA (for NTA-administered grants), and NORC, the firm contracted by MCA-N to collect the VTGF baseline data and the first wave of the VTGF follow-up data.

3. Analysis Plan

a. Intent-to-Treat (ITT) Estimates

Given the use of random assignment, the basic method to estimate impacts is to compare the mean outcomes of the treatment and control groups at follow-up. However, we intend to use regression models to estimate impacts because they provide greater analytic flexibility. The regression adjustment will enable us to account for design characteristics such as separate random assignment by training and for any differences in baseline characteristics that may arise by chance, as well as to improve statistical precision.²¹

We will therefore estimate impacts of the offer of VTGF funding on key outcomes using the following ordinary least squares (OLS) model:

$$(1) \quad Y_{ij} = \alpha + \beta T_{ij} + \gamma X_{ij} + \delta_j + \varepsilon_{ij}$$

where Y_{ij} is the outcome (for instance, measures of receipt of any vocational training, employment, or earnings) for individual i who applied for funding in training j ; T_{ij} is a binary variable that is equal to one for those randomly assigned to the offer of VTGF funding (treatment) and zero for those who were not (control); X_{ij} is a set of control variables that are time-invariant or are measured at baseline (such as gender, baseline education level, or baseline income); δ_j is a set of binary indicators, one for each training j in which random assignment was conducted;²² and ε_{ij} is a

²¹ We will report both unadjusted estimates and estimates that are regression adjusted by including relevant control variables. If the random assignment is successful, the two sets of estimates should be similar, but the adjusted estimates should be more precise.

²² The inclusion of δ_j in the model effectively means that the overall impact is estimated as the variance weighted average of training-specific impacts. In addition, because the probability of selection into the treatment group (which

random error term.²³ The estimate of the parameter, β , is the regression-adjusted estimate of the impact of the offer of VTGF on the outcome measure—also referred to as the “intent-to-treat” (ITT) impact.

b. Treatment on the Treated (ToT) Estimates

Motivation for ToT Estimates

In the basic analytic approach described above, we focus on estimating the impact of VTGF on those the subactivity “intends to treat.” However, it is likely that some training applicants randomly assigned to the treatment and control groups will not comply with those assignments: some treatment group members may not take up the offer of funding, and some control group members may end up receiving training.²⁴ Policymakers are likely to be interested in impacts on those who are treated—those who actually benefitted from training—in addition to impacts on those the subactivity intends to treat. The estimated impacts on those who actually received training, regardless of which random assignment group they are in, are known as the “treatment on the treated” (ToT) impacts. Next, we describe these ToT impacts and how we will estimate them in greater detail.

Estimating ToT Impacts

To estimate ToT impacts, we will adjust the ITT impacts by dividing them by the difference in the rate of training receipt between the treatment and control groups. For example, if 90 percent of the treatment group and 10 percent of the control group receive training, we would divide the ITT estimates by the difference in these rates: 80 percent (or 0.8). The ToT estimate would therefore be larger than the ITT estimate. Intuitively, the adjustment accounts for the fact that some in the treatment group do not receive training (and have a zero effect), while some in the control group receive training (and hence have a non-zero effect). Therefore, the simple comparison of those assigned to treatment and control (the ITT estimate) underestimates the impacts for those actually receiving training (the ToT estimate) and must be inflated to recover this effect. In practice, we will conduct these ToT estimates in an instrumental variables (IV) regression framework (Angrist et al.

(continued)

depends on the number of applicants and the number of slots) is likely to differ by training, we will apply appropriate weights in the analysis that account for the variation in selection probabilities. Use of the weights will increase the flexibility of our estimates—for example, these weights will enable us to report regression adjusted and non-adjusted results.

²³ In the case of binary outcomes like training receipt or employment, equation (1) is termed a linear probability model. Although probit or logit models are often used for binary outcomes, we prefer the linear probability model because it is easier to interpret and relies on weaker parametric assumptions. In practice, the probit or logit and linear probability models generally yield similar results for the types of marginal effects that we are estimating here (Angrist and Pischke 2008; Wooldridge 2002). Nevertheless, we will explore the robustness of our results to estimates that use these nonlinear models.

²⁴ This “noncompliance” with treatment assignment among the control group could arise in the VTGF evaluation because of control group members reapplying for additional trainings, or training provider using control group members on an ad-hoc basis as replacements for no-shows. We discuss the mechanisms of noncompliance among the control group further in Section C.5 of this chapter.

1996), which will enable us to control appropriately for covariates and estimate correct standard errors.

Types of ToT Estimates

We intend to present several types of ToT impacts for this evaluation (Table III.3 summarizes the impact estimates we will present and shows the corresponding comparison or counterfactual applicable for each estimate):

- **ToT estimates for enrollment and completion.** By adjusting for differences in enrollment between the treatment and control groups, we can obtain the average impact of enrollment in training. Similarly, by adjusting for differences in completion, we can obtain the average impact of completing training. We will assess whether these ToT estimates are necessary by examining enrollment and completion rates in the treatment and control groups. If noncompliance with the offer of training is low—that is, nearly all those in the treatment group enroll in training, and few of those in the control group enroll in training—the ToT estimates will be essentially equivalent to the IIT estimates. Similarly, if nearly all those who enroll in VTGF training complete the training, ToT estimates for enrollment and completion will be essentially equivalent; therefore, we will examine completion rates to assess whether separate ToT estimates for completion are necessary.²⁵ If so, the ToT estimate for completion will generally be higher than that of enrollment, because the former will include the limited impacts on those who dropped out of training. However, impacts on enrollment are arguably more policy relevant in the Namibian context, because they capture the reality of dropout.
- **ToT estimates for VTGF training and any training.** The ToT estimates for enrollment or completion could be conducted based on two definitions of receipt of training (measured as enrollment or completion): (1) receipt of VTGF-funded training, or (2) receipt of *any* vocational training. Using the first definition gives the ToT impact of receiving VTGF-funded training relative to not doing so. Those who do not receive VTGF training might not receive any training *or* could receive other trainings through other funding sources (such as government loans). If access to training through other funding sources is common, so that non-VTGF recipients are still trained, the estimated impacts of VTGF training may be small.²⁶ However, because MCC is interested in understanding the “value added” of the VTGF program relative to the existing funding sources, these impacts of receiving VTGF training are the most directly relevant for addressing MCC’s interest. Using the second definition gives the ToT impact of

²⁵ A key identifying assumption for the IV estimates to be valid is that the randomized offer of VTGF funding must have had no impact for those who do not enroll in or complete VTGF-funded training. This assumption cannot be directly verified from the data, and could be particularly problematic for ToT estimates of the impacts of training completion, because those offered VTGF funding could partially complete training and this could affect their outcomes. If many of those who enroll in VTGF funded trainings do not complete the trainings, so that separate ToT estimates for completion are required, we will explore the robustness of our IV estimates to those from other methods. These quasi-experimental matching methods involve identifying “potential completers” from the control group that are similar to completers in observed characteristics that are associated with completion, and using these to estimate the counterfactual (Wood et al. 2011).

²⁶ If estimated impacts are small, we will use the IIT estimates of impacts on the probability of receiving any training in equation (1) to explore whether high rates of training among the control group are driving the small impacts.

receiving *any* vocational training relative to not receiving it. These ToT estimates are likely to be of policy interest because they identify the impact of vocational training in Namibia more broadly.

In our analysis, we intend to produce both sets of ToT estimates: for the impact of VTGF trainings relative to other options, and the impact of vocational training relative to no vocational training. In addition, we will present separate estimates for enrollment and completion (if necessary) for each of these estimates, resulting in four different ToT estimates (Table III.3).

c. Time Frame of Analysis and Exposure Period

We will conduct the impact analysis based on outcomes measured approximately 12 months after the end of VTGF-funded training.^{27,28} This is a typical period used to evaluate impacts of vocational training programs in the (limited) existing impact evaluation literature on these programs (see, for example, Card et al. 2011 and Attanasio et al. 2011). We considered a longer initial follow-up period to allow for more time for outcomes to manifest, but we determined that this would be risky due to the increased possibility of recall error and sample attrition. In addition, a 12-month initial follow-up period would enable us to follow up with nearly all VTGF-funded trainings by fall 2015, thus including them in the analysis sample for the final evaluation report, due in mid-2016.

Table III.3. Types of Impact Estimates for the VTGF Evaluation

Estimate	Relevant Comparison	Estimation Approach
Offer of VTGF funding		
Intent to treat (ITT)	Those offered VTGF funding versus those who are not	OLS
Receipt of VTGF training		
Treatment on the treated (ToT) for enrollment	Those who enroll in VTGF training versus those who do not (can include no training or non-VTGF training)	IV
Treatment on the treated (ToT) for completion	Those who complete VTGF training versus those who do not (can include no training or non-VTGF training)	IV
Receipt of any training		
Treatment on the treated (ToT) for enrollment	Those who enroll in <i>any</i> vocational training versus those who do not	IV
Treatment on the treated (ToT) for completion	Those who complete <i>any</i> vocational training versus those who do not	IV

IV = instrumental variables; OLS = ordinary least squares.

²⁷ Due to the transition of the evaluation to Mathematica, the follow-up survey was delayed beyond 12 months for the first few VTGF trainings. In addition, because all training starting or ending at approximately the same time are surveyed together for logistical reasons, the follow-up period may also extend slightly beyond 12 months in other cases.

²⁸ In most cases, the end of VTGF funding is a clearly defined point in time. However, an exception is the VTGF grant for the Wolwedans Foundation—this grant funds levels 1 and 2 for all randomly-assigned applicants, but the best level 2 graduates are also offered funding for level 3. Because these level 3 students are purposefully selected, it is difficult to identify their counterparts in the control group. Therefore, level 2 will be viewed as the end of VTGF-funded training for this provider, and progression to level 3 will be included in the analysis as further training.

However, an important drawback to a 12-month follow-up period is that many trainees may have continued on to further training after the end of VTGF funding (with the same provider or with other providers), so that their labor market outcomes may not have fully manifested. Although we will measure impacts on a measure of “productive activities”—which includes employment and further training—to account for this (see Chapter IV), it may still dampen our estimated impacts on employment and income. Therefore, we would ideally like to conduct a second follow-up round, approximately 24 to 36 months after the end of VTGF-funded training (the exact timing would be determined in consultation with MCC).²⁹ This would provide a unique and valuable opportunity to assess longer-term impacts on labor market outcomes, which would be highly informative for policymakers and address an important gap in the existing literature.

d. Subgroup Analysis

Exploring the variation in impacts by subgroups is an important issue for the evaluation (research question 7). There are many subgroups that we could consider, provided that they are defined using time-invariant characteristics or characteristics defined at baseline. However, analyzing many subgroups increases the likelihood that spurious significant differences will be found by chance (this is known as the “multiple comparisons” problem; see Schochet 2008). Therefore, our primary analysis will focus on three key subgroups of particular interest to MCC and especially relevant in the Namibian context: gender, pre-training household income (as a proxy for disadvantage), and language group.³⁰ We also will explore variation in impacts by other subgroups (such as those defined by age or level of formal education) but will be cognizant of the multiple comparisons issue when interpreting these results. The impacts for a particular subgroup can be evaluated by restricting the estimation sample accordingly or by including appropriate interaction terms in the regression models specified above. As noted below, the precision of our subgroup estimates may be limited by small sample sizes, especially for small subgroups (for example, a small language group).

e. Additional Analyses

We will conduct the following additional analyses to enable us to answer some of the key research questions:

- **Descriptive analyses.** We will use the quantitative data for the impact analysis to describe the pattern of employment for those who completed VTGF-funded training (research question 4c). We will describe the distributions of the time to their first job post-training, whether any (or all) of their jobs since training were in their field of training, the average (or longest) tenure for which they held a job, and job mobility as measured by the number of jobs held post-training. This analysis will enable us to better understand the employment-related experiences of the trainees.³¹

²⁹ Regardless of the timing, this would require a modification of our current evaluation contract and extensive efforts to limit sample attrition over this extended period by staying in regular contact with the sample members.

³⁰ Our analysis of language groups will follow the practice and categorizations of other surveys in Namibia, such as the National Household Income and Expenditure Survey conducted by the Namibian Statistical Agency.

³¹ Some of these outcomes—specifically, the time to first job since VTGF training and employment in field of VTGF training—are not relevant for the control group. For these outcomes, we will rely exclusively on descriptive

- **Bounding estimates of impacts on productivity.** The impact of the VTGF program on productivity, as reflected by the earnings of those employed, may be informative regarding the extent to which vocational training increased the skills of trainees. However, simply comparing the earnings of those employed in the treatment and control groups may lead to biased estimates of the impact on productivity, because the composition of those employed is likely to differ between the two groups—especially because the program is hypothesized to affect the probability of employment. For example, it could be that VTGF funding enables less motivated individuals in the treatment group to complete training and enter employment, but only the most motivated in the control group are employed. In this case, simply comparing the earnings of those employed in the two groups would provide an estimate on productivity that partly reflects the difference in composition of those employed (the control group is more motivated), and would be biased downward. To address this and inform research question 6, we will attempt to impose an upper and lower bound on the productivity effect using methods from the labor economics literature (Lee 2008). Intuitively, these methods assume that the VTGF program induced some additional individuals in the treatment group into employment relative to the control group and that these are driving the compositional differences. By assuming that these additional individuals are the highest earning or lowest earning members of the treatment group, one can obtain lower and upper bounds on the impacts on productivity by “trimming” the highest or lowest earners from the treatment sample and re-estimating impacts. Depending on the parameters in the data, this can lead to tightly estimated bounds on the impact on productivity, although we will not be able to estimate the exact impact.
- **Analyzing differences in impacts across training providers.** To answer research question 8, we will estimate training provider-specific impacts and standard errors using modified versions of the regression models above (for each provider, the estimate will include all the VTGF trainings offered by that provider).³² We will then use meta-analysis regression techniques (Higgins and Thompson 2002) to estimate simple regressions of training provider-specific impacts on a limited set of relevant training provider characteristics. We will estimate both binary relationships (regressing impacts on one characteristic) to assess simple correlations, and multivariate relationships (regressing impacts on multiple characteristics) to attempt to isolate correlations accounting for other factors.³³ Relevant factors could include the type of training provider (public or private), features of the training (such as the length of job attachment, or the offer of job placement services), measures of the type of population served (using applicant

(continued)

statistics for the treatment group. However, we will have measures of tenure and job mobility in the treatment and control groups and will therefore estimate impacts on those outcomes as a secondary analysis, in addition to presenting the simple descriptive statistics for the treatment group.

³² Because we expect provider-specific impacts to be relatively imprecise due to small sample sizes, it is unlikely that we will be able to precisely estimate differences in impacts between specific providers.

³³ Because only 16 training providers are receiving VTGF funds, we will be limited in the number of factors we can include in a multivariate model. Therefore, we will focus on including independent characteristics that have strong binary correlations with the provider-specific impacts in these multivariate models.

characteristics such as years of education), trainee perceptions of quality, and other features of training provider management. Although the estimated relationships will be purely correlational and may reflect some unobserved differences across providers, they will provide useful suggestive evidence of the characteristics of providers that offered the most effective trainings.

4. Statistical Power

To determine the size of the effects that we will be able to detect given our anticipated sample size, we computed minimum detectable impacts (MDIs)—the smallest impacts that our design will be able to statistically distinguish from zero. The MDIs depend critically on our assumptions about follow-up sample sizes, key parameters (in particular, the regression R-squared), the power with which we would like to detect effects (typically 80 percent), and the variance of the outcome (which, for binary outcomes, depends crucially on the baseline level of the outcome). For ToT estimates, the MDIs also depend on assumptions about the degree of compliance with random assignment (enrollment estimates) and the fraction of those enrolled who complete training (completion estimates). Table III.4 shows MDIs for the key outcomes of employment and monthly income and notes the relevant assumptions and their motivation. To the extent possible, we calculated these MDIs using parameter estimates obtained from the baseline data received to date.³⁴

Under the current assumptions, the evaluation will be powered to detect ITT impacts of approximately 6.1 percentage points for employment (23 percent of the estimated control group mean of 26 percent) and N\$153 for monthly income (32 percent of the estimated control group mean of N\$475). The MDIs for the ToT estimates (which apply to the impact of VTGF trainings rather than any training) are higher than those for the ITT estimates, because they account for noncompliance. Because not all those who enroll in training will complete it, the ToT MDIs for training completion will be higher than those for enrollment.

As mentioned earlier, we are also interested in separately analyzing impacts for certain subgroups—for example, those defined by gender and language group. For a subgroup comprising approximately half of the full sample (such as women), the MDIs are about 40 percent higher than for the full sample. For a smaller subgroup comprising one-quarter of the full sample (such as a certain language group), the MDIs are double those of the full sample. This suggests that our ability to estimate precise subgroup-specific impacts may be limited, especially for small subgroups.

³⁴ The sample sizes we used in the MDI calculations are based on the latest information available from MCA-N. This information suggests that we would exclude from the follow-up survey: (1) trainings in which random assignment was not conducted (typically because the number of applicants did not exceed the number of slots); (2) three of the four COSDEC Benguela (Luderitz) intakes, which severely violated random assignment (we assume that random assignment will be correctly implemented for the fourth and final intake, which will enable us to include it in the study); and (3) the two final intakes for NamWater, which would require us to extend the follow-up survey well beyond the evaluation period. In addition, a handful of trainings have very large control groups; surveying a sample of control group members from these trainings at follow-up rather than the full group has a limited effect on the MDIs while substantially reducing the resources required for data collection. We therefore assume that we will draw a random sample of 150 trainees in trainings that have more than 150 available (we will account for this sampling in the analysis through sampling weights).

Table III.4. Minimum Detectable Impacts for the VTGF Impact Evaluation

	Sample Size			MDI (as percentage of baseline mean)	
	Treatment	Control	Total	Employment (percentage points)	Income (N\$)
ITT Impacts					
Full sample	930	1,170	2,100	6.1 (23%)	153 (32%)
Subgroup (50 percent)	465	585	1,050	8.6 (33%)	217 (46%)
Subgroup (25 percent)	233	293	525	12.2 (47%)	306 (65%)
ToT Impacts (Enrollment)					
Full sample	930	1,170	2,100	7.0 (27%)	176 (37%)
ToT Impacts (Completion)					
Full sample	930	1,170	2,100	7.8 (30%)	196 (41%)

Notes: MDI calculations assume a two-tailed test with a 95 percent confidence level and 80 percent power. We assume that the control group employment rate is 26 percent, control group mean monthly income is N\$475, and control group standard deviation of monthly income is N\$1,100 based on the VTGF baseline data received to date (using midpoints of income categories reported in the baseline survey). Based on information received from MCA-N, we expect to identify approximately 930 treatment cases and 1,240 control cases for the follow-up survey. Because we also expect approximately 6 percent of the controls to not be eligible for our sample because they have applied to more than one training (as describe below, we will only count an individual's first application for funding in the analysis), this leaves approximately 1,170 controls. Using these sample sizes of 930 treatment cases and 1,170 control cases, we then assume a follow-up survey response rate of 75 percent in the treatment group and 65 percent in the control group. The lower assumed response rate in the control group is because we expect their contact information to be poorer and because they may be more reluctant to respond, given that they were not selected for the program. We also assume that baseline covariates explain 10 percent of the variation in outcomes; we had earlier assumed that baseline income would explain a larger fraction of the variation, but the baseline data to date suggest that almost three-quarters of individuals report zero baseline income, limiting the ability of baseline income to explain 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) will eventually enroll in VTGF-funded training.

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

5. Adherence to Randomly Assigned Control Status

Ideally, all those assigned to the treatment group would have taken up the offer of VTGF funding, and no individuals assigned to the control group would have received funding. However, in practice, some deviations from this ideal situation (“noncompliance” with treatment status) have and will continue to occur. Noncompliance by the control group was not expected when the evaluation was designed, but could arise in practice under some circumstances. In this section, we describe these circumstances, and our plans to deal with them. As discussed earlier, in the impact analysis section, we will address noncompliance among the control group (as well as noncompliance among the treatment group) through our ToT estimates.

- **Control applicants reapply for additional trainings.** In some cases, applicants for VTGF funding who participated in random assignment but were not selected into the treatment group reapplied for additional VTGF-funded trainings (typically for additional

trainings offered by the same provider). Some of these control group members may have therefore been subsequently selected for the treatment group in other trainings. Our current understanding is that this was limited to approximately 6 percent of applicants.³⁵ In the data collection and analysis, we will treat these individuals according to their original randomly assigned status from the first training they applied. That is, our IIT analysis will be based solely on first time applicants for each training, and noncompliance through application to additional trainings will be accounted for in the ToT estimates. Our approach of focusing on first-time applicants has the advantages of (1) preserving the integrity of the random assignment for first-time applicants (because first-time applicants are a subgroup of all applicants, random assignment should be valid for them even if the random assignment included some who were applying again); (2) simplifying the data collection (because the timing of data collection is linked to training and individuals will be clearly linked to a particular training); and (3) avoiding other complex analytic adjustments (such as additional reweighting to account for different probability of selection for those who applied more than once).

- **Training providers use replacements.** Because not all those initially selected for VTGF funding took up the offer, training providers typically selected replacements from the control group to fill the available funding slots, using one of three methods. The first method was to work down the randomly ordered list of applicants initially assigned to the control group until enough treatment group members were obtained. In this case, we will treat the replacements as part of the initially assigned group—conceptually, this is equivalent to the training provider having decided to select more people in the first place. The second method was to re-randomize the control group to select the additional numbers required to fill the available slots. In this case, we will include the result of this re-randomization round in determining treatment status, but will account for the different probabilities of selection by reweighting. The third method was to select replacements from the control group in an ad-hoc manner (to the best of our knowledge, this has happened in very few cases so far). In these few cases, we will rely on the original random assignment, and we will address the noncompliance through the estimation of ToT impacts.

D. VTGF Subactivity: Approach to the Implementation Analysis

The qualitative implementation analysis for the VTGF evaluation will rely on information collected from relevant stakeholders (see Chapter IV for details of the data collection). These data will include reviews of administrative records and in-depth interviews with MCA-N, the 16 training providers who received VTGF grants, the GOPA consultants who supported the NTA in implementing the RPL and employer-provided training pilots, and NTA staff (including those responsible for managing VTGF grants and the two pilots). We will also conduct focus groups or

³⁵ An exception is the Benguela (Luderitz) COSDEC, in which nearly half of the initial applicants who were assigned to the control group for the first three intakes were included in the randomization for subsequent intakes. We understand that many of these re-randomized controls did not actually apply to subsequent intakes, but were included at the initiative of the training provider. Because of these severe violations of random assignment, we intend to exclude the first three intakes of this training provider from the VTGF analysis (we intend to include the fourth and final intake, provided that random assignment is implemented correctly). Given the large number of applicants and trainings, we expect the effect on statistical power to be limited.

interviews with trainees (depending on whether trainings are still ongoing) and interviews with control group members in all training providers. This will provide additional insights on experiences with training, other training opportunities, and experiences in finding employment and applying their skills beyond the information gathered in our limited quantitative surveys. In addition, we will interview employers of VTGF trainees to better understand the employment process and how trainees are implementing their skills in the labor market. We will also review the Compact Completion Report (CCR) that MCA-N will produce toward the end of the Compact, which will include an assessment of the implementation of the VTGF subactivity from the MCA-N perspective. We will triangulate the data from these data sources by systematically coding data and categorizing and sorting these data appropriately to identify key themes and patterns in the responses that will inform the key research questions.

Illustrative themes we will focus on as part of the implementation analysis include the following:

- How the VTGF grants, and RPL and employer-provided training pilots were implemented in comparison to original plans, and reasons for deviations from plans
- Key implementation successes and factors for success
- Key implementation challenges faced and how these were addressed
- Changes made by training providers (for example, in their trainings or job placement processes) as a result of the VTGF bidding process
- Training providers' perceptions of the VTGF grants
- The role of the VTGF grants in increasing access to training relative to existing opportunities
- Experiences and satisfaction of trainees with training
- The labor market experiences of trainees and the control group, including the extent to which their skills match the needs of the market
- Employers' perceptions towards the grant-funded trainees
- Employer's perceptions towards the RPL program
- Trainees' expectations about future training, employment, and earnings
- The capacity of the NTA to manage grant funds and how this evolved

E. VTGF Subactivity: Limitations and Challenges

Although our design offers the best possible opportunity to provide rigorous evidence to inform the key research questions, it has some limitations and may face certain challenges. Here, we discuss these potential limitations and challenges, and our plans to address them:

- **Limited control groups for some trainings.** In some cases, the number of applicants for a particular training may have been only slightly higher than the number of slots,

resulting in a small control group for that training (or, in a few cases, no control group).³⁶ If this situation was widespread, our ability to precisely estimate impacts through the random assignment design would be limited. However, our review of the random assignment data from the 25 trainings in which randomization has been conducted to date suggests that, although seven of the trainings had a control group of fewer than five people, the applicants in these trainings only accounted for around 5 percent of the total sample. Therefore, we do not expect the exclusion of these trainings to have a substantial effect on sample sizes and, hence, on statistical precision.

- **Differential response rates to the follow-up survey.** Substantially different responses for the treatment and control groups at follow-up, or a low overall response rate, could invalidate the balance between the treatment and control group introduced by random assignment. For example, it could be that response rates in the treatment group are high, but the people in the control group with poor labor market outcomes are transient and difficult to contact. In that case, the impact estimates might reflect the differences in characteristics for the follow-up group—in the example, members of the control group with better outcomes would be more likely to respond, dampening the estimated impacts. Similarly, even if response rates are similar, a high overall rate of nonresponse might raise concerns that nonrespondents are systematically different between the treatment and control groups, leading to bias. For the VTGF follow-up tracer survey, we do expect nonresponse to be a challenge because of the common practice of changing cell phone numbers frequently and because of inaccurate contact information for applicants (especially for the control group). We will attempt to address this in both the data collection and analysis. In the data collection, we will use several methods to ensure high response rates in the follow-up survey, as described in Chapter IV. As part of the analysis, we will assess whether the follow-up treatment and control samples are equivalent in observed baseline characteristics. If there are significant differences, we will control for these observed differences through regression-adjusted estimates or use nonresponse weights to make respondents more similar to the (balanced) original sample. If low response rates are driven by a small number of providers (perhaps due to variation in the quality of contact information), we could also present some estimates excluding these providers (although with some loss of generalizability).
- **Substitution of work experience for training by the control group.** Our impact estimates involve comparing the outcomes of the treatment and control groups in the 12-month post-training window, testing the hypothesis that the treatment group experiences better labor market outcomes in this window. However, if members of the control group use the training period to gain work experience that can substitute for formal training, this could benefit their subsequent employment and income outcomes—resulting in small estimated impacts. Although the impact estimates would still be valid, it is important to explore whether this is happening, so that we can interpret the results correctly. Therefore, we will ask respondents about employment in the period of training and will compare these patterns between the treatment and control groups.

³⁶ Because the overall impact estimates effectively combine training-specific impacts, trainings with no control group will not contribute to these estimates. Trainings with a small control group will only contribute if we can follow up with at least some of the control group; however, the variance of the estimates for these trainings will be high, so their contribution to the overall estimate will be low.

- **Limited ability to estimate differences in impacts by training characteristics.** The impacts of VTGF-funded trainings could vary based on the characteristics of the training provider (for example, whether they are public or private) or the trainings themselves (for example, the curriculum or duration of training). As described above, we will be able to provide estimates of correlations between training-specific impacts and training characteristics using meta-analysis regression techniques (Higgins and Thompson 2002). However, these estimates should not be viewed as rigorous; the evaluation was designed to rigorously test for the overall effectiveness of training, and not the relative effectiveness of different types of training.
- **Possibly limited external validity.** The ITT impact estimates apply to a specific group: those who met MCC and the VTGF training provider's requirements and were interested in applying for VTGF funding. Similarly, the ToT estimates apply to a subset of this group: those who actually enrolled in or completed training. Therefore, the estimates do not necessarily represent impacts on broader populations, such as unemployed youth in Namibia, that could be drawn into vocational training by further increases in access. To assess the possible extent of external validity, we will compare the characteristics of our sample (such as education, language group, and income) to those of a similarly aged out-of-school population in the National Household Income and Expenditure Survey. If the two samples are similar in observed characteristics, this would increase our confidence that the impact estimates are more broadly applicable. Even so, our estimates apply to the VTGF funding program as implemented by MCC in the current vocational landscape in Namibia. Future reforms to the vocational training sector (for example, a broader shift toward demand-driven trainings throughout the vocational training sector through the NTF) or general changes to the economy in Namibia could affect the impacts of future programs. The generalizability of the VTGF impact estimates to the impacts of the NTF is also unclear. Despite the similarities between the two subactivities, there are potentially also some differences between them (which we described in Chapter I) that suggest that the NTF impacts could differ. As part of the implementation analysis of the NTF subactivity, we will explore the similarities and differences between it and the VTGF subactivity, which will provide some suggestive evidence of the extent to which the impacts on trainees are likely to be similar.

F. NTF Subactivity: Key Research Questions

The evaluation of the NTF subactivity will address three key research questions, and related subquestions, which we have grouped into categories:

Establishment and Initial Operations

1. Was the establishment of the NTF levy collection, distribution, and reporting system implemented as planned?
 - a) How did actual implementation compare to planned implementation, and what were the reasons for any deviations from plans?
 - b) What were the main challenges to implementation and how were these addressed?
 - c) How was the transition from the piloting to the fully operational stage of the LCDRS managed? What were the successes and challenges related to the transition? Was it well-functioning when handed over to NTA for management?

- d) What was the NTA's readiness to manage the NTF, and to what extent was this influenced by their involvement with the VTGF?
2. How is the NTF levy collection and distribution system operating in practice compared to the specifications outlined in the regulatory framework?
 - a) What is the compliance rate?
 - b) What enforcement mechanisms are in place?
 - c) How does the NTF measure and respond to market demand for skills?
 - d) What types of providers apply for training?
 - e) How is levy funding allocated in practice across different disbursement streams and specific providers within each stream?
 - f) How is levy funding allocated in practice across different sectors and employers of different sizes?
 - g) Does the NTA assess the quality of training offered and, if so, how? How do assessments or perceptions of quality affect the allocation of funding, if at all?
 - h) What types of reports are generated through the LCDRS and what is the perceived utility of those reports?
 - i) What are the costs of administering the system?
 - j) What is the role of the private sector and civil society in the system?
 - k) What are stakeholder's perceptions of the performance of the LCDRS, and how and why have these changed over time?
 - l) What are the ongoing successes and challenges?

Perceived Sustainability

3. What are the stakeholder perceptions of sustainability of the NTF levy collection and distribution system?
 - a) Is the system perceived to be financially sustainable?
 - b) Is the system perceived to be sustainable in terms of organization and management?
 - c) Are employers satisfied with the role of the NTF, with respect to in-service training and employees trained in key priority areas?

These research questions focus on three important junctures in the evolution of the NTF's levy collection, distribution, and reporting system. These are the system's establishment (question 1), initial operations (question 2), and likely sustainability (question 3). The first question focuses on analyzing key implementation strengths and weaknesses during the establishment phase, including the critical transition from the pilot phase to full operations. Understanding these strengths and weaknesses and how any challenges were overcome can inform future program implementation and sustainability.

The second question explores the operations of the LCDRS and will focus on the levy collection and distribution components. Analysis of the collection system will focus on compliance among eligible employers and the enforcement mechanisms used by the NTA, including the extent to which these mechanisms are applied. Analysis of the distribution system will focus on allocation of funding across the disbursement streams for which it was intended, how employers and training providers interact with the system to request funding, and how funding decisions are made. More broadly, we will examine how civil society players in the vocational training sector (such as other donors) perceive the system, as well as the perceptions of other stakeholders, such as employers of different sizes and types. We will also assess the ongoing successes and challenges of the LCDRS, which has implications for its future sustainability.

Given the timing of the evaluation, our ability to assess the sustainability of the LCDRS and address the third research question—which is critical to eventually achieving the outcomes described in the logic model—will be limited. In particular, although the levy collection system is expected to begin operations in April 2014, the distribution component will not be fully functioning until one year later, in 2015—only a few months before we will be collecting data on the performance of, and experience with, the distribution component. Nevertheless, early impressions of sustainability could provide insights into whether longer-term sustainability is likely. Therefore, we will investigate early perceptions of sustainability across a range of stakeholders, including the NTA, employers, and training providers. The sustainability analysis will focus on financial sustainability (fund collection and disbursement) and on whether the NTF has the organizational and management capacity to ensure that the LCDRS operates as envisaged. We will also focus on employers' satisfaction with their interactions with the NTF, as well as their intended engagement with the LCDRS in the future—a key assumption in the logic model.

G. NTF Subactivity: Overview of the Performance Evaluation and Key Challenges

The NTF subactivity will be evaluated through a performance evaluation, which will consist of a qualitative implementation analysis that will address all the key research questions. As noted earlier, this will focus on the establishment of the LCDRS, its initial operations, and stakeholder perceptions of the levy systems' sustainability. The performance evaluation will rely on many data sources, including administrative data and in-depth interviews with NTA officials, MCA-N staff, the GOPA consultants who are supporting the NTA in establishing the LCDRS, employers who are interacting with the levy system, and other affected stakeholders (such as training providers). We also intend to conduct interviews with, and observations of, the ISCs, as well as observations of NTA operations (Chapter IV provides additional details on the key stakeholders from which we will collect data and the types of data we will collect). We will also review the implementation analyses in the CCR produced by MCA-N.

The performance evaluation will again triangulate the data from these data sources, identifying similarities and differences in perspectives. Illustrative themes the performance evaluation will focus on include the following:

- How the NTF subactivity was implemented in comparison to original plans, and reasons for deviations from plans
- Key implementation successes and factors for success
- Key implementation challenges faced and how these were addressed

- The role of the GOPA consultants in developing the LCDRS and the nature of their interactions with the NTA and other stakeholders
- The role of the ISCs in identifying key priority areas for training, and how this process works in practice
- Perceptions of stakeholders toward the LCDRS and their interactions with the system (especially employers)
- How and to what extent the NTA applied the lessons learned from the VTGF in managing the NTF
- How the NTF levy system is being managed in the post-Compact period
- What has been done to develop the authority and capacity of the NTA to bring the elements of the vocational training system into alignment
- Perceptions of sustainability of the NTF levy system, including financial, managerial, and organizational sustainability

The performance evaluation of the NTF subactivity will also face some challenges. These include:

- **Timing of implementation and analysis.** The evaluation schedule may not allow us enough time to observe and assess a mature and fully operational LCDRS. Although we will rely on perceptions of LCDRS operations in the future, these may turn out to be inaccurate.³⁷
- **Lack of quantitative data on impacts.** Given the lack of quantitative data of a representative sample of trainees or potential trainees across Namibia, it will be difficult to assess the quantitative impacts of the NTF subactivity—even in a nonrigorous manner. Given the similarities between the VTGF and NTF subactivities, the results from the VTGF impact evaluation may provide some evidence on the possible impacts of the NTF subactivity—although, as noted earlier, the VTGF results may not be directly generalizable to the NTF.

H. COSDEC Subactivity: Key Research Questions

The evaluation of the COSDEC subactivity will address six key research questions, and related subquestions, which we have grouped into categories:

Implementation Analysis

1. Was the COSDEC subactivity implemented as planned?
 - a) How did actual implementation compare to planned implementation, and what were the reasons for any deviations from plans?

³⁷ To capture the operations of a fully mature LCDRS would require an additional round of data collection approximately 12 to 18 months after the last currently planned round. This would extend far beyond our current contract, although we are open to exploring the possibility of including an additional round if MCC is interested.

- b) What were the main challenges to implementation and how were these addressed?

Trainee Outcomes

2. To what extent did the COSDEC subactivity increase the availability of training?
 - a) What types of trainings did COSDECs offer, how was this determined, and how did this affect training accessibility?
 - b) What percentage of trainees completed the different levels of trainings offered?
3. How did COSDEC training affect the employment outcomes of trainees?
 - a) What was the pattern of employment for trainees?
 1. How long did it take for trainees to find employment?
 2. To what extent did trainees enter self-employment versus paid employment?
 3. To what extent did trainees find employment in their field of training?
 4. What was the pattern of job tenure for trainees?
 5. What was the pattern of job mobility for trainees?
 - b) What was the role of SME support in the effects of the trainings on self-employment?
 - c) To what extent were increases in employment likely to be sustained?
 - d) To what extent did COSDEC trainees engage in further training?
4. How did COSDEC training affect the earnings and income of trainees?
 - a) What were the patterns of earnings and income for trainees?
 - b) To what extent are increases in earnings and income likely to be sustained?
5. Did the effects of COSDEC trainings on employment and earnings vary by trainee characteristics?

COSDEC Management

6. How were the new and renovated COSDECs managed?
 - a) What management practices were being applied, and were they likely to change in the future?
 - b) Are the new COSDECs financially sustainable?
 - c) Did COSDECs made progress toward adoption of unit standards and accreditation, and did this affect “articulation”?

The first research question will analyze the implementation of the COSDEC subactivity, including both the construction and technical assistance components. In addition to analyzing these components separately, the implementation analysis will explore the interaction between them, because they are intended to work together to achieve the outcomes outlined in the logic model. The

implementation analysis will also provide important context for the interpretation of the analysis of the remaining research questions.

The next four questions (questions 2 through 5) focus on the trainings offered by the new COSDECs, and seek to describe both the trainings offered and the outcomes of trainees. Question 2 looks at the supply and demand aspects of training offered. On the supply side, we will investigate the types of trainings offered (including national courses, local courses, and community outreach courses) and their content, how the mix of courses was determined to meet community needs and market demand, the extent to which the new and renovated COSDECs increased access to training in the areas served, and how the subactivity affected the admissions process and the types of individuals who applied and enrolled. On the demand side, we will explore the extent of demand (or oversubscription) for various types of trainings, and completion rates among those who enroll. Questions 3 and 4 link to the outcomes in the logic model, and seek to describe the employment, earnings, and income patterns of trainees. In addition to describing these patterns, we will explore related issues that are important mechanisms in the logic model: the role of the new SME units in self-employment and challenges faced in this regard, the extent to which COSDEC graduates continue to further training, and the extent to which post-training employment and earnings patterns are likely to be sustained. Question 5 looks at the variation in the patterns of employment and earnings by trainee characteristics, which could be informative regarding the types of trainees who are benefitting the most from training.

The final research question seeks to describe the management practices implemented in the new COSDECs, which are likely to affect the outcomes of trainees. We will focus on the types of practices that were included in the technical assistance component on the subactivity, and the extent to which these were implemented and are likely to be sustained. The sustainability analysis will also look at financial sustainability (for example, by exploring COSDECs' funding plans for maintenance and operations of the new buildings, equipment, and SME units). As described in Chapter I, the technical assistance component includes support for COSDECs to develop a plan to be registered and also accredited within the national qualifications framework, thus integrating them into the national vocational training system. This is an important part of the program logic, because this may enable articulation of COSDECs with the VTCs and other institutions of higher education and training, thereby creating opportunities for COSDEC trainees to continue to further training beyond the basic levels offered at the COSDEC. Therefore, we will explore the extent to which COSDECs were able to make progress toward registration and accreditation and whether trainees were able to progress to further training (especially in the VTCs).

I. COSDEC Subactivity: Overview of the Performance Evaluation

To answer the key research questions, we will conduct a detailed performance evaluation. This will include a qualitative implementation analysis that will triangulate information from different sources, as well as a quantitative outcomes analysis based on a tracer survey of COSDEC trainees.³⁸

³⁸ Although we explored the possibility of conducting a rigorous impact evaluation of the effects of COSDEC trainings by identifying a comparison group of nontrainees, we determined that this was not feasible, for several reasons. First, although demand for COSDEC places exceeds the supply, implementing a selection mechanism such as random assignment to form a comparable control group would be difficult, given that each COSDEC is managed independently and follows its own selection approach. Second, even if COSDECs did implement such a selection mechanism, those who are not admitted often reapply for subsequent intakes, contaminating the control group. Restricting controls from reapplying for a period of time to avoid this situation would not be compatible with the mission of the COSDECs to

The qualitative implementation analysis will inform all the research questions, and the outcomes analysis will be particularly useful for questions 2 through 5 (Table III.5 summarizes the approach for each question). Our primary motivation for including a quantitative outcomes analysis is that existing quantitative evidence on the labor market outcomes of COSDEC trainees is extremely limited. Although the COSDEF does conduct occasional tracer surveys, these have very low response rates. This evidence would therefore be highly valuable to the COSDECs, COSDEF, and policymakers in Namibia. In addition, we hope that by working closely with the COSDEF in the design and conduct of the tracer survey, we can improve their capacity to conduct similar, high quality tracer surveys in the future. Next, we describe the analytic approach to the outcomes analysis and qualitative implementation analysis.

Table III.5. COSDEC Research Questions and Evaluation Approach

Research Question	Performance Evaluation Approach
1. Was the COSDEC subactivity implemented as planned?	Qualitative implementation analysis
2. To what extent did the COSDEC subactivity increase the availability of training?	
a. What types of trainings did COSDECs offer, how was this determined, and how did this affect training accessibility?	Qualitative implementation analysis
b. What fraction of trainees completed the different levels of trainings offered?	Quantitative outcomes analysis
3. How did COSDEC training affect the employment outcomes of trainees?	
a. What was the pattern of employment for trainees?	Quantitative outcomes analysis
b. What was the role of SME support in the effects of the trainings on self-employment?	Qualitative implementation analysis; quantitative outcomes analysis (use of SME unit)
c. To what extent were increases in employment likely to be sustained?	Qualitative implementation analysis; quantitative outcomes analysis (employment after 12 months)
d. To what extent did COSDEC trainees engage in further training?	Quantitative outcomes analysis (enrollment in further training)
4. How did COSDEC training affect the earnings and income of trainees?	
a. What were the patterns of earnings and income for trainees?	Quantitative outcomes analysis (earnings and income)
b. To what extent are increases in earnings and income likely to be sustained?	Qualitative implementation analysis; quantitative outcomes analysis (earnings and income after 12 months)
5. Did the effects of COSDEC trainings on employment and earnings vary by trainee characteristics?	Quantitative outcomes analysis disaggregated by trainee characteristics
6. How were the new and renovated COSDECs managed?	Qualitative implementation analysis
Note:	For conciseness, research questions are only broken down by subquestion where the approach varies by subquestion. The qualitative implementation analysis will inform all research questions, even those addressed primarily by the quantitative outcomes analysis.

(continued)

reach those in most need of training. Third, given the wide catchment areas of the COSDECs, identifying a plausible comparison group for a quasi-experimental design would be challenging and costly.

J. COSDEC Subactivity: Analytic Approach to the Performance Evaluation

1. Outcomes Analysis

a. Descriptive Analysis

Without a counterfactual, the quantitative outcomes analysis based on tracer survey data will be descriptive. For each of the relevant outcomes—which will largely mirror those for the VTGF evaluation and are described in Chapter IV—we will describe the distributions in detail. We will present both graphical and numerical descriptions of these distributions for COSDEC trainees as a whole, and for groups of trainees defined by course type (national or local programs, or course content), or by COSDEC.

We will also describe the baseline characteristics of trainees to the extent possible, using information gathered from the trainee application form (for example, education level and gender) and time-invariant information from the tracer survey (for example, language group). Because we anticipate having to contact trainees at some point between their enrollment and the follow-up survey to stay in touch and update their contact details to achieve high response rates at follow-up, we are exploring the possibility of asking a small number of questions during this contact. If this is feasible, we will also be able to describe additional baseline characteristics of trainees (such as baseline training level, employment, and income), which will help COSDECs get a better sense of their intake pool.³⁹

b. Time Frame and Period of Exposure

The tracer survey will document the training outcomes, employment, and income of COSDEC enrollees in the 12-month period since the completion of training. Similar to the VTGF analysis, this 12-month period was selected to balance the desire of observing long-term outcomes with the risk of increased sample attrition and the need to report results within the time frame of the evaluation. Again, we note that a 12-month follow-up period is typical in the vocational training literature. Although a second, longer-term follow-up would also be valuable, we determined that the resources for a second round would be better allocated to the VTGF evaluation (if at all), given the more rigorous design for that evaluation.

c. Subgroup Analysis

To answer research question 5, we will break down our descriptive statistics of key outcomes by subgroup. Unlike the VTGF evaluation, however, we will have relatively limited information on baseline characteristics because we do not intend to conduct a full baseline survey. Therefore, characteristics for subgroup analysis will have to be drawn from the trainee application form, the limited questions that we may ask during our contact calls with trainees, and any characteristics measured at follow-up that are time-invariant. Relevant characteristics will include gender, language group, baseline education level (all available from either the application form or the follow-up survey) and (only if available from the contact call), baseline employment and income.

³⁹ Because of the purely descriptive nature of this analysis, we determined that the resources required to conduct a full baseline survey would not be justified.

d. Sample Sizes

We expect each new COSDEC to accommodate approximately 150 trainees in the mid 2014 intake, the intake we will focus on for the outcomes analysis (as described in Chapter IV). If this happens, we expect to be able to include all trainees in the tracer survey so that we can estimate parameters for the entire population of interest. If the number of trainees is larger than expected, or if the costs of the survey exceed the available resources, we will consider drawing a random sample from each COSDEC. If we do so, we will ensure that our sample sizes are sufficient to provide precisely estimated parameters (for example, precisely estimated means), both for the overall sample and for subgroups. For example, with a sample of 100 per COSDEC, we would be able to estimate a 95 percent confidence interval of approximately 2 percentage points either side of the estimated mean employment rate. If sampling is required, we would likely allocate the sample proportionally among the different types of programs (national and local) within each COSDEC, to ensure a representative sample of trainees.

2. Qualitative Implementation Analysis

As mentioned above, we will complement the quantitative outcomes analysis with a qualitative implementation analysis of the COSDEC subactivity. This implementation analysis will focus on the technical support provided to the COSDEF and the COSDECs, as well as the new COSDEC infrastructure, and will triangulate information from a variety of data sources. Key data sources will include in-depth interviews with MCA-N, the Transtec staff who provided technical support, and the COSDEF staff directly supporting the COSDECs. To provide a more detailed understanding of their experiences with COSDEC training and on the labor market and their expectations for the future, beyond the more limited quantitative information we will collect for the outcomes analysis, we will also gather qualitative information through focus groups and interviews with COSDEC trainees. In addition, we will conduct in-depth interviews with COSDEC staff, including center heads and instructors, which will enable us to better understand the management and operations of the new COSDECs and the sustainability of their operations (Chapter IV describes the data sources we will rely on in further detail). Similar to the other subactivities, we will also review the implementation analysis in the CCR.

Illustrative themes we will focus on as part of the implementation analysis include the following:

- How the COSDEC subactivity was implemented in comparison to original plans, and reasons for deviations from plans
- Key implementation successes and factors for success
- Key implementation challenges faced and how these were addressed
- The role of Transtec in providing technical support, and its interactions with the COSDEF and COSDECs
- The nature of COSDEC interactions with community stakeholders, including employers and potential trainees
- The extent to which COSDECs made progress toward NQA accreditation, and whether this has affected articulation of trainees to further training
- The use and role of the SME units in the new COSDECs

- Experiences and satisfaction of trainees with training
- The labor market experiences of trainees, including the extent to which their skills match the needs of the market
- Trainees’ expectations about future training, employment, and earnings
- Management and instructional practices in the new COSDECs
- Perceptions of sustainability of the new COSDECs, including financial and managerial sustainability

K. COSDEC Subactivity: Limitations and Challenges

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

- **Absence of a counterfactual.** Our design for the COSDEC evaluation does not include a counterfactual—that is, we cannot determine what the labor market outcomes of trainees would have been without COSDEC training. We emphasize that our outcomes analysis is purely a descriptive exercise and is *not* an estimate of the impact of the COSDEC trainings. That would require a rigorous impact evaluation, which we determined was not feasible in this context.
- **Sustainability of outcomes.** For reasons noted above, we intend to conduct only a single round of the tracer survey (12 months after the completion of training). Consequently, we will have limited ability to assess the long-term sustainability of employment and earnings patterns that we observe. We will have to rely on the patterns observed within the time frame of the evaluation (for example, whether employment is sustained within the 12 months post-training period), combined with qualitative evidence on perceptions of sustainability.
- **High attrition rates.** Our ability to provide quantitative evidence on trainee outcomes that is generalizable to the COSDEC trainee population depends on achieving high response rates. Otherwise, there may be a concern that only certain types of trainees—likely those with better outcomes—are appearing in the data and driving the findings. We will use several strategies to ensure high response rates, including working closely with the COSDEF and the COSDECs to ensure that correct contact information is captured as part of the application form, and contacting trainees between enrollment and follow-up to stay in touch and update their contact information. In the analysis, we will also be able to get a sense of a lower bound on some outcomes by assuming a “worst case” scenario—for example, by assuming that all those who do not respond are unemployed.

L. Cost Analysis

The evaluation of the vocational training activity will provide a valuable opportunity to investigate the costs of training provision in Namibia, on which—in our understanding—there is limited existing data. Specifically, we could use information about training provider costs that were required as part of training providers’ VTGF applications in a cost analysis. This will rely exclusively on cost data, and would be a primarily descriptive analysis of the costs of different types of trainings and the main components of these costs. We could also conduct exploratory analyses that compare

costs for different types of training providers (based on provider characteristics such as whether they are public or private, level of the training, trainee perceptions of quality, or other features of training provider management). To successfully conduct the cost analysis, training providers will have to provide cost data for different types of costs (for example, labor cost, overhead, stipend to trainees) by training. It is possible that such cost data may not be readily available from all training providers who received funding under the VTGF subactivity.

If possible, we also intend to conduct cost analyses for the other two subactivities. For the NTF subactivity, we could use data from the unit cost estimator tool used to measure the cost of trainings, which is being developed by the NTF with GOPA's support. As a first step, we will seek to better understand this tool and how it is used, as part of the implementation analysis. For the COSDSEC subactivity, we could conduct a similar cost analysis as for the VTGF subactivity, provided that the necessary cost data are made available to us (these data may have been generated as part of Transtec's technical support in budgeting). Again, given the limited information currently available on training costs, this information would be valuable to the COSDEF and policymakers in Namibia.

IV. DATA SOURCES AND OUTCOMES

A. Quantitative Data Sources and Outcomes

We intend to collect quantitative survey data for the evaluation of the VTGF and COSDEC subactivities. Here, we describe the key data sources we will use to obtain these data, as well as the key outcomes for the analysis.

1. VTGF Data Source

We will obtain quantitative data for the evaluation of the VTGF subactivity through a longitudinal tracer survey of all qualified VTGF applicants who were randomized for the evaluation, including VTGF-funded trainees and control group members.⁴⁰ This tracer survey will be administered at two points for each applicant: once at baseline and once at follow-up:

- **The baseline survey** is administered after random assignment, ideally close to the start of training. The survey instrument was developed by MCA-N, MCC, the previous evaluator, and NORC, and covers such topics as respondents' demographic information, educational background, current employment, current income, and contact information. The baseline tracer survey is being implemented locally by Survey Warehouse, the data collection firm subcontracted by NORC. Information from the baseline will be used primarily to verify baseline equivalence of the treatment and control groups, provide control variables that will increase the precision of the impact estimates at follow-up, and define subgroups for analysis (for example, based on pretraining household income).
- **The follow-up survey** will be administered approximately 12 months after completion of the training program to which respondents first applied. This survey is being developed by Mathematica and will be implemented by the Survey Warehouse-NORC data collection team.⁴¹ For the follow-up survey, Mathematica plans to modify the baseline survey by collecting training, employment, and earnings histories since the baseline survey and adding questions regarding health behaviors. The information from the follow-up survey will be used to estimate the impacts of the VTGF subactivity.

We will use our experience in locating hard-to-reach individuals in longitudinal surveys and work with the Survey Warehouse team to find creative ways to reach both treatment and control group members and ensure high survey response rates. Current plans rely on the use of applicants' preferred telephone numbers, but we will also use secondary contact information (of family members or close friends) or email addresses (where available) gathered in the baseline survey to contact respondents for follow-up. For the longer-term follow-up, if conducted, we will consider

⁴⁰ As mentioned in Chapter III, we may select a random sample of applicants from trainings with large numbers of applicants if collecting data from all applicants will exceed the data collection budget (the total number of applicants is currently unknown).

⁴¹ MCA-N has contracted with NORC to collect follow-up tracer survey data until the end of Compact in September 2014. During this period, the NORC-Survey Warehouse team is expected to collect follow-up data for a little more than 500 respondents. MCC will finalize plans for follow-up tracer survey data collection after the Compact ends.

offering incentives, such as free cell phone airtime, to encourage respondents to update their contact details periodically (we will discuss with MCC and MCA-N whether it is appropriate to do so).

2. COSDEC Data Sources

The data source for the quantitative outcomes analysis for the COSDEC subactivity will be a tracer survey of COSDEC trainees in the cohort enrolling in training in July 2014.⁴² This will include all trainees who enroll in national programs that start in July 2014 and in local center-based programs that start between July and December 2014; we will exclude those who enroll in local outreach programs, because we believe it would be difficult to obtain applicant information for the trainees in these programs (instead, we will cover these programs as part of the qualitative implementation analysis). Given the evaluation design, we will survey these trainees at a single point in time, approximately 12 months after the completion of their training. The survey will also cover those who enrolled in but did not complete training (surveyed 12 months after their expected completion) to enable us to explore completion rates and reasons for dropout.

To ensure that we are able to contact trainees for the survey, we will work closely with the COSDEF and the COSDECs to gather accurate contact information from enrollees as part of the application process. A limited amount of data (for example, gender and education level) might also be collected during the application process, and will be useful to describe the characteristics of trainees.

Mathematica will develop the COSDEC tracer survey, which we anticipate will be a modified version of the VTGF follow-up survey and will include similar topics, such as demographics, educational background, post-training employment, and earnings. We expect to work with Survey Warehouse on collecting data through the COSDEC follow-up survey, because they will be collecting similar data as part of the VTGF follow-up survey. We also intend to use similar locating and tracking strategies to those that will be used for the VTGF survey to ensure high response rates.

3. Outcomes from VTGF and COSDEC Tracer Surveys

The VTGF impact analysis and the COSDEC outcomes analyses will focus on outcomes in four key domains: (1) training receipt, (2) employment and productive engagement, (3) earnings and income, and (4) health behaviors. We will collect and analyze data for a rich set of outcome measures in each domain. However, for the VTGF impact analysis in particular, we must be mindful of the statistical problem of “multiple comparisons.” This arises when estimating impacts on a large number of outcomes; at least a few of the estimates are likely to be statistically significant by chance, even if no true impacts occurred (Schochet 2008). Therefore, we follow the recommendations of Schochet (2008) and differentiate between a limited number of primary outcomes and a larger number of secondary outcomes. We will approach the findings for secondary outcomes with more caution, assessing whether the findings are supported by statistically significant impacts on the primary outcomes or a credible pattern of statistically significant impacts on the secondary

⁴² Although a limited set of trainings has already started in some of the new COSDECs, it is expected that all the new COSDECs will be fully equipped and operational by mid 2014. Given this, as well as the time required to prepare the trainee application form that we will use to identify applicants for the tracer survey and the need to obtain follow-up data within the period of the evaluation, we decided to focus on the July 2014 cohort.

outcomes. Because the analysis for the COSDEC evaluation will be descriptive, the distinction between primary and secondary outcomes is less relevant.

As described below, we intend to measure most key outcomes in the employment and productive engagement, as well as the earnings and income domains, in a 12-month post-training window. Although it would be possible to measure these outcomes since the start of training (or, for VTGF, since the random assignment date), we focus on the 12-month post-training window for several reasons. First, we believe it would be difficult to gather accurate detailed retrospective information regarding employment and earnings over a longer period of time (for some of the longer VTGF trainings, the period since random assignment could be as long as 30 months with a 12-month post-training follow-up). Second, the evaluation has been designed to measure comparable outcomes across trainings of different durations by focusing on the same post-training window. Measuring outcomes since the start of training would lead to different periods of evaluation across trainings. Third, this approach follows that in the literature on vocational training programs in developing countries (see, for example, Card et al. 2011 and Attanasio et al. 2011). Nevertheless, as mentioned earlier, we do intend to obtain simple measures of employment during the training period (focusing on the number and duration of jobs, without delving into details of job characteristics or earnings), to help interpret our results. For example, this will enable us to explore whether the VTGF control group members substitute work experience for trainings, which could dampen the impacts of the VTGF grant.

The key primary and secondary outcomes, organized by domain, include the following (see Table IV.1):

- **Training receipt.** For the VTGF evaluation, our assessment of the impacts of VTGF funding on training receipt will include binary measures for enrollment in, and completion of, any vocational training since the random assignment date (where completion is defined as passing the end of training assessment). Secondary outcomes will include a measure for the highest vocational training level passed (as a proxy for skills and knowledge gained) and binary measures of completion of different types of training (categorized by skill area or groups of skill areas, such as traditionally male-dominated trades). For the COSDEC evaluation, the key primary outcome in this domain will be the completion of COSDEC training by those who enroll.
- **Employment and productive engagement.** The primary measure of employment will be a binary measure of whether an individual held any paid job (including self-employment) in the 12-month period after the completion of training. Secondary measures of employment, such as the amount of time in employment and current employment, will provide additional evidence on that person's labor market experience in the 12-month post-training period. The primary measure of productive engagement will be a binary measure of whether an individual held any job *or* engaged in vocational training in the 12 months after the completion of training. This will help account for the fact that VTGF funding or COSDEC training might encourage recipients to train further, so that impacts on employment alone may be misleading (that is, employment may actually decrease in the short term). We will also explore additional secondary outcomes, analogous to the employment measures.

Table IV.1. Primary and Key Secondary Outcomes for the VTGF and COSDEC Evaluations

Domain	Primary Outcomes	Key Secondary Outcomes
Training Receipt	<ul style="list-style-type: none"> Enrollment in any vocational training since the random assignment date (VTGF evaluation only) Completion of any vocational training since the random assignment date (VTGF evaluation only) Completion of COSDEC training (COSDEC evaluation only) 	<ul style="list-style-type: none"> The highest vocational training level passed Completion of different types of training (categorized by skill area)
Employment and Productive Engagement	<ul style="list-style-type: none"> Any paid job (including self-employment) held in the 12-month post-training period Any paid job held or engaged in further education or training in the 12-months post-training period 	<ul style="list-style-type: none"> Any formal job during the post-training period Self-employment during the post-training period Total amount of time that an individual was employed during the post-training period Average days per month and hours per week worked Any job held at the time of the follow-up survey Number of jobs held during the post-training period Average or maximum job tenure during the post-training period (censored at survey) Total amount of time an individual was productively engaged Productive engagement at the time of the follow-up survey
Earnings and Income	<ul style="list-style-type: none"> Average monthly earnings in 12-month post-training period 	<ul style="list-style-type: none"> Individual income in the month prior to survey Household income in the month prior to survey
Health Behaviors		<ul style="list-style-type: none"> HIV/AIDS knowledge Parenthood in previous 2 years

- Earnings and income.** Our primary measure in the earnings and income domain will be average monthly earnings in the approximately 12 month post-training period. This will be computed based on wages for those in paid employment, and as profit (positive or zero) for those in self-employment. To avoid issues of selection into employment, these measures will be unconditional and will take the value of zero for those who are not employed. Because the ultimate goal of the VTGF and COSDEC subactivities is to improve individual and household income, we will also measure these as additional outcomes in the domain. This is particularly important because the earnings measures may mask other changes in income—for example, transfers from other household

members could decrease after an individual starts to earn higher wages, limiting impacts on total income (although overall welfare would presumably still increase).

- **Health behaviors.** Most of the VTGF-funded and COSDEC trainings follow NTA unit standards, which include HIV/AIDS modules. Attending these trainings and receiving this information may have an impact on knowledge concerning HIV/AIDS—important issues in view of the high prevalence of HIV/AIDS in Namibia. Improved knowledge of safe sex, combined with the fact that trainees are occupied with training and are likely to be focusing on their future careers, may reduce the rate of unplanned parenthood among the young women and men attending training (although this effect could work in the opposite direction if the training introduces them to potential partners). We will therefore include HIV/AIDS knowledge and parenthood as additional secondary outcomes. These outcomes are not directly targeted by the intervention, but we would analyze them to explore potential unintended impacts.

B. Qualitative Data Sources

We plan to collect qualitative data as part of the evaluation of each of the three subactivities to enable us to answer specific research questions and to add depth of understanding to our quantitative analyses. For each subactivity, we will draw on a variety of qualitative data sources to enable us to triangulate relevant information. We will collect these data through a mix of data collection methods, including in-depth interviews, focus groups, observations, and reviews of documents and other administrative data. Interviews, focus groups, and observations will be conducted by local researchers. Reviews of documents and other administrative data will be conducted primarily by Mathematica staff (although the local researchers may assist by gathering these documents).

We intend to conduct two rounds of qualitative data collection across all subactivities: the first in April and May 2014 and the second in October 2015. The first round, which is close to the end of the Compact, will enable us to gather qualitative data while some key stakeholders (such as external consultants and MCA-N staff) are still available, and to best understand implementation of the subactivities during the Compact. The second round will enable us to gather evidence on the longer-term evolution of the interventions post-Compact. As Table IV.2 shows, the subactivities will be at different stages in each round of data collection; however, we determined that collecting the data for each round simultaneously across subactivities would serve the needs of the evaluation and would be the most efficient approach, given the overlap in data sources and fixed costs of data collection.

Table IV.2. Status of Subactivities at the Time of Qualitative Data Collection

Subactivity	First Round (Q2, 2014)	Second Round (Q4, 2015)
VTGF Subactivity	Most VTGF-funded trainings completed, although some still under way; RPL and employer-provided pilots completed	All VTGF-funded trainings completed at least one year earlier (up to three years earlier)
NTF Subactivity	Levy collection system recently implemented	Levy collection system in place for over a year; levy distribution and reporting system implemented six months earlier
COSDEC Subactivity	All new COSDECs recently fully operational	All new COSDECs fully operational for almost two years

Protocols for collecting the qualitative data will be developed by Mathematica, guided by the logic models presented in Chapter I. The first round of data collection will be conducted by a local

team identified and hired by MCA-N, with oversight support from Mathematica. After the end of the Compact, Mathematica is likely to be directly responsible for data collection, and will work with an in-country partner to collect data. The evaluation team members will travel to Namibia for training, pretesting or piloting of protocols, and the start of data collection for each round of data collection. Data will be coded and analyzed by the evaluation team as they are submitted by the data collection partner based in Namibia. We will use NVivo or similar qualitative data analysis software as a proven tool to help identify themes across many diverse respondent groups and data collection methods. This will enable us to develop a key set of qualitative findings that will take into account similarities as well as differences in perspectives across different respondent groups, providing a comprehensive picture of the implementation of each subactivity.

Next, we describe the qualitative data sources we will use for the evaluation of the three subactivities (see Table IV.3 for a summary). We may identify additional data sources upon review of documents and engagement with stakeholders during the evaluation.

- **Interviews with MCA-N staff.** We will conduct interviews with MCA-N staff—including monitoring and evaluation as well as education sector staff—regarding the implementation of all three subactivities and their perceptions of successes and challenges. MCA-N staff will be interviewed in the first round only, while the Compact is still in force.
- **Observations of, and interviews with, NTA staff.** NTA staff will be interviewed for both the VTGF and NTF evaluations. For the VTGF evaluation, we will focus on the NTA’s management of VTGF grants and their experiences with the RPL and employer-provided training pilots. For the NTF evaluation, we will focus on the establishment and initial implementation of the LCDRS and the development of the NTA’s capacity to manage the system. We will identify the relevant NTA staff for interviews in close consultation with MCA-N. In addition, we intend to conduct observations of typical day-to-day operations at the NTA, during which we will gather information from the observation itself and from informal interviews with staff.
- **Interviews with technical consultant staff: GOPA and Transtec.** We will interview GOPA staff who are providing technical assistance to the NTA as part of the VTGF evaluation (focusing on support for management of SLAs with training providers and the RPL and employer-provided training pilots) and the NTF evaluation (focusing on support for establishing the LCDRS). As part of the COSDEC evaluation, we will interview the Transtec staff who are providing capacity-building support to the COSDEF and COSDECs. All consultant staff will be interviewed as part of the first round of data collection, while they are still available in the country.
- **Interviews and focus groups with trainees and control group members.** We will gather qualitative data from trainees and control group members for the VTGF evaluation and trainees for the COSDEC evaluation. Where trainees are still enrolled in training, we will conduct focus groups; where they have already graduated, we will conduct interviews. These focus groups and interviews will enable us to explore the training and labor market experiences of trainees (and the control group) in more detail and in a less constrained manner than is possible in our quantitative tracer survey.
 - For the VTGF evaluation, we expect a mixture of focus groups and interviews with trainees across training providers in the first round, depending on whether trainings are ongoing or have been completed (we

intend to gather data in all 16 VTGF training providers if resources allow, and expect approximately 12 focus groups). We will also conduct interviews with control group members from each training provider, as well as interviews with individuals assessed through the RPL pilot. The second round of data collection for the VTGF evaluation will focus exclusively on interviewing trainees and control group members to enable us to learn more about their long-term labor market experiences and the sustainability of any impacts of the trainings.

- For the COSDEC evaluation, we will conduct focus groups with trainees in a mixture of national and local courses across all seven new COSDECs in the first round of data collection. We will conduct interviews with the same cohort of trainees in the second round, which will take place after trainees have graduated, to link to our outcomes analysis for these trainees.
- **Training provider staff.** We will interview training provider staff as part of both the VTGF and COSDEC evaluations.
 - For the VTGF evaluation, we intend to interview heads or other key administrative staff in all participating training providers in the first round of data collection to capture their experiences with the VTGF grants.
 - For the COSDEC evaluation, we will interview COSDEC center heads and instructors in all the new COSDECs in both rounds of data collection to understand the implementation of the subactivity, as well as the operations of the COSDECs and how they are changing. We will also interview COSDEC staff, to explore their interactions with the technical consultant (first round) and role in supporting the new COSDECs (both rounds).
- **Employers.** We will conduct interviews with a different set of employers for each of the three evaluations.
 - For the VTGF evaluation, we will interview employers of VTGF-funded trainees, which we will identify through training providers (first round only), focusing on their hiring processes and the extent to which the skills of graduate trainees are meeting their needs. This will include interviews with employers of RPL candidates, to assess their perceptions of the RPL program.
 - We will identify and interview employers in the area served by the COSDEC to gather information on their awareness and perceptions of the COSDEC (first round) and their experiences with hiring COSDEC graduates (second round).
 - For the NTF evaluation, the first round will include interviews with employers who will be subject to the LCDRS, focusing on their consultation as part of setting up the LCDRS and their initial perceptions of the system. The second round will include interviews with similar employers, including those who are participating in the LCDRS (paying the levy and requesting reimbursement) and those who are not.
- **ISCs.** ISCs play an important role in identifying key priority areas for training to inform the funding decisions of the NTF (some ISCs also identified key areas for the VTGF

grants and RPL pilots). We therefore intend to conduct interviews with ISC members and, if possible, observe their work, to learn more about the input that they give into the funding system and how the members of the ISC interact.

- **Other relevant stakeholders.** We will also conduct interviews with several other key stakeholders, primarily as part of the NTF evaluation. These include donors, development partners, relevant ministry officials, and training providers who applied for funding under the LCDRS (both successful and unsuccessful applicants). In addition, we will interview RPL recipients to better understand implementation of the full RPL program.
- **Review of administrative records.** A systematic review of administrative records will contribute to the evaluations of all three subactivities. These data will include planning documents, implementation documents, monitoring and evaluation data related to implementation, and any outcome data identified or tracked. For example, data on VTGF grant applications will contribute to our understanding of how funds were allocated and will inform the VTGF cost analysis (if conducted), NTF records will provide evidence on the extent of and compliance with the training payroll levy, and COSDEC records will provide information on the training conducted.
- **Review of the CCR.** The CCR, produced by MCA-N towards the end of the Compact, will include implementation analysis for all three subactivities from the MCA-N perspective. We will review the report and use it as appropriate once it is available.

Table IV.3. Plans for Qualitative Data Collection

Data Source	VTGF Evaluation	NTF Evaluation	COSDEC Evaluation
MCA-N	R1	R1	R1
NTA	R1	R1, R2	
GOPA	R1	R1	
Transtec			R1
Trainees/control group	R1, R2		R1, R2
Training providers	R1		
COSDEF			R1, R2
COSDECs			R1, R2
Employers	R1	R1, R2	R1, R2
ISCs		R1, R2	
Other stakeholders		R1, R2	

Note: R1 is round 1 of data collection (Q2, 2014); R2 is round 2 of data collection (Q4, 2015).

V. EVALUATION ADMINISTRATION AND MANAGEMENT

Given the complexity of this multicomponent evaluation, careful management of the evaluation and evaluation timeline is essential. In this section, we summarize several administrative issues relevant to the conduct of the evaluation and present a detailed timeline for evaluation activities.

A. Summary of Institutional Review Board Requirements and Clearances

Mathematica is committed to protecting the rights and welfare of human subjects by obtaining approval from an IRB for relevant research activities. For this study, these activities include the VTGF tracer survey (baseline and follow-up), COSDEC tracer survey (follow-up), and the two rounds of qualitative data collection.

IRB approval for the VTGF baseline tracer survey was handled by NORC. NORC applied for IRB clearance for the VTGF baseline tracer survey; the IRB determined that the research was exempt from federal policy for the protection of human research subjects.

Mathematica is expected to handle IRB clearance for the remaining research activities, including subsequent waves of the VTGF follow-up tracer survey, the COSDEC tracer survey, and the qualitative aspects of the evaluation. We intend to seek IRB clearance from the Health Media Lab IRB, with which we have an existing relationship; we will request exempt status for the planned activities.

B. Preparing Data Files

Mathematica will deliver three packages of quantitative data to MCC: baseline and follow-up data for the VTGF study, and follow-up data for the COSDEC study. Each package will consist of three separate, well-documented, Stata data sets: a raw data file, a clean file suitable for internal use with complete, non-anonymized data, and a clean file suitable for public use with anonymized data. The baseline data will be delivered within 12 months of the completion of data collection, and the follow-up data will be delivered within 18 months of the completion of data collection. Each package will include documentation in the form of standardized metadata that is compliant with the international Data Documentation Initiative (DDI) and Dublin Core Metadata Initiative (DCMI) standards, as specified in MCC's data documentation and anonymization guidelines (Millennium Challenge Corporation 2012). We will also update MCC's metadata template, which is based on the International Household Survey Network's (IHSN) Metadata Editor.

The public-use data sets will maintain adequate privacy of survey respondents, differing from the internal ones in the number and content of the variables available. To minimize the risk of disclosing sensitive information or the identity of survey respondents, the public versions will anonymize data files by omitting some variables and replacing others with recoded or grouped data. These measures are designed to retain the usefulness of the data while simultaneously preserving the privacy of survey respondents.

Documentation that will accompany the data sets in each package includes user manuals, annotated questionnaires, and analysis programs. All will be delivered in Word and PDF formats. Stata variable names will be marked on copies of the survey instruments as an aid to data users. Because not all data elements, such as text variables, will be included on the public-use files, the

annotated questionnaires will be customized so that only the public-use variables will appear on public-use versions of the questionnaires.

C. Dissemination Plan

To ensure that the results and lessons from the evaluation reach a wide audience, we will work with MCC to increase the visibility of the evaluation and findings targeted to the vocational education sector, particularly policymakers and practitioners. During the first phase of the evaluation, we will release outreach materials based on our final design report, which will include a non-technical summary of the evaluation and answers to frequently asked questions to inform and engage stakeholders in the evaluation process. We will ensure these materials are distributed to MCA-N, the NTA, and other representatives of the GRN. We will also present the findings from the VTGF baseline report to MCC in Washington, DC, and, if possible, to key stakeholders in Namibia. The VTGF baseline report, with a nontechnical summary, will be available online.

Likewise, we will conduct presentations based on the VTGF and COSDEC follow-up analysis and the performance evaluation analyses. The presentations will be made in Washington, DC, and Namibia so that local stakeholders (such as the NTA, the COSDEF, and the heads of the COSDECs) can be made aware of the results and learn from them. The final evaluation reports, with a nontechnical summary, will be available online on the MCC website within six months of the drafts being submitted.

Given the limited existing evidence on the impacts of vocational training programs, we expect the broader research community to have strong interest in the findings. To facilitate wider dissemination of findings and lessons learned, we will collaborate with MCC and other stakeholders to identify additional forums—conferences, workshops, and publications—to disseminate the results and encourage other donors and implementers to integrate the findings into their programming.

D. Evaluation Team: Roles and Responsibilities

Mathematica's tightly knit team brings together strong design, data collection, and evaluation expertise, as well as extensive knowledge of vocational education programs. Our core team consists of a program manager, two senior analysts, and a vocational training expert, assisted by a junior analyst. **Dr. Arif Mamun** leads the evaluation team as program manager. He oversees the design and implementation of the evaluation, assuming primary responsibility for coordinating deliverables and for ensuring the on-time completion of tasks within budget and at the highest level of quality. He will also take the lead on updating work plans and stakeholder communication. **Dr. Evan Borkum**, an expert in quantitative evaluation methods, led the development of the evaluation design, and will lead the analysis and reporting for the evaluation. **Dr. Kristen Velyvis** is leading the survey and qualitative data collection activities. She will oversee design of data collection instruments and protocols, as well as maintain regular communication with local data collection partners. She also will oversee production of public-use data sets and documentation. **Dr. John Middleton**, an expert in vocational education issues in southern Africa and other developing country contexts, serves as the vocational training expert. He provides context-specific knowledge that helped shape the evaluation design and will help interpret the results. **Mr. Lucas Heinkel**, a former Peace Corps volunteer in Namibia who has helped design and implement education studies, assists senior team members. **Dr. Anu Rangarajan**, a leading expert in impact evaluation design and implementation, will review key deliverables prepared by the core team to assure quality. As discussed with MCC, we will also explore the possibility of engaging an in-country research assistant to complement the team.

E. Evaluation Timeline and Reporting Schedule

The timing of the evaluation activities will correspond to that of implementation (Figure V.1). The VTGF baseline survey began soon after applicants were randomly assigned for the first VTGF-funded training, and will end within about one month of random assignment for the last training. The 12-month follow-up survey for the VTGF subactivity began in March 2014, and the COSDEC 12-month follow-up survey will begin in the third quarter of 2015. Both will be conducted approximately 12 months after the training each respondent applied for is completed.⁴³ As described in Chapter IV, we also intend to conduct two rounds of qualitative data collection that will cover all subactivities. The first will be conducted in the second quarter of 2014, and the second toward the end of 2015. The timing of these rounds will enable us to gather the needed data from key stakeholders within the Compact period (first round) and to finalize our reports within the evaluation period (second round).

Our findings will be summarized in several memos and reports. By the end of the third quarter of 2014, we will provide an implementation analysis report based on the analysis of the first round of qualitative data and the CCR prepared by MCA-N. In the second quarter of 2015, we will prepare a memo summarizing the characteristics of COSDEC trainees that we will include in our tracer survey (which will include those in national and local programs), using data from the COSDEC application form. In the second quarter of 2015, we will produce a baseline report describing the data from the VTGF baseline, which will have been completed for all VTGF-funded trainings by this time. Finally, in 2016, we will prepare follow-up reports for all subactivities based on our follow-up surveys (for the VTGF and COSDEC subactivities) and the second round of qualitative data collection (all subactivities).⁴⁴

⁴³ The additional longer-term VTGF follow-up survey conducted 24 to 36 months after the end of training that was discussed in Chapter III is not included in the schedule presented here. If conducted, the data collection, analysis, and reporting schedule for this second follow-up would be similar to that for the first follow-up, only shifted back by 12 to 24 months.

⁴⁴ Given the timeline of data collection (especially the follow-up VTGF and COSDEC surveys), it will likely be necessary to extend the evaluation contract for a few months beyond the current end date to complete all the required reporting tasks. We intend to request a contract modification from MCC to allow for this extension.

Figure V.1. Timeline for the Evaluation

	2011		2012				2013				2014				2015				2016		
	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
VTGF Subactivity																					
Implementation			VF									VS		VE							
Quantitative Data Collection – Baseline																					
Quantitative Data Collection – Follow-up																					
Qualitative Data Collection																					
Reporting													RI			RB					RF
NTF Subactivity																					
Implementation												NL	NG				ND				
Qualitative Data Collection																					
Reporting													RI								RF
COSDEC Subactivity																					
Implementation									CM	CM	CE/ CF	CT/ CF	CF	CF							
Quantitative Data Collection – Follow-up													CR	CR							
Qualitative Data Collection																					
Reporting													RI			RM					RF

Note: The Figure only shows implementation that overlaps with evaluation activities; implementation started earlier in most cases.

VF = end of first VTGF training, VS = start of last VTGF training, VE = end of last VTGF training;

NL = levy implementation begins, NG = GOPA staff leave, ND = levy distribution begins;

CM = moved to new buildings, CE = received equipment, CF= end of trainings included in follow-up survey, CT = Transtec staff leave, CR = recruit trainees for survey;

RI = implementation analysis report, RB = baseline report, RM = baseline memo, RF = follow-up report

REFERENCES

- Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press, 2008.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, vol. 91, no. 434, June 1996.
- Asian Development Bank. "Viet Nam: Vocational and Technical Education Project." Performance Evaluation Report. Manila, Philippines: ADB, 2013.
- Attanasio, Orazio, Adriana Kugler, and Costas Meghir. "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics*, vol. 3, no. 3, 2011, pp. 188–220.
- Betcherman, Gordon, Amit Dar, and Karina Olivas. "Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries." Washington, DC: World Bank, 2004.
- 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.
- Borkum, Evan, John Middleton, and Arif Mamun. "Summary of Literature on Evaluations of Vocational Training Programs." Memorandum submitted to MCC. Princeton, NJ: Mathematica Policy Research, July 18, 2013.
- Card, David, Jochen Kluge, and Andrea Weber. "Active Labour Market Policy Evaluations: A Meta-Analysis." *Economic Journal*, vol. 120, no. 548, 2010, pp. F452–F477.
- Card, David, Pablo Ibararán, 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. 2013. "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.
- GOPA Consultants. "Final Report on Levy Implementation Framework: Executive Report." Submitted to the Millennium Challenge Account Namibia, Windhoek, Namibia: June 2013.

- 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.
- Higgins, Julian, and Simon G. Thompson. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine*, vol. 21, no. 11, 2002, pp. 1539–1558.
- 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.
- Johanson, Richard K. "A Review of National Training Funds." Social Protection and Labor Discussion Paper No. 0922. Washington, DC: World Bank, 2009.
- Johanson, Richard K., and Arvil Van Adams. *Skills Development in Sub-Saharan Africa*. Washington, DC: World Bank, 2004.
- Kelly, Alan, Richard Willing, and Marivic Puzon. "The Philippines-Australia Technical and Vocational Education Project (PATVEP)." AusAID Evaluation No. 4. Canberra, Australia: AusAID, 1998.
- Lee, David S. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, vol. 76, no. 3, 2009, pp. 1071–1102.
- Maitra, Pushkar, and Subha Mani. "Learning and Earning: Evidence from a Randomized Evaluation in India." Monash Economics Working Paper 44-12, December 2012, Monash University, Department of Economics.
- Millennium Challenge Corporation. "MCC-Financed Data Documentation and Anonymization Guidelines." Washington, DC, March 2012. Available at http://www.app.ma/uploads/pdf/MONITORING%20AND%20EVALUATION/MCC%20Data%20Documentation%20and%20Anonymization%20Guidelines%20_April2012_%20_2_.pdf.
- Namibia Statistics Agency. "Poverty Dynamics in Namibia: A Comparative Study Using the 1993/94, 2003/04 and the 2009/10 NHIES Surveys." Windhoek, Namibia: November 2012a.
- Namibia Statistics Agency. "Namibia Household and Expenditure Survey 2009/10." Windhoek, Namibia: 2012b.
- Schochet, Peter. *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations*. (NCEE 2008-4018). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2008.
- Schochet, Peter Z., John Burghardt, and Sheena McConnell. "Does Job Corps Work? Impact Findings from the National Job Corps Study." *American Economic Review*, 2008, vol. 98, no. 5, 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: 9, 2013.
- U.S. Agency for International Development. "State of the Field Report: Examining the Evidence in Youth Workforce Development." Washington, DC: USAID, 2013.
- U.S. Agency for International Development. "USAID/Namibia Country Strategic Plan FY 2004–2010." Washington, DC: USAID, 2003.
- Wood, Robert G., Quinn Moore, and Andrew Clarkwest. "BSF's Effects on Couples Who Attended Group Relationship Skills Sessions: A Special Analysis of 15-Month Data." Princeton, NJ: Mathematica Policy Research, 2011.
- World Bank. "Country Partnership Strategy for the Republic of Namibia for the period FY2014–FY2017." June 26, 2013.
- Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.



MATHEMATICA
Policy Research

www.mathematica-mpr.com

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

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research

