Appendix A: Research questions and activity grid

June 9, 2014

The following grid shows the original four research questions identified by ESDC in the Evaluator Terms of Reference (numbered 1 through 4 and highlighted in the first column). Additional research questions, along with notes providing the motivation behind them, were added by SRDC and are labelled A to D in the first column as well. All research questions are then fleshed out in column 2, followed by indicators, data sources and collection methods.

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
A. Is the proposed pilot project a pay-for-success Social Impact Bond model? Note: ESDC has been presenting the pilot publicly as a way to test elements of a SIB model.	What are the defining characteristics of pay-for-success SIB models, and does the proposed model possess these characteristics?	Key components of pay-for-success, social impact bond models are present (e.g., risk borne by private investor, private investor provides up-front capital for program with social or environmental benefit, evidence of potential for private and social returns on investment, government pays only for success outcome(s) achieved) *COMPLETED*	▪ Literature review ▪ SRDC produced PPT with findings, presented March 2014	▪ Policy analysis	SRDC
1. How effective are performance-based models supported by social finance to increase the LES skills of low-skilled Canadians?	What are participant *outcomes* following the training, in terms of LES skill gains?	IALS scores at baseline and at the end of training	▪ Participants	▪ LES assessments in workplace	SRDC
	Are skill gains maintained in the longer term?	IALS scores 12 months past the end of training	▪ Participants	▪ LES assessments in workplace	SRDC

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
	What is the *impact* of the model on LES skills? Would skill gains have been realized in the absence of the training provided in the tested model?	Difference between IALS scores of participants and comparison group at end of training, and 12 months beyond. Note: SRDC has proposed a staggered design, in which workers at selected sites receive training while those in other selected sites do not initially receive training but act as the comparison group. Twelve months after the program group has completed training – the required follow-up period for testing – the comparison group members will commence training.	▪ Participants ▪ Comparison group	▪ LES assessments in workplace	SRDC
	What is variation of *impacts* across participant subgroups?	Program-comparison group differences by subgroups of interest (e.g., occupation, baseline literacy level, gender, immigrant status) if a subgroup sample size is sufficiently large (over 30 observations) for reliable estimates and exploratory analysis. Statistical power of testing subgroup differences may not be sufficient to detect small to medium size impact differences.	▪ Participants ▪ Comparison group	▪ LES assessments ▪ Baseline participants (and comparison group) survey	SRDC

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
B. How effective is the model in changing workers' overall performance at work? Note: Answering this question is paramount to **employers'** willingness to invest in such LES training in the future.	What are the effects of the training on attitudes, behaviours, well-being, and other attributes? What are the impacts? Are they maintained in the longer term?	▪ Changes in pre-post indicators over time, e.g., competence in completing work tasks, health and well-being, safety incidents, absenteeism, measures associated with overall productivity ▪ Impacts: program-comparison group differences on same indicators ▪ Long-term: Post-post outcomes/impacts	▪ Participants ▪ Employers/supervisors ▪ Comparison group ▪ Employers/supervisors	▪ Participants surveys (pre- and post-training) ▪ Employers/supervisors survey (pre and post) ▪ Comparison group survey (pre and post) ▪ Long-term: all surveys post-post	SRDC
2. What factors contribute to successful models?	How do impacts relate to business needs, training context, employer and employee engagement, receptivity to learning, etc.?	▪ Participants (and comparison group): Degree of performance and skill gaps at baseline; trust, motivation and engagement ▪ Firms: Business Needs, support for training ▪ Training received by supervisors	▪ Participants ▪ (Comparison group) ▪ Employers ▪ Service providers	▪ All of the above ▪ Business and occupational needs analysis ▪ PMIS	SRDC AWES
	What are the incentive effects created by the reimbursement mechanism?	▪ Employee enrollment and attendance; attrition; ILP ▪ Employer adoption of ES practices in workplace	▪ Service providers ▪ Participants ▪ Employers	▪ PMIS	AWES
	Do impacts vary according to content and/or dosage of the intervention?	Evidence of adherence to model and consistency of outcomes across sites and time	▪ Service providers ▪ Participants ▪ Employers	▪ Participant surveys ▪ Employer surveys ▪ Key informant interviews	SRDC

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
C. What are the relative costs and benefits of this model, from multiple perspectives?	What are the returns from investments in LES training for workers, firms, government and society? How can social impacts best be estimated?	Costs of training: delivery, release time, other opportunity costs Financial benefits to participants: increased employment and earnings? Financial benefits to firms: increased employee performance and productivity; reduced absenteeism; increased retention Financial benefits to government t: reduced transfer payments, increased taxes Financial costs to government: cost of subsidy Non-financial benefits to participants: self-efficacy, confidence, health and well-being Other benefits to society? Note: With the availability of a comparison group, benefits to participants and employers will be directly measurable and not have to be based on pre-post observations. Estimates of government benefits will be extrapolated from available cross-sectional evidence as it is assumed that it will not be possible to use actual data on earnings and EI use beyond the duration of the project.	▪ Employers ▪ Participants ▪ Government administrative data ▪ Literature review	▪ Analysis of employers' records ▪ Analysis of admin data for government transfers and tax data ▪ Monetization of non-financial benefits ▪ Estimation of costs-benefits over longer term	SRDC

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
3. What are the minimum rates of return for employers to be willing to invest in this training?	What was the ROI to firm(s) under this model? Willingness of participating firm(s) to pay for the training once the project is completed?	▪ ROI results ▪ Employer satisfaction with training and continued financial support ▪ Evidence that indicators of success are relevant, directly related to business needs ▪ Evidence that indicators are measured objectively Note: Indicators will only be available for participating employer(s), quite likely only one firm. In order to answer the research question, analyses would be required for multiple firms. SRDC proposes a revealed preferences design for this, in which a substantial number of employers (100-200) would be tested using a laboratory experiment on their willingness to make training investments at variable rates of return. This is not in the scope of the current project.	▪ Participating employer(s) ▪ *PROPOSED:* Larger sample of non-participating employers	▪ Key informant interviews ▪ *PROPOSED:* Revealed preferences research using laboratory experimental design	SRDC
4. How do employers perceive this model and what motivates them to invest in training?	What motivates employers to invest in LES training?	▪ Employer perceptions of risk-reward, and returns on investments in training ▪ Evidence that training can be successfully implemented, supported in workplace ▪ Evidence of participating employers' preferences for investment following training	▪ Employers	▪ Key informant interviews ▪ Employer surveys	SRDC

Research question	Specific questions	Indicators	Data sources	Data collection methods	Resp.
D. Does the pay-for-success SIB model yield better outcomes than the traditional model of full government support? Note: Although this question is not included in the **Evaluator's terms of** reference, it has come back regularly in discussions with OLES as one of the main motivation behind the pilot project.	What is the theory behind the notion that a SIB model should yield superior outcomes? What are the 'traditional models' against which the SIB model is being compared, and how do they differ? How does the ROI to participants, firms, government and society compare in each model?	Note 1: An ideal evaluation model would have a separate program group receiving workplace literacy training through the traditional pay for service model. The efficacy of the pay-for-success SIB model would then be measured as the differences of outcomes between the pay-for-success program group and the pay-for-service treatment group. However, since a pay-for-service program group is not available for AWES Skilling UP, an alternative non-experimental benchmark sample of traditional pay for service workplace literacy training is needed to act as a comparator. The UPSKILL program group is a candidate for this, since many of the outcome measures for the two projects are common. However, since UPSKILL was implemented in a different setting (location and sector) for a different group of workers, some adjustments would be required to construct a comparable benchmark. Furthermore, differences in the nature of the interventions used in the AWES and UPSKILL models could reduce comparability.	▪ Literature ▪ Participants ▪ Employers ▪ Service Providers	▪ Comparative analysis ▪ Key informant interviews	SRDC

Appendix B: Summary table of outcomes and returns for PFS projects

Activity	Output	Outcome	Indicator/Returns
Design model	Design document agreed by partners	Identified measurable indicators/strong predictors of socially desirable outcomes (measurement of success) and the value of the socially desirable outcomes Government pays only if intervention is successful Risk is borne by private investor Service Provider has access to significant funds not otherwise available to scale up intervention Service Provider has more flexibility in delivering intervention; can more readily use expertise and focus on outcomes as opposed to outputs	ROI for government and employer commensurate with participant success Employer-as-investor can increase their ROI based on their own behaviour (i.e., supporting training during and after)
Identification of intervention with evidence of success and potential for scaling up	Analysis of previous intervention results	Larger number of participants achieve skill gains at minimum or higher level than previously	Participants achieve success outcome (i.e., LES skill gains) at minimum, to trigger reimbursement Larger gains bring larger reimbursement
Design model to attract employer(s) as funder	Financial contract between employer-funder and Service Provider	Established a pay-for-success reimbursement structure based on the social value of the success Training is funded Employees receive wages while attending training	SP receives funds for training up-front Employer provides release time for frontline employees, supervisor/managers Employer extends ES training funds for other employees More employers are attracted to invest in ES training

Activity	Output	Outcome	Indicator/Returns
Design reimbursement formula to incentivize employer(s) to support training at all levels	MOU between employer-funder and SP with agreed procedures, the structure of monitoring and management of training: ▪ the training addresses the needs of the business (e.g., private returns on investment to employers) ▪ the skills gaps of workers (private returns on investment to workers) ▪ the probability of success (public returns on investment / private returns on investment to employers)	Identify optimal parameters of training (such as the conditions, success factors, content of training, and the mode of delivery) Facilitate creation of the optimal environment for training delivery Larger skill gains are realized than previous projects	Employer encourages/mandates attendance at training Employer supports frontline employees, and supervisors and managers to attend training Employer makes appropriate adaptations to workplace so that employees can utilize new skills
Evaluate PFS funded ES training	Continuous optimization of training quality Analysis and Report	PFS funded model with employer-as-investor shows higher employee skill and performance gains than traditional funding Adopting basic skills training to low skills workers as a human resource strategy Making use of available government funding (such as Canada Job Grant) to supplement provision of basic skills training to low skills workers even if the private return of investment is low	Participant skill gains are higher than previous projects Participant performance gains are better Employee productivity increases LES training as integrated human resource strategy Making use of government funding appropriately

Appendix C: ESSF baseline survey

Thank you for helping us evaluate this program. Your information is kept confidential. It is used for research purposes only.

1. Please enter your ESSF ID

Case ID: _ _ _ - _ _ _

First Name:

_____

Last Name:

_____

How you feel about the training you are about to start

2. Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
I am looking forward to taking this training.	☐	☐	☐	☐	☐
I am motivated to do the best I can in this course.	☐	☐	☐	☐	☐
People close to me support me in taking this course.	☐	☐	☐	☐	☐
I think the training will increase my chances of getting a good job.	☐	☐	☐	☐	☐

3. Why are you taking this training? Please check all that apply.

☐ Referred by case manager
☐ To help me get into a technical or occupational training program
☐ To help me get a job, or a better job
☐ To explore different career options

- ☐ To advance my career
- ☐ To find out how my skills match up with what jobs require
- ☐ To improve my speaking and listening skills
- ☐ To improve my problem solving skills
- ☐ To improve my reading and writing skills
- ☐ To improve my math skills
- ☐ Other, please specify:

  _____

- ☐ Don't know

Your education and training

4. Have you taken any education, training programs, or courses in the past year? Please select all that apply.

- ☐ Career planning or job search workshop
- ☐ Workplace or on-the-job training
- ☐ Job or work-related skills training outside of the workplace
- ☐ General language or literacy courses, not directly related to jobs or work
- ☐ Courses for personal interest, such as hobby/leisure, volunteer activities, or to improve my general level of education
- ☐ Other education or training, please specify:

  _____

- ☐ I have not taken any training programs or courses in the last year → SKIP TO QUESTION 7

5. Were any of the education, training programs, or courses you took in the last year part of a program of studies towards a degree, diploma or certificate?

- ☐ No → SKIP TO QUESTION 7
- ☐ Yes → Continue

6. Please specify the degree, diploma, or certificate:

- ☐ A high school diploma or equivalent
- ☐ Trade or vocational diploma or certificate
- ☐ Apprenticeship certificate
- ☐ Community college or CEGEP diploma or certificate
- ☐ University degree
- ☐ Other diploma, degree, or certificate, please specify:

  _____

7. Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
a. I am more likely to get a better job if I do some learning	○	○	○	○	○
b. Learning new things makes me more confident	○	○	○	○	○
c. Getting qualifications takes too much effort	○	○	○	○	○

Your employment history

8. Are you currently receiving Employment Insurance?

○ Yes → SKIP TO QUESTION 11
○ No → Continue

9. Have you received Employment Insurance at any time in the last 3 years?

○ Yes
○ No

10. Are you currently receiving one of the following:

a. Income Assistance? ○ Yes ○ No
b. Assistance for persons with disabilities? ○ Yes ○ No

11. Which of the following barriers might interfere with your ability to find or keep a job? Please select all that apply.

☐ Difficulty with English
☐ Learning disability
☐ Legal issues
☐ Limited work experience
☐ Lack of child care support
☐ Transportation issues
☐ Physical disability, injury or illness
☐ Housing problems

☐ Family member health
☐ Education
☐ Drug or alcohol problems
☐ Lack of job hunting skills
☐ Family issues
☐ Credentials not recognized in Canada
☐ Other, please specify:

_____

12. What is your current employment status?

○ Currently working part-time or full-time → SKIP TO QUESTION 15
○ Not working → Continue

13. Have you ever had any paid employment?

○ No → SKIP TO QUESTION 23
○ Yes → Continue

14. How long have you been out of work?

○ Less than 3 months
○ Between 3 and 6 months
○ More than 6 months but less than one year
○ More than one year to less than 3 years
○ More than 3 years to less than 5 years
○ More than 5 years

15. What is the total number of months you were employed over the last three years (36 months)? If you are not sure about the exact number of months, please give us your best guess.

Number of months you had a job, in the past 36 months:

_____

The next few questions ask about your current job if you are currently working, or your most recent job if you are currently out of work. If you currently have (or most recently had) more than one job, answer for the job at which you work (or worked) the most hours.

16. When did you start this job? Please give us your best guess if you can't remember the exact date.

Month _____ Year _____

17. Are you still working at this job?

○ Yes → SKIP TO QUESTION 19
○ No → Continue

18. When did this job end? Please give us your best guess if you can't remember the exact date.

Month _____ Year _____

19. On average, how many paid hours per week do (or did you) you work at this job? Please include paid overtime hours if you usually work paid overtime.

_____ hours per week

20. How many weeks per month do you (or did you) usually work at this job?

☐ 1 week
☐ 2 weeks
☐ 3 weeks
☐ 4 weeks

21. How often do you (or did you) experience high levels of stress at this job?

☐ Once a month or less
☐ A few times a month
☐ A couple of days a week
☐ Almost every day

22. Please select ONE of the options below to tell us how much you are paid (or were paid) in this job before taxes and other deductions. Check ONE box only.

☐ $_____ per hour
☐ $_____ per week
☐ $_____ every two weeks
☐ $_____ per month
☐ Other (please specify): $ _____ per _____
☐ Don't know

Your career and job search activities

23. Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
a. I have not really decided what my career objectives should be yet	○	○	○	○	○
b. I have a strategy for achieving my career goals	○	○	○	○	○
c. I know what I need to do to reach my career goals	○	○	○	○	○

24. For each statement, please indicate how confident you are that you can successfully do each of the following activities.

	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
a. Accurately assess how well your abilities are suited for the kind of work you want to do	○	○	○	○	○
b. Find information about occupations you are interested in	○	○	○	○	○
c. Find out the employment trends for an occupation over the next ten years	○	○	○	○	○
d. Find out about the average yearly earnings of people in an occupation	○	○	○	○	○

	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
e. Talk with a person already employed in the field you are interested in	○	○	○	○	○
f. Find information about education or training programs in the field you are interested in	○	○	○	○	○
g. Select one occupation from a list of potential occupations you are considering	○	○	○	○	○
h. Select one education or training program from a list of potential programs you are considering	○	○	○	○	○
i. Choose a career that will fit your abilities and interests	○	○	○	○	○
j. Identify employers, firms, institutions relevant to your career possibilities	○	○	○	○	○
k. Change jobs if you did not like your job	○	○	○	○	○
l. Determine the steps to take if you are having trouble with an aspect of your job	○	○	○	○	○
m. Identify some reasonable occupation or career alternatives if you are unable to get your first choice	○	○	○	○	○

25. Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
a. I have a clear idea of the type of job I want	○	○	○	○	○
b. I have very clear job search objectives	○	○	○	○	○
c. I have a clear idea of the type of company I want to work for	○	○	○	○	○
d. It is not very clear to me where I should be looking for a job	○	○	○	○	○

26. For each statement, please indicate how confident you are that you can successfully do each of the following activities.

	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
a. Use social networks to obtain job leads	○	○	○	○	○
b. Prepare resumes that will get you interviews	○	○	○	○	○
c. Impress interviewers during employment interviews	○	○	○	○	○
d. Make "cold calls" that will get you a job interview	○	○	○	○	○
e. Conduct information interviews to find out about careers and jobs that you are interested in pursuing	○	○	○	○	○

	No confidence at all	Very little confidence	Moderate confidence	Much confidence	Complete confidence
f. Communicate your skills and experience in a way that will attract the interest of employers	○	○	○	○	○
g. Plan and organize a weekly job search schedule	○	○	○	○	○
h. Find out where job openings exist	○	○	○	○	○
i. Use a variety of sources to find job opportunities	○	○	○	○	○
j. Search for and find good job opportunities	○	○	○	○	○

Background information

27. What language do you speak most often at home?

○ English
○ French
○ Other, please specify:

_____

28. Do you speak any other languages on a regular basis at home?

○ No
○ Yes, English
○ Yes, French
○ Yes, Other, please specify:

_____

29. What is your current marital status?

- ○ Married
- ○ Common law
- ○ Single, never married
- ○ Separated, divorced, or widowed

30. How many children under the age of 18 live in your household?

Please include all children who usually live with you, including those who may be away attending school, travelling or in hospital.

Number of children in your household _____ → If none, **SKIP TO QUESTION 32**

31. How many of these children are under the age of 6?

Number of children under 6 years old in your household _____

32. How many adults 18 years of age or older, including yourself, live in your household?

Please include all adults who usually live with you, including those who may be away attending school, travelling or in hospital.

Number of adults in your household, including yourself _____

33. Of all the people in your household, including yourself, how many are currently employed?

Number of people in your household who are employed, including yourself _____

34. What was your household income in the last 12 months, before taxes and other deductions?

Include all sources of income (for example, wages, alimony, investments, Employment Insurance, social assistance, grants, scholarships, income earned outside of Canada, etc.). Please give us your best guess if the exact figure is not known.

- ○ Less than $10,000
- ○ $10,000 to less than $20,000
- ○ $20,000 to less than $30,000
- ○ $30,000 to less than $40,000
- ○ $40,000 to less than $50,000
- ○ $50,000 to less than $60,000
- ○ $60,000 to less than $70,000
- ○ $70,000 to less than $80,000
- ○ $80,000 to less than $90,000
- ○ Greater than $90,000

Information about your use of reading, math and other skills

35. How often do you do each of the following activities outside of work?

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
a. Do math (such as for household budgets, bills, bank accounts or credit cards)?	○	○	○	○	○	○
b. Write notes, letters, or e-mails?	○	○	○	○	○	○
c. Read or use information from books - fiction or non-fiction?	○	○	○	○	○	○
d. Use a library or visit a bookstore?	○	○	○	○	○	○

Your health and activities

36. Does a physical condition or health problem reduce the amount or kinds of activities you can do at work, at home, or for recreation?

Not at all	Rarely	Sometimes	Often
○	○	○	○

37. Does an emotional condition or health problem (such as feeling depressed or anxious) reduce the amount or kinds of activities you can do at work, at home, or for recreation?

Not at all	Rarely	Sometimes	Often
○	○	○	○

Your opinions about life in general

38.

	Very dissatisfied 1	2	3	4	5	6	7	8	9	Very satisfied 10
Using a scale of 1 to 10 where 1 means "very dissatisfied" and 10 means "very satisfied", how do you feel about your life as a whole right now?	○	○	○	○	○	○	○	○	○	○

39. Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
My life is determined by my own actions.	☐	☐	☐	☐	☐
I don't really plan for the future because things change so much.	☐	☐	☐	☐	☐
Since I can't affect the future, it doesn't really matter what I do.	☐	☐	☐	☐	☐

40. If you lost your wallet or purse that had $200 in it, how likely do you think it is that it would be returned with the money still in it if it was found by:

	Very unlikely 1	Somewhat unlikely 2	Neither likely nor unlikely 3	Somewhat likely 4	Very likely 5
A neighbour who lives close by	☐	☐	☐	☐	☐
An employee at a local business	☐	☐	☐	☐	☐
A total stranger	☐	☐	☐	☐	☐

Please feel free to tell us your comments about the survey such as questions that were difficult to answer, or confusing.

_____

_____

_____

If you have any questions about the survey, please phone SRDC at 1-866-896-7732.

THANK-YOU!

Appendix D: Skilling UP baseline survey

Please enter your case ID, first name, and last name

Case ID: _ _ _ - _ _ _

First Name:

_____

Last Name:

_____

SECTION A > How you feel about the training you are about to start

Please read the statement and check the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
A) I am looking forward to taking this training.	☐	☐	☐	☐	☐
B) I am motivated to do the best I can in this course.	☐	☐	☐	☐	☐
C) My supervisor supports me in taking this course.	☐	☐	☐	☐	☐
D) I think the training will make me and my co-workers more productive.	☐	☐	☐	☐	☐
E) I think my employer will share any benefits of improved productivity with the workers.	☐	☐	☐	☐	☐
F) I think my employer will change the way my work is organized or performed to allow me to use the skills I learned in this training.	☐	☐	☐	☐	☐

Why are you taking this training? *Please check all that apply.*

- ☐ Required by employer
- ☐ To help me do my job better
- ☐ To advance my career
- ☐ To prepare me for further training or education
- ☐ To improve my speaking and listening skills
- ☐ To improve my problem solving skills
- ☐ To improve my reading and writing skills

☐ To improve my math skills
☐ Other (specify): _____
☐ Don't know

SECTION B > Your education and training

Have you taken any education, training programs, or courses in the past year? *Please select all that apply.*

☐ Career planning or job search workshop
☐ Workplace or on-the-job training
☐ Job or work-related skills training outside of the workplace
☐ General language or literacy courses, not directly related to jobs or work
☐ Courses for personal interest, such as hobby/leisure, volunteer activities, or to improve my general level of education
☐ Other education or training, please specify: _____
☐ I have not taken any training programs or courses in the last year → **GO TO QUESTION 7**

Were any of the education, training programs, or courses you took in the last year part of a program of studies towards a degree, diploma or certificate?

☐ Yes → **GO TO QUESTION 6**
☐ No → **GO TO QUESTION 7**

Please specify the degree, diploma, or certificate:

☐ A high school diploma or equivalent
☐ Trade or vocational diploma or certificate
☐ Apprenticeship certificate
☐ Community college or CEGEP diploma or certificate
☐ University degree
☐ Other diploma, degree, or certificate, please specify: _____

Please read the statement and check the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
A) I am more likely to get a better job if I do some learning.	☐	☐	☐	☐	☐
B) Learning new things makes me more confident.	☐	☐	☐	☐	☐
C) Getting qualifications takes too much effort.	☐	☐	☐	☐	☐

SECTION C > Information about your job

When did you start working for this company? *Please enter the approximate date if you can't remember the exact date.*

Month _____ Year _____

When did you start working in your current position at this company? Please enter the approximate date if you can't remember the exact date.

Month _____ *Year* _____

On average, how many paid hours per week do you work in your job? Please include paid overtime hours if you usually work paid overtime.

_____ hours per week

How many weeks per month do you usually work in your job?

- ☐ 1 week
- ☐ 2 weeks
- ☐ 3 weeks
- ☐ 4 weeks

How often do you experience high levels of stress on the job?

- ☐ Once a month or less
- ☐ A few times a month
- ☐ A couple of days a week
- ☐ Almost every day

Here is a list of things that sometimes cause stress for people at work. For each item, please check how often it causes stress for you.

	Once a month or less	A few times a month	A couple of days a week	Almost every day
A) Not having enough time to finish my work	☐	☐	☐	☐
B) Not always having the skills I need to do my job well	☐	☐	☐	☐
C) Not being informed in advance of important decisions made at work	☐	☐	☐	☐
D) Not receiving all the information I need to do my job well	☐	☐	☐	☐
E) Not having my work recognized and appreciated	☐	☐	☐	☐
F) Ineffective teamwork	☐	☐	☐	☐
G) Needless repetition of job tasks	☐	☐	☐	☐
H) Another source of work stress we haven't listed above, please specify: _____	☐	☐	☐	☐

Please indicate how much you agree or disagree with each statement by checking the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
A) In my job, I can work effectively as part of a team.	☐	☐	☐	☐	☐
B) In my job, I feel accepted by other employees.	☐	☐	☐	☐	☐
C) My job helps me with my specific career goals.	☐	☐	☐	☐	☐
D) I am able to succeed in my job even when the work is challenging or difficult.	☐	☐	☐	☐	☐
E) When doing my job, I sometimes find it hard to keep up with what is expected of me.	☐	☐	☐	☐	☐
F) I get all the training I need to do my job well.	☐	☐	☐	☐	☐
G) I get quite anxious in my job.	☐	☐	☐	☐	☐
H) I don't think I have much control over how well I do in my job.	☐	☐	☐	☐	☐

Please read the statement and check the appropriate box, where 1 means "very dissatisfied" and 7 means "very satisfied".

	Very dissatisfied 1	2	3	4	5	6	Very satisfied 7
A) Your pay	☐	☐	☐	☐	☐	☐	☐
B) The opportunities to use your skills and experience	☐	☐	☐	☐	☐	☐	☐
C) The opportunities to use your own initiative and make decisions	☐	☐	☐	☐	☐	☐	☐
D) Your job security	☐	☐	☐	☐	☐	☐	☐
E) Support from your supervisor or manager	☐	☐	☐	☐	☐	☐	☐
F) The opportunities for career growth and promotion	☐	☐	☐	☐	☐	☐	☐
G) The opportunities for learning new things and developing your own abilities	☐	☐	☐	☐	☐	☐	☐
H) All in all, how satisfied are you with this job?	☐	☐	☐	☐	☐	☐	☐

Please select ONE of the options below to tell us how much you are currently paid in your job (before taxes and other deductions). Check ONE box only.

☐ $_____ per hour
☐ $_____ per week
☐ $_____ every two weeks
☐ $_____ per month
☐ Other (please specify): $_____ per _____ (Text field)

SECTION D > Background information

What is the language you speak most often at home?

- ☐ English
- ☐ French
- ☐ Other (specify) _____

Do you speak any other languages on a regular basis at home?

- ☐ No
- ☐ Yes, English
- ☐ Yes, French
- ☐ Yes, Other (specify) _____

What is your current marital status?

- ☐ Married
- ☐ Common law
- ☐ Single, never married
- ☐ Separated, divorced or widowed

How many children under the age of 18 live in your household? Please include all children who usually live with you, including those who may be away attending school or travelling.

Number of children in your household _____

How many adults 18 years of age or older, including yourself, live in your household?

Please include all adults who usually live with you, including those who may be away attending school or travelling.

Number of adults in your household, including yourself _____

Of all the people in your household, including yourself, how many contribute to the total income of your household?

Number of people, including yourself, who contribute to household income _____

What was your household income in the last 12 months, before taxes and other deductions? Include all sources of income (for example, wages, alimony, investments, Employment Insurance, social assistance, grants, scholarships, etc.). Please give us your best guess if the exact figure is not known.

- ☐ Less than $10,000
- ☐ $10,000 to less than $20,000
- ☐ $20,000 to less than $30,000
- ☐ $30,000 to less than $40,000
- ☐ $40,000 to less than $50,000
- ☐ $50,000 to less than $60,000
- ☐ $60,000 to less than $70,000
- ☐ $70,000 to less than $80,000
- ☐ $80,000 to less than $90,000
- ☐ Greater than $90,000

SECTION E > Your opinions about life in general

Please read the statement and check the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5
A) My life is determined by my own actions.	☐	☐	☐	☐	☐
B) I don't really plan for the future because things change so much.	☐	☐	☐	☐	☐
C) Since I can't affect the future, it doesn't really matter what I do.	☐	☐	☐	☐	☐

If you lost your wallet or purse that had $200 in it, how likely do you think it is that it would be returned with the money still in it if it was found by:

	Very unlikely 1	Somewhat unlikely 2	Neither likely nor unlikely 3	Somewhat likely 4	Very likely 5
A) A neighbour who lives close by	☐	☐	☐	☐	☐
B) An employee at a local business	☐	☐	☐	☐	☐
C) A total stranger	☐	☐	☐	☐	☐

Please provide us with any comments that you have about the survey such as questions that were difficult to answer or questions that were confusing.

If you have any questions about the survey, please phone SRDC at 1-866-896-7732.

THANK-YOU!

Appendix E: SRDC PowerPoint presentation to OLES – March 2014

Social Research and Demonstration Corporation

- In our 24th year of evaluating social program models in real-world settings
- Work spans the fields of adult education and training, pre-school to post-secondary education, labour market integration, health, income security, among others
- Research methodologies from large-scale demonstration projects with multiple partners, to program evaluation, policy reviews and experimental economics
- Extensive expertise in essential skills demonstration projects, including Upskill and Workplace Training Program
- Produced comprehensive reports on application of performance-based funding models and social finance approaches to the field of adult education and training

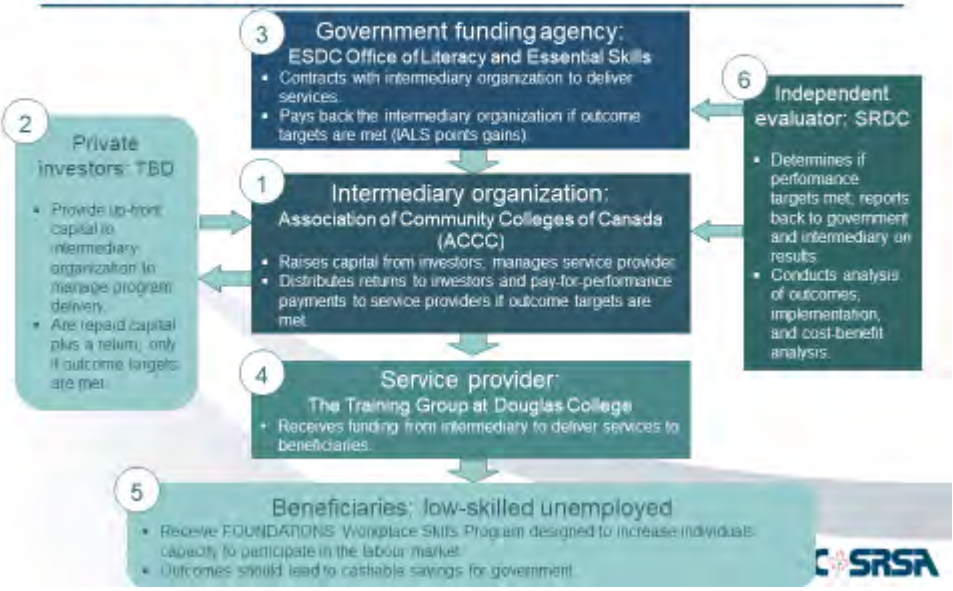
SRDC SRSA

SRDC role in *Evaluation of SIB approach to Essential Skills training*

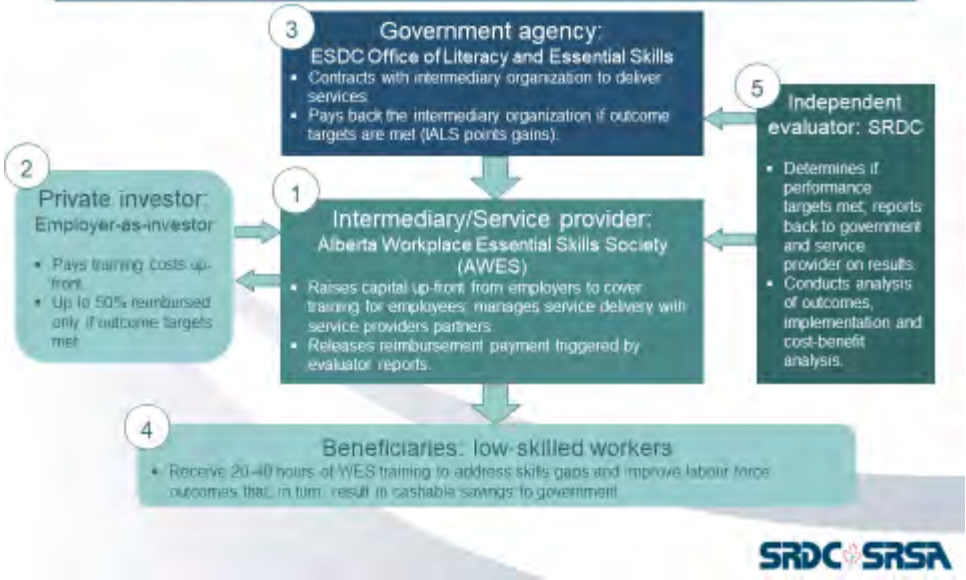
- Contracted to develop an overarching evaluation framework and conduct independent evaluations of two pilot projects sponsored by federal government
- Support project proponents (ACCC and AWES) through collaboration on design and implementation
- Responsible for overseeing funder reimbursement mechanisms

SRDC SRSA

The ACCC pilot



The AWES pilot



SRDC SRSA

Both pilot projects include elements of a Social Impact Bond approach

- Activity generates a social dividend and economic return to investor
  - Social and economic benefits associated with a more skilled workforce
  - Returns on training investment (ACCC)/increased productivity and profitability (AWES)
- Private investors pay full costs of intervention up front
  - Essential skills training to employed (AWES) or unemployed (ACCC)
- Intervention addresses a well-identified social/environmental problem or goal
  - Labour market vulnerability of low-skilled workers
- Government funding is tied to measurable results
  - Reimbursement is made only if training is successful, defined as 25-point gain on IALS literacy scale
- Service providers paid on basis of results achieved (under consideration)
  - Payment/bonus contingent on literacy point gains
- Presence of cashable savings for government
  - Increased employment/retention reduces EI and SA costs; higher earnings lead to tax revenue increases; higher skills should reduce WC costs
- Risk borne by private sector
  - If desired outcomes are not achieved, private investors bear 100% of costs

SRDC SRSA

Some departure from SIB approach

- In existing Social Impact Bonds, government funding is directly associated with cashable savings to government
  - In the Essential Skills pilot projects government funding in the form of reimbursements for training is made on the basis of intermediate outcomes, i.e. literacy points gains
- The employer-as-investor is a new variant on a Social Impact Bond
  - In the AWES model, the private investor is motivated not by return on *capital investment* per se, but on economic returns from increased productivity as well as reimbursement of training *expenses*
  - The employer-as-investor is incentivized to encourage employee attendance at training, facilitate adoption and promote use of new skills in the workplace in order to achieve and retain learning gains

SRDC SRSA

Research questions

- How effective are performance-based models supported by social finance to increase the literacy and essential skills of low-skilled Canadians?
- What factors contribute to successful models?
- What are the minimum rates of return for investors and employers to be willing to invest in such models?
- How do investors and employers perceive social finance and what motivates them to invest in such models?

SRDC SRSA

Evaluation components

- Outcome analysis
  - Measure skills gains
  - Labour market and other outcomes of interest for participants and employers
  - Counterfactual data for impact analysis
- Implementation analysis
  - What delivery factors are associated with success?
  - What is the investor and employer engagement and feedback?
- Cost-benefit analysis
  - Measure costs of the training to private and government funders
  - Benefits to participants, funders, and employers
  - Monetize impacts to estimate return on investment for workers, employers, government and society as a whole

SRDC SRSA

Outcomes of interest

- Literacy and essential skills gains
- Labour market outcomes:
  - Employment
  - Wages and hours
  - Retention and advancement
- Performance metrics:
  - Employers – retention, lower supervisory costs, number of injuries
  - Participants – absenteeism, use of materials, health and safety practices
- Non-financial outcomes:
  - Behavioural – job search, use of ES, volunteerism
  - Psychosocial – future orientation, motivation and encouragement at work, self-efficacy
- Social benefits:
  - Cost savings and net benefits to government: EI, SA, and increased taxes
  - Externalities: e.g. better health, social cohesion

SRDC SRSA

In progress

- Working with AWES and ACCC to finalize the project designs and develop detailed implementation workplans
- Fleshing out the evaluation framework, detailed research questions and proposed data sources
- Preparing materials on Upskill impacts and returns on investment to assist partners in engaging employers or investors
- Analyzing IALS points gains data from previous essential skills projects to support calculation of reimbursement formula
- Exploring environments and administration of LES assessments

SRDC SRSA

Appendix F: Estimating outcomes of Essential Skills interventions – October 2014

Table of contents

1. Introduction

In January 2014, the federal Office of Literacy and Essential Skills contracted SRDC to conduct independent evaluations of two Essential Skills (ES) training projects. For the first time in Canada, a Social Impact Bond (SIB) approach to funding ES training is being tested, whereby private investors pay for the training up-front, and are repaid by the government if the training is successful in achieving pre-established outcomes.

The pilot project led by Colleges and Institutes Canada (CICan) proposes to enroll 400 unemployed lower-skilled Canadians to receive FOUNDATIONS, an established ES training program delivered by Douglas College. Private investor(s) will receive their investment plus a financial return of up to 10% if training is successful. Alberta Workforce Essential Skills Society (AWES) is leading a second pilot project, in which the investor is a private sector employer(s) who will be reimbursed up to 50% of training costs if the training achieves target outcomes.

The risk to investors of investing in a social impact bond corresponds to the probability distribution of potential outcomes. How can the risk of potential outcomes be assessed? In reviewing pay-for-success schemes, a priori expectations of the range of potential outcomes is typically based on evidence – range and expected mean-median outcomes – from similar interventions in the past. Reliability of the estimates depends on the extent to which similar interventions have been delivered, and how accurately the success outcomes have been measured. Earlier work by SRDC[1] analyzed data from three previous ES projects – FOUNDATIONS, Workplace Training Program (WTP), and UPSKILL – and presented findings on minimum expected gains.

Fully understanding and quantifying the range of potential outcomes makes reimbursement formulation difficult. In the case of social impact bonds, how should a bond issuer price the bond associated with risk when the risk is not readily quantified? This technical analysis aims to take the analysis of demonstrated skill gains a step further by shedding light on quantifying the probability distribution of all potential outcomes of the two Essential Skills interventions. This is done by estimating the individual impact distribution of SRDC's UPSKILL project.

There are two major challenges involved in examining potential outcomes of an Essential Skills intervention. First, the theoretical linkage of individual impact distribution and the potential outcomes have not been formally established. Second, even with a carefully implemented randomized controlled trial, individual impact distribution is difficult to estimate. Once these two challenges are overcome and the probability distribution of potential outcomes is estimated, the risk of potential outcomes of a similar Essential Skills intervention can be assessed. At the time of writing, the quantitative risk analysis in this paper is the first in the field of social impact bonds.

UPSKILL was a large scale randomized controlled trial (RCT) of delivering workplace Essential Skills training to nearly 1500 workers in participating Canadian firms across eight provinces. It was funded by the Office of Literacy and Essential Skills (OLES) of the Employment and Social Development Canada (ESDC) to test the efficacy of paid Essential Skills training tailored to address

[1] Technical Paper: Analysis of 25-point skill gains in Essential Skills projects, SRDC, May 2014.

the needs of businesses and workers. Launched in 2010 as a pan-Canadian research and demonstration project, UPSKILL utilized an experimental design methodology in order to provide the most reliable measures of the impacts of LES training in the workplace. A total of 88 firms in the Accommodations and Food Services sector, primarily hotels, were randomly assigned to a program group where employees were offered a maximum of 40 hours of Literacy and Essential Skills training on-site during working hours, or to a control group whose employees did not receive the training. Random assignment efficiently controls for all factors, other than the UPSKILL training, which could affect employee and employer outcomes. Thus program-to-control group comparisons provide reliable measures of the impact of UPSKILL training.

UPSKILL collected a vast amount of project data for analysis including:

- three sets of literacy assessments and surveys of participants: at baseline, post-training, and post-post-training (6-18 months following);
- supervisor and management surveys; and
- on-the-job performance assessments by independent observers.

The combination of UPSKILL's large sample size, RCT design, and comprehensive data makes the project results reliable, and compelling. The findings demonstrated that the workplace LES training had large positive impacts on workers' skills, job performance, and a range of economic and social outcomes for workers and firms. Participants in the training experienced literacy skill gains in the range of 11-18 points in the post- and post-post assessments. Along with improvements in skill scores, the percentage of participants achieving the literacy skills level required in their job increased substantially; for the average employer with 15 employees, 3 additional workers met the literacy requirement of their job following UPSKILL training.[2]

Significant gains in job performance were also observed among UPSKILL program group members including a greater breadth of service quality, improved relations with customers, and increased task efficiency. At the same time, Essential Skills training led to an increase of over 12 percentage points in the number of employees achieving industry certification standards of job performance, compared to the changes observed in the control group.

A benefit-cost analysis also revealed that firms in the program group experienced gains in revenue, cost savings from increased productivity, and reductions in hiring costs; their net benefit was $577 per participant, for an average return on investment of 23 per cent.

Importantly, the study also found that the pattern of impacts was heterogeneous, varying among firms and workers in ways that have important implications for the design and delivery of effective ES training programs. Larger impacts were observed among immigrants, older workers, Aboriginals, workers with low initial skills level, and workers from low-income families. Training hours were also associated with a larger skills gain. The heterogeneity of impacts, combined with

[2] For detailed information on UPSKILL methodology and results see Gyarmati, D., Leckie, N., Dowie, M., Palameta, B., Hui, T.S., Dunn, E., and Hébert, S. *UPSKILL: A Credible Test of Workplace Literacy and Essential Skill Training*. SRDC Technical Report, August 2014.

large sample size and detailed data collection, makes UPSKILL a good candidate for suitable and reliable source of information on potential outcomes of similar Essential Skills interventions.

This paper describes the technical analysis undertaken to use UPSKILL data to estimate individual impact distributions and propose risk profiles for SIB investment in the two ESDC-funded training projects. Section 2 below discusses the theoretical linkage of heterogeneous individual impacts and the risk of potential program outcomes of a future program. It is followed by a section on estimating individual impact distribution of Essential Skills interventions (using UPSKILL data) and the distribution of potential program outcomes of a future essential skill program similar to that of UPSKILL. Section 4 demonstrates the application of the estimated distribution of potential program outcomes in a financial analysis of scenarios of social impact bond reimbursement for the two projects. The paper concludes with a brief discussion of the contribution of the analysis as well as the limitations of the methodology.

2. Heterogeneous individual impact and risk (probability) of potential outcomes

Heterogeneous impacts are not uncommon among learning interventions; some participants experience larger impacts than others, as seen in the UPSKILL project results. Similarly, subgroup variations have been demonstrated in various labour market interventions. Since any outcome of a program is a function of the program's impact on participants, the risk associated with any program outcome is directly related to the variation of impacts.

The reasons for impact variation are not completely understood, though many conjectures explaining impact variation are plausible. UPSKILL identified four primary types of factors that appear to affect the training impact:

- Implementation and delivery-based factors such as the expertise of the instructors, content customization, dosage, and timing;
- Demographics and socioeconomic background of the learners;
- Learning gaps of participants, and depth of organizational needs of their employers; and
- Training readiness and learner engagement; firms that are committed to training – either through direct expenditures and or other incentives for workers – are associated with higher impacts of LES training.

All of the above factors may be observable either directly, or indirectly through proxy characteristics. Thus program delivery agents may be able to influence program outcomes by creating the conditions for success, for example, by recruiting participants and matching them with instructors most likely to achieve higher impacts, by customizing training content to the specific needs of the learners and employers, and so on.[3]

However, the ability to influence program outcomes is not unlimited even when all factors influencing impacts are observable. For example, the supply of "good" instructors is not unlimited; and the population has a limited number of people who would experience high impacts of the program.

In contrast to observable influencers of outcomes, a second type of impact variation is completely random and/or unobservable. For example, any skills assessment is subject to measurement errors, and the resulting scores contain measurement errors. As well, unobservable variation in characteristics of individual learners, such as their motivation, is thought to affect learning outcomes. During program delivery and recruitment, these unobservable or random factors are not likely to be controllable.

[3] See "Conditions for success" in *UPSKILL: A Credible Test of Workplace Essential Skills Training Summary Report*. SRDC, August 2014.

Separating a program's impact into its controllable and uncontrollable portions is important for assessing the range of the potential outcomes of the intervention. If the probability distribution of the controllable portion of potential program impacts is known, it is possible to calculate the proportion of the population that may achieve a certain level of impact through various implementable means, such as assigning particular instructors to matched learners in a training and education program. In any case, having a low proportion of the population that can achieve a desired level of impact suggests that the outcome was difficult to achieve, while a high proportion reflects that it was relatively easy to reach the desired outcome. Therefore, the probability distribution of the controllable part of potential program impacts on individuals summarizes the difficulties and possibilities of achieving certain impacts. Since a program's outcome is the group level measurement of participants, the range of possible program outcomes is a function of the possible individual impacts. The probability distribution of controllable individual impacts is thus sufficient to identify the probability distribution of program outcomes.

In technical terms, if a program's potential impact on an individual i, D_i, is expressed by its controllable and uncontrollable portions, it is

(1) $D_i = O_i + U_i$,

where O_i is the controllable portion while U_i is the uncontrollable portion. For simplicity of the analysis, it is assumed that the uncontrollable portion of a program's potential impact has a zero mean and that it is uncorrelated with the controllable portion of individual impact.[4] It follows that the mean of O_i is the same as the mean of D_i and the variance of O_i is not larger than the variance of D_i. In other words, the variation of the controllable portion of impacts is bounded by the variation of the program's potential impact.

Equation (1) also implies that the distribution of the controllable portion of impacts is bounded by the distribution of individual impacts. The probability (or risk) of achieving a minimum level of the controllable portion at z by selecting combinations of intervention environment and observable characteristics of participants is not greater than the probability of achieving the same minimum level of individual impact, z. That is:

(2) $\mathrm{K}(O_i \geq z) \leq \Gamma(D_i \geq z)$,

where Γ and K are the probability distribution functions of individual impacts D_i and controllable portion of impacts O_i, respectively.

The property presented in equation (2) is very important in establishing the linkage between individual impact distribution and the risk of producing certain levels of outcomes by the program

[4] The controllable part of the impact could be a function of observable characteristics (that the program delivery agent can control for), and the parameters of this function can be estimated by a regression when individual impact is observed. However, for the purpose of assessing risk without observing individual level impact nor the controllable characteristics, the relevant information is the degree of variation of impacts. Whether the regression estimates are biased or not is irrelevant to the research question.

delivery agent. If the individual impact distribution is known, it is possible to identify the maximum risk involved (or the difficulties involved) in producing any specific level of outcomes.

If it is further assumed that the potential distribution impacts, the controllable portion of impacts and the uncontrollable portion of impacts are all normally distributed, and that the ratio of the variance of the controllable portion to the variance of the potential individual impacts is φ, then the identification of the mean and standard deviation of potential impacts will lead to the identification of the exact distribution of the controllable portion of impacts. In turn, this will lead to the identification of the exact distribution of program outcomes.

3. An estimation model of individual-level impact distribution

The individual level of impact is defined as the difference between the observed individual outcome after the intervention and the counterfactual outcome had the individual not been exposed to the intervention. Because the counterfactual outcome is unobservable, the impact of a program on an individual is therefore also unobservable and the distribution of individual-level impact cannot be directly estimated.

However, in the case of a randomized controlled trial, when the outcomes of individuals in both the treatment and control groups are measured, the distribution parameters of individual-level impact can be indirectly estimated. In theory, in a randomized controlled trial with a large sample, the distribution of a measured outcome among the control group is a consistent estimate of the counterfactual distribution of the outcome among members of the treatment group had they not been exposed to the intervention. The systematic differences between the two distributions of the outcome is thus attributed to the impact of the intervention. Using empirical distributions of an outcome of the treatment and control groups, it is possible to establish the distribution of individual impacts that would be required in order to explain the differences in the observed outcome distributions.

The UPSKILL project, as a large-scale randomized controlled trial of workplace Essential Skills training, provides a suitable source of information to estimate the individual-level impact distribution of Essential Skills training. The measured skill gain in document use of each participant in the UPSKILL treatment and control groups is a common measure used in Essential Skills training projects, and UPSKILL has demonstrated that the training produced positive average impacts on the skill gain of the treatment group. At the same time, since UPSKILL was delivered in multiple locations by different instructors with many different participants, and for different business needs and skill gaps, it is not unreasonable to assume that the UPSKILL delivery was drawing randomly from the population of various implementable measures that reflect the population distribution of controllable and uncontrollable impacts.

It is pertinent to note that participation in UPSKILL was voluntary; participants could officially withdraw from the program at any time without penalty. Similarly, participants could attend the training offered but block out the content if it was deemed unsuitable, or not useful. Since there were no adverse consequences for officially or unofficially withdrawing from the training, it follows that a rational participant would withdraw if the training were evidently harmful. In other words, even if the *potential* impact of the training on an individual could be zero or negative, it is expected in that situation that the participant would withdraw or ignore it. An important assumption of this analysis is that it is reasonable to expect that the program's *actual* impact on skill gain would only

be positive, i.e. it is not expected that a participant's Essential Skills deteriorated or remained static because of the training.[5]

The intervention's *actual* impact on an individual can be written as d_i:

$$(3)\ d_i = \begin{cases} D_i & if\ D_i > 0 \\ 0 & if\ D_i \leq 0 \end{cases};$$

$$(4)\ \mathrm{G}(d_i < \mathrm{z}|\mu,\sigma) = 0 \text{ if } \mathrm{z} < 0;\ \mathrm{G}(d_i < \mathrm{z}|\mu,\sigma) = \Gamma(D_i < z|\mu,\sigma) = \Phi\left(\frac{\mathrm{z}-\mu}{\sigma}\right), \text{if } \mathrm{z} \geq 0;$$

where G is the cumulative density function of individual impacts, Φ is the cumulative density function of the standard normal distribution, μ is the mean potential impact and σ is the standard deviation of potential impacts.[6,7]

With the assumptions and specification on the distribution of actual individual impacts, the distribution of the program group member skill gain can be expressed as a function of the distribution of the counterfactual skill gain in the absence of the program, and the distribution of impact:

$$(5)\ \mathrm{F^P}(\mathrm{s} < \mathrm{k}|\text{treated}) = \mathrm{F^P}(\mathrm{s} < \mathrm{k}|\text{not treated}) - \int_{-\infty}^{\mathrm{k}} \mathrm{f^P}(\mathrm{c}|\text{not treated})\mathrm{G}(\mathrm{d} > \mathrm{k} - \mathrm{c})\,\mathrm{dc} + \varepsilon(\mathrm{k})$$

where $\mathrm{F^P}$ is the cumulative density function of skill gains of program group members, $\mathrm{f^P}$ is the corresponding probability density function, and ε(k) captures the unknown residual transition from the not-treated state to the treated state. Equation (5) specifies the stochastic process of the treatment effect: it simply states that the impact of the program is shifting some participants with low skill gains to the level of higher skill gains, and that the treatment contributes to the change in distribution of the measured outcome.

$\mathrm{F^P}(\mathrm{s} < \mathrm{k}|\text{treated})$ can be estimated directly with the empirical density function using program group data in an experiment. While $\mathrm{F^P}(\mathrm{s} < \mathrm{k}|\text{not treated})$ and $\mathrm{f^P}(\mathrm{c}|\text{not treated})$ are not observable and not directly estimated, they can be estimated indirectly with empirical density functions using control group observations, $\mathrm{F^C}(\mathrm{s} < \mathrm{k}|\text{not treated})$ and $\mathrm{f^C}(\mathrm{c}|\text{not treated})$, respectively in an experiment. That is:

$$(6)\ \mathrm{F^P}(\mathrm{s} < \mathrm{k}) = \mathrm{F^C}(\mathrm{s} < \mathrm{k}) - \int_{-\infty}^{\mathrm{k}} \mathrm{f^C}(\mathrm{c})\mathrm{G}(\mathrm{d} > \mathrm{k} - \mathrm{c}|\mu,\sigma)\,\mathrm{dc} + \varepsilon(\mathrm{k}).$$

At any level of outcome, *k*, the first two terms of the right hand side of equation 6 represent the "simulated" cumulative density if the control group were subjected to the intervention, while the left hand side is the actual cumulative density of those in the treatment group who were subjected

5 In the case of UPSKILL, quality control measures were implemented to ensure that the training was beneficial to participants and their employers, and it is believed that the negative impact scenario happened rarely, even if it were possible.

6 Although the assumption of normal distribution is a strong assumption, it is not unreasonable in the context of literacy and essential skills. The scales of measuring literacy and essential skill are linked to the assessments used in the International Adult Literacy Survey which were designed to produce normally distributed scores in the population.

7 Without loss of generality, μ is likely to be greater than zero, and so μ is the "median impact" under the distribution. However, σ is a measure of variation but it is not the "standard deviation" of the actual impact, because of the truncation of non-positive values.

to the intervention. The estimation of the parameters of impact distribution (μ, σ) involves locating the estimates such to minimize the difference between the simulated cumulative density and the actual cumulative density function.

Estimation of impact distribution using UPSKILL data

With the UPSKILL data, skill gains of each participant is binned into 36 categories of 5 point skill change in document use (with the exception of the bottom group – up to -86 points – and the top group of 90 points or more in skill change) to approximate the empirical density functions $\hat{F}^C$ and $\hat{f}^C$.

For each category of skill change, *k*, the deviation of the actual and simulated distributions is

$$(7)\ e(k) = \hat{F}^P(k) - [\hat{F}^C(k) - \sum_{c<k} \hat{f}^C(c) G(d > k - c|u,s)] = \hat{F}^P(k) - [\hat{F}^C(k) - \sum_{c<k} \hat{f}^C(c) \left(1 - \Phi\left(\frac{k-c-u}{s}\right)\right)],$$

where u and s are the estimators of mean, μ, and standard deviation, σ, of the potential impact distribution, respectively. In actual estimation, numerical minimization of the sum of squared deviations of equation (7) is conducted. Since the results of this analysis are used in the financial analysis of the reimbursement structure of the two SRDC social finance projects, in which the main outcome of interest is the proportion of participants with 25 or more points gain in skills, the numerical minimization is also subjected to an additional constraint such that the simulated cumulative density is exactly the same as the actual cumulative density at the 25-plus point gain. It is found that UPSKILL's *potential* impacts on the essential skill of document use are distributed with a mean and standard deviation of 7.09 and 12.89 points respectively.

Identifying the risk of achieving certain levels of program outcomes using UPSKILL data

This section demonstrates four possible scenarios of simulated outcomes and risks based on the proportion of controllable vs uncontrollable impact variation. The scenarios were produced to investigate how the potential outcomes and their associated risks vary under different unknown scenarios. The extreme scenario in scenario #1 – where 100% of the impact variation is controllable – provides the upper bound of risk, but is not likely to be realistic. Scenarios 2, 3, and 4 with various levels of controllable impact establish the bounds of more realistic risk scenarios.

It is important to note that SRDC's earlier analysis of 25-point gains was based on microdata from three projects; two were workplace-based (UPSKILL and WTP) and one was for unemployed (FOUNDATIONS). SRDC's proposed benchmarks for the ES SIB projects set the targets based on the appropriate dataset(s) for each pilot; AWES was based on UPSKILL and WTP, and the CICan project based on FOUNDATIONS. Because the current analysis is based on UPSKILL – a workplace ES training program – comparison with the AWES pilot project parameters is a closer fit. Thus, the scenarios in this section refer to the proposed benchmark outcome measure of 40% of the participants achieving gains of 25 points or more, which was the AWES benchmark based on the earlier SRDC analysis of 25-point gains. The findings from the risk profile analysis are nonetheless applicable to both the AWES and CICan projects, and both are included in the Section 4 analysis.

The risk scenarios are based on an assumption that the program delivery agent adopts a strategy of recruitment and implementation such that the controllable portion of impact is at the level Z. The distribution of actual impact is then a shifted distribution of the uncontrollable portion of impact, i.e.

$$(8)\ d_i = \begin{cases} Z + U_i & if\ Z + U_i > 0 \\ 0 & if\ Z + U_i \leq 0 \end{cases}$$

With (8), it is possible to simulate the cumulative density of an outcome:

$$(9)\ \ddot{F}(k|Z) = \hat{F}^C(k) - \sum_{c<k} \hat{f}(c)\ddot{G}(d > k - c|O = Z, u, s_U),$$

where $\ddot{G}$ is the distribution function of actual impact given that the controllable portion of impact is Z, the estimated mean level of the controllable portion of impact is u, and the estimated standard deviation of the uncontrollable portion of impact is s_U. The cumulative density of an outcome is non-decreasing in the controllable portion of impact Z. Equation (9) also identifies that it may be possible to achieve any cumulative density of the outcome at a particular level of k by identifying the required level of controllable impact.

The probability that the program delivery agent can choose the controllable portion of impact at the minimum level Z is given by $\widehat{K}(O_i \geq Z|u, s_O)$. Therefore, as long as the ratio of standard deviations between the controllable portion and the potential impact is known, it is possible to identify the required level of impact to achieve any outcome and the associated risk of achieving that level of impact.

Scenario 1: completely controllable impacts

Since u and s are the estimated parameters of the potential impact distribution, the associated potential impact distribution is also the distribution bound of the controllable portion of impacts as noted in (2). That is, at the extreme $\varphi = 1$, the standard deviation of the controllable portion of impacts is $s_O = s$, and the standard deviation of the uncontrollable portion of impacts is $s_U = 0$. Under this extreme scenario, if the program delivery agent adopts a strategy of recruitment and implementation that is equivalent to choosing the controllable portion of impact at the level Z, the cumulative density function of realizable impacts becomes:

$$(10)\qquad \ddot{G}(d_i < Z) = 0;\ \ddot{G}(d_i \leq Z) = 1.$$

With (10), it is possible to simulate the cumulative density of the outcome:

$$(11)\qquad \ddot{F}(k|Z) = \begin{cases} 0 & if\ k \leq Z \\ \widehat{F^C}(k - Z) & if\ k > Z \end{cases}.$$

For SRDC's SIB projects, one of two key outcome indicators of success is the proportion of participants with 25-point or more skill gains in document use. In theory, with equation (11), the required level of impact to achieve each given percentage point of participants with 25 plus point skill gains can be identified (denoted as Z_F below). However, the empirical cumulative density function used in actual estimation is a discrete function (for every 5 points of skill change to simplify the estimation process), and the simulated cumulative density function (11) is also discrete in the required levels of impacts.

The probability of achieving a specific percentage point of participants with 25-plus point skill gains is:

(12) $$\hat{k}(Z_F|u, s_O) = \phi(^{Z_F - u}/_{S}).$$

Table 1 presents potential outcomes and their associated risk under the completely controllable impacts scenario using UPSKILL data. The leftmost column contains all the plausible percentages of participants with 25-point or more skill gains in document use, if UPSKILL were replicated and it were possible to precisely identify and implement the conditions for those participants with particular levels of skill gain. The second column presents the required minimum level of skill gains within each group, in order to achieve the corresponding percentage achieving 25 points or more. For example, in order to achieve 47 per cent or more participants with a 25-point skill gain, the delivery agent should expect to recruit participants and create conditions such that the minimum skill gain for all participants is 20 points. The third column is the marginal probability of achieving the outcome. Compared to achieving 47 per cent or more participants with a 25-point skill gain, the probability of achieving 43 per cent of more participants with a 25-point skill gain is increased by only 11.17 percentage points. The fourth and fifth columns present the corresponding expected mean and median skill gain of the participants at each plausible outcome.

Table 1 Probability Distribution of Potential Essential Skills Intervention Outcomes – Completely Controllable

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
29	0	29.01%	3.42	4
33	5	14.46%	8.60	9
38	10	15.40%	13.51	14
43	15	14.13%	18.43	19
47	20	11.17%	23.31	24
62	25	7.61%	28.21	29
63	30	4.46%	33.03	34
67	35	2.26%	37.71	39
70	40	0.98%	42.35	44
75	45	0.37%	47.10	49
84	50	0.12%	51.71	54
85	55	0.03%	56.33	59
88	60	0.01%	60.61	64

Source: Calculations based on UPSKILL assessments of participants' document use skills.

When the impact variation is completely controllable, in order to achieve at least 67% of participants with 25-point or more gain, the program delivery agent has to recruit all the available

participants and use instructors and environment factors to achieve 35-point or more skill gains for all participants. The probability of that happening is 2.26%. It is obvious that the probability of achieving 70 per cent of more participants with 25-point or more skill gains is extremely small, at less than 2%.

Scenario 2: most of the potential impact variation is controllable

In this scenario, 90 per cent of the variation of the potential impact is from the controllable portion of impacts. This is similar to Scenario 1, but more realistic, since it is basically impossible to control all potential factors of impacts.

When $\varphi = 0.9$, the variance of the controllable portion of impacts is $s_O^2 = 0.9\, s^2$, and the variance of the uncontrollable portion of impacts is $s_U^2 = 0.1\, s^2$. If the controllable portion of impact is at the level Z, the remaining variation in potential impact comes from the uncontrollable portion. Therefore, the cumulative density function of realizable impacts becomes:

$$\ddot{\mathrm{G}}(d_i > \mathrm{z}|\mathrm{O} = \mathrm{Z}, \mathrm{u}, \mathrm{s_U}) = \begin{cases} 1 - \Phi\left(\frac{\mathrm{z}-\mathrm{Z}}{\mathrm{s_U}}\right) & \text{if } \mathrm{z} - \mathrm{Z} \geq 0 \\ 1 & \text{if } \mathrm{z} - \mathrm{Z} \leq 0 \end{cases} \tag{13}$$

With equation (12), the cumulative density function of the outcome is:

$$\ddot{\mathrm{F}}(\mathrm{k}|\mathrm{Z}) = \hat{\mathrm{F}}^{\mathrm{C}}(\mathrm{k}) - \sum_{\mathrm{c}<\mathrm{k}} \hat{\mathrm{f}}(\mathrm{c})\left(1 - \Phi\left(\frac{\mathrm{k}-\mathrm{c}-\mathrm{Z}}{\mathrm{s}\sqrt{0.1}}\right)\right), \tag{14}$$

Again, it is possible to identify the required level of impact Z to achieve any percentage of participants with a particular level of skill gains by solving equation (14). The probability of achieving a specific percentage point of participants with 25-plus point skill gains is given by:

$$\hat{\mathrm{k}}(Z_F|u, s_O) = \phi\left(\left.Z_F - u\right/ s\sqrt{0.9}\right). \tag{15}$$

Table 2 presents potential outcomes and the associated risk under the mostly controllable impacts scenario using UPSKILL data.

Table 2 Probability Distribution of Potential Essential Skills Intervention Outcomes – Mostly Controllable

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
29	-5.33	15.49%	3.95	4.42
30	-2.43	6.33%	5.03	5.84
31	-0.66	4.49%	6.11	7.29
32	0.77	3.94%	7.19	8.43
33	2.03	3.71%	8.28	9.50
34	3.22	3.62%	9.36	10.71
35	4.36	3.58%	10.44	11.89

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
36	5.47	3.57%	11.52	12.95
37	6.57	3.56%	12.60	13.95
38	7.65	3.53%	13.66	14.97
39	8.71	3.47%	14.71	16.19
40	9.76	3.36%	15.73	17.26
41	10.78	3.21%	16.73	18.22
42	11.75	3.00%	17.69	19.11
43	12.68	2.76%	18.60	20.00
44	13.54	2.50%	19.45	21.00
45	14.35	2.25%	20.23	21.85
46	15.10	2.01%	20.97	22.58
47	15.80	1.81%	21.65	23.24
48	16.46	1.63%	22.30	23.84
49	17.08	1.49%	22.90	24.42
50	17.68	1.37%	23.49	25.00
51	18.26	1.27%	24.05	25.68
52	18.82	1.19%	24.61	26.30
53	19.38	1.13%	25.15	26.88
54	19.94	1.08%	25.70	27.43
55	20.51	1.04%	26.25	27.97
56	21.09	1.01%	26.82	28.50
57	21.69	0.99%	27.40	29.06
58	22.33	0.99%	28.02	29.66
59	23.01	0.99%	28.68	30.40
60	23.76	1.00%	29.40	31.24
61	24.59	1.02%	30.20	32.09
62	25.54	1.05%	31.11	33.00
63	26.63	1.07%	32.16	34.00
64	27.88	1.04%	33.35	35.24
65	29.25	0.96%	34.65	36.75
66	30.70	0.82%	36.01	38.14
67	32.16	0.66%	37.38	39.50
68	33.62	0.51%	38.74	41.09
69	35.04	0.39%	40.07	42.53
70	36.39	0.28%	41.33	43.78
71	37.64	0.20%	42.51	44.96
72	38.78	0.15%	43.58	46.26
73	39.82	0.11%	44.56	47.32
74	40.79	0.08%	45.46	48.23

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
75	41.69	0.06%	46.30	49.05
76	42.55	0.05%	47.10	49.87
77	43.38	0.04%	47.88	50.82
78	44.20	0.03%	48.64	51.70
79	45.03	0.02%	49.42	52.52
80	45.89	0.02%	50.21	53.32
81	46.81	0.02%	51.06	54.16
82	47.81	0.01%	51.98	55.16
83	48.97	0.01%	53.04	56.46
84	50.35	0.01%	54.30	57.82
85	52.10	0.01%	55.87	59.44
86	54.30	0.01%	57.81	61.81
87	56.97	0.00%	60.12	64.32

Source: Calculations based on UPSKILL assessments of participants' document use skills.

Scenario 3: half of the potential impact variation is controllable

When $\varphi = 0.5$, the variance of the controllable portion of impacts is $s_O^2 = s^2/2$, and the variance of the uncontrollable portion of impacts is $s_U^2 = s^2/2$. That is, half of the variation of the potential impact is from the controllable portion of impacts.

If the controllable portion of impact is at the level Z, the cumulative density function of the outcome is:

$$\ddot{\mathrm{F}}(\mathrm{k}|\mathrm{Z}) = \hat{\mathrm{F}}^{\mathrm{C}}(\mathrm{k}) - \sum_{\mathrm{c}<\mathrm{k}} \hat{\mathrm{f}}(\mathrm{c})\left(1 - \Phi\left(\frac{\mathrm{k}-\mathrm{c}-\mathrm{Z}}{\mathrm{s}\sqrt{0.5}}\right)\right), \tag{16}$$

Again, it is possible to identify the required level of impact Z to achieve any given percentage of participants with a particular level of skill gains by solving equation (16). The probability of achieving a specific percentage point of participants with 25-plus point skill gains is given by:

$$\hat{\mathrm{k}}(Z_F|u, s_O) = \phi\left(\left.Z_F - u\right/_{s\sqrt{0.5}}\right). \tag{17}$$

Table 3 presents potential outcomes and the associated risk under the half-controllable impacts scenario using UPSKILL data.

Table 3 Probability Distribution of Potential Essential Skills Intervention Outcomes – Half Controllable

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
29	-13.23	1.29%	3.95	4.33
30	-7.72	3.92%	5.02	5.18
31	-4.66	4.66%	6.07	6.75
32	-2.38	5.06%	7.11	8.08
33	-0.52	5.27%	8.14	9.30
34	1.10	5.36%	9.14	10.46
35	2.55	5.36%	10.13	11.56
36	3.87	5.28%	11.10	12.61
37	5.09	5.14%	12.05	13.61
38	6.24	4.96%	12.98	14.59
39	7.33	4.75%	13.89	15.57
40	8.36	4.51%	14.79	16.51
41	9.35	4.26%	15.67	17.43
42	10.31	4.00%	16.53	18.32
43	11.23	3.73%	17.37	19.19
44	12.13	3.47%	18.21	20.05
45	13.01	3.21%	19.03	20.91
46	13.87	2.95%	19.84	21.74
47	14.71	2.71%	20.64	22.56
48	15.55	2.47%	21.44	23.38
49	16.37	2.25%	22.23	24.19
50	17.19	2.04%	23.01	25.00
51	18.01	1.84%	23.80	25.82
52	18.83	1.65%	24.58	26.63
53	19.65	1.47%	25.37	27.43
54	20.47	1.31%	26.16	28.25
55	21.30	1.15%	26.96	29.07
56	22.13	1.01%	27.76	29.91
57	22.98	0.88%	28.58	30.77
58	23.85	0.76%	29.41	31.63
59	24.73	0.65%	30.25	32.50
60	25.62	0.55%	31.11	33.39
61	26.54	0.46%	31.99	34.31
62	27.48	0.38%	32.88	35.25
63	28.44	0.31%	33.80	36.22
64	29.43	0.24%	34.73	37.20
65	30.43	0.19%	35.69	38.20

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
66	31.46	0.15%	36.66	39.23
67	32.51	0.11%	37.66	40.29
68	33.58	0.08%	38.66	41.36
69	34.66	0.06%	39.68	42.44
70	35.76	0.04%	40.71	43.52
71	36.86	0.03%	41.74	44.62
72	37.97	0.02%	42.78	45.75
73	39.09	0.01%	43.82	46.86
74	40.21	0.01%	44.86	47.98
75	41.35	0.01%	45.92	49.11
76	42.50	0.00%	46.98	50.27
77	43.66	0.00%	48.05	51.44

Source: Calculations based on UPSKILL assessments of participants' document use skills.

Scenario 4: potential impact variation is mostly uncontrollable

Comparing Scenario 3 to Scenario 2 shows that the range of realistic potential outcomes shrinks substantially when the proportion of variance that is controllable decreases from 0.9 to 0.5. At the other extreme when φ is almost zero, the potential impacts are nearly uncontrollable and random, and a large-scale program similar to UPSKILL is going to mimic the results of UPSKILL with almost no risk. This is, however, a moot point since it is not applicable to the real world. But what if a very high proportion of impact variation is not controllable? Say 90%?

When $\varphi = 0.1$, the variance of the controllable portion of impacts is $s_O^2 = 0.1s^2$, and the variance of the uncontrollable portion of impacts is $s_U^2 = 0.9s^2$.

The cumulative density function of the outcome is:

(18) $$\ddot{\mathrm{F}}(\mathrm{k}|\mathrm{Z}) = \hat{\mathrm{F}}^{\mathrm{C}}(\mathrm{k}) - \sum_{\mathrm{c}<\mathrm{k}} \hat{\mathrm{f}}(\mathrm{c})\left(1 - \Phi\left(\frac{\mathrm{k}-\mathrm{c}-\mathrm{Z}}{\mathrm{s}\sqrt{0.9}}\right)\right),$$

Again, it is possible to identify the required level of impact Z to achieve any given percentage of participants with a particular level of skill gains by solving equation (18). The probability of achieving a specific percentage point of participants with 25-plus point skill gains is given by:

(19) $$\hat{\mathrm{k}}(Z_F|u, s_O) = \phi\left(\left.Z_F - u\right/ s\sqrt{0.1}\right).$$

Table 4 presents potential outcomes and the associated risk under the scenario when only one tenth of impacts is controllable using UPSKILL data.

Table 4 Probability Distribution of Potential Essential Skills Intervention Outcomes – Mostly Uncontrollable

Simulated Percentages (Rounded) of UPSKILL Participants with 25-point or More Skill Gain	Required Level of Controllable Portion of Skill Gain	Probability	Expected Mean Skill Gain	Expected Median Skill Gain
29	-18.75	0.00%	3.94	4.29
30	-11.84	0.00%	4.98	4.93
31	-8.10	0.01%	6.00	6.34
32	-5.36	0.10%	6.99	7.66
33	-3.15	0.49%	7.97	8.87
34	-1.26	1.43%	8.93	10.01
35	0.42	3.06%	9.88	11.09
36	1.94	5.22%	10.82	12.12
37	3.33	7.55%	11.74	13.12
38	4.64	9.55%	12.65	14.10
39	5.87	10.86%	13.54	15.06
40	7.04	11.28%	14.43	16.03
41	8.16	10.84%	15.31	16.97
42	9.24	9.74%	16.19	17.90
43	10.28	8.23%	17.05	18.81
44	11.30	6.57%	17.91	19.71
45	12.29	4.98%	18.77	20.61
46	13.26	3.60%	19.62	21.49
47	14.21	2.48%	20.46	22.37
48	15.15	1.63%	21.31	23.25
49	16.08	1.03%	22.15	24.12
50	17.00	0.62%	22.99	25.00
51	17.91	0.36%	23.84	25.87
52	18.82	0.20%	24.68	26.74
53	19.73	0.10%	25.53	27.62
54	20.64	0.05%	26.38	28.50
55	21.54	0.02%	27.23	29.39
56	22.46	0.01%	28.09	30.29
57	23.37	0.00%	28.95	31.19

Source: Calculations based on UPSKILL assessments of participants' document use skills.

Ex ante analysis of the risk of potential outcomes

As shown in the scenarios above, it is apparent that the spread of risk is directly related to the ratio of variance of controllable impacts to the variance of all potential individual impacts. For example, the probability of achieving 65% of participants with 25-points or more gains in document use

score is 4.45% if 90% of impact variation is controllable, but 0.71% if only half of the impact variation is controllable, and almost 0% if only 10% of the impact variation is controllable. The higher the proportion of impact variance is controllable, the larger the spread and risk of potential outcomes.

This result seems counterintuitive. How can a more controllable situation be riskier than a less controllable situation? It should be noted that uncontrollable individual impacts are distributed symmetrically around zero. By the law of large numbers, the proportion of participants with certain levels of skill gain because of the uncontrollable impact is the same as the proportion of participants with the same level of skill loss because of the controllable impact.[8] That is, the uncontrollable portion of individual impact is not affecting the potential outcomes of the sample (as measured by the proportion of participants with 25-point or more skill gains) substantially in a large sample. As a result, the only major source of potential outcome variation originates from the strategies adopted by the program delivery agent to affect the impacts. The higher the variance of the controllable portion of impact is, the higher the variation or risk of the potential outcomes.

Without knowing the value of the ratio of controllable impact variance to the total impact variance, it is impossible to pinpoint the exact risk of a proposed intervention. However, the scenarios presented above provide bounds of the risk involved. This is sufficient for financial analysis for pay-for-success structures. Since a program delivery agent may increase this ratio somewhat by dedicating more resources to identify high impact conditions, the exact value of the ratio may not be of importance in designing a SIB reimbursement formula. Instead, assuming a high ratio of variance of controllable impacts and providing reward to program delivery agents of achieving a higher outcome may provide the incentives required to improve the program. The ex ante analysis of the risk of potential outcomes should therefore focus on the situation when impact variation is mostly controllable – the case in Scenario 2.

As shown in Tables 1 to 4, a higher ratio of controllable impacts also implies a higher expected mean/median skill gain. Given that one of the two key outcome indicators for the SIB reimbursement formulae is a median skill gain target to reduce the risk of delivery partners focusing efforts on participants more likely to succeed, it is proposed that the formulae take the lesser risk scenario (such as Scenario 4) such that the median skill gain target is more realistically achievable.

[8] Skill loss in this sense refers to natural depreciation of human capital, or measurement error; it is not expected to be a result of the training.

4. Application of the model in designing the reimbursement formulae of the two SIB projects

In order to attract investors, the pilot projects need to establish realistic performance targets. This section applies the distribution of document use impacts from UPSKILL to provide a financial analysis of the proposed reimbursement formulae.

Based on the two-factor reimbursement formulae presented in SRDC's earlier analysis of 25-point gains, OLES has determined that performance targets should include the following two factors:

1. Minimum percentage of participants achieving 25-point (or more) skill gains; and
2. Median skill gain for the group of participants.

Skill measurements will be conducted at baseline (before the intervention), right after the intervention (post-test) to measure the skill gains, and 12 months following the end of training (post-post-test) to measure skill gain maintenance.

Inclusion of the median skill gain as one of the performance targets ensures that skill gain is not simply concentrated around the few top performers, but shared by many participants. This safeguards against the possibility of delivery agents spending less time or attention on participants less likely to succeed in favour of those more likely to achieve targets. If the median skill gain target is realistic and the delivery agents are not engaged in strategic behaviour, focusing on the benchmark percentage of participants to achieve 25-point (or more) skill gains should automatically qualify to meet the associated median skill gains required for the benchmark scenarios.[9] (See the mean and median skill gains associated with certain proportions of the population achieving 25-point gains or more, as shown in the rightmost two columns of Tables 1 to 4.)

Both pilot projects require that the reimbursement formulae include some payment based on measurement of skill gains 12 months following the end of the treatment, to ensure that skill gains realised right after the intervention are sustained, and not a result of random measurement error, nor strategic behaviour on the part of participants or delivery agents. However, based on well-documented experience in other projects, sample attrition is expected to be substantial 12 months following the end of the intervention, such that it will be difficult to accurately measure the maintenance of skill gains. The skill maintenance requirement is thus interpreted as finding sufficient evidence of that sampling error was not a main contributor to measured skill gain right after the intervention, because the 12-month measurement is outside the margin-of-error of the post-training assessment. If the typical level of type I error at 5 per cent is adopted, the interpretation suggests that there is only a 5 per cent probability that the 12-month measurement demonstrates evidence of failing to maintain skills.

[9] The benchmark median calculated in this analysis is 16 points for AWES, down from 18 points in the benchmark analysis that included data from the Workplace Training Program, in which participants demonstrated higher skill gains overall.

Since it is reasonable to assume that individual impact variance is not completely controllable, the next part of the analysis will consider the outcomes and risks under Scenarios 2, 3, and 4. OLES has provided guidance on the desirable outcomes and the bounds of reimbursements in each of the two SIB projects. SRDC's analysis found under it is possible to identify reimbursement structure and thresholds that are achievable and attractive to investors.

Alberta Workforce Essential Skills Society social impact bond project

Because the AWES pilot project is expected to be similar to UPSKILL in terms of the duration and workplace-based focus of the ES training, information from UPSKILL serves as a good reference for the potential outcomes of the AWES pilot project. Indeed, it is expected that type of employers and participants recruited for AWES are likely to be similar to that of the UPSKILL, such that UPSKILL's control group data serves as a good representative distribution of skill gains in the absence of the intervention under AWES.

Table 5 presents the range of potential outcomes and the associated probability under the three scenarios. The table simply summarizes the findings presented previously in Tables 2 to 4.

Table 5 Probability Distribution and Expected Median Skill Gain of Potential Outcomes – AWES

	Probability			Expected Median Skill Gain		
Simulated Percentages (Rounded) of Participants with 25 Point or More Skill Gain	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
29	15.49%	1.29%	0.00%	4.42	4.33	4.29
30	6.33%	3.92%	0.00%	5.84	5.18	4.93
31	4.49%	4.66%	0.01%	7.29	6.75	6.34
32	3.94%	5.06%	0.10%	8.43	8.08	7.66
33	3.71%	5.27%	0.49%	9.50	9.30	8.87
34	3.62%	5.36%	1.43%	10.71	10.46	10.01
35	3.58%	5.36%	3.06%	11.89	11.56	11.09
36	3.57%	5.28%	5.22%	12.95	12.61	12.12
37	3.56%	5.14%	7.55%	13.95	13.61	13.12
38	3.53%	4.96%	9.55%	14.97	14.59	14.10
39	3.47%	4.75%	10.86%	16.19	15.57	15.06
40	3.36%	4.51%	11.28%	17.26	16.51	16.03
41	3.21%	4.26%	10.84%	18.22	17.43	16.97
42	3.00%	4.00%	9.74%	19.11	18.32	17.90
43	2.76%	3.73%	8.23%	20.00	19.19	18.81
44	2.50%	3.47%	6.57%	21.00	20.05	19.71
45	2.25%	3.21%	4.98%	21.85	20.91	20.61
46	2.01%	2.95%	3.60%	22.58	21.74	21.49

Simulated Percentages (Rounded) of Participants with 25 Point or More Skill Gain	Probability			Expected Median Skill Gain		
	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
47	1.81%	2.71%	2.48%	23.24	22.56	22.37
48	1.63%	2.47%	1.63%	23.84	23.38	23.25
49	1.49%	2.25%	1.03%	24.42	24.19	24.12
50	1.37%	2.04%	0.62%	25.00	25.00	25.00
51	1.27%	1.84%	0.36%	25.68	25.82	25.87
52	1.19%	1.65%	0.20%	26.30	26.63	26.74
53	1.13%	1.47%	0.10%	26.88	27.43	27.62
54	1.08%	1.31%	0.05%	27.43	28.25	28.50
55	1.04%	1.15%	0.02%	27.97	29.07	29.39
56	1.01%	1.01%	0.01%	28.50	29.91	30.29
57	0.99%	0.88%	0.00%	29.06	30.77	31.19
58	0.99%	0.76%	0.00%	29.66	31.63	
59	0.99%	0.65%	0.00%	30.40	32.50	
60	1.00%	0.55%	0.00%	31.24	33.39	
61	1.02%	0.46%	0.00%	32.09	34.31	
62	1.05%	0.38%	0.00%	33.00	35.25	
63	1.07%	0.31%	0.00%	34.00	36.22	
64	1.04%	0.24%	0.00%	35.24	37.20	
65	0.96%	0.19%	0.00%	36.75	38.20	
66	0.82%	0.15%	0.00%	38.14	39.23	
67	0.66%	0.11%	0.00%	39.50	40.29	
68	0.51%	0.08%	0.00%	41.09	41.36	
69	0.39%	0.06%	0.00%	42.53	42.44	
70+	1.11%	0.12%	0.00%	43.78	43.52	

Source: Calculations based on UPSKILL assessments of participants' document use skills.

It is expected that the AWES pilot project will not perform worse than the control group of UPSKILL (where they received no UPSKILL intervention), where the minimum realizable (rounded) proportion of participants with 25-point or more skill gains in document use was about 29 per cent. Depending on the scenario of controllable impact variation, potential outcomes are likely to be between 29 to 60 points. With the distributions of potential outcomes, it is possible to estimate the mean and standard deviation of rates of return to the employers under various payment structures.

SIB repayment schemes are typically not binary, i.e. based on simply achieving a success outcome or not, but rather are based on benchmark success outcomes, with variable payments according to the level of 'success' achieved. For the pilot projects, benchmark outcomes (chart a. below) were set based on the analysis of 25-point gains from previous Essential Skills training interventions.

Chart b. proposes a variable payment structure with minimum and maximum payments made both below and above the benchmark.

a. Benchmark payment structure for AWES:

Benchmarks	Meet post-test skill attainment	Meet post-post-test skill attainment and maintenance
▪ Median gain ≥16 points ▪ Percentage with 25-point gain ≥ 40%	Reimbursement of 45% of employer's training costs	Additional 5% of employer's cost

b. Proposed variable payment structure for AWES:

Target outcomes	Meet post-test skill attainment	Meet post-post-test skill attainment and maintenance
▪ Median gain ≥ 4 points ▪ Percentage with 25-point gain ≥ 30%	Reimbursement of 30% of employer's training costs	Additional 2% of employer's cost
▪ Median gain ≥16 points ▪ Percentage with 25-point gain ≥ 40%	Reimbursement of 40% of employer's training costs	Additional 3% of employer's cost
▪ Median gain ≥20 points ▪ Percentage with 25-point gain ≥ 45%	Reimbursement of 42.5% of employer's training costs	Additional 4% of employer's cost
▪ Median gain ≥25 points ▪ Percentage with 25-point gain ≥ 50%	Reimbursement of 45% of employer's training costs	Additional 5% of employer's cost

In order to calculate the rate of return to AWES employers, the cost and net profit figures from the UPSKILL project were used. The following are the calculation assumptions:

- Expected mean impact on the percentage with 25-point or more skill gain: 11.8 points,[10]
- Cost of Delivery (C): $2,863 per participant,
- Natural rate of participant with 25+ point skill gain: 28.51%,
- Net benefit (productivity gain) to firm from the impact: $3,112 (or $263.73 per percentage point),
- Net revenue to the government: $1,006 (or $85.25 per percentage point),
- Net benefit (income gain) to participants: $828 (or $70.17 per percentage point),

[10] The expected mean impact on the percentage with 25-point skill gains is slightly different from the UPSKILL published figure of 12.6 percentage points because of minor difference in the treatment of skill gain of a few observations at the low end of the skill distribution because of artifacts of the scoring mechanism in assessments.

- Probability of failing to maintain skill gains in post-post-test: 5%,[11]
- Risk-free rate: 0.03%.[12]

Table 6 presents the financial indicators of investing in AWES, under the two proposed payment structures and the scenarios of controllable impacts. Under any circumstance, the mean rate of return (within a year) is over 35%, which would likely be considered respectable from an investment perspective. The variable payment structure is superior for employers under any scenario simply because of the lower downside risk from partial reimbursement for outcomes under the 40 per cent. The higher mean rates of return and lower standard deviations result into higher Sharpe ratios (the ratio of risk premium over risk) for the variable payment structure, compared to that of the simple benchmark.

In finance, Sharpe ratio (the excess return beyond the risk-free interest rate divided by the standard deviation of the rate of return) is an indicator commonly used to assess the risk-return profile of an investment. In historical calculations, annualized Sharpe ratio typically varies from 0.4 to 0.8.[13]

For AWES, Sharpe ratios range from 0.285 to 0.821 under the simple benchmark payment structure, and from 0.326 to 1.316 under the variable payment structure. Even under the most risky scenario (when most of individual impact variation is controllable), the variable payment structure offers a risk-return profile comparable to that of traditional investments in the financial market.

11 Failing the maintenance requirement is equivalent to finding evidence to suggest that what is observed at the post-test is simply due to sampling error. Therefore, the probability of failing the maintenance requirement is assumed to be the same as level of significance.

12 US 3-month T-bill rate as of July 31, 2014.

13 The annualized Sharpe ratio for Warren Buffett's investments, from 1976-2011, is 0.76. The overall stock market performance is only 0.39. Source: Frazzini, Andrea, Kabiller, David, and Pedersen, Lasse Heje (2013): "Buffett's Alpha". Yale University Department of Economics Working Paper. Retrieved from http://www.econ.yale.edu/~af227/pdf/Buffett's%20Alpha%20-%20Frazzini,%20Kabiller%20and%20Pedersen.pdf.

Table 6 Financial Performance of Investing in AWES

Financial Indicators	Proposed Benchmark Payment Structure			Proposed Variable Payment Structure		
	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
Employers						
Mean Rate of Return	35.07%	37.59%	43.97%	39.38%	43.87%	51.83%
Standard Deviation of Rate of Return	123.16%	95.62%	53.54%	120.68%	90.70%	39.73%
Sharpe Ratio	0.285	0.393	0.821	0.326	0.483	1.316
Government						
Mean Net Revenue (per participant)	$407.42	$350.53	$168.74	$284.01	170.72	($56.44)
Standard Deviation of Net Revenue (per participant)	$553.15	$443.18	$499.11	$545.53	$398.13	$183.93
Participants						
Mean Benefit (per participant)	$859.49	$862.54	$862.69	$859.49	$862.54	$862.69
Standard Deviation of Benefit (per participant)	$778.35	$567.73	$248.73	$778.35	$567.73	$248.73

On average, the net revenue implication to the government is positive under the simple benchmark payment structure and slightly above/below 0 under the variable payment structure. With the expected impact on participants' benefits taken into consideration (well over $800 per person on average), the average net cost (if any) seems justifiable.

Colleges and Institutes Canada social impact bond project

SRDC expects that the participants recruited for the CICan will be similar to those in previous FOUNDATIONS projects, for which micro-data were used in SRDC's earlier analysis of 25-point gains. However, without a comparison group, there is no reliable source of counterfactual information upon which to estimate the distribution of individual impacts of that Essential Skills intervention on unemployed participants. Thus, the distribution of skill gains of participants for the CICan project in the absence of the Essential Skills intervention is simulated by making use of the empirical distribution of FOUNDATIONS' program participants, equation (7), and the obtained

distribution of individual impacts from UPSKILL. Once the simulated distribution of skill gains of the counterfactual is obtained, the range of potential outcomes and the associated probability can be assessed under the various assumptions of controllable impacts used in the AWES analysis. Table 7 presents the range of potential outcomes and the associated probability under the three scenarios.

For the purpose of the financial analysis of the repayment structure for the CICan social impact bond project, it is assumed that the potential impacts on document use are distributed with the same mean and standard deviation as in UPSKILL (at 7.09 and 12.89 points respectively).

Table 7 Probability Distribution of Potential Outcomes – CICan

	Probability			Expected Median Skill Gain		
Simulated Percentages (Rounded) of Participants with 25-point or More Skill Gain	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
36	15.82%	1.38%	0.00%	16.06	16.04	16.04
37	4.17%	2.28%	0.00%	16.71	16.66	16.66
38	3.12%	2.64%	0.00%	17.35	17.28	17.28
39	2.68%	2.88%	0.01%	17.99	17.90	17.90
40	2.45%	3.06%	0.04%	18.63	18.53	18.53
41	2.31%	3.20%	0.13%	19.27	19.15	19.15
42	2.22%	3.30%	0.34%	19.92	19.79	19.79
43	2.17%	3.38%	0.73%	20.56	20.43	20.43
44	2.13%	3.44%	1.34%	21.20	21.07	21.07
45	2.11%	3.48%	2.20%	21.83	21.71	21.71
46	2.10%	3.50%	3.26%	22.47	22.36	22.36
47	2.09%	3.50%	4.46%	23.10	23.01	23.01
48	2.09%	3.49%	5.67%	23.73	23.67	23.67
49	2.08%	3.46%	6.77%	24.36	24.33	24.33
50	2.08%	3.43%	7.65%	25.00	25.00	25.00
51	2.08%	3.37%	8.22%	25.64	25.67	25.67
52	2.07%	3.31%	8.44%	26.27	26.35	26.35
53	2.06%	3.24%	8.29%	26.90	27.03	27.03
54	2.05%	3.15%	7.83%	27.54	27.71	27.71
55	2.03%	3.06%	7.11%	28.17	28.40	28.40
56	2.01%	2.96%	6.23%	28.81	29.10	29.10
57	1.98%	2.85%	5.26%	29.45	29.80	29.80
58	1.96%	2.73%	4.30%	30.10	30.50	30.50
59	1.92%	2.61%	3.39%	30.75	31.22	31.22
60	1.89%	2.48%	2.58%	31.40	31.93	31.93
61	1.85%	2.35%	1.90%	32.06	32.65	32.65

Simulated Percentages (Rounded) of Participants with 25-point or More Skill Gain	Probability			Expected Median Skill Gain		
	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
62	1.81%	2.21%	1.35%	32.71	33.38	33.38
63	1.76%	2.07%	0.93%	33.38	34.12	34.12
64	1.71%	1.94%	0.62%	34.05	34.86	34.86
65	1.66%	1.80%	0.40%	34.73	35.61	35.61
66	1.60%	1.66%	0.24%	35.42	36.37	36.37
67	1.55%	1.52%	0.14%	36.11	37.14	37.14
68	1.49%	1.39%	0.08%	36.81	37.92	37.92
69	1.43%	1.26%	0.05%	37.52	38.70	38.70
70	1.36%	1.13%	0.02%	38.24	39.50	39.50
71	1.30%	1.01%	0.01%	38.97	40.31	40.31
72	1.23%	0.89%	0.01%	39.72	41.14	41.14
73	1.16%	0.78%	0.00%	40.48	41.97	41.97
74	1.10%	0.68%	0.00%	41.25	42.82	42.82
75	1.03%	0.59%	0.00%	42.04	43.69	43.69
76	0.96%	0.50%	0.00%	42.84	44.58	44.58
77	0.89%	0.42%	0.00%	43.66	45.49	45.49
78	0.82%	0.35%	0.00%	44.50	46.42	46.42
79	0.76%	0.29%	0.00%	45.37	47.37	47.37
80	0.69%	0.24%	0.00%	46.26	48.35	48.35
81	0.63%	0.19%	0.00%	47.17	49.36	49.36
82	0.56%	0.15%	0.00%	48.12	50.41	50.41
83	0.50%	0.11%	0.00%	49.09	51.48	51.48
84	0.45%	0.09%	0.00%	50.12	52.60	52.60
85	0.39%	0.06%	0.00%	51.18	53.77	53.77
86	0.34%	0.05%	0.00%	52.28	54.99	54.99
87	0.29%	0.03%	0.00%	53.44	56.27	56.27
88	0.24%	0.02%	0.00%	54.67	57.62	57.62
89	0.20%	0.01%	0.00%	55.98	59.05	59.05
90+	0.58%	0.02%	0.00%	57.37	57.37	60.58

Source: Calculations based on UPSKILL assessments of participants' document use skills.

Based on the simulated distribution, the minimum realizable (rounded) proportion of participants with 25-point or more skill gains in document use is about 36 per cent. Depending on the scenario of controllable impact variation, potential outcomes are likely to be between 36 to 70 points. With the distributions of potential outcomes, it is possible to estimate the mean and standard deviation of rate of returns to the investors under different payment structures.

Charts a. and b. below illustrate the benchmarks (from 25-point gains analysis) and proposed variable payment structure for the CICan project.

a. Benchmark payment structure for CICan:

Benchmarks	Meet post-test skill attainment	Meet post-post-test skill attainment and maintenance
▪ Median gain ≥25 points ▪ Percentage with 25-point gain ≥ 50%	Reimbursement of 100% of investor's initial investment, PLUS 5% return on investment	Additional 5% return on investment

b. Proposed variable payment structure for CICan:

Target outcomes	Meet post-test skill attainment	Meet post-post-test skill attainment and maintenance
▪ Median gain ≥16 points ▪ Percentage with 25-point gain ≥ 36%	Reimbursement of 90% of investor's initial investment	Additional 1% return on investment
▪ Median gain ≥18 points ▪ Percentage with 25-point gain ≥ 40%	Reimbursement of 95% of investor's initial investment	Additional 2% return on investment
▪ Median gain ≥21 points ▪ Percentage with 25-point gain ≥ 45%	Reimbursement of 97.5% of investor's initial investment	Additional 2% return on investment
▪ Median gain ≥25 points ▪ Percentage with 25-point gain ≥ 50%	Reimbursement of 100% of investor's initial investment	Additional 3% return on investment
▪ Median gain ≥25 points ▪ Percentage with 25 point gain ≥ 55%	Reimbursement of 100% of investor's initial investment PLUS 2.5% return on investment	Additional 4% return on investment
▪ Median gain ≥25 points ▪ Percentage with 25-point gain ≥ 60%	Reimbursement of 100% of investor's initial investment PLUS 5% return on investment	Additional 5% return on investment

In order to calculate the rate of return to social bond investors to the CICan project, some cost and net profit figures from the UPSKILL project were used as reference. Note that the targeted participants of the CICan are unemployed, while UPSKILL's participants were all employed, and the training took place in the workplace context. As a result, the costs and benefits of the CICan project are likely significantly different from that of UPSKILL. However, in the absence of another reliable reference for the project's costs and benefits, UPSKILL's figures are used as a reference. Regardless of the assumptions on the exact per participant costs and benefits, the calculations of the potential rates of return for the investors are unaffected since reimbursement is relative to the original cost. The following are the calculation assumptions:

- Expected mean impact on the percentage with 25-point or more skill gain: 11.8 points,
- Cost of Delivery (C): $2,529 per participant,[14]
- Natural rate of participants with 25+ point skill gain without the intervention: 35.16%,
- Net revenue to the government (without accounting for potential savings on income assistance): $1,006 (or $85.25 per percentage point),
- Net benefit (income gain) to participants: $828 (or $70.17 per percentage point),
- Probability of failing to maintain skill gains in post-post-test: 5%,
- Risk-free rate: 0.03%.

Table 8 presents the financial indicators of investing in the CICan project under the proposed payment structures and the scenarios of controllable impacts. Under any of the scenarios, the mean rate of return (within a year) is negative for the benchmark payment structure. In other words, investors cannot expect any return by investing in the project if the reimbursement structure has no partial reimbursement under the 50 per cent threshold. A person or company funding the initial cost of the training in this case would more likely be a philanthropist than a rational investor, and thus the risk-return profiles of either payment structure are not relevant in this type of decision.

With lesser investment loss in the range of 36% to 50%, the mean rates of return become positive under any of the scenarios for the variable payment structure. The Sharpe ratios range from 0.171 to 1.023 under the proposed variable payment structure. Under the most-risky scenario (when most of individual impact variation is controllable), this variable payment structure offers a risk-return profile below that of the traditional investment market. However, the risk-return profiles are more acceptable if less impact variation is controllable (scenarios 3 and 4). Note that based on the assumption of net revenue to the government (without accounting for savings in income assistance), the expected mean net revenue is about -$1,100. This is a relatively small cost compared to typical cost savings when participants find employment and are no longer receiving income assistance. Moreover, based on the reimbursement structure, the risk of paying for an ineffective program is completely eliminated for the government.

[14] The use of the workplace ES training in UPSKILL to estimate delivery costs for this project is done with caveats: training will not take place in a workplace setting, and learners will be under- or unemployed; organizational needs analysis is not likely to be a factor, but the Foundations training may cost more reflecting a more intense intervention.

Table 8 Financial performance of investing in CICan project

Financial Indicators	Proposed Benchmark Payment Structure			Proposed Variable Payment Structure		
	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)	Most (90%) of impact variation controllable (Scenario 2)	Half (50%) of impact variation controllable (Scenario 3)	Little (10%) of impact variation controllable (Scenario 4)
Investors						
Mean Rate of Return	-42.43%	-37.44%	-17.62%	1.31%	2.49%	3.41%
Standard Deviation of Rate of Return	54.81%	54.34%	47.49%	7.52%	5.96%	3.30%
Sharpe Ratio	-0.775	-0.69	-0.372	0.171	0.412	1.023
Government*						
Mean Net Revenue (per participant)	$51.60	($76.32)	($577.53)	($1,054.71)	($1,086.04)	($1,109.34)
Standard Deviation of Net Revenue (per participant)	$764.55	$852.59	$960.16	$1,053.81	$761.41	$321.11
Participants*						
Mean Benefit (per participant)	$1,240.84	$1,239.45	$1,239.47	$1,240.84	$1,239.45	$1,239.47
Standard Deviation of Benefit (per participant)	$1,009.33	$739.75	$326.80	$1,009.33	$739.75	$326.80

Note: *Figures of net benefits to participants and net revenue to government are sensitive to the assumptions on the project's costs and benefits.

5. Conclusion

Social finance represents a new funding model of social programs that aligns social priorities with private incentive in the capital market. Quantitative finance has been developing quickly in the last 40 years for the private capital market, and various quantitative indicators have become the daily language of the financial market. The ability of social finance proponents to speak to the financial market by presenting the needed quantitative information is key to the success of applying social finance in public programs. Although initiated to inform the repayment formulae for the aforementioned SIB projects, this technical analysis aims to help spur the further development of quantitative financial analyses for the development of social impact bonds and other forms of social finance.

There are four major contributions of this analysis. First, it establishes the theoretical linkage between distribution of individual impacts and the distribution of potential program outcomes; in so doing, it solves the issue of quantifying risk of program impacts, by means of individual impact distribution. Second, it pioneers a way of identifying the distribution parameters of program impacts at the individual level through stochastic process simulation. Third, the analysis demonstrates that it can identify the bounds of program implementation risk with the distribution of individual impacts; at the time of writing, this type of quantitative risk analysis is the first known in the field of social impact bonds. Finally, the results of the analysis are used to propose variable SIB repayment structures for each of the two ESDC Essential Skills pilot projects.

Based on this analysis, the proposed repayment structures would appear to be reasonably financially attractive – in terms of mean rate of return and Sharpe ratio – to social impact bond investors when compared to the traditional investments in stock market. However, it is important to note that financial indicators, although straightforward and commonly-applied, have limitations. For example, the Sharpe ratio is not a valid measure of risk-return profile if the distribution of program outcomes is not a normal distribution. Indeed, the assumption of normal distribution of the potential individual impacts is a strong assumption that may not represent the actual distribution. With the probability distribution of potential program outcomes better established, more financial analyses using different financial indicators would be possible. However, the structure of the stochastic process of program treatment effect is likely valid and it should not be difficult to extend the current model to other parametric distribution assumptions.

Because the main aim of the estimation of impact distribution was to conduct quantitative risk analysis of the financial repayment structure, other potential analytical components were not in scope; the sampling errors of the estimated impact distribution were not estimated, and the robustness of the estimation model has not been investigated. As noted earlier, the application of UPSKILL data to the FOUNDATIONS model is not as close a fit as it is with the AWES project.

Limitations of the methodology notwithstanding, the results of this technical analysis are intended and anticipated to inform and help support decision-making of project partners in devising appropriate SIB reimbursement formulae for the Essential Skills pilot projects.

Appendix G: Analysis of IALS 25-point gains in Essential Skills training projects – May 2014

Table of contents

Background

In January 2014, the federal Office of Literacy and Essential Skills contracted SRDC to conduct independent evaluations of two Essential Skills (ES) training projects. For the first time in Canada, a social finance approach to funding ES training is being tested, whereby private investors pay for the training up-front, and are repaid by the government if the training is successful in achieving pre-established outcomes.

The pilot project led by the Association of Canadian Community Colleges (ACCC) proposes to enroll 400 unemployed lower-skilled Canadians to receive *Foundations*, an established ES training program delivered by Douglas College. Private investor(s) will receive their investment plus a financial return, if training is successful. Alberta Workforce Essential Skills Society (AWES) is leading a second pilot project, in which a private sector employer(s) will be reimbursed 50% of training costs if the training achieves target outcomes.

ACCC and AWES submissions to OLES proposed literacy and essential skills (LES) gains of 25 points, measured on a standard scale, as the success outcome to trigger reimbursements. Subsequent discussions with OLES have since yielded more detailed parameters for reimbursement. These include:

1. Average of 25 point skills gains as measured on the IALS scale
2. Gains can be on one skill domain of the three (reading, document, numeracy)
3. Proportions can be used for calculation (e.g., X% of participants achieving 25-point gains)
4. Skill gains are to be measured at end of training to trigger a first repayment, and again 12 months later to trigger a bonus repayment for retention of skill gain.

Respecting these parameters, it is incumbent on partners to propose reimbursement formulae that are feasible, accountable, and based on accurate measurement of success outcomes, in keeping with best practices in pay-for-success (PFS) social finance projects. To support ACCC and AWES in refining the reimbursement formulae, SRDC conducted an analysis of literacy and essential skills gains resulting from previous Essential Skills (ES) training interventions. Specifically, the purpose of the analysis was to gather data to assess the probability that participants in each pilot project will succeed in achieving 25-point gains on the IALS scale. The findings could then inform the refinement of the performance targets to be negotiated with funders.

Review of IALS point gains in other ES training projects

To begin, SRDC identified ES training projects in settings and with populations similar to the ones proposed, where pre- and post-training LES assessments had been undertaken. Foundations, Upskill, Workplace Training Program (WTP), and the ACCC National Framework project all included a baseline and post-training assessment of literacy and essential skills. Upskill and WTP participants were low-skilled workers, perhaps more closely resembling the AWES pilot project participants, while the Foundations model is targeted to unemployed persons and has been selected as the model for the ACCC pilot. The National Framework project sample included both students, and employed persons; only the group of employed individuals was retained as a comparator to the employed sample for the AWES project.

Initially, SRDC reviewed available data and did preliminary analysis on all three literacy domains. Of the core literacy skills, workplace training curricula tend to emphasize document literacy, as it is salient in the majority of work settings. It is not surprising, therefore, that document literacy is measured in all projects – unlike reading, which was not included in WTP or Upskill assessments. Although all four projects also measured numeracy, document literacy is the primary domain that is reported on, and is the focus of this analysis as well.

Data sourced from respective project reports are presented in the following table to provide a snapshot of average document literacy gains, generally measured right at the end of training, or soon after.

Table 1 Average literacy gains (document use) achieved in completed ES training projects

PROJECT	POPULATION	AVERAGE SKIILL GAINS (Document use) USE	NOTES
Upskill program group (20-24 hrs)	Workers	First post-training = 12.8 Second post-training = 16.4	Higher skill gains associated with lower initial skills Skill gains and performance gains continue to increase over time (e.g., impacts for those with assessments >12 months are 23 points)
Upskill control group	Workers	First post-training = 2.2 Second post-training = -1.8	
WTP (26 hrs)	Workers	19	Lower skill workers appear to show higher gains
ACCC National Framework (20-40 hrs)	Workers, Levels 1&2	18 (range 11-20)	Level 3 learners appear to have lower gains than Levels 1&2; some variation in lower levels
Foundations (2-12 weeks)	Unemployed and underemployed	DOCUMENT: 31+ (PROSE: 25+) (NUM: 28+)	Lower skill workers appear to show higher gains

Among the three projects offering training for employed persons, average skill gains for document use were in the range of 12-19 points, or less than the 25-point average gains proposed for the SF pilots. On the other hand, Foundations, generally a more intensive intervention than the other three, yielded higher skill gains, clearing the average 25-point gain proposed for the SF pilot projects.

What do these findings suggest? If the SF pilots are set to reproduce the intensity and quality of training of these benchmark pilots and if the target groups and characteristics of the participants are both similar to the benchmark models, then it is reasonable to expect similar results at the end of the day. To the extent the ACCC project is modeled on the Foundations project, the average 25-point target gain looks achievable. However, results could differ if the compositions of the sample for the ACCC pilot turns out to be different than the original Foundations benchmark. If, for instance, participants for the ACCC pilots include a larger proportion of individuals with characteristics associated with low skill gains, then the average score gains achieved under Foundations may not be reached. These sample composition issues will need to be addressed in order to determine whether or not the target has been met for reimbursement purpose. Not to do so could create a perverse incentive at the time of recruitment.

The other clear message from Table 1 is that an average 25-point gain appears unlikely for workplace training models. While increasing the intensity of training to achieve higher point gains could be contemplated, employers have clearly indicated in past projects that they were not willing to free up their workers for longer periods. Revisiting the parameters for reimbursement, an alternative approach would then be to exclude part of the sample in the calculation of gains and determine a threshold of participants who would need to meet the 25-point gain mark.

What should that threshold be? Going back to the benchmark projects, we can calculate the percentage of participants that achieved 25 points or more in each project and use this percentage as the target threshold. We can also look at the characteristics of participants that made them more likely to achieve the 25-point gain. As expanded below, this information will be useful in making adjustments for sample composition.

Analysis

SRDC conducted an analysis of the microdata sets for Foundations, Upskill and WTP, as benchmark projects. The analysis first looked at the percentage of trainees in each benchmark project who achieved the 25-point gain proposed target. Results are presented on the last lines of Tables 2 to 4 below and are described under each table.

Then, SRDC conducted a series of logistic regressions to determine the influence of various participant characteristics on the probability of achieving a 25-point skill gain or more. The first step in conducting these logistic regressions was to perform a sensitivity analysis to identify dataset variables that were good predictors for skill gains. For all three benchmark datasets, the identified predictors included baseline literacy level, gender, education, age and immigrant status/home language. Additionally, Upskill and WTP included Aboriginal status, and low-income status.[1]

Findings from the regression analysis clearly demonstrated the importance of baseline literacy level as a predictor of skills gains. Regressions results in Appendix A contain the odds ratios of the score gains by project, indicating that those with starting skills at low level 1 were anywhere from 2 to 4 times more likely to achieve 25-point gains than those with upper level 1. They were also 3 to 7 times more likely to achieve these gains that their counterparts with low level 2. A similar pattern is observed among those within level 2. For instance, those with starting skills at low level 2 were 1.5 to 5 times more likely to achieve 25-point gains that those at upper level 2. Results presented in the next section are reported using a breakdown of baseline literacy levels to further illustrate the strength of the relationship between this participant characteristic and resulting literacy gains.

[1] Regression coefficients, odds ratios, standard errors, and tests of significance were calculated for each benchmark model and are found in Appendix A. For each model, variables that were shown to have statistical significance on the probability of achieving the 25-point skill gain were considered predictors.

Results

Document score gains for the total samples in each benchmark project, broken down by baseline literacy level, are shown below; similar tables for each of the potential predictor variables are found in Appendix B. Note that for each project, the total sample (N) is the participants in the microdata sets for whom we have both a pre- and post-test literacy score recorded.

Table 2 Foundations – Document use, total sample[2]

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0	--	10	-15	-12.4
Upper 2	12	30.0	30.5	40	10	4.4
Lower 2	38	41.8	37	91	18	17.7
Upper 1	115	57.2	42	201	27	29.4
Lower 1	49	73.1	48	67	36	38.8
Total	214	52.3	41	409	25	24.8

Highlights:

- Over half (52%) of the 409 Foundations participants were successful in achieving at least a 25-point gain
- Skill gain clearly and uniformly increases inverse to baseline literacy levels; note that this is the case for all projects reported here
- The average gain for all participants – whether calculated as median or mean – was 25 points
- The regression analysis revealed that in addition to baseline literacy, variables that are significant for predicting a 25-point gain include gender, immigrant status, and age (see Appendix A).

[2] A small discrepancy between the findings of this analysis and other program documents could be attributable to the use of a subset of the cumulative microdata set

Table 3 Upskill – Document use, program group only

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	4	8.5	42.5	47	-17	-14.3
Upper 2	19	25.0	41	76	-10	-16.5
Lower 2	30	37.5	25	80	25	14.3
Upper 1	76	52.1	44	146	25	29.3
Lower 1	26	68.4	34	38	28	38.8
Total	155	40.1	38	387	18	12.8

Highlights:

- 40% of the Upskill program group participants achieved document skill gains of 25 points or more on the first follow-up assessment[3]
- Median gain of total sample is 18 points
- As with Foundations, baseline literacy level and points gains are inversely related
- Predictive variables in addition to baseline literacy level include immigrant status, age, Aboriginal status, and low-income status. (Note: unlike Foundations, gender was not a predictive variable; note that Aboriginal sample size is small; physical disability appeared significant in sensitivity testing but the sample size was too small to include it as a predictive variable.)

[3] Program group members who completed only the second follow-up assessment are excluded here; program group members completing an assessment >12 months after training showed higher skill gains.

Table 4 Workplace Training Program – Document Use, total sample

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	2	6.1	26	33	-2	-14.8
Upper 2	17	27.9	41	61	5	5.4
Lower 2	43	34.7	54	124	11	15.7
Upper 1	125	39.9	48	313	18	31.3
Lower 1	31	51.7	51	60	26	31.7
Total	218	36.9	49	591	17	18.1

Highlights:

- Similar to Upskill (at 40%), 37% of WTP participants achieved 25-point gains or higher in document use
- Average gains (median/mean) were 17-18 points
- Similar to both Foundations and Upskill, a clear (inverse) relationship is demonstrated between baseline skill levels and skill gains
- Predictive variables, in addition to baseline skill levels, are gender, Aboriginal status, and low-income status.

Recommendations

In light of the analysis of the results achieved in previous projects, SRDC is proposing a three-step process to determine the reimbursement to investors/employers.

- First, establish fixed targets for each project prior to implementation.
- Second, at the time of calculating skill gains, weight individual score gains to reflect the composition of the enrolled samples.[4]
- A third step would be to validate the use of the predictive variables in the weighting calculations, by measuring again the impact of various participant characteristics on the skill gains achieved by participants in the two SF pilots.

We discuss each of the steps in more detail below.

Establishing fixed targets

It is recommended that the reimbursement formulae be based on two performance targets:

1. A pre-established minimum proportion of participants achieving a gain of 25 points or more in document literacy, as per OLES' parameter #3 above; and
2. A pre-established minimum median point gain over the total sample.

Having these dual targets meets the funder requirements while ensuring that the performance of all participants is taken into account in triggering reimbursement payments. This is desirable not only from a financial accountability perspective, but also for transparency as it reduces the real or perceived risk of creaming in delivery to focus on participants with higher probability of achieving 25-point gains.

It is recommended that the same performance targets be applied to both reimbursements – the one at the end of training, and the retention reimbursement 12 months later.

For the **ACCC pilot project**, which is based on the Foundations benchmark model, reaching a minimum proportion of 50% of participants achieving gains of 25 points or more looks achievable. As a second target, a median 25-point gain for the total sample seems feasible as well, although not by a wide margin.

For the **AWES pilot project** (Skilling Up), evidence from WTP and Upskill shows that setting a minimum 40% of the sample achieving 25 points or more in order to trigger reimbursement would be achievable. For the second performance target, based again on the analysis of the benchmark models, an intervention of 20-30 hours is unlikely to produce average point gains of 25 points across the whole sample in a population of lower-skilled workers. Instead, the benchmark analyses

[4] Precedent for applying regression adjustment in the calculation of performance measures is found in many PFS projects including some in the employment training field such as Job Corps.

points to a minimum median gain of 18 points in document use to be more realistic for this intensity of intervention.

Adjust weighting of individual results based on composition of enrolled sample[5]

As noted above, the achievement of performance targets is partially determined by the composition of the sample. For example, in the benchmark model for Foundations, gender was a predictive variable for learning gains. As shown in the two tables below, median gains for men were 18 points, compared to 27 points for women; and 56% of women attained gains of 25 points or more, compared to 41% of men. Noting that the ratio of men to women in the Foundations benchmark was about 1:3 (111:298), the enrolled sample for the ACCC pilot would need to have the same male:female ratio in order to expect the same overall average results. As this is unlikely to happen by coincidence, to compensate for different gender ratios, scores for individuals would be re-weighted accordingly. For example, if the ratio of males to females turned out to be 1:1 in the ACCC pilot project, female scores would be assigned a weight of 3/2, and males 1/2 in order to come up with the same probability for achieving an overall score of 25 points or more. Re-weighting would be done in a stepwise fashion for each of the predictive variables, as determined by the composition of the enrolled sample.

Table 5 FOUNDATIONS – Document Use, MEN

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0	--	4	11.5	5.0
Upper 2	4	22.2	39	18	13	3.9
Lower 2	9	33.3	40	27	16	16.6
Upper 1	25	50.0	47	50	24.5	26.7
Lower 1	8	66.7	58.5	12	49.5	42.3
Total	46	41.4	48	111	18	21.4

[5] The term “enrolled sample” refers to the population of participants who complete pre- and post-training PRIME assessments, upon which overall skill gains are to be calculated.

Table 6 FOUNDATIONS – Document Use, WOMEN

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0	--	6	-22.5	-24.0
Upper 2	8	36.4	29	22	6.5	4.7
Lower 2	29	45.3	37	64	20.5	18.2
Upper 1	90	59.6	41	151	28	30.2
Lower 1	41	74.6	44	55	35	38.0
Total	168	56.4	39	298	27	26.1

The same procedure would be applied to the results of the AWES project, except that the weighting procedure would apply to the predictive variables that were identified in the Upskill and WTP benchmark projects. This would be done in a step-wise fashion for each of the predictive variables.[6]

For both projects, sample re-weighting may be necessary in the calculation of the retention reimbursement, as attrition could affect the composition of the sample.

Validate pilot project training model with benchmark model

Given that the proposed training model for the ACCC project is the same as that of the benchmark model (Foundations), it is reasonable to assume it will have similar, differential, impacts on subgroups such as the example for gender above. However, the results should be validated by comparing the differential results obtained during the pilot project to those of the benchmark model. If Foundations has the same differential effect on males and females in the ACCC pilot as in the benchmark, then gender would be included in the calculation of skill gains according to the weighting scheme above. However, if it does not, re-weighting by gender would not be done, because the benchmark regression results are not replicated in the pilot project and could instead be attributed to sample size/composition/attrition or a host of other variables related to training content or delivery.

The same validation procedure would apply for the AWES pilot project.

[6] Because of small sample sizes for Aboriginal status in the benchmark models, we suggest the use of a threshold whereby weighting is only applied if the AWES project sample reaches 16-18% Aboriginals, which is about twice the prevalence as in Upskill and WTP.

Regression coefficients and significance

FOUNDATIONS: Determinants of document use gain of >= 25 points

Parameter	Odds Ratios	Coeff.	Standard Errors	P-value
Intercept		0.149	0.555	0.788
Low Level 1	8.135	2.096	0.478	<.0001***
High Level 1	4.093	1.409	0.404	0.001***
Low Level 2	2.032	0.709	0.427	0.097*
Immigrant	0.401	-0.914	0.272	0.001***
Male	0.585	-0.535	0.253	0.034**
HS or less	0.770	-0.261	0.265	0.324
Education – missing	0.737	-0.306	0.475	0.520
25 to 34	0.673	-0.396	0.427	0.354
35 to 44	0.670	-0.400	0.418	0.338
45 to 54	0.465	-0.765	0.425	0.072*
55+	0.388	-0.948	0.589	0.107

Upskill: Determinants of document use gain of >= 25 points

Parameter	Odds Ratios	Coeff.	Standard Errors	P-value
Intercept		-2.5801	0.7184	0.000***
Low Level 1	186.003	5.2258	0.7652	<.0001***
High Level 1	37.088	3.6133	0.604	<.0001***
Low Level 2	25.767	3.2491	0.6286	<.0001***
High Level 2	5.635	1.7291	0.6163	0.005***
Immigrant	0.337	-1.0876	0.3028	0.000***
Immigrant – Missing	1.606	0.4738	0.9517	0.619
Male	1.5	0.4054	0.2958	0.171
Sex – missing	1.802	0.589	1.0178	0.563
Aboriginal	0.43	-0.8444	0.5643	0.135
Aboriginal – Missing	0.291	-1.2359	0.4823	0.010**
HS or less	1.347	0.2975	0.2751	0.279
Education – missing	2.437	0.8907	0.9464	0.347
Under 25	1.461	0.3794	0.5001	0.448
25 to 34	2.912	1.0687	0.4386	0.015**
35 to 44	1.206	0.1873	0.4152	0.652
45 to 54	1.279	0.2459	0.4215	0.560
Age – missing	0.852	-0.1601	0.9416	0.865
Income < 40k	0.525	-0.6443	0.3044	0.034**
Income missing	0.476	-0.7417	0.3873	0.056*
Physical health below norm	0.547	-0.6033	0.3715	0.104
Physical health – missing	0.626	-0.4677	0.2945	0.112
Married	1.719	0.5419	0.2984	0.069*
Married – missing	2.513	0.9215	0.7729	0.233

WTP: Determinants of document use gain of >= 25 points

Parameter	Odds Ratios	Coeff.	Standard Errors	P-value
Intercept		-2.315	0.766	0.0025***
Low Level 1	20.955	3.042	0.787	0.0001 ***
High Level 1	11.978	2.483	0.746	0.0009 ***
Low Level 2	8.757	2.170	0.761	0.0044 ***
High Level 2	5.948	1.783	0.792	0.0244**
Aboriginal	0.422	-0.862	0.407	0.0342 **
Aboriginal – Missing	0.896	-0.110	0.428	0.7979
Home Language Not English	0.770	-0.262	0.250	0.2953
Language – Missing	0.710	-0.343	0.695	0.0532 *
Male	0.623	-0.474	0.243	0.0511*
Sex – missing	3.070	1.122	0.966	0.2458
No HS diploma	1.151	0.141	0.237	0.552
Education – missing	1.242	0.217	0.471	0.6456
25 to 34	1.526	0.423	0.280	0.1304
35 to 44	1.370	0.315	0.322	0.3283
45 to 54	0.944	-0.058	0.300	0.8468
55+	0.925	-0.078	0.380	0.838
Income < 40k	0.591	-0.526	0.231	0.0231 **
Income missing	0.563	-0.575	0.366	0.1162

Skill gains by potential predictive variables

FOUNDATIONS – Document Use, whole sample

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0	--	10	-15	-12.4
Upper 2	12	30.00	30.5	40	10	4.4
Lower 2	38	41.76	37	91	18	17.7
Upper 1	115	57.21	42	201	27	29.4
Lower 1	49	73.13	48	67	36	38.8
Total	214	52.32	41	409	25	24.8

Cutoff score needed for a mean gain of 25 points or greater = 290 and below (bottom 99.0% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 297 and below (100% of the sample)

FOUNDATIONS – Document Use, men

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	4	11.5	5.0
Upper 2	4	22.22	39	18	13	3.9
Lower 2	9	33.33	40	27	16	16.6
Upper 1	25	50.00	47	50	24.5	26.7
Lower 1	8	66.67	58.5	12	49.5	42.3
Total	46	41.44	48	111	18	21.4

Cutoff score needed for a mean gain of 25 points or greater = 264 and below (bottom 88.3% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 240 and below (bottom 72.1% of the sample)

FOUNDATIONS – Document Use, women

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0	--	6	-22.5	-24.0
Upper 2	8	36.36	29	22	6.5	4.7
Lower 2	29	45.31	37	64	20.5	18.2
Upper 1	90	59.60	41	151	28	30.2
Lower 1	41	74.55	44	55	35	38.0
Total	168	56.38	39	298	27	26.1

Cutoff score needed for a mean gain of 25 points or greater = 297 and below (100% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 297 and below (100% of the sample)

FOUNDATIONS – Document Use, Immigrants

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	3	-19	-9.0
Upper 2	1	9.09	26	11	0	-8.9
Lower 2	18	36.00	36.5	50	15.5	15.1
Upper 1	57	47.50	38	120	23	23.2
Lower 1	24	66.67	44	36	31	35.0
Total	100	45.45	38	220	21	21.2

Cutoff score needed for a mean gain of 25 points or greater = 231 and below (bottom 77.3% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 235 and below (bottom 79.6% of the sample)

FOUNDATIONS – Document Use, non-Immigrants

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	7	-11	-13.9
Upper 2	11	37.93	32	29	20	9.4
Lower 2	20	48.78	37	41	22	20.9
Upper 1	58	71.60	47	81	36	38.8
Lower 1	25	80.65	50	31	42	43.2
Total	114	60.32	42	189	30	29.0

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>295 and below (100% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>295 and below (100% of the sample)</u>

FOUNDATIONS – Document Use, 45+ years old

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	--	--	0	--	--
Upper 2	1	10.00	37	10	4.5	-1.0
Lower 2	7	28.00	37	25	12	13.5
Upper 1	24	46.15	38.5	52	22.5	22.0
Lower 1	14	73.68	43	19	35	38.5
Total	46	43.40	38.5	106	19.5	20.8

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>235 and below (bottom 78.3% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>235 and below (bottom 78.3% of the sample)</u>

FOUNDATIONS – Document Use, under 45 years old

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	7	-19	-13.7
Upper 2	6	26.09	30.5	23	4	0.5
Lower 2	20	39.22	37.5	51	16	15.7
Upper 1	71	59.17	42	120	28	30.1
Lower 1	24	70.59	48.5	34	41	37.9
Total	121	51.49	42	235	25	23.9

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>275 and below (bottom 97.0% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>297 and below (100% of the sample</u>

Upskill – Document Use, whole program group sample

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	4	8.51	42.5	47	-17	-14.3
Upper 2	19	25.00	41	76	-10	-16.5
Lower 2	30	37.50	25	80	25	14.3
Upper 1	76	52.05	44	146	25	29.3
Lower 1	26	68.42	34	38	28	38.8
Total	155	40.05	38	387	18	12.8

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>249 and below (bottom 64.3% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>274 and below (bottom 87.9% of the sample, i.e. everyone below level 3)</u>**

Upskill – Document Use, Immigrant

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	20	-19.5	-20.2
Upper 2	3	8.82	41	34	-29.5	-36.3
Lower 2	15	28.30	25	53	11	9.8
Upper 1	33	40.74	31.5	81	25	21.4
Lower 1	14	58.33	28	24	25	36.2
Total	66	31.13	25	212	11	7.0

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>222 and below (bottom 46.7% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>249 and below (bottom 74.5%% of the sample)</u>

Upskill – Document Use, non-Immigrant

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	4	14.81	42.5	27	-13	-10.0
Upper 2	16	39.02	40.5	41	9	1.7
Lower 2	16	64.00	41	25	29	25.8
Upper 1	42	68.85	57	61	44	40.5
Lower 1	11	84.62	35	13	34	44.8
Total	89	53.29	50	167	27	21.0

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>274 and below (bottom 83.8% of the sample, i.e. everyone below level 3, everyone who doesn't score at the ceiling)</u>**

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>298 and below (100% of the sample)</u>

Upskill – Document Use, household income < $40K

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	2	9.52	42.5	21	-20	-12.6
Upper 2	9	21.43	39	42	-25	-24.9
Lower 2	12	30.76	25	39	9	8.1
Upper 1	39	48.15	42	81	25	27.2
Lower 1	13	65.00	27	20	25	36.3
Total	75	36.95	36	203	14	9.5

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>232 and below (bottom 61.1% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>257 and below (bottom 80.8% of the sample)</u>

Upskill – Document Use, household income >= $40K

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	1	5.56	29	18	-16.5	-18.9
Upper 2	9	50.00	55	18	26.5	10.6
Lower 2	15	51.72	27	29	25	25.9
Upper 1	23	63.89	50	36	30.5	34.6
Lower 1	9	100	43	9	43	44
Total	55	50.00	44	110	25	20.4

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>274 and below (bottom 83.6% of the sample, i.e. everyone below level 3, everyone who doesn't score at the ceiling)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>298 and below (100% of the sample)</u>

Upskill – Document Use, Aboriginal

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
Total	9	45.00	32	20	22	8.8

Upskill – Document Use, non-Aboriginal

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
Total	137	41.77	39	328	21	13.4

WTP – Document Use, whole sample

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	2	6.06	26	33	-2	-14.8
Upper 2	17	27.87	41	61	5	5.4
Lower 2	43	34.68	54	124	11	15.7
Upper 1	125	39.94	48	313	18	31.3
Lower 1	31	51.67	51	60	26	31.7
Total	218	36.89	49	591	17	18.1

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>212 and below (bottom 44.2% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>194 and below (bottom 25.2% of the sample)</u>

WTP – Document Use, men

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	4	10.5	5.3
Upper 2	1	11.11	50	9	2	-4.1
Lower 2	6	21.43	55	28	10.5	11.8
Upper 1	22	34.92	43.5	63	18	20.0
Lower 1	6	66.67	49	9	28	32.0
Total	35	30.97	50	113	11	16.5

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>211 and below (bottom 39.8% of the sample)</u>

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>204 and below (bottom 36.3% of the sample)</u>

WTP – Document Use, women

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	1	5.00	25	20	-5	-22.7
Upper 2	15	34.88	41	43	11	8.8
Lower 2	29	40.85	55	71	14	18.5
Upper 1	85	43.81	50	194	19.5	24.5
Lower 1	16	41.03	48	39	21	25.0
Total	146	39.78	49.5	367	16	19.0

Cutoff score needed for a <u>mean</u> gain of 25 points or greater = <u>212 and below (bottom 44.4% of the sample)</u>.

Cutoff score needed for a <u>median</u> gain of 25 points or greater = <u>NONE</u>

WTP – Document Use, Aboriginal

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	1	-82	-82
Upper 2	2	50.00	47	4	26.5	23.0
Lower 2	2	16.67	42.5	12	-7	0.2
Upper 1	5	26.32	56	19	18	19.5
Lower 1	1	33.33	27	3	9	7.3
Total	10	25.64	47	39	10	10.4

Cutoff score needed for a mean gain of 25 points or greater = 199 and below (bottom 33.3% of the sample). Cutoff score needed for a median gain of 25 points or greater = NONE

WTP – Document Use, non-Aboriginal

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	1	4.35	25	23	-2	-15.3
Upper 2	14	30.43	38.5	46	4.5	5.6
Lower 2	30	36.59	57.5	82	12	18.8
Upper 1	96	43.44	50.5	221	19	24.5
Lower 1	19	46.34	51	41	22	27.7
Total	160	38.74	52.8	413	16	19.4

Cutoff score needed for a mean gain of 25 points or greater = 212 and below (bottom 43.6% of the sample). Cutoff score needed for a median gain of 25 points or greater = NONE

WTP – Document Use, household income < $40K

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	0	0.00	--	14	-13	-19.6
Upper 2	10	30.30	38.5	33	3	5.9
Lower 2	15	27.78	55	54	9.5	12.3
Upper 1	65	37.79	49	172	17	19.8
Lower 1	18	54.55	51	33	27	31.7
Total	108	35.29	49.5	306	11	16.5

Cutoff score needed for a mean gain of 25 points or greater = 211 and below (bottom 41.5% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 192 and below (bottom 21.2% of the sample)

WTP – Document Use, household income >= $40K

Starting Level	>= 25 pt. gain frequency	>= 25 pt. gain percent	Median gain for those >=25 pts	N	Median gain overall	Mean gain overall
3 or higher	1	20.00	25	5	0	0.2
Upper 2	5	31.25	50	16	4.5	6
Lower 2	14	43.75	52	32	22	20.8
Upper 1	34	58.62	49.5	58	32.5	35.2
Lower 1	1	11.11	45	9	5	8.6
Total	55	45.83	50	120	20.5	24.0

Cutoff score needed for a mean gain of 25 points or greater = 279 and below (bottom 96.7% of the sample)

Cutoff score needed for a median gain of 25 points or greater = 255 and below (bottom 87.5% of the sample)

Appendix H: Memorandum Re: Analysis of 25-point skill gains in Essential Skills training projects – May 2014

To: Cindy Messaros, AWES; Terry Anne Boyles, ACCC; and partners for Social Finance pilot projects
C.c. Cathy Wynn, Office of Literacy and Essential Skills

From: Jean-Pierre Voyer and Sheila Currie, SRDC

Date: May 6, 2014

Re: Analysis of 25-point skill gains in Essential Skills training projects

---

The attached technical paper presents findings from an analysis of the literacy point gains achieved in previous Essential Skills training projects. The analysis was conducted to support the refinement of the repayment formulae in the Social Finance pilot projects, for which a minimum 25-point gain on the IALS scale is stipulated as a target for triggering reimbursement. The paper is intended for internal circulation only, for the use of partners and the Office of Literacy and Essential Skills as funder.

We begin with identification of projects with pre- and post-training LES assessments, in settings and with populations similar to the ones proposed for the pilot projects. Aggregate data from program documents or publications were reviewed for Foundations, Upskill, Workplace Training Program (WTP), and the ACCC National Framework project. We then conducted a series of regression analyses on program microdata files for Upskill, WTP and Foundations, in order to detect observable characteristics associated with the probability of achieving 25-point gains. The results suggest it may be difficult to achieve average 25-point gains in the full sample, particularly in a workplace context.

We recommend that the formulae be based on two performance targets: an established proportion of participants achieving a gain of 25 points or more in document literacy, plus an established minimum average point gain over the total sample. Having these dual targets meets the funder requirements, and ensures that the performance of all participants is counted in triggering reimbursement payments, not just the proportion who achieve 25 points. This is desirable not only from a financial accountability perspective, but also for transparency as it reduces the real or perceived risk of creaming in delivery to focus on participants with higher probability of achieving 25-point gains.

The findings suggest specific performance targets for consideration:

ACCC: Minimum 50% of participants achieving a 25-point gain or more in document literacy; PLUS minimum median 25-point gain for the total sample.

AWES: Minimum 40% of participants achieving a 25-point gain or more in document literacy; PLUS minimum median 18-point gain for the sample.

The targets are based on evidence from the benchmark projects, and include a mechanism for validating assumptions about predictive variables based on the actual sample enrolled in the pilot projects.

We hope this analysis will be helpful to advancing the pilot project designs, and look forward to discussing these and other project matters with you in the coming weeks.