Impact Evaluation (IE) Concept Note

Moving Youth Away from The Market for Crime: Interventions in The Honduras Safer Municipalities Project

i2i Dime

World Bank

The publication of this study has been made possible through a grant from the Jobs Umbrella Trust Fund, which is supported by the Department for International Development/UK AID, and the Governments of Norway, Germany, Austria, the Austrian Development Agency, and the Swedish International Development Cooperation Agency.

© 2017 International Bank for Reconstruction and Development / The World Bank.

1818 H Street NW, Washington, DC 20433, USA.
Telephone: 202-473-1000; Internet: www.worldbank.org.

Some rights reserved
This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Nothing herein shall constitute or be considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

Rights and Permissions

This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) http://creativecommons.org/licenses/by/3.0/igo. Under the Creative Commons Attribution license, you are free to copy, distribute, transmit, and adapt this work, including for commercial purposes, under the following conditions:

Attribution—Please cite the work as follows: i2i Dime, "Impact Evaluation (IE) Concept Note: Moving Youth Away from The Market for Crime: Interventions in The Honduras Safer Municipalities Project". World Bank, Washington, DC. License: Creative Commons Attribution CC BY 3.0 IGO.

Translations—If you create a translation of this work, please add the following disclaimer along with the attribution: This translation was not created by The World Bank and should not be considered an official World Bank translation. The World Bank shall not be liable for any content or error in this translation.

Adaptations—If you create an adaptation of this work, please add the following disclaimer along with the attribution: This is an adaptation of an original work by The World Bank. Views and opinions expressed in the adaptation are the sole responsibility of the author or authors of the adaptation and are not endorsed by The World Bank.

Third-party content—The World Bank does not necessarily own each component of the content contained within the work. The World Bank therefore does not warrant that the use of any third-party-owned individual component or part contained in the work will not infringe on the rights of those third parties. The risk of claims resulting from such infringement rests solely with you. If you wish to re-use a component of the work, it is your responsibility to determine whether permission is needed for that re-use and to obtain permission from the copyright owner. Examples of components can include, but are not limited to, tables, figures, or images.

All queries on rights and licenses should be addressed to World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.
Images: © World Bank China. Further permission required for reuse.

Impact Evaluation (IE) Concept Note

Moving youth away from the market for crime: interventions in the Honduras Safer Municipalities Project

Honduras

P152314

June 30, 2015

Keywords: [1] Laboratory, Individual Behavior (C91); Field Experiments (C93); Behavioral Microeconomics: Underlying Principles (D03); Education and Economic Development (I25); Welfare, Well-Being, and Poverty (I3); Measurement and Analysis of Poverty (I32); Illegal Behavior and the Enforcement of Law (K42); Human Resources, Human Development, Income Distribution, Migration (O15) **(R)**

[1] Please refer to JEL classification codes http://papers.ssrn.com/sol3/displayjel.cfm.

TABLE OF CONTENTS

IE PROFILE INDICATORS

No.	Indicator	Description
1	IE code	P152314
2	IE Title	Safer Municipalities Project IE
3	IE TTL	Marcus Holmlund
4	IE Contact Person	Marcus Holmlund, DECIE
5	Region	LCR
6	Sector Board/Global Practice	GSURR
7	WBG PID (if IE is evaluating a WBG operation)	P130819
8	WBG Project Name (if IE is evaluating a WBG operation)	HN Safer Municipalities
9	Project TTL (if IE is evaluating a WBG operation)	Marcelo Fabre
10	Intervention	Labor market readiness/insertion program targeting two categories of at-risk youth
11	Main Outcomes	Criminality and antisocial behavior, soft skills, mental health, education, labor market outcomes
12	IE Unit of Intervention/Randomization	Individual
13	Number of IE Units of Intervention	450 individuals in each of two target populations
14	IE Unit of Analysis	individual
15	Number of IE Units of Analysis	900 individuals in each of two target populations
16	Number of Treatment Arms	2 treatment arms for each of two target populations
17	IE Question 1 (Treatment Arm 1)	What is the impact of an intervention combining vocational training, cognitive behavioral therapy, and a temporary job opportunity on the outcomes of interest among at-risk youth?
18	Method IE Question 1	Random assignment at individual level
19	Mechanism tested in IE Question 1	Package
20	IE Question 2 (Treatment Arm 2)	What is the impact of an add-on individual mentoring scheme on the outcomes of interest?
21	Method IE Question 2	Random assignment at individual level
22	Mechanism tested in IE Question 2	Mentoring
23	IE Question 3 (Treatment Arm 3)	n/a
24	Method IE Question 3	n/a
25	Mechanism tested in IE Question 3	n/a
25	Gender-specific treatment (Yes, No)	No
27	Gender analysis (Yes, No)	No
28	IE Team & Affiliations	In alphabetical order by last name: - Marco Castillo, Associate Professor, George Mason University - Laura Chioda, Senior Economist, LCRCE - Benjamin Feigenberg, Assistant Professor, University of Illinois - Marcus Holmlund, Economist, DECIE Ragan Petrie, Associate Professor, George Mason University
29	Estimated Budget (including research time)	USD 1,291,000
30	CN Review Date	June 2016
31	Estimated Timeframe for IE	June 2015 – December 2017
32	Main Local Counterpart Institution(s)	Fondo Hondureño de Inversión Social; Instituto de Desarrollo Comunitario, Agua, y Saneamiento (Honduran Social Investment Fund; Institute for Community Development, Water, and Sanitation

1. EXECUTIVE SUMMARY

(1 page)

Describe the proposed IE in non-technical language in one paragraph or less. This could be an abstract of your IE. Include broad motivation/background and policy/research contribution. (E,R) Present IE questions and main outcome(s) the intervention aims to affect. Briefly explain how you are proposing to test your main evaluation question(s).

The study described in this concept note is an experimental impact evaluation of a training and jobs program targeting at-risk youth in high-violence municipalities in Honduras. The program combines vocational training, soft skills training through cognitive behavioral therapy, and temporary job placements, and targets two groups of at-risk youth: (i) youth currently in secondary school grades 7-9 and (ii) youth that are no longer in school and are weakly attached to the labor market. Participants in the first group will be offered a school-break internship program (November-January), while participants in the second group will be offered a more intensive training and labor insertion program lasting up to one year. An add-on mentoring scheme will be offered to a random sub-set of participants in both programs.

With a homicide rate of 90.4 per 100,000 population, Honduras is the most violent country in the world. Young men – of whom less than one in two graduate secondary school and who, regardless, face very limited labor market opportunities – are disproportionately affected: 63% of murder victims are men aged 15-34, and homicide rates for youth in their early 20s are above 300 per 100,000. Socioeconomically disadvantaged young men are particularly susceptible to gangs involved in drug trafficking and other illegal activities.

The impact evaluation study described in this concept examines the protective role of a labor market intervention in moving youth away from the market for crime towards legal, individually and socially productive opportunities. Evidence on supply-side programs aiming to enhance the technical/vocational skills of individual job candidates through education and on-the-job training in high-violence contexts is disappointing, suggesting little benefit beyond limited/rare short-term positive impacts. However, literature examining interventions targeting soft skills and personality traits suggest these have the potential to augment and sustain short-term impacts observed in more traditional job-training interventions, but there is virtually no evidence of combined soft-and-hard skills approaches from low- and middle-income countries. The work described here is designed to help fill this gap.

Using a randomized controlled study design, we will study whether the training and jobs program – vocational training, cognitive behavioral therapy, and a temporary job – prevents delinquency, promotes positive behaviors, improves mental well-being, and reduces drop-out (for youth still in school) or improves post-intervention labor market outcomes (for youth no longer in school). Furthermore, individual mentors will be randomly assigned to a subset of program participants to study whether this additional component strengthens the effects observed through the main intervention and/or improves their sustainability.

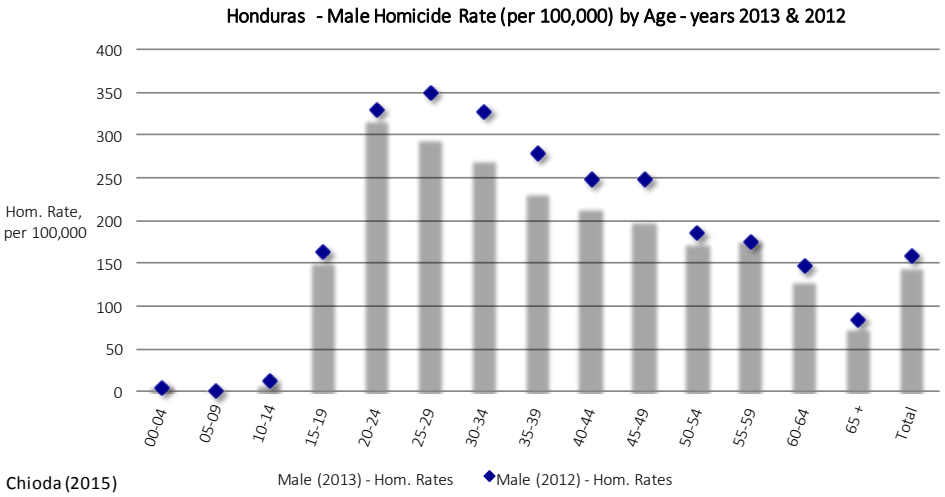
This study has been designed and will be implemented in close collaboration with the Safer Municipalities Project team and data, evidence, and knowledge generated throughout will be used to inform design modifications and eventual scale-up of the interventions described here to other high-crime municipalities in Honduras (contingent on positive results). Additionally, we expect that the experience of the Safer Municipalities Project will be relevant to other countries in the region, in particular neighboring El Salvador and Guatemala which also suffer from very high rates of violence and crime.

2. Background and Key Institutional Features

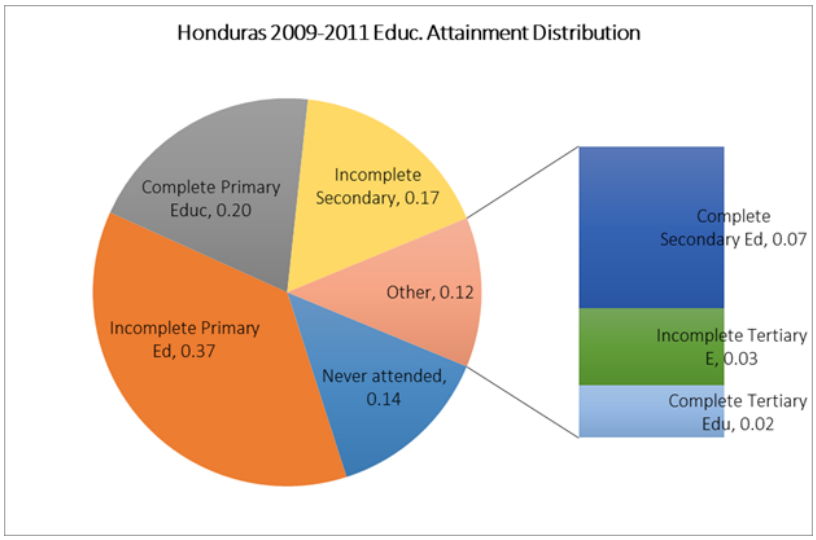
(1 page)

Present an overview of the local context. Identify and define the problem: what is the policy/research problem this IE is proposing to study? Which groups are affected by the problem? Describe the intervention whether existing or new, implementing organization, institutional setting and any important consideration. Describe the intervention geographic/demographic scale and scope: Does it represent the "mode" of delivery in the country? (R, E)

With a homicide rate of 90.4 per 100,000 population, Honduras is the most violent country in the world (UNODC Homicide Statistics, 2011). The number of murders per year far exceeds the WHO threshold for endemic violence (10/100,000) and even for conflict (30/100,000). Violence in Honduras is concentrated in urban areas (65% of homicides occur in 5% of municipalities), and 94% of homicide victims are male, with men aged from 15 to 34 years accounting for 63% of the total (World Bank 2012; see the figure below). The homicide rate has more than doubled in the past decade: in 2005, it was 37 per 1,000,000.



Drug trafficking is the main factor associated with increasing levels of violence, and its negative impacts are enabled and exacerbated by socio-economic and governance factors. Socioeconomic factors include poverty, social inequalities, and lack of opportunities. In Honduras, half of secondary school-age youth do not attend school (see graph below), and only 35% graduate in the expected year. For those that do graduate, job prospects are limited, which drives youth away from the legal labor market towards gangs and crime. Migration to the United States in search of better opportunities leads to family disintegration, placing already disadvantaged youth further at risk.



At the macroeconomic level, the annual economic costs of violent crime are estimated at 10% of the country's GDP, nearly $900 million, and the value of disability-adjusted life years lost to violence is estimated at 1.3% of GDP. With costs to business are estimated at 3.9% of firm sales, crime and violence are cited by businesses as one of the main constraints to productivity and growth (World Bank 2012).

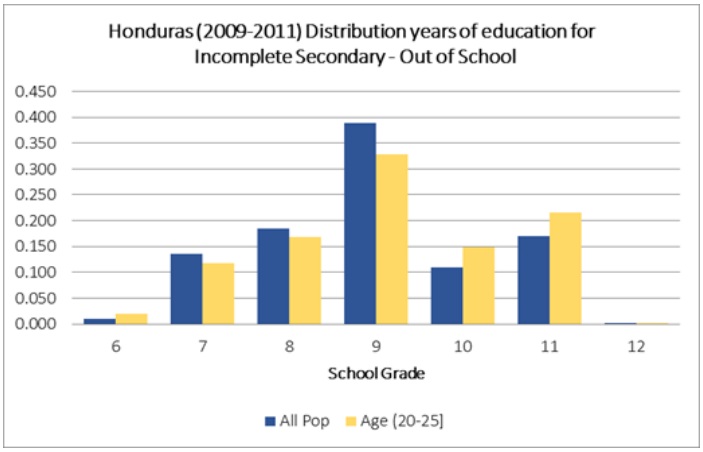
Against this backdrop, the Project Development Objectives of the Safer Municipalities Project (SMP) – currently the only active World Bank lending operation to directly target violence and crime – are to (i) improve the capacities of national and local authorities in violence prevention; (ii) address risk factors of crime and violence in selected municipalities, and (iii) improve the country's capacity to respond promptly and effectively to an eligible emergency. SMP is implemented by the Honduras Social Investment Fund (Fondo Hondureño de Inversion Social; FHIS) within the Institute for Community Development, Water, and Sanitation (Instituto de Desarrollo Communitario, Agua, y Saneamiento; IDECOAS[2]).

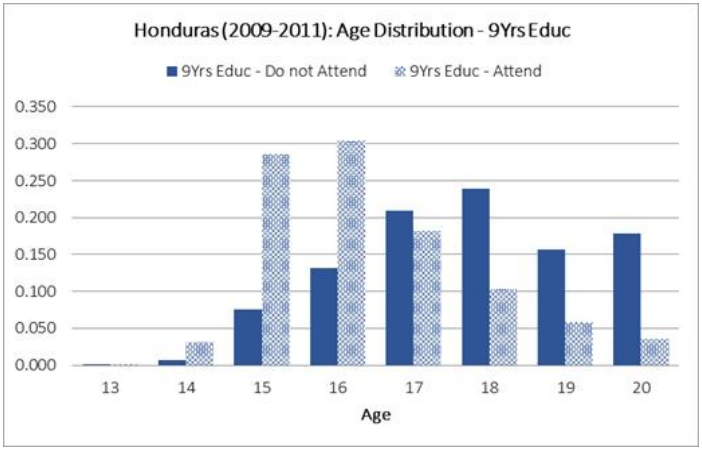
SMP focuses on areas with high concentrations of crime and violence.[3] Specifically, the project is implemented in selected urban "clusters" in three medium-sized municipalities (population less than 250,000) which all have homicide rates above the national average of 90.4 per 100,000 population. Each cluster comprises three geographically proximate (often contiguous) neighborhoods, and there is one cluster per municipality with an average population of about 11,500 persons per cluster. The selected municipalities vary in their economic profiles: La Ceiba is known for tourism, Choloma is an industrial center with several textile assembly plants, and El Progreso's economy is built on commerce. Together, they account for about 11% of 2014 homicides for the entire country (Honduras Violence Observatory[4]).

Within the context of the SMP, the IE described in this concept note examines the protective role of a labor market intervention in moving youth away from the market for crime towards legal, individually and socially productive opportunities in a setting where the target population suffers from one or more of the following: limited human capital (hard and soft skills); limited legal labor market opportunities; potentially attractive criminal market opportunities; a culture of machismo and violence; and a biological propensity towards risky and delinquent behavior.

[2] IDECOAS has the same status as a ministry.
[3] Municipal selection criteria are detailed in World Bank 2012.
[4] iudpas.org/observatorio

To build evidence to address this challenge, the proposed IE studies the effects of a training and jobs program on citizen security and related education and labor-market outcomes. The program targets two groups of at-risk youth: (i) youth currently in secondary school grades 7-9 and (ii) youth that are no longer in school nor strongly attached to the formal labor market. The target populations were chosen taking into consideration some key stylized facts about dropouts in Honduras. The transition from lower to upper secondary – 9th grade in particular – is critical to explain dropout rates (see first graph below). Furthermore, consistent with Honduras' elevated grade retention rate, those dropping out are on average 16.8 years old, or approximately 2 years older than students on the ideal pace of grade progression by the end of their 9th grade (see second graph below). Attachment to the labor force for these low skilled youth is often precarious, short-term, informal, and characterized by frequent transitions between employment and unemployment (Székely and Karver, 2014).





Participants in the first group (youth currently in school grades 7-9) will be offered an internship program during the school break (November-January), while participants in the second group (youth no longer in school nor strongly attached to the labor marker) will be offered a more intensive training and labor market insertion program lasting up to one year. An add-on mentoring scheme will be offered to a sub-set of participants in both programs. It is expected that in the first year of the program's implementation in each of the SMP's municipalities, 150 youth from each group will participate in their respective program (450 youth total from each target population group).

3. Literature Review (E)

(1 page or less)

Describe most relevant literature/scientific background specifically linked to your problem/evaluation question(s).

The link between labor markets and anti-social behavior is complex. On the one hand, labor market participation incapacitates potential perpetrators of crime and violence and provides an alternative source of income. It also generates protective factors through building social networks, strengthening social identity, and creating bonds with one's community. On the other hand, higher income may increase victimization (due to change in routines and/or additional available income), and participation in the labor market may interfere with investments in education and human capital more broadly. Additional income could be spent on crime-inducing goods such as drugs and alcohol which increase the likelihood of criminal behavior. In single-parent households, the need to work may lead to youth being left largely unsupervised for large amounts of unstructured time, which is a risk factor for delinquency and other risky behavior.

Attachment to the labor market appears to play an important role for more "vulnerable" (i.e. at higher risk of engaging in antisocial behavior) subgroups of the population: youth and unskilled workers (Gould, Weinberg, and Mustard, 2002; Machin and Meghir, 2004; Lin, 2008). In particular, antisocial behavior responds to incentives set by the labor market (employment opportunities, wages). The evidence also suggests that crime and work may coexist: having a job per se may not be a sufficient to deter crime. Quality of employment matters (e.g., stability, formality, opportunity for wage growth and advancement). For instance, throughout the lifecycle, perpetrators in Mexico are employed at higher rates than the rest of the population; however the number of workers employed in the formal sector consistently predicts declines in the homicide rate (Chioda 2014). And in Brazil, rates of formal job creation for young men (who are more likely to be both victims and perpetrators of crime and violence) are systematically related to lower homicides, whereas labor market opportunities for adults or women have no predictive power. These findings are consistent with the notion that early and low-quality attachment to the labor market may be a risk factor for future criminal and violent behavior and underscore the challenge of moving youth out of the market for crime.

Few job programs measure the impact on antisocial behavior, such that the evidence is rare, especially in developing countries. Schochet et al. (2008b) show that the opportunity cost of illegal activities increases as productive activities are made available to program participants in the US, and Blattman et al. (2014) draw similar conclusions from a program in Uganda.[5] Intensive residential and non-residential programs in the US that target at-risk youth show some promise and appear to achieve their positive impacts through human capital accumulation and high school completion (Schochet et al. 2008a and Raphael 2011). However, gains seem to be mostly concentrated among older, lower-risk individuals.

Evidence from training programs in middle-income countries that target youth of low socioeconomic status (but do not focus on crime/violence impacts) is consistent with that emerging from the US experience. Card et al. (2011) and Ibarraran et al. (2011) present weak evidence of improved earnings for youth participating in a job training program in the Dominican Republic, conditional on working (though they do not find impacts on employment). Attanasio, Kugler, and Meghir (2009) experimentally evaluate Colombia's Jovenes en Accion program, and find positive impacts on paid employment and a large positive impact on earnings (12-15% increase), with larger effects concentrated among women (lower-risk).

There is growing evidence on the importance of interventions targeting "soft" skills and personality traits among adolescents and young adults. Adolescence and young adulthood are key stages for brain development, with a high degree of malleability and brain plasticity. The picture emerging from U.S. interventions aimed at improving the likelihood that children succeed in school and in the labor market suggests that soft skills, as well as intelligence, are key determinants of behavior and outcomes (Heckman and Kautz 2012; Tough 2012; Heckman, Pinto, and Savelyev 2013). Cook et al. (2014) and Heller et al. (2013) provided disadvantaged youth in Chicago with non-academic support aimed at teaching soft skills through cognitive behavioral therapy, and found that the intervention resulted in a 0.64 standard deviation improvement in math scores compared to the control group and a 0.48 standard deviation improvement compared to the distribution of the US as a whole. This is equivalent to a 15% increase in the rank of the test score distribution. Heller (2014) experimentally evaluates a summer jobs program (One Summer Plus) for disadvantaged high-school youth in Chicago, and finds that violence decreases by 43% over 16 months (3.95

[5] Blattman et al. (2014) report a lower probability of re-joining armed forces, but find no impact on the likelihood of engaging in other illegal activities.

fewer violent-crime arrests per 100 youth), and that this decline occurred largely after the eight-week intervention ended. This intervention combined short-term employment (summer jobs) with cognitive behavioral therapy[6] and individual mentors for participants. Heller et al. (2015) provide further evidence on the role that automatic behavior plays in explaining disparities in youth outcomes like delinquency and dropout, highlighting how disadvantaged youth are more likely to face circumstances that will translate impulsive and automatic responses into negative outcomes. Blattman, Jamison, and Sheridan (2015) find that eight weeks of group cognitive behavioral therapy for criminally-engaged men in Liberia reduced acts of crime and violence by 20-50%, with impacts lasting at least one year when accompanied by an unconditional cash grant.

Though the evidence to date is suggestive of areas for further research, at present we have very limited well-identified causal evidence on mechanisms to (i) prevent at-risk youth from entering into criminal activities and (ii) encourage youth who are already engaged in such activities to move towards socially and economically productive opportunities. Furthermore, existing evidence is primarily from high-income countries. The impact evaluation research proposed in this concept note will contribute to the existing literature on protective factors for at-risk youth by experimentally studying the causal impacts of a jobs program (that combines technical and soft skills training, including cognitive behavioral therapy) with temporary job placements, and an add-on individual mentoring scheme.

4. Policy Relevance

(1/2 page or less)

Assess the extent to which the study may influence policy and institutional capacity at the national, regional, and international level. Explain how you plan to track the policy influence of your study (see Appendix on DIME indicators of IE influence on program/policy).

Between 2000 and 2006, Honduras adopted a *mano dura*, or "tough-on-crime" approach. This, however, was accompanied by an increase in violence, and this sparked a gradual move towards a more comprehensive approach combining law enforcement and prevention resulting in the 2011-12 National Citizen Security and Coexistence Policy under which the World Bank-supported Safer Municipalities Project was created.

Still, Honduras holds the unenviable title of most violent country in the world, with a homicide rate of 90.4 per 100,000 in 2012 (UNODC Homicide Statistics), well above the WHO's "conflict" threshold of 30 homicides per 100,000. Violence in Honduras is concentrated in urban areas, with 65% of homicides occurring in 5% of municipalities. Most homicide victims are male (94%), with men aged 15 to 34 years old accounting for 63% of the total. An average of three young persons are murdered each day – more than 1,000 per year.

The World Bank's last Country Partnership Strategy (CPS) for Honduras, for FY12-14, has improving citizen security as its first objective. The impact evaluation study proposed here is designed specifically to inform this agenda in Honduras, and is well placed also to contribute policy-relevant knowledge to other countries in Latin America and the Caribbean – particularly to Honduras' neighbors in Central America. More broadly, evidence and knowledge produced through the work proposed here will be relevant to the work of the Bank's Social, Urban, and Rural Development Global Practice and to other development partners focusing on urban crime and violence.

The study described in this concept note was designed to respond to country-level policy needs from the outset. It is the result of a collaborative process involving the Safer Municipalities Project Implementation Unit at IDECOAS-

[6] Cognitive behavioral therapy is short-term psychotherapy which takes a hands-on and practical approach to problem solving. It aims to change patterns of thinking or behavior that underlie people's actions, and thus to change the way they think and act.

FHIS. This process began in March 2014 at the Impact Evaluation 4 Peace workshop, jointly implemented by DIME, the Bank's Fragility, Conflict, and Violence unit, and the Bank's Latin America and the Caribbean Citizen Security Team.[7]

To date, this collaborative process has influenced policy/program design in three ways. First, policy design for youth employment and crime prevention is being rationalized through drawing on international experience with and evidence on hard and soft skills training youth employment programs, and through structuring the proposed Safer Municipalities Project employment program for at risk youth with the specific objective of establishing its casual impact and understanding key mechanisms through which the Program does or does not work. Second, this process entails the introduction of structured learning through multiple treatment arms for each of our target population groups, as is further discussed below. Third, as highlighted in the literature review above the employment program and the accompanying impact evaluation study proposed here builds on international evidence which suggests that interventions targeting soft skills and personality traits are necessary complements to more traditional technical/vocational and labor insertion programs, particularly with a view to achieving positive outcomes which persist beyond the short run.

5. THEORY OF CHANGE (E)

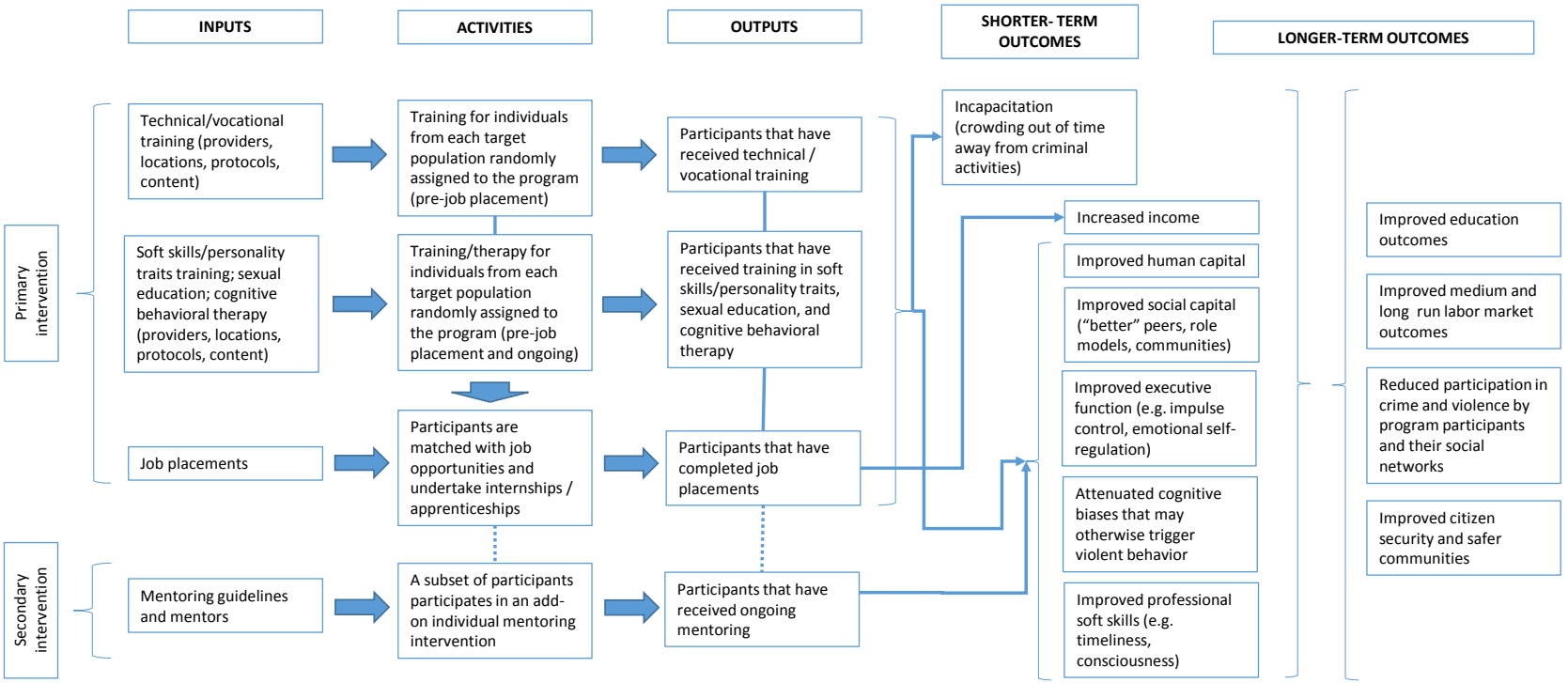
(1 Figure and 2-3 paragraphs)

This impact evaluation study is designed to identify whether increasing the opportunity cost of engaging in criminal behaviors can lead to longer-term behavior change (moving away from antisocial activities) and improved education and labor market outcomes which persist beyond the intervention. We propose to experimentally evaluate an intervention which combines technical and life skills training and cognitive behavioral therapy with temporary job placements, and targets two groups of at-risk youth: (i) youth currently in secondary school grades 7-9 and (ii) youth that are no longer in school nor strongly attached to the legal labor market. Participants from the first group will be offered a school-break ("summer") internship program from November-January, while participants from the second group will be offered a more intensive training and labor market insertion program lasting up to one year. An add-on mentoring scheme will be offered to a subset of participants in both programs.

The fundamental theory of change underlying the intervention is that crime and violence reduction and reintegration into the productive sectors of society requires changing both the subjective and objective opportunity costs of crime by increasing access to and suitability for the legal labor market and reducing anti-social behaviors that are typically associated with criminal activity. A simplified program logic – from inputs, activities, and outputs; to shorter and longer-term outcomes – is presented in the figure on the next page.

[7] See http://tinyurl.com/ie4peace for more details on the workshop.

Theory of Change



The theory of change relies on four key assumptions. First, it must be the case that training and employment opportunities provided through the intervention induce human capital gains sufficiently large that the employability of the targeted population increases as a result of the intervention. Second, it must be the case that local labor markets are sufficiently well-functioning that such human capital gains in turn drive increases in local employment rates and earnings, rather than simply leading to displacement of those with stronger ex-ante labor force attachment. Importantly, violation of this second assumption will not undermine the capacity to identify program impacts among the targeted (i.e. treatment vs. control) population but will drive negative spillovers that affect the nature of community-level impacts and the aggregate welfare gains associated with the intervention. Third, while cognitive behavior therapy and soft-skills development programs have been shown to positively impact targeted youth in the context of youth labor market programs in the United States, evidence on how such programs impact at-risk youth in a setting such as Honduras is essentially non-existent. Consequently, we must assume that previously estimated impacts in the U.S. setting have sufficient external validity to be relevant in the Honduran setting. Fourth, any additional benefit conferred by the secondary mentoring intervention relies critically on quality of mentors, and so we must assume that mentors selected for the program will have sufficient capacity and motivation to positively impact the human and social capital of their mentees.

6. HYPOTHESES/EVALUATION QUESTIONS (E,R)

(1/2 page)

List the hypotheses derived from your theory of change. List the main evaluation question(s) to be addressed by the proposed study. Methods to answer sub-questions on heterogeneous treatment effects and spillovers should be described in the methods section. Describe how the evaluation questions were derived.

As illustrated in the theory of change diagram, the proposed Safer Municipalities At-Risk Youth Employment Program is a multi-pronged intervention characterized by three sets of complementary mechanisms which are linked to the reduction of antisocial behavior:

1. Increased access to legal labor market opportunities (in the form of employment opportunities, vocational training, or summer jobs) is linked to several theoretical mechanisms that can serve as protective factors against antisocial behavior:
    - A mechanical incapacitation effect that crowds out time that could otherwise be allocated to criminal activities;
    - An income effect that renders the “need” for crime less prominent;
    - A human capital effect, whereby training and employment provide skills and experience valued in the labor market, thereby raising expected future (legal) labor market wages;
    - The accumulation and improvement of social capital, in terms of better peers, role models, and communities.

2. Life-skills training and cognitive behavioral components play complementary roles by:
    - Improving executive function (e.g. impulse control, emotional self-regulation);
    - attenuating cognitive biases that may trigger violent behavior;
    - encouraging professional skills development (timeliness, consciousness, etc.).

3. Mentoring acts as a reinforcement/reminder mechanism to strengthen both
    - social capital (pro-social peers, role models, links to communities);
    - human capital in terms development of soft skills and executive functions.

Therefore, the program is designed to affect not only the opportunity cost of crime, which should be understood as including not only wages but also the cost of foregoing non-pecuniary attributes of jobs (social capital components). It also affects the "production function" of decision-making by targeting cognitive biases, impatience and self-control.[8]

The fundamental question is whether the initiatives described herein are successful at curbing the criminal involvement and violent behavior of youth and at inducing firmer attachment to the labor force for the older group and more years of education for the younger group. Further, we ask whether mentoring enhances these effects.

Specifically, we ask:

I. REDUCED FORM IMPACTS ON CRIME AND ANTISOCIAL BEHAVIOR

1a. Does the combination of vocational training + cognitive behavioral therapy (CBT) + jobs prevent delinquency (as measured by encounters with the police, number of crimes) and change behavioral outcomes (self-reported and observed by employers/teachers/households) for both youth in school and youth out of school?

1b. Does the combination of vocational training + CBT + jobs improve mental well-being and PTSD-related outcomes (along either the intensive or extensive margins, or both)?

2. Does the additional mentoring component strengthen the results in (1a) and (1b) by either increasing the magnitude of the effects and/or by rendering them more persistent (sustainability)?

II. POTENTIAL MECHANISMS

1. **Incapacitation**. Does the treated group experience declines in delinquency and criminal involvement *during the intervention*? In particular, if the exact time and date of crimes/incidents are recorded, we could quantify the importance of incapacitation effects.
    a. Does this incapacitation effect represent a net decline in the number of incidents/crimes?
    b. Alternatively, do the interventions simply displace crimes/delinquency, such that the same number of incidents are recorded among the treated group, but take place during hours and on days outside of training/employment?
    c. Medium/long Run: Do decreases in crime persist beyond the intervention suggesting movement out of the market for crime, or are these temporary suggesting short-term displacement of crime?

2. **Human Capital Accumulation**. As compensation for summer jobs, the younger group will receive grants that cover school costs during the subsequent school year, to induce participants to continue their education. Do we observe lower dropout rates for treated children relative to the control group? Do we see improvements along other educational outcomes?
    a. Medium Run: Do declines in dropouts persist beyond the school year following the summer program? For instance, are youth who participated during the summer between 7th and 8th grades less likely to drop out in the 9th grade than their peers in the control group?

3. **Soft skills, impulsivity and impatience** – Does the treatment alter the composition of committed crimes by reducing the severity of crimes/incidents committed? For instance, do crimes committed among the

[8] The Beckerian framework regarding the supply of criminality does not require full rationality and is consistent with boundedly rational decision making

treated group de-escalate from violent crime to property crime relative to the control group? If so, this might suggest that the cognitive behavioral component is effective at curbing harmful impulses.

a. Similarly, does treatment have measurable effects on soft skills, impatience, and impulsivity?

4. **Social Capital** -- Do the interventions induce changes in the composition of participants' social networks? For instance, do treated youth switch out of "low-quality" anti-social peers in favor of more pro-social peers?

As in the case of reduced form impacts, we will consider both how the primary intervention (vocational training + CBT + jobs) and the secondary intervention (add-on mentoring scheme) affect the impacts linked to the mechanisms described above.

7. Main Outcomes of Interest (E,R)

(1 table)

Briefly list and define main outcomes of interest (primary and secondary/intermediate) as in Table 1. Further details on how the outcomes will be measured/collected will go in the data collection section.

Table 1. Main Outcomes of Interest

The Summer Jobs Initiative refers to the set of interventions targeting at-risk youth currently in grades 7-9. The Temporary Jobs Initiative refers to the set of interventions targeting at-risk youth currently not in school nor strongly attached to the legal labor market.

Outcome Type	Outcome Name	Definition	Measurement Level
Primary (Summer Jobs Initiative + Temporary Jobs Initiative)	Arrests	Individual arrest records linked to study participants (Control +Treatment)	Individual; administrative data
Primary (Summer Jobs Initiative + Temporary Jobs Initiative)	Self-reported criminality and victimization	Frequency and severity of self-reported involvement in crime perpetration or victimization	Individual; household surveys
Primary (Summer Jobs Initiative + Temporary Jobs Initiative)	Employer/Mentor/ Household/Teacher reports on conduct and behavior	Behavioral outcomes linked to soft skills, as well as self control; punctuality and attendance	Individual; Employer and mentor report cards, household surveys, and school-based surveys
Primary (Summer Jobs Initiative + Temporary Jobs Initiative)	Well-Being	Mental Health, executive functions, and PTSD measurements	Individual; household surveys
Primary (Summer Jobs Initiative)	Educational outcomes	Academic performance, academic attendance, academic promotion, and reported school-based behavioral incidents	Individual; household surveys and school-provided administrative data

Primary (Temporary Jobs Initiative)	Labor Market outcomes	Information on earnings and employment history, including perceived employment security and opportunity	Individual; households surveys
Secondary (Summer Jobs Initiative + Temporary Jobs Initiative)	Reported crimes and clearance rate	Number of crimes reported to local police jurisdictions (categorized by crime severity and clearance rate)	Cluster-level; administrative data
Secondary (Summer Jobs Initiative + Temporary Jobs Initiative)	Pro-social behaviors	Information on social connectedness, frequency of social activity, time use, etc.	Individual; households surveys
Secondary (Summer Jobs Initiative + Temporary Jobs Initiative)	Real-time crime reporting	An electronic system will be established to catalogue the geography of reported crimes in real time (rates can then be compared to self-reports and official statistics)	Cluster-level; collected using crime mapping tools
Secondary (Summer Jobs Initiative + Temporary Jobs Initiative)	Network mapping and network outcomes	Information on friendship relationships and antisocial behavior of individuals in participants' networks. This information will be constructed using self-reported network linkages within neighborhoods.	Individual surveys
Secondary (Summer Jobs Initiative + Temporary Jobs Initiative)	Social norms, including trust levels, crime reporting, economic opportunity, and perceived public safety	Surveyed individuals will be asked about perceptions of public safety, economic opportunity, etc., and individuals will be asked hypotheticals about returning of found valuables, likelihood of crime reporting, etc.	Cluster-level measures; collected via randomly sampled households, includes the use of payoff-based field experiments where applicable

8. EVALUATION DESIGN AND SAMPLING STRATEGY (E,R)

(2 pages or less)

Present the main features of the proposed evaluation design to address the evaluation question(s). Describe precisely the identification strategy (e.g., trial design including clustering, factorial, stratification details) for each evaluation question. Report all inclusion/exclusion criteria to define the target population/population studied, providers, settings, and clusters (as relevant). Report any ethical issues that may arise concerning the evaluation design and the sampling strategy (not related to data collection).

Description of the interventions

The Safer Municipalities Project Employment Program targets two groups of at-risk youth. The **first is youth currently studying in 7th, 8th, or 9th grade** who will be provided with a Summer Jobs Initiative[9] (total duration: 2-3 months). This group is targeted due to the high incidence of school dropout observed after 9th grade. Once a youth drops out of the formal schooling system, s/he is extremely unlikely to re-enroll in school and faces very limited employment prospects, increasing his/her risk for and exposure to crime, hence an important objective of the Summer Jobs Initiative is for participants to return to school. Therefore, in order to avoid incentivizing dropout by increasing labor market attachment among program participants, Summer Jobs Initiative participants will be compensated through grants that cover school costs during the subsequent school year. Potential participants will be targeted primarily through in-school advertisement of the program. Additional channels through which to recruit for the program include youth centers, churches, and community leaders.

The second target group comprises youth aged 16 and above that are currently not studying nor strongly attached to the labor market.[10] This group will be targeted through a longer-term Temporary Jobs Initiative offering internships of about 6 months, preceded by 3-6 months of training. To discourage youth from dropping out of school in order to access this program, participation will be limited to persons that have not been enrolled in school for at least the past one year. Though preliminary discussions with various groups in the project's target areas suggest recruitment should not be a major challenge, we anticipate that it will be more challenging to recruit a sufficient number of participants for this group and are proposing an extensive advertising campaign using channels such as posters and flyers, youth centers, churches, community leaders, and word-of-mouth.

The Initiative for each group of at-risk youth comprises two basic elements, training and a temporary job. **Training** for both groups will initially focus on life skills and will include a small group-based cognitive behavioral therapy component that encourages professional skills development (timeliness, reliability, use of appropriate workplace language, etc.). Furthermore, training will be designed to help participants develop a more positive worldview, improve executive function (e.g. impulse control, emotional self-regulation) and attenuate cognitive biases that may trigger violent behavior (e.g. through improving participants' ability to read social cues and interpret the intensions of others). Participants in the Temporary Jobs Initiative (those neither in school nor working) will receive additional technical training related to specific jobs offered through the initiative.

Jobs for both the Summer Jobs Initiative and the Temporary Jobs Initiative will be sourced primarily from the private sector, with a view to the eventual sustainability of both Initiatives. By working with private sector employers who are willing to pay interns without subsidies to defray their costs, we will substantially reduce the extent to which the

[9] The long school break in Honduras takes place between November and January. However, we refer to the program targeting youth in schools as the "Summer" Jobs Initiative for consistency with existing literature.

[10] The precise meaning of strong attachment is still to be defined. For the time being, however, we suggest this be defined as having a stable job over the last four months.

pilot initiatives make hiring participants more desirable than would typically be the case and so generate impacts that are unsustainable in the private market. The Safer Municipalities Project will hire an "Operator" in each of its municipalities, who will work with the local chamber of commerce and industry associations to secure jobs for the program. A "guarantee of good behavior" may be offered by the mayor's office on behalf of program participants, and furthermore certain firms may be eligible for a tax incentive by participating in these initiatives as part of corporate social responsibility. All jobs offered through the initiative will comply with Honduran labor law, which restricts the types of jobs that may be done by persons aged less than 18 years.

A third element of both Initiatives, which will be offered to some but not to all participants, is a complementary **mentoring** intervention. Here, individual participants are matched with local role models tasked with promoting/reinforcing life-skills development, encouraging goal setting, and providing labor market and educational guidance. Mentors will remain actively involved with selected program participants after the termination of both the Summer Jobs Initiative and Temporary Jobs Initiative.[11]

We have yet to finalize the details and procedures surrounding the recruitment of mentors. Two options have been discussed with the project team. Mentors could be recruited either by leveraging existing community leaders and youth organizations—and their knowledge of the community— or by direct recruitment among university students who might be selected based on their interest in working with at-risk youth. Mentors will receive training in CBT following a task shifting/sharing model, in which lay counselors with no mental health experience deliver the intervention.[12] Mentors will serve to reinforce the group-based CBT delivered as part of the primary intervention, and will also foster supportive and pro-social relationships with at-risk youth.

[11] The mentoring component of our program is being designed to replicate the large benefits identified for participants in the Chicago One Summer Plus program. In that program, youth received part-time summer employment and adult mentors, and researchers found that violent crime arrests declined 43% for these youth in response (relative to a randomly-assigned Control group).

[12] Following this model, medical evaluations of CBT interventions have yielded encouraging results in low- and middle-income countries, including Zambia, Pakistan, Uganda, DRC, and India (Murray, 2013).

The key features of the Summer Jobs Initiative and Temporary Jobs Initiative are summarized in the following table:

	Summer Jobs Initiative	Temporary Jobs Initiative
Target population	Youth currently in school grades 7-9	Youth that are neither studying nor strongly attached to the labor market
Program duration	2-3 months	9-12 months
Initial training	1-2 weeks of training in life skills and group cognitive behavioral therapy	3-6 months of training in life skills and group cognitive behavioral therapy + vocational training for specific occupations
Types of work	Full or part-time internships over a period of 7-11 weeks	More specialized apprenticeships over a period of 3-6 months
Ongoing training	Life skills and group cognitive behavioral therapy; sexual education including prevention of gender-based violence	Life skills and group cognitive behavioral therapy; sexual education including prevention of gender-based violence; Additional technical training
Add-on individual mentoring component	A subset of program participants are matched with personal mentors to provide personalized life coaching	

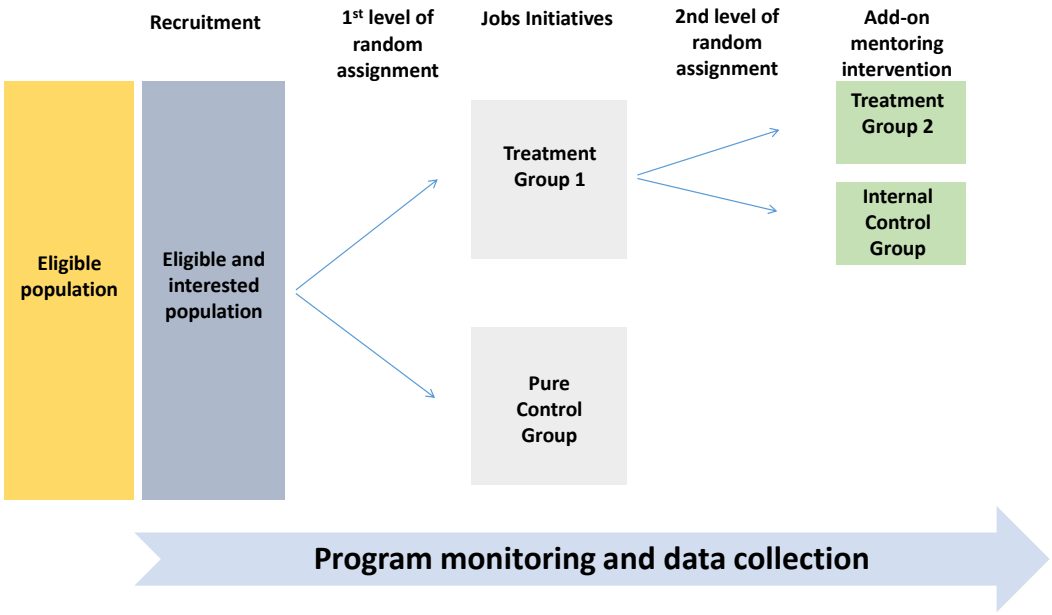
Identification Strategy

This evaluation will take place in selected urban "clusters" in three medium-sized municipalities with homicide rates above 90 per 100,000 inhabitants (World Bank 2012). The municipalities vary in their economic profiles: La Ceiba is known for tourism, Choloma is an industrial center with several textile assembly plants, and El Progreso's economy is primarily based on commerce. The municipalities and clusters were selected by the Safer Municipalities Project.

We will evaluate separately the impacts of the Summer Jobs Initiative and the Temporary Jobs Initiative. For both evaluations, we will employ a common design.

In both cases, our study design relies on over-subscription to the Initiative. This oversubscription will allow us to randomly allocate eligible participants to treatment and control groups to isolate the causal impacts of each Initiative on educational, labor market, and behavioral outcomes. We will stratify random assignment by cluster in order to increase the precision with which we can examine impact heterogeneity based on geographical setting/labor market features. In addition, to gain insight on the mechanisms and on the policy features most effective in ensuring sustained effect over time we will randomly allocate mentoring services to a subset of those selected for program participation.

For each Initiative, this design will create a pure control group of individuals receiving no treatment, and two treatment arms. In the first treatment arm, participants will receive the basic package of interventions under each initiative (T1: training + CBT + job). In the second treatment arm, participants will also receive the additional mentoring intervention (T2: training + CBT + job + mentoring). This experimental design will allow us to evaluate the impacts of each Initiative, with and without the additional mentoring intervention. The experimental design is illustrated in the figure below.



Our evaluation design depends on a high level of interest in, and high take-up of, the intervention. First, as described above, oversubscription provides the foundation for our experimental identification strategy as this provides the basis for random assignment to either the first or second treatment or control groups. Furthermore, high take-up will greatly improve our capacity to detect significant program impacts, particularly if we wish to exploit heterogeneity in baseline characteristics in order to identify, for instance, how program impacts vary based on educational and work history. Based on conversations with the Safer Municipalities Project team, we anticipate that the employment opportunities offered through this intervention will be sufficiently attractive to targeted individuals to ensure that program take-up is high, and we will further quantify this interest and desirable intervention features through an initial youth census (see Section 9 on Data Collection), which will likewise allow us to quantify the extent to which outreach efforts of both Initiatives were successful in raising interest amongst target populations.

Another key parameter is program attrition (i.e. the fraction of treatment individuals who do not continue participating after initial program enrollment). Our discussions thus far indicate that attrition concerns will likely be limited as well given the financial/professional attractiveness of the program being evaluated.

The benefits of being in the treatment group are assumed to be substantial, at least in the short term while targeted youth are employed and receive financial support (in the form of a stipend for the low labor force attachment intervention and in the form of school fees for the in-school population). This may be a source of tension for those randomly assigned to the control group, and for this reason it may be important to implement a rollout design where those initially assigned to the Control group will subsequently become eligible to participate in the program to which they applied. Such an approach has been discussed with the Safer Municipalities Project team, but of course its desirability and feasibility depends on initial experiences with the two jobs initiatives.

8.1 Treatment and Control Groups

Provide specific description of features of each control and treatment arm (one paragraph per arm).

Summer Jobs Initiative: The target group comprises youth currently studying in 7th, 8th, or 9th grade. The total duration of the Summer Jobs Initiative is 2-3 months. This group is targeted due to the high incidence of school dropout observed after 9th grade. Our study design relies on over-subscription to the Initiative. This oversubscription will allow us to randomly allocate eligible participants to Control, Treatment 1, and Treatment 2 groups to isolate the causal impacts of each treatment on the education and criminality outcomes described in detail in Section 7.

Control	Treatment 1: training + CBT + job	Treatment 2: training + CBT + job+ mentoring
Individuals randomized into the control group will be asked to complete household surveys and will be offered either the Treatment 1 package of training + CBT + jobs or the Treatment 2 package of training + CBT + jobs+ mentoring in Year 2.	Participants will be offered summer jobs lasting 2-3 months and will be compensated through grants that cover school costs during the subsequent school year. Jobs will be sourced primarily from the private sector. In addition to providing short-term employment, the Initiative includes a training component that will initially focus on life skills and will include a small group-based cognitive behavioral therapy component that encourages professional skills development (timeliness, reliability, use of appropriate workplace language, etc.). Furthermore, training will be designed to help participants develop a more positive worldview, improve executive function (e.g. impulse control, emotional self-regulation) and attenuate cognitive biases that may trigger violent behavior (e.g. through improving participants' ability to read social cues and interpret the intensions of others).	Individuals in Treatment 2 will receive the summer jobs and training described in the preceding column. In addition, individual participants will be matched with local role models tasked with promoting/reinforcing life-skills development, encouraging goal setting, and providing labor market and educational guidance. Mentors will remain actively involved with selected program participants after the termination of the Summer Jobs Initiative.

Temporary Jobs Initiative: The target group comprises youth aged 16 and above that are currently not studying nor strongly attached to the labor market. The Temporary Jobs Initiative will offer these individuals internships of about 6 months, preceded by 3-6 months of training. To discourage youth from dropping out of school to access this program, participation will be limited to persons that have not been enrolled in school for at least the past year. Our study design relies on over-subscription to the Initiative. This oversubscription will allow us to randomly allocate eligible participants to Control, Treatment 1, and Treatment 2 groups to isolate the causal impacts of each treatment on the labor market and criminality outcomes described in detail in Section 7.

Control	Treatment 1: training + CBT + job	Treatment 2: training + CBT + job + mentoring
Individuals randomized into the control group will be asked to complete household surveys and will be offered either the Treatment 1 package of training +CBT + jobs or the Treatment 2 package of training +CBT + jobs + mentoring in Year 2.	This group will receive paid internships lasting approximately 6 months. In addition to the internship, the Initiative includes a three to six month training component that precedes the internship and that will initially focus on life skills and will include a small group-based cognitive behavioral therapy component that encourages professional skills development (timeliness, reliability, use of appropriate workplace language, etc.). Furthermore, training will be designed to help participants develop a more positive worldview, improve executive function (e.g. impulse control, emotional self-regulation) and attenuate cognitive biases that may trigger violent behavior (e.g. through improving participants' ability to read social cues and interpret the intensions of others). Participants in the Temporary Jobs Initiative will receive additional technical training related to specific jobs offered through the initiative before the start of individual internships, and possible additional external training during the course of their internship.	Individuals in Treatment 2 will receive the internships and training described in the preceding column. In addition, individual participants will be matched with local role models tasked with promoting/reinforcing life-skills development, encouraging goal setting, and providing labor market and educational guidance. Mentors will remain actively involved with selected program participants after the termination of the Temporary Jobs Initiative.

8.2 Sample Size Calculations

Present the sample size estimates. Describe how the sample size was determined, including the sampling frame, and main assumptions including Minimum Detectable Effect (MDE), variance estimates, intra-cluster correlation, and units per cluster (if applicable).

We anticipate first estimating aggregate average treatment effects within each Initiative by pooling Treatment 1 and Treatment 2 in order to compare those individuals who receive training + CBT + jobs to those assigned to the Control group. In a specification that pools the sample across Treatment 1 and Treatment 2 within each Initiative, we would require 393 Treatment and 393 Control individuals in order to have 80% power to identify effects that are 0.2 standard deviations in magnitude. This calculation is based on the assumption that standard errors need not be

clustered above the individual level given that randomization will take place at the individual level.[13] To facilitate interpretation of the 0.2 standard deviation effect size in reference to specific outcomes of interest, suppose that the baseline share of the target population that has been arrested in the past twelve months is 0.4. Then, a 0.2 standard deviation impact corresponds to a decline in the arrest rate of 0.1, or roughly 25%. As another example, if the baseline number of annual school absences in the target population for the Summer Jobs Initiative is 30 with a standard deviation of 10, then we will be able to detect changes in days absent greater than 2.

Our proposed sample size includes 450 Control individuals within each Initiative in addition to 225 individuals assigned to Treatment 1 and 225 individuals assigned to Treatment 2 within each Initiative. This sample size provides 85% power to detect treatment effects of the magnitude described above. In addition, the proposed sample size provides approximately 55% power to detect impacts of the same magnitude in a comparison between Treatment 1 and Treatment 2 outcomes within each Initiative. Although this estimate suggests that we may be somewhat underpowered to detect differential outcomes across treatment arms, we propose to increase power by increasing the frequency of measurement (and thereby reducing the standard errors associated with outcomes of interest). We will achieve increased frequency by (1) making use of administrative data that is available with greater frequency than survey measures (2) conducting both baseline and endline surveys in order to increase the anticipated signal/noise ratio associated with outcomes measurement.

9. DATA COLLECTION (E,R)

(1 page if basic, 1-2 pages if include all sections for registration and ethical clearance)

Outcomes will be measured at the individual, household, and community levels. When possible, administrative data will be used to analyze program impacts in order to reduce research costs and improve capacity for long-run outcomes measurement.

For the Summer Jobs Initiative, targeting youth that are currently in school, educational outcomes (attendance, achievement, promotion, disciplinary infractions) will be measured using administrative records provided by schools. Through a partnership with local municipalities, individual arrest records and cluster-level crime reporting and clearance rates will be obtained. While data on arrests and crime reporting will provide valuable information related to criminal justice involvement of study participants and local residents, such data is imperfect and incomplete in that it will only reflect reported crimes and arrests made (rather than the actual distribution of crime and individual involvement). Therefore, it will be important to supplement this administrative police data with additional survey measures.

In particular, supplemental individual and household measurements will be collected through baseline and endline surveys that will be conducted with the full universe of study participants as well as randomly sampled non-participants who reside in the study communities. These surveys will include comprehensive modules related to self-reported criminal activity and victimization, perceptions of the acceptability of criminal behavior (social norms), perceptions of safety and crime, and youth aspirations. Additional survey modules will focus on education and labor market outcomes (in order to supplement the administrative data described above), as well as mental health and executive function. Household members of study participants will also be asked questions related to study participant behavior and "soft skills" development. A social networks module will also be included and will be used to develop social network maps and characterize program spillovers. Specifically, the social networks module will collect information on friendship relationships and antisocial behaviors of those in study participants' self-defined

[13] This assumption is discussed in more detail in Section 10, and we propose to test the sensitivity of estimated standard errors to the assumed error covariance structure when conducting data analysis.

social networks. Field-based ("real money") experiments may also be incorporated into individual and household surveys in order to measure, for instance, norms of trust and trustworthiness, willingness to engage in risk sharing, etc.

To supplement administrative educational outcomes, we intend to conduct brief teacher surveys that will be used to characterize Summer Jobs Initiative participants' behavioral changes, including development of soft skills such as self-control. For participants in both Initiatives, we also propose to collect employer and mentor "report cards" related to these outcomes in order to improve our capacity to characterize behavioral mechanisms driving any observed causal impacts.

Finally, we propose to develop a real-time crime reporting system whereby community members can safely and privately report, via cell phone texts, crimes, potentially risky situations, or potentially risky areas. This system would provide (approximate) GPS coordinates along with the nature of indicated insecurity. Heat maps could then be produced in real time to monitor how crime is changing and to provide local government agents with information on locations of insecurity; these heat maps would also provide an alternative, continuous measure of localized crime. In conjunction with administrative data, these heat maps could be analyzed geographically to examine how crime patterns change over the course of the intervention as a function of where program participants live.

9.1 Quantitative Instruments

Describe how primary and secondary outcomes (from section 7) will be measured, their timing and frequency.

As discussed above, administrative data on individual arrests and cluster-level crime reporting and clearance rates will be provided regularly by municipal authorities. Administrative data sources will also be used to measure educational and labor market outcomes.

Individual and household surveys will be conducted for all study participants (i.e. both Treatment and Control individuals) prior to the start of the intervention and after twelve months. The timing of the baseline and endline surveys will ensure that relevant outcomes can be measured before program delivery begins for the Treatment population and before Control individuals become eligible for program participation. As discussed, the household survey will include questions related to the following topics: self-reported victimization and criminality, mental health and executive function, labor market outcomes, social capital (social connectedness, frequency of social activity, etc.), perceptions of public safety, social norms, and household member perceptions of study participants' behavioral changes. Standard best practices to reduce non-response rates and attrition, such as multiple home visits at staggered time of day, will be employed.

A natural starting point for the modules on soft skills are the World Bank's Skills Towards Employability and Productivity Program (STEP) and the ILO's School-To-Work-Transition Surveys (SWTS). These instruments are designed to measure cognition, soft skills and employability in low- and middle-income countries for workers and youth in school, respectively. Existing skills surveys will be complemented with instruments for direct and indirect measurement of executive functions, which include the Behavior Rating Inventory of Executive Function (BRIEF) and the Motivated Strategies for Learning Questionnaire (MSLQ), respectively. We will also consider adapting commonly used instruments to measure emotional well-being – such as the Post-Traumatic Stress Diagnostic Scale, Trauma Symptom Inventory, and the Life Events Checklist – specifically for at-risk youth in FCV contexts.

We will also incorporate teacher surveys and employer/mentor "report cards" that characterize participants' development of soft skills, such as self-control, persistence, and positive outlook. Teacher surveys will be conducted at baseline and endline while employers/mentors will be asked to complete brief "report cards" on a monthly basis.

9.2 Management of Data Quality

Describe methods used to enhance the quality of measurements (e.g., multiple observations, training of surveyors), electronic data collection, protocols for quality assurance.

The availability of administrative data related to educational, labor market, and crime and violence outcomes will significantly improve the precision of estimated treatment effects by increasing frequency of measurement (relative to survey-based outcomes). We anticipate that administrative data on outcomes of interest will be available at a quarterly frequency. We also intend to utilize electronic data collection methods, which will improve data quality by reducing data entry errors and by allowing for data quality checks to be conducted in real time (i.e. daily).

9.3 Ethical Issues

Describe if this IE will require ethical approval, informed consent procedures, and important ethical considerations related to data collection.

We will collect informed consent for all household surveys conducted with study participants. We will also put in place a stringent data encryption and security plan to ensure that no individually identifiable information is released. We will seek ethical clearance from an accredited institutional review board in Honduras and/or the United States prior to the start of all primary data collection activities.

We anticipate that the benefits associated with Treatment assignment will be financially substantial for the target population, and so we have proposed a roll out design that ensures all study participants are eventually offered these same benefits contingent on the continuation of either or both Jobs Initiatives.

9.4 Qualitative Instruments

Provide a description of all qualitative instruments (if applicable).

In order to improve our understanding of the key factors that constrain local economic opportunity and lead to high rates of crime and violence, we propose to conduct a series of focus groups at various phases of the project. These focus groups will initially help us to refine our quantitative instrument (i.e. the baseline survey) and will subsequently aid in the program process evaluation. Through this process evaluation, we will aim to more precisely identify the mechanisms driving estimated program impacts (or the lack thereof).

9.5 IE Implementation Monitoring System (R)

Describe the IE implementation monitoring system, particularly, what specific indicators and system will be used to follow up the studied population, their treatment participation, treatment actually delivered and received based on activities, and outputs (see the theory of change section).

Full-time on-site program managers hired by the Operator responsible for program implementation in each municipality will receive regular (monthly) updates on individual program participation (i.e. attendance, punctuality, behavioral measures, etc.) for both the Summer Jobs Initiative and Temporary Jobs Initiative. Program managers will liaison with local educational administrators as well as study participant employers and mentors in order to monitor closely the fidelity of the IE implementation (i.e. that all study participants are assigned and receive program services based on their IE Control/Treatment 1/Treatment 2 status). Program managers will also follow up regularly with randomly chosen study participants in order to ensure that their own reported program involvement matches that reported by organizational partners.

The program manager in each municipality will be supervised by an Employment Program focal point hired by the Safer Municipalities Project team. This person will be supported by a Field Coordinator (Short Term Consultant) hired by the World Bank task team to act as a permanent liaison between the impact evaluation research team and the Safer Municipalities Project team.

10. Data Processing and Analysis

10.1 Model Specification for Quantitative Data Analysis

Describe the statistical method(s) that will be used to compare groups for primary and secondary outcomes (the specific equation should be included), any transformations to quantitative data. Specify whether the standard errors will be clustered or corrected. Specify what IE parameter of interest will be estimated (e.g., ITT, TT, MTE, LATE). Describe how you plan to address multiple hypothesis testing.

For the sake of brevity, this section focuses on the estimation of the reduced form impacts described in sections 6. That is, what are the effects of the combination of vocational training, CBT, and jobs on the following outcomes:

- antisocial behavior (delinquency, aggression, adherence to expected codes of conduct as observed in day-to-day interactions);
- measures of mental health and the incidence of PTSD.

We further ask whether the additional mentoring component strengthens these impacts either by increasing their magnitude or by rendering them more persistent.

In general, the discussion below does not need to distinguish between the two populations under study here, namely youth in school and those out-of-school. While the outcomes of interest may differ across the two groups (as do details about the treatment), the quantitative analysis will be very similar, because of the symmetric structures of the experiment.

Parameters of interest: the intent-to-treat (ITT) and treatment-on-the-treated (TOT) parameter estimates of intervention effects are of interest from both academic and policy perspectives.

- Treatment-on-the-treated (TOT) corresponds to a comparison between treatment and control groups, for those individuals who received treatment (that is, individuals who are assigned to training but do not receive it are excluded from the treatment group). Depending on the source of the self-selection, the estimates can represent an upper or lower bound of the potential effectiveness of the intervention.[14]
- Randomized controlled trials (RCTs) can suffer from noncompliance and missing outcomes if those assigned to treatment opt out of treatment, or those assigned to the control group manage to receive treatment, or others simply attrit from the sample. For instance, with the offer of training some youth may sign up, but not attend, nor report for work. Measuring outcomes only for those who report to work or attend training may introduce a bias if, say, only the most motivated participants show up for training. To overcome selection, we can compare the outcomes of the group assigned to treatment to those assigned to the control group, delivering an estimate of the intent-to-treat (ITT) parameter.
    - In other words, ITT estimates abstract from noncompliance, deviations from protocol, withdrawal, and anything that happens after randomization. Hence, ITT results will represent a more realistic

[14] If people actually receiving treatment are positively selected in terms of their motivation and behavior, we may observe better labor market outcomes, but because they are also less likely to commit crime, we may also observe smaller impacts on antisocial behavior and crime.

estimate of the effect of the intervention, given that individuals can only be assigned eligibility for the intervention, but not to treatment itself.

We are of course also interested in heterogeneity of the impacts across the entire population. The distribution of treatment effects can be recovered by quantile methods, delivering Quantile Treatment Effects (QTE), which represent the impact of the intervention on the *q*th (conditional) quantile of outcomes. That is, we compare the *q*th quantile of the treated group and *qth* quantile of the control group, conditional on individual characteristics. Realistically, QTE estimates may not be precisely estimated given our sample size, unless we increase the frequency of the measurement and/or we expand sample size by introducing a phase-in design.

Estimating equations. Given random assignment, Z_i, of individual *i* to treatment, the following equation may be estimated by OLS to recover an estimate of the ITT parameter:

(1) $Y_i = \alpha + \beta_i Z_i + \pi X_i + \varepsilon_i$

where Y_i represents some outcome of interest such as antisocial conduct or dropping out of school, and X_i is a vector of individual characteristics. The parameter of interest is β, the average effect of assignment to the treatment group on outcomes.

In turn, the TOT is estimated based on the following equation:

(2) $Y_i = \alpha + \gamma_i T_i + \pi X_i + \varepsilon_i$

where T_i indicates whether individual *i* received vocational training, CBT, and jobs. Given the concerns described above regarding self-selection into treatment, the equation is estimated by instrumenting receipt of treatment (T_i) with assignment to the treatment group (Z_i).

In the current context, note that we also dispose of a baseline measurement which we can exploit in a difference-in-differences framework. Concerns about selection into treatment are mitigated by way of individual fixed effects, α_i, which account for the unobserved and time-invariant characteristics of individuals that might be correlated with their treatment status:

(3) $Y_{it} = \alpha_i + \theta_i(T_i \times P_t) + \gamma T_i + \delta P_t + \pi X_{it} + \varepsilon_{it}$,

where P_t is an indicator for the period that follows treatment (i.e. the non-baseline period). Estimating this equation delivers an estimate of $E(\theta_i)$, the average improvement in outcomes among program beneficiaries between the baseline and post-treatment periods relative to the corresponding change in outcomes among the control group.

Finally, as discussed earlier, a random subset of participants in the treatment groups will also be assigned to mentors, who will reinforce life-skills development and provide advice regarding the labor market and educational choices. We ask the question whether the outcomes of these beneficiaries differ from those of participants in the treatment group who were not assigned mentors. Letting M_i be an indicator for assignment to a mentor, the estimating equation analogous to (2) above may be written

(4) $Y_i = \alpha + \gamma_i T_i + \varphi_i M_i + \pi X_i + \varepsilon_i$

where φ_i may be interpreted as the additional benefit to individual *i* of mentoring above and beyond receipt of training and job since, for those who are assigned a mentor, the program's effect on their outcomes is $\gamma_i + \varphi_i$.

Standard Errors. There are two main reasons for clustering standard errors in RCT settings. Either a macro treatment is administered to micro-units, which mechanically introduces correlation within units of treatment,[15] or outcomes, behaviors, and decisions to engage in a given behavior may be subject to peer effects (i.e. in a given neighborhood, social network, class) so that the independence assumption across observations is violated.

The former reason for clustering is less relevant in the current context given the nature of the treatment that individuals will be receiving. Not only will the randomization take place at the individual level, but some dimensions of treatment (i.e., jobs will expose youth to separate experiences as will mentoring) generate individual-level variation. There may be a rationale for clustering at the training class-level, and we plan on testing the robustness of our results to various assumptions concerning the errors structure. However, at least theoretically, the rationale for doing so is weaker.

More relevant to the problem at hand is the likelihood that the outcomes of interest are subject to social interactions (e.g., class behaviors and interactions, crime and violence, etc.) and that all are influenced by common factors/opportunities at the neighborhoods levels. The more appropriate level of clustering would be the network of peers. If we believe that the decision to enroll in the program and geographical proximity are determinants of networks, then a less precise yet conservative level of clustering would be at the neighborhood-training class level for the out-of-school subjects and at the neighborhood-school-class level for those in school.

Subgroup analysis. Mindful of multiple testing problems, heterogeneity of impact analysis will be carried out parsimoniously on the following four dimensions, which are relevant to scaling up and policy design:

- *Age*. Soft skills and personality traits appear to be amenable to policy interventions outside of early childhood, which is the typical window of opportunity identified for cognitive ability (Cunha, Heckman, Lochner, and Masterov, 2006; Heller et al., 2013; Heckman, Stixrud, and Urzua, 2006).[16] The focus here on age groups in- and out-of-school will permit to gain insight into the age gradient of effects of the program and their persistence/sustainability over time.
- *Gender.* Gender subgroup analysis will also be conducted. Recent research indicates that girls may be more sensitive to these types of interventions and may benefit more in terms of well-being. However, given their lower baseline delinquency, we expect the interventions to have a smaller effect on crime and violence among girls.
- *School outcomes & criminal history*. Documenting differential effects according to pre-program criminality and educational outcomes will be important in terms of targeting and program design. Heterogeneous effects along this dimension may point to high returns among those at higher risk, or the need to boost the intervention for them.
- *Mentors*. The effectiveness of programs that seek to develop human capital may vary considerably across individual program providers or across settings. Understanding the effect of mentor characteristics, such as their human capital and soft skills, is particularly relevant for scaling-up.

Spillover and peer effects. As highlighted earlier, many behaviors are influenced by social interactions, including criminality and decisions to participate in the program. Mapping individuals' social networks and recording the outcomes of network members will permit to measure the program's spillovers. Furthermore, self-selection into the program and individual randomization will generate variation in within-network treatment intensity (i.e. how many individuals within one's network are treated). This will allow us to evaluate whether the intensity of treatment within

[15] For instance, a new malaria drug is randomized at the village level, and all individuals in a treatment village receive the same drug.

[16] In particular, executive functions and self-control regulation remain malleable into the mid 20s.

network affects the behavior of network peers and even identify "central/pivotal" players within networks. The magnitude of peer and spillover effects is critical to the cost-benefit calculations of this type of intervention.

Baseline balance tests. We plan on conducting a census and a baseline which will provide detailed information on socio-economic status including education outcomes, labor market history, age, household composition, income proxies, etc. Our baseline will also include measures of personality traits, mental health, and self-reported antisocial behavior and past criminal records/history. This set of information will inform not only the selection into the program as a function of the degree to which youth may be at risk, but will also allow us to control for non-random compliance, attrition, etc. Furthermore, as mentioned above, we also plan to map youths' social network and collect a parsimonious set of outcomes on members of these networks. This will allow us to identify individuals who play pivotal roles in influencing the decision to select into the program. Also, the peers' outcomes will offer additional characteristics to control for threats to the integrity of the protocol.

Correcting for sources of bias. In order to mitigate concerns related to randomization imbalance, we propose to stratify the randomization by geographical cluster. Moreover, conducting the baseline survey prior to the randomization will eliminate concerns related to endogenous non-response at baseline.

Nonetheless, we must still address potential concerns related to endogenous take-up and non-random attrition/non-response.

- With regards to endogenous take-up, we will rely primarily on intent-to-treat estimates, which are derived based on randomized treatment assignment rather than actual program participation. While endogenous take-up will attenuate intent-to-treat estimates, these attenuated estimates will accurately reflect the impacts we would predict from program expansion (see Section 11 for additional discussion of external validity). Importantly, the availability of baseline data will allow us to statistically test whether the baseline characteristics of non-participants assigned to treatment differ from the characteristics of those assigned to treatment. Evidence of statistically significant differences as a function of program take-up will suggest that the population of compliers is not representative of the target population and so treatment-on-the-treated estimates should be interpreted with particular caution.

To deal with attrition or non-response, we will test whether the baseline characteristics of attriters/non-responders differ significantly from the characteristics of other study participants. To the extent that these characteristics do differ significantly, we propose to test robustness based on the following approaches:

- re-weighting observations so that the sample of responses matches the full sample in terms of baseline measures (this procedure relies on the assumption that non-response/attrition is correlated with observables but not unobservables),
- logical imputation whereby missing observations are assumed to result from unobserved arrests and so correspond to cases of school dropout or loss of employment;

Details on planned methods for data entry, and for handling missing data, imputations are provided in the appendix.

Study registry. We will register this study in the AEA RCT Registry (www.socialscieneregistry.org) prior to the completion of baseline data collection.

We include detail on **multiple hypothesis testing** and **data coding, entry, and editing in Appendix 1.**

11. STUDY LIMITATIONS AND RISKS (E)

(1/2 page)

Provide an assessment of risk and threat to internal validity (related to previous section). Discuss issues related to external validity, particularly (i) representativeness of the sample; (ii) representativeness of the institution(s) delivering the intervention, and (iii) feasibility that the intervention can be scaled up.

Randomization into individual treatment assignment will be stratified by geographical cluster in order to ensure balance along this dimension. The randomization will be conducted with a stable seed in order to ensure replicability, and the baseline survey will be conducted before treatment assignment to eliminate any concerns that assignment may itself influence survey responses (or response rate). Additional threats to internal validity related to multiple inference, clustering of standard errors, and missing data/attrition are discussed in more detail in Section 10. In that section, we discuss specific practices that will be employed to mitigate these threats. In order to reduce threats to internal validity related to treatment non-compliance, we will rely on on-site program managers who are tasked with monitoring compliance to treatment assignment, reporting issues related to program fidelity, and conducting random visits with program participants to ensure accuracy of organizational reporting.

An important issue in the recruitment process, particularly for potential Temporary Jobs Initiative participants who are no longer in school and have weak labor force attachment, is self-selection into the program. Since more motivated individuals will be more likely to seek employment opportunities (all else equal), it is critical that we advertise broadly enough that we attract "low-motivation" individuals who are induced to participate by our advertisement campaign but may be differentially impacted by the intervention (relative to "high-motivation" participants). Ensuring that estimated treatment effects are not driven primarily by "high-motivation" individuals (who are not representative of the target population) will improve external validity by reducing anticipated differences between the impacts estimated in the initial evaluation and the impacts that would be associated with a scaled-up version of the program.

In terms of institutional representativeness, the intervention has been designed with scalability and external validity in mind. In particular, reliance on private sector employers who do not receive direct compensation for the hiring of program participants will ensure that the types of employers who are induced to participate will be largely representative of the employers that would choose to participate in a scaled-up version of the program. Importantly, if the proposed intervention identifies economically significant impacts associated with treatment assignment, we are prepared to work in conjunction with the Honduran government to quickly bring the program to scale. High rates of unemployment and school dropout within the target population of Honduran youth ensure that finding eligible participants in a scaled-up version of the intervention should not be an issue.

12. IE MANAGEMENT (E,R)

12.1 EVALUATION TEAM AND MAIN COUNTERPARTS

Provide list of all IE team members with their position, affiliation, and responsibilities (including lead researcher, other research team members, and all project staff involved in the IE work, and main implementing agency counterparts).

Table 2. IE Team and Main Counterparts

Name	Role	Organization/Unit
Laura Chioda	Lead Investigator, IE Co-TTL	WBG/LCRCE
Marcus Holmlund	Principal Investigator, IE TTL	WBG/DECIE
Marco Castillo	Principal Investigator	George Mason University
Benjamin Feigenberg	Principal Investigator	University of Illinois
Ragan Petrie	Principal Investigator	George Mason University
Marcelo Fabre	Safer Municipalities Project TTL	WBG/GSURR
Rocío Calidonio	Safer Municipalities Project Co-TTL	WBG/GSURR
Zunilda Martel	Safer Municipalities Project Coordinator	IDECOAS-FHIS

12.2 Work Plan and Deliverables

Table 3. Milestones, Deliverables, and Estimated Timeline

Milestones	Deliverables	Completion Date
Peer-reviewed Concept Note	Methodology note	June 2015
Data collection plan and pilot	TORs Questionnaires	August -November 2015
Data collection (Baseline)	Cleaned data Dictionaries	January-February 2016
First data analysis	Presentation Data file Do files Baseline report	March-August 2016
Implementation of intervention aligned to evaluation	Rollout plan Monitoring reports verifying treatment and control status	February 2016 – February 2017
Follow-up data collection plan	TORs Questionnaire	August -November 2016
Data collection (Follow-up)	Cleaned data Dictionaries	January-February 2017
Final report and policy notes	Technical note Policy note Data file Do files	August 2017
Dissemination of findings	Presentations	Following all data analyses and reports

12.3 BUDGET

(1 paragraph)

Present total budget disaggregated by staff time, data collection, and travel. Include all sources of funding, both Bank-executed and client-executed (BB resources, trust fund and grants, FBS, EFO, project financing for the IE, such as data collection, and other client financing). Estimate and include all research/staff time (not only the time charged).

Table 4. Total Budget per Category by FY

Category	FY15 ($ x 1000)	FY16 ($ x 1000)	FY17 ($ x 1000)	FY18 ($ x 1000)	FY19 ($ x 1000)	TOTAL ($ x 1000)	% of TOTAL
Staff	**24**	**47**	**47**	**47**	**47**	**212**	**16%**
STC (PIs, RAs/Field Coordinators)	**12**	**81**	**67**	**68**	**84**	**312**	**24%**
- *Principal Investigators**	*12*	*36*	*21*	*21*	*36*	126	*10%*
- *Field Coordinator / Research Assistant*	-	*45*	*46*	*47*	*48*	186	*14%*
Data Collection (Total)	**-**	**391**	**28**	**259**	**-**	**678**	**53%**
- *Baseline*	-	*321*	-	-	-	*321*	*25%*
- *Ongoing*	-	*70*	*28*	*28*	-	*126*	*10%*
- *Endline*	-		-	*231*	-	*231*	*18%*
Travel	**7**	**32**	**11**	**28**	**11**	**89**	**7%**
TOTAL	**43**	**551**	**153**	**402**	**142**	**1291**	**100%**

** 50% cost sharing, i.e. financed directly through external sources*

The total estimated budget for this IE is $1,291k over FY15-FY19, including the design phase leading up to this concept note. Excluding costs financed directly through cost sharing with external sources, the total estimated budget is $1,228k. Of this, 55% is for data collection activities, and 45% is for research and analytical services (time, travel, and other expenses for the research team).

We anticipate that data collection will be covered through the Safer Municipalities Project budget, and this has been discussed with the project team. The scale and scope of data collection will be adjusted to resource availability.

For the remaining $550k for research and analytical services, to date we have secured $310k from the following sources: (i) i2i Trust Fund grant of $190k; (ii) Jobs Multi-Donor Trust Fund grant of $300k, with approximately $100k allocated to this project; and (iii) $20k of FY16 WPA from the CMU.

To cover the remaining $240k, needed primarily to finance FY18 and FY19 activities, we have applied to the Knowledge for Change Program III ($200k requested, with approximately $65k allocated to this project), and to the Joint ILO-WBG Research Program ($300k requested, with approximately $150k allocated to this project). We will continue fundraising from suitable sources, such as the Global Innovation Fund, until fully adequate resources are available.

13. PLAN FOR USING DATA AND EVIDENCE FROM THE STUDY

(1 paragraph)

Describe communication, participation, and dissemination strategy (potential users of findings, media channels) at all stages of the IE (design, baseline analysis, mid-corrections, follow-up analysis, and final results).

This study has been designed and will be implemented in close collaboration with the Safer Municipalities Project team, and data, evidence, and knowledge generated throughout the IE process will be communicated first-hand to

the project team. This will be used to inform design modifications and eventual scale-up of the interventions described here to other high-crime municipalities in Honduras (contingent on positive results).

Additionally, we expect that the experience of the Safer Municipalities Project will be relevant to other countries in the region, in particular neighboring El Salvador and Guatemala which also suffer from very high rates of violence and crime. We will work with the World Bank's Central America Country Management Unit and Social, Urban and Rural Development and Resilience Global Practice to disseminate evidence to potentially interested parties within Central America.

Study findings at all stages will be widely disseminated through Development Impact Evaluation (DIME) and other World Bank Group channels (e.g. the Chief Economist's Office for Latin America and the Caribbean, GSURR), including multi-country workshops, seminars, and various print and electronic media. Findings will be published in one or more World Bank working papers and submitted to peer-reviewed journals.

REFERENCES

Attanasio, O. P., Kugler, A. D., & Meghir, C. (2009). Subsidizing Vocational Training for Disadvantaged Youth in Developing Countries: Evidence from a Randomized Trial. *IZA Discussion Papers* .

Benjamani, Krieger, & Yekutieli. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 491-507.

Blattman, C., Fiala, N., & Martinez, S. (2014). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *The Quarterly Journal of Economics*, 697-752.

Blattman, C., Jamison, J., & Sheridan , M. (2015). Reducing Crime and Violence: Experimental Evidence on Adult Noncognitive Investments. *NBER Working Paper*.

Card, D., Ibarraran, P., Regalia, F., Rosas-Shady, D., & Soares, Y. (2011). The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics*.

Chioda, L. (2014). Revisiting some stylized facts about crime and violence. *Conference on Violence and Policing in Latin American and US Cities.* Stanford.

Cook, P., Dodge, K., Farkas, G., Fryer, G., Guryan, R., Ludwig, J., . . . Steinberg, L. (2014). The (Surprising) Efficacy of Academic and Behavioral Intervention with Disadvantaged Youth:. *Economic Research Working Paper*.

Cunha, F., Heckman, J., Lochner, L., & Masterov, D. (2006). Interpreting the Evidence on Life Cycle Skill Formation. *Handbook of the Economics of Education*.

Gould, E., Weinberg, B., & Mustard, D. (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997 . *Review of Economics and Statistics* .

Heckman, J. J., & Kautz, T. (2012). Hard Evidence on Soft Skills. *Labour Economics, 4*, 451-464.

Heckman, J. J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review, 103*(6), 2052-86.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and non-cognitive abilities on labor market outcomes and social behvior. *Journal of Labor Economics*.

Heller, S. (2014). The effects of summer jobs on disadvantaged youth. *Unpublished Manuscript*.

Heller, S., Pollack, H., Ander, R., & Ludwig, J. (2013). Preventing youth violence and dropout: A randomized field experiment. . *NBER Working paper*.

Heller, Shah, Guryan, Ludwig, Mullainathan, & Pollack. (2015). Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago. *NBER Working Paper*.

Horowitz, J., & Manski, C. (2000). Nonparametric analysis of randomized analysis of randomized experiements with missing covariates and outcome data. *Journal of the American Statistical Association*, 77-84.

Kling, J., Katz, L., & Liebman, J. (2007). Experimental Analysis of Neighbourhood effects. *Econometrica*, 83-119.

Lee, D. (2009). Training, Wages, And Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economics Studies*, 1071-1102.

Lin, M.-J. (2008). Does Unemployment Increase Crime? Evidence from US Data 1974-2000. *The Journal of Human Resources*.

Little, R., & Rubin, D. (2002). *Statistical Analysis with Missing Data.* Hoboken: John Wiley and Sons.

Machin , S., & Meghir, C. (2004). Crime and Economic Incentives. *The Journal of Human Resources*.

Murray, Dorsey, Skavenski, Kasoma, Imasiku, Bolton, . . . Cohen. (2013). Identification, modification, and implementation of an evidence-based psychotherapy for children in a lo-income country: the use of TF-CBT in Zambia. *INternational Journal of mental health systems*.

Raphael , S. (2011). Improving Employment Prospects for Former Prison Inmates: Challenges and Policy. In Cook, Ludwig, & McCrary, *Controlling Crime: Strategies and Tradeoffs.* Chicago: Chicago University Press.

Schochet, P. (2008). Statistical Power for Random Assignment Evaluations of Education Programs. *Journal of Educational and behavioral statistics*.

Schochet, P., Burghardt, J., & McConell, S. (2008). Does Job Corps Work? Impact Findings from the National Job Corps Study. *American Economic Review*, 1864-1886.

Szekely, M., & Karver, J. (2014). *"Youth out of school and out of work in Latin America: a cohort analysis".* Mimeo.

Tough, P. (2012). *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character.* Mariner Books.

UN Office on Drugs & Crime. (2011). *2011 Global Study on Homicide.* Vienna: UNODC.

APPENDIX 1. ADDITIONAL DETAIL ON STUDY DESIGN

1. MULTIPLE HYPOTHESIS TESTING

Researchers typically conduct multiple hypothesis tests to answer key impact evaluation questions. Tests are conducted to assess intervention effects for multiple outcomes, for multiple subgroups of individuals, and sometimes across multiple treatment alternatives. In such instances, separate t-tests for each outcome or subgroup are often performed to test the null hypothesis of no impacts, where the type I error rate (statistical significance level) is typically set at $\alpha = 5$ percent for each test. However, when the hypothesis tests are considered together, the combined type I error rate could be considerably larger than 5 percent. At the same time, statistical procedures that correct for multiple testing typically result in hypothesis tests with reduced statistical power, i.e. the probability of rejecting the null hypothesis given that it is false. Stated differently, these adjustment methods reduce the likelihood of identifying real differences between the contrasted groups.

There is disagreement as to the use of multiple testing procedures and the appropriate tradeoffs between type I error and statistical power (Schochet, 2008). However, the following two approaches are those that seem to deliver the better trade-off between power and type I error. The first approach reduces the number of tests being conducted. This method avoids p-value adjustments, which generally reduce the power of any given test, at the cost of limiting the scope of hypothesis testing. The second approach maintains the number of tests but adjusts the p-values accordingly. This method allows for an arbitrarily large number of tests, but the power of each specific test can fall as the number of tests grows.

- Approach 1: We plan to limit the total number of hypotheses being tested, by first choosing a specific set of outcomes based on a priori assessments of their importance and then implementing summary index tests for our three broad outcome areas (antisocial behavior, mental well-being, and school/economic outcome for the population in school and out of school respectively). These indexes combine multiple measures to reduce the total number of tests conducted.
- Approach 2: we adjust the p-values on the summary index tests to reflect the fact that we will still test multiple indexes.[17] Specifically, we will control the family wise error rate (FWER) — the probability of rejecting at least one true null hypothesis— using the free step-down resampling method and the two-stage false discovery rate (FDR)-control procedure.[18] When reporting results for specific outcomes, we will control the FDR, the proportion of rejections that are "false discoveries" (type I errors). FDR control is well suited to exploratory analysis because it allows a small number of type I errors in exchange for greater power than FWER control.

2. DATA CODING, ENTRY, AND EDITING (E)

Describe planned methods for data entry, and for handling missing data, imputations.

Robustness of the estimates to possible missing data will be assessed using one or more of the following three approaches:

[17] We will not rely on Bonferroni-type corrections, which rest of the assumption that all tests are independent of one another. In practical applications, this is often not the case. Depending on the correlation structure of the tests, the Bonferroni correction could be extremely conservative, leading to a high rate of false negatives (i.e. reject the null too often).

[18] Benjamini, Krieger & Yekutieli (2006) show that the two-step FDR sharpens the original formulation of FDR control as long as p-values are either independent or positively correlated.

- Follow Kling, Liebman, and Katz (2007) we could assign the relevant treatment or control group mean to youth with missing values on any element. This approach assumes elements of our outcome index are ***missing completely at random.***
- ***Inverse probability weighting***. We could use only non-missing records, but re-weight the data so the distribution of baseline characteristics in the selected-observed sample is representative of the population of interest. This approach assumes that the data are missing at random in ways that are related to youth observable characteristics but not unobserved determinants of outcomes.
    - Inverse probability weighting can be applied to the attrition problem in panel data.
- ***Imputation***. A different approach to missing data is to try to fill in the missing values, and then analyze the resulting data set as a complete data set. Little and Rubin (2002) provide review of accessible imputation and multiple imputation methods. The most appropriate imputation method will be a function of the observed patterns in missing data. The team will rely on Imputation to test robustness of results.
    - Imputation often leads to incorrect inference because of inconsistent variance estimation. For instance, in linear regression, the estimated variance tends to be too small.

In light of our sample, size nonparametric approaches followed in Horowitz and Manski (2000) and Lee (2009) are unlikely to yield meaningful and tight bounds