

Experience report

Program evaluation of Juventud y Empleo (JE): A critical appraisal

Avaliação do programa Juventud y Empleo (JE): Uma revisão crítica

Bernardo Menescal Ferreira da Silva¹ 

¹Eberhard-Karls-Universität Tübingen, Tübingen, Deutschland

Bernardo Menescal Ferreira da Silva, pardo, mestrando em Políticas Públicas e Mudança Social na Universidade de Tübingen.

HOW TO CITE: Silva, Bernardo Menescal Ferreira. (2023). Program evaluation of Juventud y Empleo (JE): A critical appraisal. *Revista Brasileira de Avaliação*, 12(2 spe), e121223. <https://doi.org/10.4322/rbaval202312012>

Abstract

Like the policies they purport to assess, evaluation itself has goals and can thus also be object of appraisal. By scrutinizing an evaluation, we gain insights into its applicability to actual decision-making, which might help curb premature excitement or disappointment with the results. Foremost, however, the critical review can provide guidance for evaluators seeking to refine their strategies. This essay presents a critical assessment of the 2011 evaluation study conducted by David Card et al. on the landmark *Juventud y Empleo* policy in the Dominican Republic. This is done, first, via a reconstruction of the implementation stages; second, the evaluation's implicit logic model is made explicit; and, finally, we reflect on the quantitative analyses/controls employed. We find that the study shines in modelling participant attrition and quantitatively estimating the bias in (re)assignment. At the same time, we suggest (a) that the inclusion of logic models in published evaluations can boost transparency, reproducibility and support its assessment and (b) a better overview of implementation hurdles can provide a more adequate, context-sensitive frame within which to weight the effect sizes of the intervention.

Keywords: Labor policy. Policy evaluation. Critical assessment.

Resumo

Assim como as políticas que pretendem avaliar, as avaliações em si também têm objetivos e, portanto, também podem ser objeto de avaliação. Ao examinar uma avaliação, obtemos informações sobre sua aplicabilidade à tomada de decisão, o que pode ajudar a conter a empolgação ou a decepção prematuras diante dos resultados. A revisão crítica é sobretudo útil, porém, ao fornecer orientações para avaliadores à busca de refinar suas estratégias. No presente ensaio deslinda-se uma avaliação crítica do estudo de avaliação de 2011 conduzido por David Card et al. sobre a política *Juventud y Empleo* na República Dominicana. Isso é feito, primeiro, mediante a reconstrução das etapas de implementação do programa; segundo, o modelo lógico implícito na avaliação é explicitado; e, finalmente, refletimos sobre as análises quantitativas e controles empregados. Observamos que o estudo se destaca ao modelar a perda de participantes e estimar quantitativamente o viés na (re)atribuição. Ao mesmo tempo, sugerimos que (a) a inclusão de modelos lógicos na avaliação, quando publicados, pode aumentar a transparência, a reprodutibilidade e apoiar sua avaliação e (b) uma visão geral melhor dos obstáculos de implementação pode fornecer um quadro mais adequado e sensível ao contexto para ponderar a magnitude dos efeitos da intervenção.

Palavras-chave: Política de emprego. Avaliação de políticas públicas. Revisão crítica.

RBAVAL supports efforts related to the visibility of African descent authors in scientific production. Thus, our publications request the self-declaration of color/ethnicity of the authors to make such information visible in the journal.

Received: May 07, 2023

Accepted: May 17, 2023

***Corresponding author:**

Bernardo Menescal Ferreira da Silva

E-mail: bmfs22@gmail.com

Instituições Parceiras: Fundação Arymax e Fundação Tide Setubal



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Introduction

I here attempt a critical assessment of a 2011 evaluation study, conducted by David Card, Pablo Ibarán, Ferdinando Regalia, David Rosas-Shady and Yuri Soares, of the landmark Dominican *Juventud y Empleo* (JE) employment programme. Part of a series of employment policy interventions financed by the Interamerican Development Bank (IDB) at the turn of the century, JE innovated by building a Randomized Control Trial (RCT)-type experimental setup into the programme from its inception. The evaluation study scrutinizes the programme quantitatively and stands out for its meticulous deployment of regression and probabilistic models to filter out potential confounders arising from group reassignment, non-follow-up, heterogeneity of implementation quality, as well as time, individual and even place-related unobservables. Of primary concern is the identification average treatment effects for the whole intervention “package” on employment outcomes, without opening the “black box” to specify mechanisms (e.g., relations of mediation and moderation).

This critical appraisal begins with an outline of the core aspects of the intervention, namely the target group, treatment and control, design features and implementation. In tracing this overview, I take a critical perspective on how intervention details are conveyed in the evaluation study. Subsequently, the evaluation itself is sketched out and subjected to critical scrutiny, starting with a discussion on the underlying (quasi-tacit) theoretical framework, followed by a reflection on the methods and analytical procedures adopted.

Intervention

Despite the emphasis on mostly “black-box” quantitative analyses, rather than on richly descriptive qualitative procedures, the authors provide a comprehensive account of the stages of the intervention, from the very first stages of service-provider selection and outreach to follow-up (see Figure 1). Most of the attention is dedicated to the intervention’s design, but implementation details are used – albeit unsystematically – to flag analytical caveats.

Low-income youth (18-29), from poor neighbourhoods, and having no more than 11 years of schooling are the chosen target population of JE. It is not clear from the description of the evaluation design, whether the program participants were selected from the total population that is in principle eligible for JE, given the characteristics above, or if the public that was made aware of and took part in the programme is in fact a *sui generis* subset of the population. It is possible that certain features of the outreach and enrollment procedure acted as entry barriers to youth who were, on paper, also meant to be targeted by the programme. Nevertheless, the 8,391 programme participants¹ – a representative subset of which is used in the quantitative analysis – match the target group description: most were aged 20-24, had an average of 9.2 years of schooling; about a third had only primary education, one fourth had access to an outdoor toilet, one-fifth was married and/or had dependents.

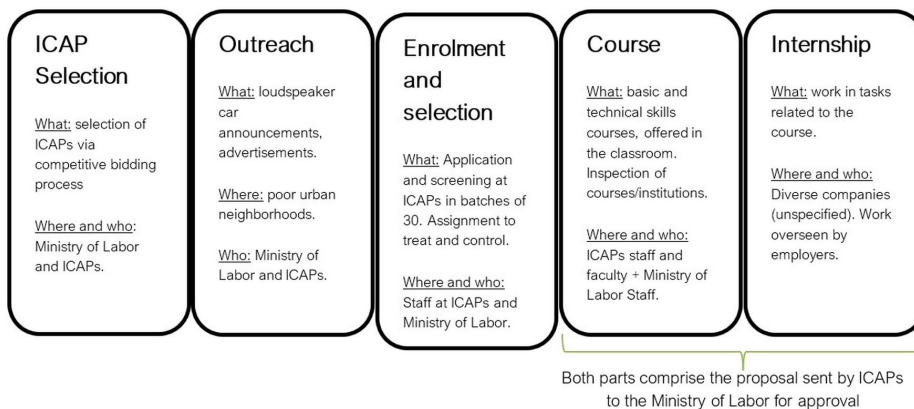


Figure 1. Stages of program implementation.

¹ By “participant” I mean “eligible applicants”, whether they were allocated to the treatment or to the control group.



Nowhere is it specified that the target group consists only of urban youth, but the description of the outreach component suggests that JE was only publicized in Dominican *cities* at a time when nearly one-third of the population still lived in rural areas (The World Bank, 2005). While this does not preclude the possibility rural youth may have heard of the program through other means, such as word-of-mouth or radio spots, the fact remains that outreach was urban-centred.

As clarified by the authors of the study, the intervention is a type of Technical and Vocational Education and Training (TVET) program composed of a two-part classroom component offered at a *Institución* – Basic Skills and Technical Skills – followed by an internship of at least two months duration – a practical component. The study specifies that the Basic Skills module was intended to bolster the “self-confidence and work habits” of participants, whereas the Technical Skills module, to be developed in consultation with the firms offering the internships, should equip them with the requisite competencies to perform well in those jobs. Not much is said, however, about the concrete syllabus, teaching methods, and learning objectives of the two modules, or on the quality of implementation across training institutions (ICAPs). Moreover, the study only reports the maximum formal duration of the classroom component (course), 350 hours, and not the workload of the *actual* courses offered. Another aspect which is not clear is the frequency of the classes, their time, and how they were distributed over time (in semesters, in a few months).

The conditions of internship are not fully elucidated in the article either. With the given information alone, we cannot tell in which sectors, for what kinds of tasks, roles, and for how many hours per week JE treatment group participants were hired. Another noteworthy implementation detail – because to some extent it challenges the premise of employability – is the fact that relatively few of the participating employers seemed to have used the internships as a screening procedure channel for recruiting new employees. Since the internships were fully subsidized, they had an incentive to let employees go and hire new interns at the end of the training period.

Despite not being a component of the training program proper, the outreach campaign and enrolment protocol that function as a sampling procedure are an inextricable part of the policy, for they could, at least theoretically, influence the make-up of the participants. Advertising of the policy, application and screening procedures as well as sampling are all portrayed in detail. Also, given that ICAPs themselves are selected via a competitive bidding process, then strictly speaking the effectiveness of this process (in turn determined by the government’s underlying capacity to deliver) is yet another design element².

Evaluation

The study purports to “report the impact” of the JE programme, focusing on four outcome variables: employment status, hours worked, monthly earnings and hourly wage. Despite the scope of the evaluation not being explicitly stated, it becomes clear through the text that it aims to (a) assess the programme as a whole (a black box) without specifying which design components are responsible for what, and (b) the focus lies on *proximal outcomes*, namely, employment effects on the programme on participants, rather than long-term outcomes or general equilibrium effects, sometimes referred in the literature as *impact* (CDC, 2007).

As much as the authors exhibit a piercing understanding of the variables at play and how they relate to one another, the absence of an explicit logic model of any kind gets in the way of conveying the rationale of the program and of the analysis to readers. It is nonetheless possible to construct a rough logic model from some empirical findings and the programs statement (see Appendix A). No social-scientific theories are mentioned and the theoretical basis, so to speak, consist of the findings of previous studies of a similar nature and these provide insight on outcomes rather than precise mechanisms. In other words, they

² That one needs to take into account in assaying the applicability / translatability of this program to other contexts. Employing the same, or a similar bidding procedure, ought to leave to the same selection outcome and, consequently – assuming a “population” of ICAPs of the same quality and in a similar distribution to the ones participating in the bidding in the Dominican Republic.



are empirical, rather than theoretical and serve to frame expectations from the analysis. Card et al. (2011) note, for example, that studies from the US indicate modest impacts of training programs, moreover dependent on *the characteristics of the participants and the type of training* (pg. 269). Specifically, part of the literature finds that women benefit more than men from training (pg. 270); voluntary programs are more effective than mandatory programs (pg. 270); interventions targeting youth tend to be less effective than those targeting adults (pg. 270); and programs of the kind implemented in Latin America tends to yield better results than in US and Europe (pg. 270).

A range of social scientific theories could be mobilized to enrich the analysis. Job search theory, which postulates that individuals have imperfect information about job opportunities, could aid specification of the labor market-related competencies and resources – “search efforts” – that improve individual employability (van den Berg & Uhlendorff, 2015). Another promising direction is human capital theory within the Vocational Education and Training Literature, where the link between job tailored training programs and greater productivity is emphasized (Wallenborn, 2010). Additionally, a rational choice theory model of employer self-interest complemented by a qualitative assessment, could aid understanding of Dominican employers’ cost-benefit calculations when choosing to hire or fire a trainee (Bevir, 2010).

Notably, in comparison to previous studies of training programs in Latin America, the study leverages the JEs Randomized Control Trial (RCT) experimental design, to carry out several distinct analyses of participant assignment, sample and group differences, and policy impact following and Intention to Treat (ITT) framework (see Table 1). An array of relevant covariates is mobilized (gender, education, proxies for socioeconomic status), and the main chosen outcome measures of employment rate (% of employed months during window of observation) and wage level are well-established indicators of the performance of training programs such as the JE, posing no issues³. Only for assessing job quality do the authors rely on the proxy of “jobs with health insurance” – at any rate useful because it correlates with formal employment, one of the goals of the JE. Data for the evaluation comes from surveys completed by the participants at the moment of application and after the program was completed. Card et al. (2011) note that most follow-up surveys were conducted nearly 6 months later than originally expected and required participants to trace back their employment and educational activities since the end (or the moment they left) the program – possibly resulting in reporting imprecisions.

Most remarkable in the study is perhaps the extensive investigation of the repercussions of assignment imperfections and non-follow-up on treatment effect estimates, and covariate controls. In addition to including regressors such as region, training institution, and others mentioned above, the authors specify a model of selection associating participant characteristics to labor market outcomes in order to estimate likely labor market outcomes for the participants that were not followed up. From this analysis they derive the insight that no-show behavior was selective, and likely to be influenced by ICAP quality alongside non observed individual characteristics.

Table 1. Summary of analyses conducted by Card et al. (2011)

Assignment	Effect
1) T-test comparison of characteristics of original and realized (excluding dropouts and including reassigned participants) treatment and control groups.	1) OLS estimation of impact of assignment to treatment with extensive controls
2) Logit model of effect of covariates and institutions on participant assignment.	2) Joint probit and OLS models of selection into group plus outcome.
3) T-test comparison of characteristics of original and realized evaluation samples.	3) OLS estimation of impact of assignment to treatment by subgroup (gender, age, education, and location)
	4) OLS estimation of impact of assignment to treatment on labor market outcomes.

³ Other measures include hours of work per week and hourly wage.



Framing the study from a broader evaluation perspective, qualitative methods would no doubt have improved our understanding of the program, although for these methods to be fruitfully applied, a richer logic model would have to be specified first. Canonical Grounded Theory (Bryant, 2020) centered around the analysis of qualitative interview data, for example, could be used to build clearer theoretical models of participants' decision to drop out of the program, or of employers' attitudes towards participants. Such models would be useful not only in clarifying the mechanisms behind course and employment transitions, but also in suggesting avenues for further quantitative analysis. As for quantitative methods, the use of a survival analysis model could have yielded valuable insights on the development of employment spells beyond the observation window, considering that data was right censored at the time of the follow-up interview (Blossfeld & Rohwer, 2019).

With respect to the presentation of study results, those are generally well described in a range of informative tables and the interpretations adhere rigorously to the provided data. A minor nuisance is the omission of p-value symbols from the tables, forcing the reader to make rough assessments of whether error intervals overlap or not. Presenting such information in comparative graphs, such as stacked dots with error margins could aid comprehension.

Conclusion

In contrast to previous (mostly observational) studies, Card et al. (2011) find that the JE program had next to no impact on participants' employment status, though it appears to have contributed to the likelihood of having a *better job*, that is, with health insurance and slightly higher wages, conditional on being employed. The authors are quick to point out that these insights should be interpreted with caution, for even the more robust experimental design of JE was marred by implementation issues, particularly the non-follow-up of dropouts, the non-random reassignment of participants from control to treatment, and quality differences among ICAPs. Although these hurdles were partially circumvented through statistical procedures, Card et al. (2011) conclude, they are no substitutes for "better implementation and simple design". In addition to finding no significant effects on employment⁴ and small positive effects on wage conditional on being employed, the authors make valuable contributions to our understanding of the effect of training quality, and to the relationship between attrition behavior and labor market outcomes.

Ultimately, I find that the study succeeds in its intention to provide a fine-grained assessment of an employment support policy in comparison to foregoing observational evaluations, using sophisticated models to counterbalance hiccups in implementation that are all too common. Framing this study against the backdrop of evidence-informed policy making, I suggest that it could be improved or supplemented by (a) one, or multiple explicit logic models of impact informed by social scientific theories such as job theory, human capital theory and rational choice theory, distinguishing employability and job quality (e.g., wage, benefits packages), (b) a more systematic presentation of the stages of implementation and instances where it fell short of program design, (c) the use of qualitative methods to probe issues of training quality, as well as drop-out and hire-and-fire decisions, on the one hand, and of quantitative survival analysis methods for investigating participants' employment histories post-program.

Financial support

None.

Conflict of interest

None.

⁴ Measured via the employment rate, defined as the proportion of post-treatment months during which an individual was employed (within the stipulated observation window of 12 months). Some effects on "employability" and on the likelihood of getting a better quality job (one with health insurance) are found.



References

- Bevir, Mark. (2010). *Encyclopedia of Political Theory*. London: SAGE Publications. <http://dx.doi.org/10.4135/9781412958660>
- Blossfeld, Hans-Peter, & Rohwer, Götz (2019). *Event history analysis with Stata* (2nd ed.). Routledge: Taylor & Francis Group.
- Bryant, Antony. (2020). The grounded theory method. In Patricia Leavy (Ed.), *The Oxford handbook of qualitative research* (pp. 166-199). Oxford: Oxford University Press. <http://dx.doi.org/10.1093/oxfordhb/9780190847388.013.15>
- Card, David, Ibararán, Pablo, Regalia, Ferdinando, Rosas-Shady, David, & Soares, Yuri. (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, 29(2), 267-300. <http://dx.doi.org/10.1086/658090>
- Centers for Disease Control and Prevention – CDC. National Center for HIV/AIDS, viral hepatitis, STD and TB prevention. (2007). *Types of evaluation*. Retrieved 2022, March 21, from <https://www.cdc.gov/std/program/pupestd/Introduction.pdf>
- The World Bank. (2005). *Población urbana (% del total) - Dominican Republic*. Retrieved 2022, March 21, from <https://datos.bancomundial.org/indicador/SP.URB.TOTL.IN.ZS?locations=DO>
- van den Berg, Gerard J., & Uhlendorff, Arne. (2015). Economic job search and decision-making models. In Ute-Christine Klehe & Edwin A. J. van Hooft (Eds.), *The Oxford handbook of job loss and job search* (vol. 1). New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199764921.013.021>
- Wallenborn, Manfre. (2010). Vocational education and training and human capital development: Current practice and future options. *European Journal of Education*, 45(2), 191-198. <http://dx.doi.org/10.1111/j.1465-3435.2010.01424.x>



Appendix A. Logic model.

Classroom component			Internship component		
Resources	Activities	Outputs	Resources	Activities	Outputs
Training facilities, materials, and faculty of ICAPs	Participants receive basic training in workplace skills	Participants with improved self-confidence for work and new competencies. Documented issue: learning activities and their rationale not specified.	Workplace, colleagues and managers, and participants with newly acquired workplace-relevant skills and greater confidence	Performance of workplace tasks, interaction with peers and social learning.	Accrual of work experience, workplace competencies, social networks, among other factors expected to boost employability.
Monthly stipend for transport	Participants receive vocational training customized to the needs of employers	Participants with relevant technical skills for the workplace. Documented issue: courses are known to be general, with little or no employer-based adaptation	Subsidized wage	ICAPs seek out employers to conclude internship agreements (enabling condition for internship)	Guaranteed internships of at least 2-months in local firms
	Participants attend the courses (enabling condition for the course effects)			Participants are encouraged to participate in the program (as opposed to seeking paid work elsewhere)	
Program outcomes					
Short-term outcomes		Intermediate outcomes		Long-term outcomes	
Participants become more employable, that is, more attractive to employers (due to acquired experience and skills), and more competent in navigating the labour market (seeking information on job opportunities, succeeding in selection procedures etc.). Consequently, they find more and better jobs (with higher wages and health insurance). Documented issue: subsidized wage lowers employers' opportunity costs of hiring new interns issued from the JE training course instead of keeping interns from the previous cohort on a firm-paid wage.		Participants achieve better livelihoods, consistently employed in the formal sector, with access to improved working conditions, including higher wages, healthcare plans and other benefits. Employers benefit from ease of finding more skilled workers (by linking up with ICAPs), reducing hiring costs and boosting efficiency and quality of service/production.		Reduction in poverty levels and increased individual well-being among youth (and older workers, as program participants age in better careers) and their dependents. Greater wealth generation owing to improved firm output through better human capital.	
Obs.: The stated aim of the JE program is to increase the employability of participants, which can be read as a short-term outcome if no time-horizon is specified. Intermediate and long-term outcomes shown here are extrapolations inspired by development theory.					