sustainability

MDPI

Article

Evaluating the Efficacy of Social Innovation Programming at Advancing Rural Development in the Context of Exogenous Shocks

Mauricio Espinoza [1,2,*], Rodrigo Rivarola [1], Ricardo Fort [1] and Joshua Fisher [3,4,*]

[1] Grupo de Análisis para el Desarrollo, Lima 15063, Peru
[2] Development Economics, Wageningen University & Research, 6708 Wageningen, The Netherlands
[3] Network for Education and Research on Peace and Sustainability, Hiroshima University, Hiroshima 739-0046, Japan
[4] Columbia Climate School, Columbia University, New York, NY 10115, USA
* Correspondence: mespinoza@grade.org.pe (M.E.); jf2788@columbia.edu (J.F.)

Abstract: In this paper, we design and deploy an experimental approach to evaluate the efficacy of a social innovation initiative implemented in rural communities situated in the highlands of Peru, which confronted the challenges posed by the COVID-19 pandemic in the midst of its implementation. Using three rounds of information collected before, during, and after participation, we examine the efficacy of the social innovation economic development approach at increasing household livelihoods. We present robust, statistically significant improvements in household economic well-being, food security satisfaction, and community outlook for participating households compared to non-participating households following program engagement. The results presented in this study suggest that the social innovation program facilitated a notable restructuring of the portfolio of household income and livelihood activities towards more lucrative and value-added pursuits. This transition manifested in increased involvement in entrepreneurial ventures and specialized labor associated with both agricultural and non-agricultural sectors while distancing from traditional agricultural and livestock-related endeavors. The income gains stemming from self-employment activities and wage labor outweigh the losses incurred from reduced agricultural and livestock earnings. Furthermore, our analysis underscores the resilience of innovative income-generating pathways in the face of the COVID-19 pandemic, outperforming traditional agrarian value chains. These findings highlight the efficacy of social innovation programming in facilitating economic development and also shed light on sustainable strategies for economic resilience amidst unforeseen challenges such as the recent global health crisis.

Keywords: social innovation; social entrepreneurship; rural development; exogenous shocks

Citation: Espinoza, M.; Rivarola, R.; Fort, R.; Fisher, J. Evaluating the Efficacy of Social Innovation Programming at Advancing Rural Development in the Context of Exogenous Shocks. *Sustainability* **2024**, *16*, 5664. https://doi.org/10.3390/su16135664

Academic Editor: Jianming Cai

Received: 13 April 2024
Revised: 11 June 2024
Accepted: 28 June 2024
Published: 2 July 2024

1. Introduction

Equitable and effective economic and sustainable development remains a persistent challenge in much of the world [1,2]. One possible reason for this is that traditional models of development programming that encourage the uptake of pre-designed programs and depend on market mechanisms to spur technological and business innovation have been argued to be maladapted to the realities of rural and internal periphery communities [3]. In response to the ongoing need for better and more effective development, social innovation models have been presented as alternatives to traditional development programs for rural areas in both academic and civil society discourse [4,5]. While there is as yet no consensus on the definition of social innovation, it is broadly understood as a collective action process that stimulates the development or expansion of a resource and service base, fosters trust and empowerment for marginalized groups, and transforms social, economic, and structural factors to enhance well-being [6–10] suggest that social innovation consists of the interconnected dimensions of satisfaction of basic human needs, changes in social relationships, and expansion of socio-political capacity and resources. In rural development programs, such processes often take the form of multi-stakeholder collaborative planning, implementation, learning, and adaptation programs that leverage the varied resources, power, and networks of participating stakeholders to create economic development and income-generating initiatives aimed at enhancing local economies, increasing income for participants, and enhancing well-being for communities [11–14].

There is a long history of academic research into social innovation, with early academic focus dating back to the early 20th century [15]. However, the historical focus has evolved over time, with an early focus on social movement and, later on, environmental action [9]. Academic and research interest in the application of social innovation to rural development has only recently gained momentum [16,17], with a recent bibliographic study highlighting that the majority of research into rural applications of social innovation has been published in the period starting after 2016 [9]. Due to the relatively recent focus on rural development, bibliographic studies [1,9] suggest that there are important focal, geographic, and demographic knowledge gaps related to rural applications of social innovation. For instance, the majority of published studies have been conducted in European and other Western contexts, with studies of rural social innovation in Latin America being underrepresented. Similarly, authors publishing work on social innovation tend to come from Europe and other Western countries, and less authorship has been published by Latin American researchers [1]. Of the published work on rural applications of social innovation, the vast majority has been theoretical and conceptual, and several studies highlight the important gap between empirical studies and robust evidence related to rural applications of social innovation [2,3,13,18–20]. These gaps are gaining important recent academic interest, with focal journals taking up the mantle of building an evidence base, such as a recent special issue of the *Journal of Rural Studies*, guest edited by [3].

Proponents of social innovation argue that they are effective vehicles to advance rural development [21,22] and suggest that social innovation models offer alternative and promising means of supporting entrepreneurship and job creation because they are co-designed by communities and supporting organizations and thus better and more closely attuned to the contextual, cultural, and structural conditions of rural and peri-urban communities. Using the case study of ecotourism in a rural community in Peru, ref. [23] has recently suggested that the community dynamics that enable effective social innovation programs can enhance community resilience to exogenous shocks like the COVID-19 pandemic. The participatory and highly discursive design of social innovation methodologies may enable participants to more readily access and deploy relevant contextual information about the development ecosystem in order to be more agile in business design and more responsive to evolving market dynamics. However, critics of social innovation approaches highlight several dilemmas that academic and practice communities have yet to resolve. Ref. [2] refers to several of these, including the lack of clear definitions and metrics to observe the social innovation risk of the concept becoming a buzzword that is functionally devoid of meaning [24,25]. Similarly, critics argue that a focus on social innovation risks public sector actors off-loading their responsibility for promoting public goods to civil society, and private sector actors [19,21] highlight that the purported dividends of social innovation are far from guaranteed, and that the success of such initiatives depends on multiple factors within a community. Such debates highlight the urgent need for rigorous experimental and empirical investigation into the effectiveness of social innovation programs in rural development. This leads to the question: *How effective are social innovation approaches to rural development at enabling job and enterprise creation in rural contexts exposed to exogenous shocks?*

That is a non-trivial question from a methodological perspective, though one that has been difficult to address. As highlighted above, the notion of 'social innovation' refers to a wide swath of participatory and stakeholder-centered approaches to development. Each approach is idiosyncratic in its design, rationale, and approach to including participants in programming and implementing programming. Without some standardization, it is difficult to devise evaluation and measurement techniques that are both robust and comparable across cases or programs. The reasons for this difficulty range from small sample sizes to limited statistical power of measuring specific interventions within a program, to variation in the techniques of implementation in treatment groups, novelty and idiosyncrasy of appropriate indicators for social innovation programs, and ultimately difficulty in matching treatment and controls [26,27]. In this study, we make important advances in these regards by presenting empirical evidence that tests the effectiveness of social innovation programming at creating jobs and enterprises and supporting the creation of an enabling entrepreneurial ecosystem in Peru using an experimental design. While our study is constrained in its use of predominantly economic indicators [28], this work is an in-road to beginning to enhance social innovation measurement of complex, systemic interventions [29]. Rather than focusing on single intervention points, we evaluate a program that takes a systemic approach to rural livelihood development and focuses on enhancing the enabling ecosystem for entrepreneurship. Our evaluation of that program enables us to circumvent some of the challenges highlighted above.

The remainder of the paper proceeds as follows. Section 2 introduces the specific social innovation initiative that aims to increase job creation and social enterprise development in Peru. The paper next describes the methodology employed to conduct an independent evaluation of the program using an experimental design. The paper then presents the key results of that evaluation. In Section 6, the paper discusses the effectiveness of the social innovation program in achieving its intended outcomes and extrapolates evidence for the broader research question.

2. Work4Progress Social Innovation Program

In the rural Andes and Amazon regions of Peru, the challenges of traditional development models have been pervasive despite decades of governmental, bilateral, and non-governmental development assistance [30–32]. These challenges were made more acute by the COVID-19 pandemic [33]. These ongoing challenges have catalyzed interest in social innovation models in Peru, with coalitions of non-governmental organizations and governmental actors working to introduce innovation in rural areas and a number of academic studies beginning to explore the effectiveness of and mechanisms underpinning social innovation programming in the region [23,34–36].

Among the various initiatives that have been developed and implemented, the Work4Progress (W4P) program has been promoted in the Peruvian provinces of Quispicanchi (Cusco) and Condorcanqui (Amazonas) by the "la Caixa" Foundation with the objective of encouraging innovation and quality employment for vulnerable rural women and young people. Globally, the program is promoted in India, Mozambique, Peru, and Colombia (https://work4progress.fundacionlacaixa.org, accessed on 12 April 2024). The Work4Progress program is designed to move beyond traditional approaches of entrepreneurship and economic development that are limited to isolated projects into social innovation platforms formed by civil society organizations, universities, the private sector, financial institutions, and public stakeholders. The theory of change underpinning the program is unique among peer social innovation programs in that it takes an explicitly systemic approach to building and strengthening the enabling ecosystem for entrepreneurship and job creation. The logic of the program assumes that by co-creating prototypes with communities that are attuned to local needs, the program can fill missing value chain linkages in the short term. Over time and through iterative co-creation processes, the program builds a portfolio of interconnected and complementary prototypes that reinforce the value chain linkages and enable the emergence of new ones to respond to needs and gaps and to capitalize on emergent opportunities. While the core of the program focus is on individual prototypes, the success of the program depends on the diversified portfolio approach in the same way that ecosystem integrity benefits from functional diversity

[37]. The sustainability of the program is thus not dependent on whether or not any single prototype succeeds but rather on the ability of the portfolio to be self-reinforcing and emergent over time. The W4P Peru Platform started in 2018, and at present, it is implemented by a consortium of Peruvian and international NGOs, including Entreculturas, Fe y Alegría Perú, CCAIJO, SAIPE, AVSI, Acción contra el Hambre, Universidad Nacional Mayor de San Marcos, Codespa, NESST, Alternativa, World Vision, Fablab Lima, Cite Textil, Caja Cusco, and Caja Huancayo, among others.

Rather than relying on single enterprises or pre-designed entrepreneurship training programs, the portfolio approach is essential to the program logic to create the diversity and complementarity required to build a supporting entrepreneurial ecosystem. The logic of the Work4Progress approach is depicted in Figure 1.

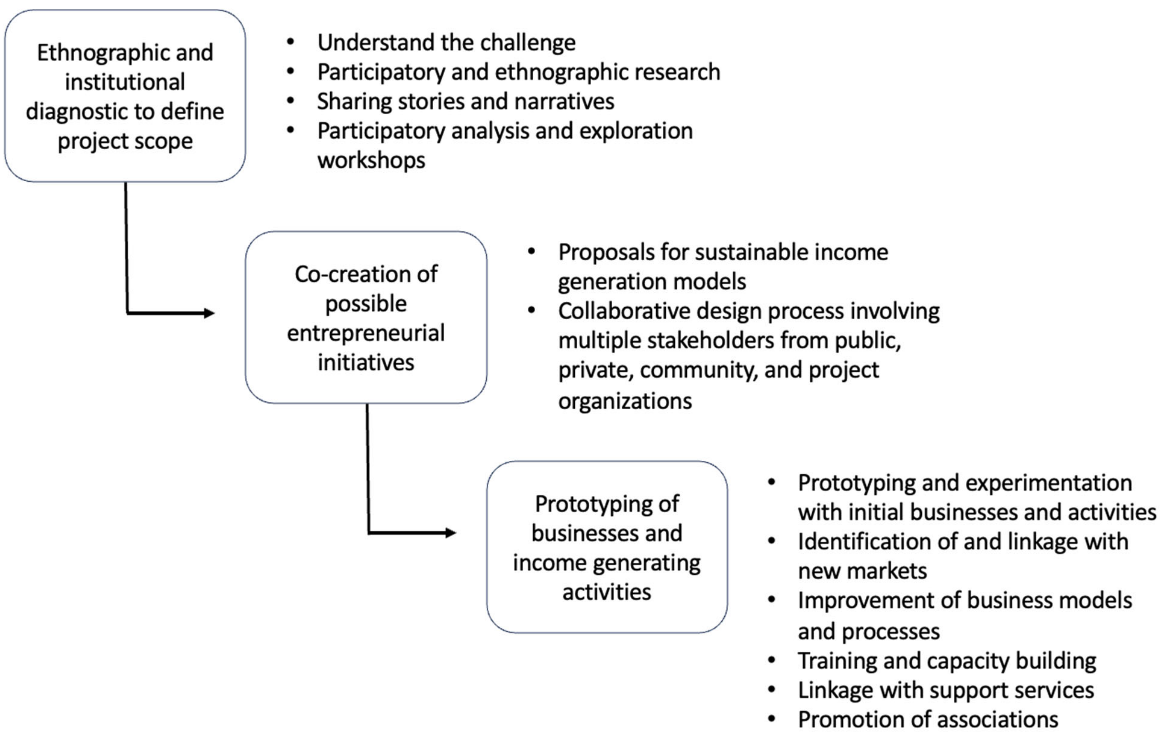

Figure 1. Work4Progress Program Logic.

As a result of the co-creation and prototyping stages, in the first phase of the project, the supporting organizations worked with 500 rural households in the target geography to identify 12 prototypes or specific initiatives to support pilot job creation and enterprise development. Because these were co-created with the participating households, the project assumed them to be adapted to local conditions and the needs and potential of the population and environment. Those prototypes ranged in scope from rural agricultural intensification and specialized product development to fortification of rural tourism and regional working groups on enterprise development. As the project progressed through the piloting of prototypes, the participating households and supporting organizations refined the operating models to identify ways to strengthen and hone the new initiatives and enterprises. Later, phase 2 of the project development expanded to include 1100 households and streamlined many of the existing prototypes into more focused initiatives. Additionally, in phase 2, emphasis was placed on the more systemic prototypes aimed at strengthening the entrepreneurial ecosystem. The list of the prototypes selected by the project is detailed in Table 1. These prototypes can be grouped into three categories. The first one is productive activity implementation or development, which refers to the identification of market niches for new products or the development of an existing one. The second one is associativity promotion, which refers to already existing productive activities where producers could benefit from associativity (for instance, by increasing the scale of production). Finally, the third one is related to labor skills development for job market insertion.

Table 1. Program's Prototypes.

Prototype Group	Prototype	Brief Description
Productive activity implementation or development	Chicken production	Chicken production (animal husbandry) to supply local restaurants
	Mushroom cultivation	Collecting wild pine mushrooms to supply local restaurants
	Tourism	Stimulate tourism to Ausangate mountain
	Artisanal textile production	Artisanal textile production from Alpaca wool
	Crafts center	Start an alpaca fiber craft exhibition center/store
	Freshly cut flower production	Production of freshly cut flowers to supply gourmet restaurants in Cusco city
	Commercial production of livestock (Guinea pig) for sale	Improvement in livestock animal husbandry processes and feed production to enhance commercialization and quality control to supply local restaurants
	Rural secondary education	Diverse projects with rural secondary students
	Andean cheese production	Produce craft cheese to supply local restaurants and households
Associativity promotion	Dairy products	Promote associativity among dairy producers to increase the scale of production
Labor skills development	Business development center	Support local entrepreneurs to set up their business
	Job placement services	Gather available job opportunities from private firms and organizations and promote them within the participant communities

The logic of the Work4Progress program assumes that local households and communities are best equipped to understand the local context for job creation and enterprise development. The project also assumes that locally derived initiatives can be successful when supporting organizations from civil society and non-governmental sectors play a facilitative role by providing funding, training, networking, and creating value chain linkages. Through ongoing monitoring and evaluation that includes participants as both holders of knowledge and active participants in analysis and adaptation, referred to by the program as *Developmental Evaluation*, prototype initiatives will have the support and information needed to effectively develop into thriving enterprises and will be able to better adapt to changing economic, political, social, and environmental conditions. However, the program itself was not poised to assess the accuracy of those assumptions and, as such, required an independent, external evaluation to validate the underlying development hypothesis.

In order to assess the effectiveness of the Work4Progress model, a Peruvian research institute was commissioned to design and conduct an independent external evaluation that consisted of an experimental design to measure the outcome and effect of the program in host communities (treatment group) compared against non-participating communities (control group). We conducted an experimental evaluation over a three-year period from 2019 to 2022, using a double differences analytical technique to estimate the program's impact on the treatment group in three observation periods. The unanticipated onset of the COVID-19 pandemic between our baseline and first observation period enabled us to simultaneously conduct a natural experiment to examine the effectiveness of this social innovation model in enabling recovery from the economic impacts of the pandemic and compare that to the post-pandemic recovery of control groups.

While numerous evaluation studies have examined rural development programming, ours is unique in two aspects. First, our research design is specifically focused on rural development models using social innovation methods. Next, the onset of the COVID-19 pandemic during the study enabled us to examine the question posed above of *how effective social innovation approaches are in advancing rural development in contexts exposed to exogenous shocks*. Together, these two aspects add needed evidence to the growing literature on the utility of social innovation and rural development programming.

3. Material and Methods

3.1. Study Design

We examine the causal impact of the program on participant outcomes by using a clustered randomized trial, which is a design where clusters of households participating in the Work4Progress program rather than households themselves are assigned at random to a treatment group. In the context of the W4P program, the intervention targeted a number of rural communities (clusters) within the program's area and provided treatment to households situated in these communities. The intervention involved both household-level and community-level initiatives, aiming to enhance job creation and foster business development in these areas.

The research team, in coordination with the program's staff, who knew the geographic context more precisely, randomly selected a group of communities to include in the treatment and other (non-participating) communities to serve as the control group. In each of these communities, the study randomly selected a subset of households to participate in the survey. The data collection comprised three rounds of surveys conducted before, during, and after program participation, enabling the measurement of changes in household characteristics and economic well-being over time.

We employed a difference-in-difference methodology to estimate the average treatment effect of the program on participating households. While our primary focus was on evaluating the program's economic impacts, we also delved into its effects on other pertinent aspects, including food security, migration plans, and community opportunities. This comprehensive approach allowed for a thorough examination of the multifaceted impacts of the program on both individual households and the broader community context.

The evaluation initially employed a pre-post design but was later adapted in response to the COVID-19 pandemic. This involved integrating an additional midline survey and extending the endline to 2022, deviating from the original plan set for 2021. The baseline survey was conducted before the program and prior to the onset of the pandemic. The midline survey was administered amidst the pandemic in October 2020, shortly after introducing the program's prototypes in the treatment communities. The endline survey was conducted over a year after the program concluded, in October 2022. The unexpected emergence of the COVID-19 pandemic between the baseline and midline surveys provided a unique opportunity to analyze changes in households' circumstances during the pandemic and their subsequent recovery compared to the baseline.

3.2. Study Area and Sample Selection

The project unfolded across three rural districts in Cusco, specifically targeting 35 Andrean communities within this area (refer to Figure 2). This constituted the universe from which communities were selected to either receive the treatment or function as a control group. The decision to allocate the intervention at the community level was influenced by the multifaceted nature of the project's prototypes, encompassing initiatives at both household and community levels. This choice also stemmed from practical considerations, including feasibility, and the need to mitigate potential spillovers that could introduce bias into the treatment effect estimates.

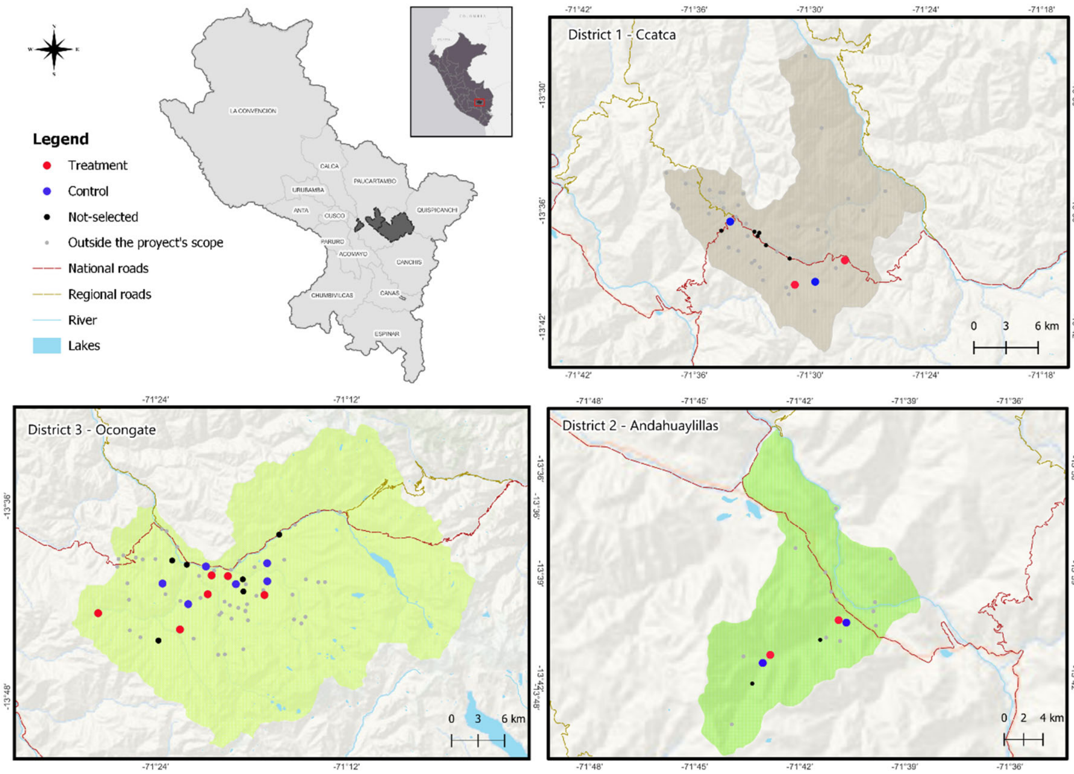

Figure 2. Study area and location of treatment, control, and non-selected communities.

We employed a random assignment procedure to allocate treatment to ten communities and designated ten others as the comparison group (refer to Figure 2 for the distribution of treatment and control communities). Randomization was performed by a member of the research team using Stata 15.0 (StataCorp, College Station, TX, USA) to allocate 20 cluster units randomly to the two study groups (treatment and control). The study budget allowed for data collection from 310 households, evenly distributed between treatment and control communities. Within each of the selected communities, households for the survey were also chosen through a random selection protocol. The three rounds of data collection were conducted by an independent local firm with no affiliations to the project. The entire process was overseen by the research team, which was comprised of the authors of this study, ensuring rigorous supervision in the data collection process.

The sample size of 310 households was primarily determined by the program's initial phase, encompassing ten communities and engaging over 500 households in total. Rather than disaggregating the sample according to the individual prototypes discussed in Table 1 above, we elected to consider the entire cohort as a single treatment group. For practical purposes, the heterogenous distribution of participants in each prototype resulted in very small samples for any single prototype; thus, a clustered approach was more suitable. More importantly, though, the project's logic assumes that the prototypes are part of a portfolio of complementary initiatives that collectively build an enabling ecosystem for entrepreneurship and job creation. While the individual prototypes are a core focus of programming, the actual model operates on the clustered or aggregated approach. As such, we focused our analytical effort on examining the entire sample as a single treatment.

Despite the small size of our sample (i.e., 310 households), our study's statistical power is capable of detecting important changes induced by the program for several reasons. Firstly, the data exhibits high quality and consistency, marked by a significant autocorrelation in outcomes over time. For instance, our primary outcome of interest, gross income per capita, exhibited a correlation of nearly 0.6 between the baseline and midline. Secondly, the data reveal a low level of intra-cluster correlation, with the intra-community correlation in income per capita reaching 0.12 within our sample. Thirdly, the sample is distributed across 20 communities, averaging 15 households per community. This distribution enhances the statistical power compared to having fewer clusters with a larger number of observations per cluster. Considering these factors, our sample is adept at detecting substantial impacts in the case under study. Assuming a 95% confidence level and 80% power and considering 20 clusters with 15 observations per cluster, the minimum size effect our study can detect is approximately 40%. This means that any observed impact exceeding this threshold can be reliably distinguished using our sample and current approach.

3.3. Estimation Method

In the presence of baseline and multiple post-intervention surveys, the treatment effect of a binary intervention can be estimated via the following difference-in-difference (DiD) specification:

$$Y_{i,t} = \alpha + \varphi G_i + \sum_{t=1}^{T} \delta_t + \beta TREAT_{i,t} + \varepsilon_{i,t} \qquad (1)$$

here, $Y_{i,t}$ is the outcome of interest for unit i in survey round t, where $t = 1$ corresponds to the baseline, and $t = 2, \ldots, T$ represents follow-up periods. α is a constant parameter, G_i denotes the treatment group (0 = control, 1 = treated), δ_t are time dummies, and $\varepsilon_{i,t}$ is the error term. $TREAT_{i,t}$ takes the value one if unit i received treatment by round t (i.e., for $t = 2, 3, \ldots, T$), and zero otherwise. The Average Treatment Effect (ATE) is captured by β, as it reflects the difference in outcome mean change between the treatment and control groups relative to baseline. Meanwhile, α is the baseline outcome mean of the control group, φ indicates the baseline mean difference between the groups, and δ_t captures the mean for the control group in each time period.

The DiD approach offers greater precision and statistical power compared to the post-intervention estimator by correcting for baseline imbalances in the outcome between the treatment and control groups. This design is especially well-suited for studies with small sample sizes and highly autocorrelated outcomes [38], where baseline adjustment can improve measurement and increase power.

Given the collection of three rounds of data (baseline, midline, and endline) and the administration of the treatment following baseline, we can adapt Equation (1) to capture the impacts of the program between surveys, as follows:

$$Y_{i,t} = \alpha_i + \varphi G_i + \delta_{midline} + \delta_{endine} + \beta_1 TREAT_{i,midline} + \beta_2 TREAT_{i,endline} + \varepsilon_{i,t} \qquad (2)$$

where $\delta_{midline}$ and $\delta_{endline}$ denote time dummies for midline and endline surveys, respectively, with $\delta_{baseline}$ serving as the omitted category. $TREAT_{i,midline}$ takes a value of one for midline and endline, and zero for baseline, while $TREAT_{i,endline}$ takes a value of one for endline and zero for baseline and midline. Within this framework, β_1 captures the intervention's effect between the baseline and midline surveys, and β_2 captures the effect between the midline and endline. The cumulative effect between the baseline and endline surveys is represented by the sum of both effects, $\beta_1 + \beta_2$.

Given that the community served as the unit of randomization for the intervention, we employed both robust and cluster-robust standard errors to assess inference. While cluster-robust standard errors account for heteroskedasticity and within-cluster error correlation, their performance is less reliable with a small number of clusters (below 30–50) due to asymptotic properties. In light of our study's data collection across 20 communities (clusters), we addressed this limitation by employing the Wild Cluster Bootstrap-t procedure. This method is recognized for providing asymptotic refinement and maintains effectiveness even with as few as six clusters [39]. Consequently, we assessed the significance of our findings by presenting both robust (Huber–White) and Wild Cluster Bootstrap-t p-values.

Additional approaches for estimating treatment effects in the presence of baseline and multiple post-intervention surveys include the Analysis of Covariance (ANCOVA), Covariate Adjustment, and the two-way ANCOVA [38,40–43]. ANCOVA can increase statistical power when outcomes are weakly autocorrelated or when dealing with changes in measurement between the baseline and follow-up, conditions that do not apply to our study. ANCOVA assumes equal baseline values for treatment and control groups, essentially corresponding to estimating Equation (2) with the assumption that $\varphi = 0$ (meaning the exclusion of the treatment group dummy G_i). Covariate adjustment involves controlling for baseline variables as additional regressors in Equation (2). This adjustment is particularly valuable when baseline covariates strongly predict the outcome [43]. Finally, two-way ANCOVA is akin to ANCOVA but introduces treatment × covariate interactions as supplementary regressors. While our primary method for estimating treatment effects is the DiD specification, we also examine the robustness of our findings across these three additional methods.

3.4. Data Collection and Baseline Balance

Data collection was performed in October, considering the agricultural calendar. This is the time when the campaign has just ended, so it is more likely that farmers remember precise details about their production, costs, and earnings. In the baseline survey, we collected data from 310 households, 156 of which were from treated communities and 154 controls. Due to attrition occurring exclusively in the control group, the sample size dropped to 308 and 301 in midline and endline, respectively. Attrition was primarily due to households moving out of the community, though overall attrition rates were low and did not significantly impact the statistical power of the sample.

Table 2 shows that treatment and control groups exhibit balance in terms of pre-treatment community-level and household-level characteristics. For the community-level figures (Panel A), we employed data from the most recent national census (2017) and found no statistical difference in key variables related to household characteristics and housing conditions. For household-level comparisons (Panel B), we used data from our baseline survey. We found a balance between the treatment and control groups concerning household characteristics (household size, education, and gender of the household head), income, ownership of assets (farm), and internet access. However, some minor differences were observed in housing conditions, particularly in aspects such as the type of water connection and the ownership of mobile phones and livestock barns. These imbalances, however, are modest and do not impact our findings, as demonstrated by the robustness of our findings in adjusting for these confounding factors (see Table A2 for results using DiD and Covariate Adjustment).

Table 2. Balance between treatment and control groups.

	Treated	**Controls**	**Difference**
Panel A: 2017 National Census (community-level)			
Number of Communities in the study area	10	10	0
Number of Communities in Andahuaylillas District	3	3	0
Number of Communities in Ccatca District	2	2	0
Number of Communities in Ocongate District	5	5	0
Total households	1369	1006	356
Total population	4874	3559	1315
Elevation (meters above the sea level)	3762	3815	−53
Household head speaks a native language	0.97	0.98	−0.01
Household head completed high school	0.18	0.15	0.03
Flooring is not mud of earth	0.07	0.05	−0.02

Water connection in the house	0.68	0.64	0.04
Water drain connection in the house	0.11	0.13	−0.01
Source of energy is electricity	0.76	0.77	−0.01
Energy used for cooking is electricity, charcoal, or gas	0.55	0.55	−0.01
Panel B: Baseline survey (household-level)			
Water connection in the house	0.61	0.51	0.11 *
Household size	4.8	4.5	0.2
Gender of household head is female	0.10	0.09	−0.01
Household head age	45.0	42.8	2.3
Household head years of education	5.1	5.3	−0.1
Household total income (PEN)	8456	9009	−550
Owns a mobile phone	0.91	0.83	0.08 *
With internet in the house	0.31	0.27	0.05
Owns a farm	0.96	0.96	0.00
Owns a livestock barn	0.32	0.22	0.11 **
Cash transfer recipient in the household	0.67	0.56	0.10 *

Notes. *t*-tests on the equality of means significance levels: ** $p < 0.05$, * $p < 0.1$.

3.5. Survey and Variables

The questionnaire included inquiries about household and members' characteristics, sources of income (including agricultural and livestock activities, entrepreneurs, and wage labor), migration perspectives, employability, women's empowerment, and welfare, among others. The use of objective indicators, such as income and its distribution by source, and subjective indicators, such as perceptions about empowerment, welfare, and migration, was key to achieving a precise and complete image of the household's initial situation and how it changed over time.

Our main outcome variables are related to income and its distribution among sources. Specifically, we calculated annual gross and net income, both total and per capita. We considered four major income sources: wage employee, self-employment, livestock, and agricultural. We also employed other outcome variables related to food security, migration, and employability. For further details on the outcome variables and other control variables, along with their descriptive statistics, please refer to Table A1 in Appendix A.

4. Results

We present the results in three sections. First, we discuss the program's effect on income, assessing both its magnitude and statistical significance between surveys. We also discuss the robustness of the findings across different estimation methods and income outcomes. Then, we examine the effects across various income sources, shedding light on the mechanisms influencing the observed economic impacts. Finally, we investigate effects on other dimensions, exploring households' perceptions of well-being and future expectations.

4.1. The Program's Average Effect on Income

We start by examining the Average Treatment Effect (ATE) of the program on income per capita, which serves as our primary economic outcome variable. Panel A in Table 3 provides an overview of the mean values of the outcome at baseline, midline, and endline surveys for both the treatment and comparison groups. The two groups exhibited comparable mean income values at baseline, averaging 2000 Peruvian Nuevo Sol (PEN)—roughly equivalent to 600 USD in July 2019—for the period spanning August 2018 to July 2019. The difference between the means of the treatment and control groups was a negligible 68 PEN. This indicates that both groups shared a comparable economic situation before the intervention took place (late 2019/early 2020), confirming the appropriateness of the control group as a reliable benchmark.

Table 3. Average Treatment Effect on Income per Capita.

Panel A: Average Gross Income per Capita (PEN)			
	Baseline	**Midline**	**Endline**
Control	2013.7	2441.2	3159.3
Treatment	1946.0	2184.6	4017.6
Difference	−67.7	−256.6	858.3
Panel B: Average Treatment Effect			
	Between Baseline and Midline	**Between Midline and Endline**	**Between Baseline and Endline**
Absolute effect (PEN)	−188.921	1114.89	925.973
robust *p*-value	[0.641]	[0.023] **	[0.059] *
wild cluster *p*-value	[0.457]	[0.027] **	[0.164]
Standardized effect (std. dev) [a]	−0.079	0.466	0.387
Standardized effect (mean) [b]	−0.094	0.554	0.460

Notes. N = 903 and r-squared = 0.064. ** $p < 0.05$, * $p < 0.1$. [a] The outcome was normalized for each round and group, such that the mean and standard deviation of the control group in the baseline are zero and one, respectively (i.e., we subtract the mean of the control group in the baseline and divide by the standard deviations). This measure is known as the coefficient of variation. [b] Similarly, the outcome was normalized by subtracting and then dividing the outcome by the mean of the control group in the baseline. This measure captures the size of the effect as a proportion of the control group mean in the baseline.

In the subsequent year (August 2019 to July 2020), average income experienced a moderate increase for both the treatment (+239 PEN) and control (+428 PEN) groups relative to baseline. However, the control group exhibited a slightly higher growth, with a 21% increase compared to the treated group's 12%. As a result, the gap between the means of the two groups expanded to 257 PEN in favor of the control at midline. Despite the treatment group undergoing intervention during this period, the anticipated tangible impacts of the program were not expected to materialize immediately, considering the recent implementation and the requisite time for their effects to become discernible. Indeed, the comparatively less favorable income growth observed among treated households might reflect the adjustment period associated with the adoption of the new economic initiatives (prototypes) facilitated by the program. Simultaneously, these income patterns may also reflect the differentiated impact of the economic and sanitary crisis triggered by COVID-19 in March 2020, potentially affecting treated households to a greater extent due to the market-oriented nature of the program's prototypes. In any case, while the pandemic may have contributed to a potential slowdown in local economies, the positive income growth observed among treated and control households at midline suggests that this phenomenon did not dramatically reduce incomes, at least in the very short term. These findings align with recent literature documenting that rural areas in Peru were less adversely affected by the pandemic compared to densely populated urban areas [44]. This is corroborated by the official poverty figures reported by the National Institute of Statistics and Information of Peru (INEI), indicating a more pronounced surge in urban poverty post-COVID compared to its rural counterpart (Urban poverty increased by 11.5 percentage points between 2019 and 2022 (from 14.5% to 26%), while rural poverty increased by 4.9% (from 40.8% to 45.7%)).

Two years later, spanning from August 2021 to July 2022, average income per capita sharply increased among treated households, revealing an impressive rise of 1833 PEN and marking an 84% growth between the midline and endline. In contrast, the control group witnessed a positive but notably less pronounced increase in mean income during this period, amounting to 718 PEN and representing a 29% growth. Consequently, the difference between the means of the two groups reversed, favoring the treated group by 858 PEN at endline. Overall, mean income increased by 57% in the control group and more than doubled in the treatment group (106%) between the periods before (baseline) and after (endline) the intervention.

Panel B of Table 3 reports the Average Treatment Effect of the program, captured by the difference in outcome mean change between the treatment and control groups across periods. A small and non-significant negative effect of −189 PEN is observed in the midline relative to baseline. But then, between the midline and endline, we find a large and robust impact that amounts to 1115 PEN. This effect is significant at a 95% confidence level, considering both robust and wild cluster standard errors. The magnitude of the impact is substantial, equivalent to a 0.47 standard deviation or a 55% mean increase compared to the control group's baseline income. As a result, over the period spanning baseline (pre-intervention) to endline (at least one year post-intervention), the average effect of the program on income per capita reached 926 PEN or 0.39 standard deviation. While this overall impact remains significant with robust standard errors, it becomes non-significant under wild cluster standard errors, likely attributed to the sample's limited statistical power.

These findings demonstrate robustness to alternative estimation methods and consistency across various income measures. In Table A2 in Appendix A, we present results from alternative estimation models, including ANCOVA, two-way ANCOVA, and Covariates Adjustment. The results remain comparable regardless of the chosen method and barely change when controlled by baseline characteristics (see the notes in Table A2 for further details on the covariates adjustment procedure). The findings also withstand scrutiny when subjected to alternative definitions of the income measure. Table A3 in Appendix A details the program's average effect considering both gross and net income, as well as total and per capita values. In all cases, the impact of the program is large and significant between the midline and endline and slightly weaker between the baseline and endline. The model with the better fit, captured by the higher r-squared, is the one that employs gross income per capita, our chosen primary outcome measure.

4.2. Induced Changes in Income-Generating Activities

The previous section documented a substantial impact of the program on participants' household income. In this section, we explore the potential driver of this outcome by disaggregating the impacts across different income sources. Specifically, we looked at the program's income impacts on agricultural and livestock activities, self-employment, and wage employment. This examination seeks to unravel the induced changes made by the program on different income-generating activities. The results are presented in Table 4.

Table 4. Average Treatment Effect across Income Sources.

	(1)	**(2)**	**(3)**	**(4)**
	Agriculture	**Livestock**	**Self-Employment**	**Wage Employment**
Dependent variable: Household's annual net income by source (PEN)				
ATE (between baseline and midline)	104.028	−782.307	402.802	96.782
robust *p*-value	[0.466]	[0.129]	[0.706]	[0.913]
wild cluster *p*-value	[0.509]	[0.054] *	[0.591]	[0.868]
ATE (between midline and endline)	−310.157	−549.853	2346.51	1786.61
robust *p*-value	[0.052] *	[0.368]	[0.040] *	[0.072] *
wild cluster *p*-value	[0.097] *	[0.218]	[0.058] **	[0.167]
ATE (between baseline and endline)	−206.129	−1332.160	2749.307	1883.388
robust *p*-value	[0.229]	[0.027] **	[0.034] **	[0.044] **
wild cluster *p*-value	[0.192]	[0.007] ***	[0.097] *	[0.146]

Observations	903	903	903	903
Households	301	301	301	301
R-squared	0.024	0.010	0.018	0.020
Treatment mean in baseline	−200.0	1788.0	2350.4	2508.1
Control mean in baseline	−269.3	1006.9	2811.4	3960.4

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The results highlight divergent impacts of the program on different income sources. The program significantly boosted income from self-employment ventures and wage employment within treated households, whereas earnings from farm and livestock activities experienced a modest decline. This suggests a restructuring in the income-generating activities of participating households, signaling a transition in the household income matrix towards more profitable and value-added sources. This includes engagements in businesses and wage work associated with processed or specialized agricultural and non-agricultural products while distancing from primary activities such as agriculture or livestock farming.

The income gains stemming from self-employment activities and wage labor outweigh the losses incurred from reduced agricultural and livestock earnings. To illustrate, comparing the baseline to endline, treated households witnessed a decrease in agriculture and livestock income by 206 PEN and 1332 PEN, respectively, in contrast to control households. The combined loss amounted to 1538 PEN. Conversely, self-employment income saw an increase of 2749 PEN, and wage income rose by 1883 PEN, resulting in a total gain of 4632 PEN. Consequently, the net income from these activities exhibited an overall gain of 3094 PEN among treated households.

4.3. Impact on Food Security and Opportunities within the Community

The program induced shifts in the livelihood strategies of participating households, leading to increased incomes. Nevertheless, a lingering question persists: Do these new strategies, which are less centered on primary agriculture, have any adverse effects on households' food security? Another pertinent inquiry involves exploring whether these emerging economic initiatives have energized the community's economic landscape and reduced out-migration patterns within the community. In this section, we delve into these considerations by examining the program's impact on food security and assessing perceptions regarding both local and external opportunities within the community.

In the first column of Table 5, we assess the program's impact on food security by querying each household about their food security status. We categorized those responding with either "always eat the food they want and in sufficient quantity" or "always eat enough food but not always the food they want to eat" as food secure. Conversely, those reporting "sometimes or often do not eat enough food" were classified as food insecure. The results indicate a substantial and positive impact of the program on food security. Between the baseline and endline, the percentage of treated households reporting food security increased by nearly 28% compared to control households. While this impact was more pronounced between the baseline and midline, it persisted even after the program concluded. These findings suggest that the program's positive effect on income translated into an improved food situation, notwithstanding the observed reduction in primary food production.

Table 5. Average Treatment Effects on Other Outcomes.

	(1)	(2)	(3)	(4)
	Reported Food Security Status	**Considers That There Are Good Job Opportunities in the Community**	**Want Children to Migrate Out of the Community**	**Participant Migration Probability (as Proportion)**
Dependent variable: Indicated in column heading				
ATE (between baseline and midline)	0.233	0.054	−0.104	−0.057
robust p-value	[0.002] ***	[0.420]	[0.086] *	[0.196]
wild cluster p-value	[0.003] ***	[0.438]	[0.257]	[0.046] **
ATE (between midline and endline)	0.044	0.083	0.051	−0.015
robust p-value	[0.529]	[0.236]	[0.411]	[0.769]
wild cluster p-value	[0.625]	[0.133]	[0.429]	[0.787]
ATE (between baseline and endline)	0.277	0.137	−0.053	−0.072
robust p-value	[0.000] ***	[0.050] **	[0.409]	[0.121]
wild cluster p-value	[0.000] ***	[0.054] *	[0.484]	[0.297]
Observations	903	902	898	903
Households	301	301	299	301
R-squared	0.071	0.016	0.007	0.030
Treatment mean in baseline	0.428	0.207	0.834	19.641
Control mean in baseline	0.654	0.205	0.801	24.487

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We also explore the program's influence on community members' perceptions of both internal and external opportunities. To address the former, we inquired households about their perceived economic prospects within the community. Simultaneously, we assessed the program's impact on households' migration plans to explore the latter aspect. The findings unveil a positive effect of the program on households' perception of economic opportunities within the community. Between the baseline and endline, treated households exhibited a notable 14% increase in this positive perception compared to the control group. However, the results also indicate that the program did not significantly alter households' migration expectations, with observed impacts not deviating statistically from zero. Although a slight but significant reduction in migration likelihood was noted between the baseline and midline, this effect dissipated after the project's conclusion. The program's limited impact on migration expectations could potentially be attributed to the proximity of the communities to the city of Cusco and mining opportunities in nearby localities.

5. Discussion

Returning to the question posed at the outset of how effective social innovation approaches are in advancing rural development in contexts exposed to exogenous shocks, the results presented above provide compelling evidence that in the context of entrepreneurship programming in the Quispicanchi department of Peru, households and communities who participated in the Work4Progress program prior to, during, and following the COVID-19 pandemic were more economically successful than control counterparts. Moreover, the program appears to have imparted social and well-being dividends to participating communities in the form of better satisfaction with levels of food security and improved perceived economic opportunities within the community. The research design and estimation techniques presented above are robust against a variety of estimation techniques (see Appendix A), lending confidence to the reported results and suggesting that participation in this social innovation program was causally linked to better social and economic outcomes for treatment households.

In terms of the specific impacts reported above, it appears that households participating in the social innovation program initially fared more poorly than their control counterparts between the baseline and midline but then experienced exponential growth in per capita income compared to the control group (Table 3). This could be due to a variety of factors. For example, the program logic (Figure 1) involves a long participatory engagement and co-creation design phase in which participants and supporting organizations tailor prototype ideas to the local context. In so doing, the prototype initiatives experience a necessary lag between ideation and becoming income-generating activities. It may be the case that once this tailored approach is implemented, it is readily poised to scale quickly. Of course, alternative factors may be at play, including the onset of or adverse impacts encountered during the COVID-19 pandemic. Exploration of such drivers was outside the scope of this study, but future work could examine the effect of such drivers in more detail.

Importantly, the results presented above demonstrate that the gains generated from economic initiatives generated through a social innovation approach are not uniform. When comparing agricultural and livestock initiatives with waged earning and self-employment (Table 4), the agrarian activities saw consistent declines in net income across the three observation periods, while non-agrarian activities experienced increased profitability. However, as we discussed earlier, the social innovation approach led to a restructuring of the portfolio of household income and livelihood activities for agrarian and agrarian-linked households that yielded net economic gains across the project cycle. This suggests that the treatment households substituted less profitable activities for new income-generating activities. What is clear from the results we provide is that households engaged in certain value chains and sectors that were tapped into by the social innovation program experienced increased profitability compared to non-participating counterparts. Whether and how the specific nuance of the social innovation approach of the Work4Progress program enabled those participating households to identify new markets and new value chain linkages could be further explored in qualitative studies of the program but is outside the scope of this research. Additionally, the study is further limited in that the efficacy of the Work4Progress methodology, as compared to other social innovation programs, fell outside the scope of this research but could be explored in future studies. As noted above, social innovation programs are inherently idiosyncratic, which presents challenges for direct comparison across programs. The systemic nature of the program at the center of our study further exacerbates that challenge [29], as there are no direct peers that provide direct corollaries for comparison.

The results presented above suggest that households that participated in the social innovation program were able to weather certain economic and social impacts of the pandemic more effectively than non-participating households. This is evidenced by the changes in per capita income between the midline and endline assessment and, to a slightly lesser extent, between the baseline and endline assessment. In evaluating any development intervention or program, a key issue is the sustainability of the reported results over time. In our study, we observed the reported results over two time periods, both occurring during or directly following the initial implementation of the program. Other studies have highlighted that the benefit of development interventions often experiences steep declines following the close of the program or a change in the focus of the supporting organizations. Our current study is unfortunately limited in its ability to speak empirically to the long-term viability of the program in question, as we did not collect data at later time periods, nor did we design the study to simulate and forecast future effects. Our study was further limited in its application of predominantly economic indicators to measure program success, whereas other studies rightly recognize that the economic impact of social innovation tells only one part of a more complex story of impact [28]. However, our study was designed to measure the impact of a program that is systemic in nature, focusing on building function diversity into the ecosystem for entrepreneurship such that program success is self-reinforcing whether any single prototype succeeds or fails.

Further, the program is designed to enable participating actors and organizations to identify and capitalize on emerging dynamics in that ecosystem to better reinforce existing value chain linkages and create new ones. The results we observed in this study suggest that, at least initially, the program is seeing success in creating reinforcing and complementary prototypes that yield value for participating households. Moreover, the results we observed suggest that actors in the ecosystem were better able to adapt to the exogenous shocks brought by the pandemic. While not definitive, this suggests that there is some basis to assume the underlying logic of the program enables actors to adapt to emerging dynamics in a manner suggested by the program's theory of change. In this sense, the sustainability and viability of the program will depend on the participant's ability to interact and create reinforcing feedback processes in the ecosystem rather than depending on the funding and implementing organizations to serve that function.

This study makes a compelling case that, at least in the target geography and time frame used for the present study, once the collaboratively designed income-generating initiatives became viable, household economic performance and social dividends began to accrue quickly and were set up to benefit on longer time scales compared to non-participating households. In so doing, the study makes a valuable contribution to the academic literature on social innovation by robustly demonstrating the impact of such programs in rural contexts that are exposed to exogenous shocks. In addition, our study makes important advances in designing experimental techniques to advance the academic study of social innovation programming. The highly participatory nature of social innovation programming leads to idiosyncratic and heterogenous entrepreneurship programming, as depicted in the variety of prototype initiatives included in the Work4Progress program (Table 1). This presents challenges in terms of precisely measuring the impact of such programs, as the sample sizes for any specific prototype tend to be small, and isolating the impact of a program in households with diversified income streams can be challenging. In addition to the academic contributions described above, our study makes a practical contribution to the field of social innovation by demonstrating robust analytical techniques to enhance monitoring and evaluation and impact evaluation of similar programs. Other studies have highlighted the mechanisms through which social innovation programs' focus on the complex interactions among networked entrepreneurs and actors can build resilience into development programs and systems [45]. They argue that rather than focusing on single interventions, a systemic approach is better positioned to advance the social impact of programming [46]. The results of our study, while limited in time and types of indicators included, suggest that rural development programs that take a systemic approach can enhance local communities and households to recover from exogenous shocks and may be able to enhance the enabling ecosystem for innovation and entrepreneurship.

6. Conclusions

The research we present demonstrates that social innovation programming can have a positive impact in rural geographies during times of exogenous shocks. Despite such impact, the challenges of rural employment, livelihood security, and food security remain persistent in the geography we investigated, suggesting that social innovation itself is no panacea for the challenges that rural populations face. However, in an era in which climate change and ecosystem degradation threaten the viability of rural livelihoods around the world, the results reported here offer hope that supporting rural communities to identify new market opportunities and new avenues for value chain linkage creation can provide needed resilience in the face of exogenous shocks. In the results presented above, it was the more innovative lines of income generation that proved better positioned to weather the COVID-19 pandemic compared to agrarian value chains and compared to non-participating households, suggesting that there is something important that the social innovation approach was able to leverage to support the target communities during this period. Future research that investigates the efficacy of different models of social innovation would make valuable contributions to knowledge on rural livelihood development. Additionally, future research that investigates demographic, ethnic, and socioeconomic disparities in social innovation programming would make valuable contributions in visualizing structural inequalities that may persist for rural populations and specific sectors of a society. These remain crucial questions that impact the effectiveness of rural livelihood support programs. However, the results presented above offer hope that by working with rural households to identify common challenges, unlock creative potential, and provide support as they design and invest in new livelihood strategies, we can unleash collective and collaborative action to confront new and ever-emerging challenges.

Author Contributions: The study was designed and led by co-authors M.E. and R.F. with support from co-author R.R. Field-data collection was supervised by M.E., R.R., and R.F. Statistical analysis was conducted principally by M.E. with analytical support by R.R. R.F. was responsible for funding acquisition. J.F. led the writing of the introduction and discussion sections of this manuscript, while M.E. led the writing of the methods and results sections of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for the study was provided under a contract between "la Caixa" Foundation and Grupo de Análisis para el Desarrollo.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. The Grupo de Análisis para el Desarrollo has an internal code of ethics that governs its members research, and permission for use of de-identified data was provided by "la Caixa" Foundation.

Data Availability Statement: Data for this study are not publicly available. Interested researchers should contact the authors at Grupo de Análisis para el Desarrollo for inquiries into aggregated and reproduction data.

Acknowledgments: The authors would like to acknowledge the coalition of Peruvian and international NGOs in the Work4Progress program who contributed to the implementation of the project that served as the basis of this study. We also thank Milagros Miguel Jurado for research assistance.

Conflicts of Interest: This study was funded through a contract between the "la Caixa" Foundation, who finance the Work4Progress Program, and Grupo de Análisis para el Desarrollo, who conducted an independent, external impact evaluation of this study. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results. However, we thank the "la Caixa" Foundation, which agreed to publish the results from the anonymized and aggregated data in an academic publication for the purpose of advancing knowledge and methods in impact evaluation of social innovation programs. Co-author Fisher is an advisor to the "la Caixa" Foundation on monitoring, evaluation, and learning and liaised with Grupo de Análisis para el Desarrollo on the design of the external impact evaluation.

Appendix A

The following three tables provide additional analytical materials to support the analysis provided in the main body of the paper. Table A1 provides descriptive statistics of the sample that underlie the difference in differences estimation models we present in the paper. Table A2 provides results of alternative estimation models for comparison against the models presented in the paper. Finally, Table A3 provides results from multiple response variables that measure household economic performance.

Table A1. Descriptive Statistics.

Variables	Baseline					
	Control	Treated	Diff.	Base-Line	Mid-Line	End-Line
Income						
Annual gross income per capita	1946.0	2013.7	−67.7	1981.1	2317.5	3572.8
Annual gross income	8456.0	9008.8	−552.8	8742.5	9477.6	14,097.6
Annual net income per capita	1549.0	1721.8	−172.8	1638.6	1964.6	2635.4

Net annual income	6515.8	7559.9	−1044.2	7056.9	8077.5	10,082.3
Income Sources						
Agricultural annual net income	−200.0	−269.3	69.4	−235.9	−128.2	−455.5
Livestock annual net income	1788.0	1006.9	781.1 *	1383.2	1655.7	1931.1
Self-employment annual net income	2350.4	2811.4	−461.0	2589.3	2388.0	3959.7
Wage employee annual net income	2508.1	3960.4	−1452.4 **	3260.8	4134.7	4643.4
Other Outcomes						
Food security (dummy)	0.428	0.654	−0.226 ***	0.545	0.734	0.787
Want children to migrate out of the community (dummy)	0.834	0.801	0.033	0.817	0.857	0.807
Considers that there are good job opportunities in the community	0.207	0.205	0.002	0.206	0.213	0.273
Migration probability (selected participant)	19.641	24.487	−4.846	22.153	26.066	23.425
Type of participant						
Adult women (dummy)	0.552	0.513	0.040	0.532	0.532	0.532
Youth (dummy)	0.280	0.372	−0.092	0.328	0.328	0.328
Adult men (dummy)	0.168	0.115	0.052	0.140	0.140	0.140
Covariates adjustment						
Water connection in the house (dummy)	0.614	0.506	0.107	0.558	0.558	0.558
Adequate floor in the house (dummy)	0.207	0.090	0.117 **	0.146	0.146	0.146
JUNTOS cash transfer recipient in the household (dummy)	0.669	0.564	0.105	0.615	0.615	0.615
Has livestock barn (dummy)	0.324	0.218	0.106	0.269	0.269	0.269

Notes. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$ indicate significance levels for a t-test on the equality of means for treatment and control groups at baseline.

Table A2. Alternative Estimation Models.

	(1)	(2)	(3)	(4)
	DiD	**ANCOVA**	**Two-Way ANCOVA**	**Covariates Adjustment**
Dependent variable: Household's annual gross income per capita (PEN)				
ATE (midline versus baseline)	−188.921	−256.595	−285.877	−188.921
robust p-value	[0.641]	[0.370]	[0.293]	[0.630]
wild cluster p-value	[0.457]	[0.755]	[0.624]	[0.457]
ATE (endline versus midline)	1114.893	1114.893	1339.161	1114.893
robust p-value	[0.023] **	[0.023] **	[0.006] ***	[0.018] **
wild cluster p-value	[0.027] **	[0.027] **	[0.038] **	[0.027] **
ATE (endline versus baseline)	925.973	858.298	1053.285	925.973
robust p-value	[0.059] *	[0.031] **	[0.012] **	[0.049] **
wild cluster p-value	[0.164]	[0.420]	[0.308]	[0.164]
Observations	903	903	903	903
Households	301	301	301	301
R-squared	0.064	0.376	0.378	0.131
Treatment mean in baseline	1946.008	1946.008	1946.008	1946.008
Control mean in baseline	2013.683	2013.683	2013.683	2013.683

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. In columns (3) and (4), we included additional regressors for controlling for baseline imbalances. These are the following baseline dummy variables: (i) water connection in the house, (ii) housing flooring is not earth or mud, (iii) the household is beneficiary of JUNTOS cash transfer, and (iv) the household owns a livestock barn. Column (3) employs an ANCOVA framework and introduces treatment × covariate interactions as supplementary regressors, while column (4) follows the standard DiD framework and introduces the four baseline control variables as additional regressors.

Table A3. Average Treatment Effects on Different Income Outcomes.

	(1)	(2)	(3)	(4)
	Gross Income per Capita (PEN)	**Total Gross Income (PEN)**	**Net Income per Capita (PEN)**	**Total Net Income (PEN)**
Dependent variable: Indicated in column headings				
ATE (midline versus baseline)	−188.921	−995.422	−31.305	−158.243
robust *p*-value	[0.641]	[0.595]	[0.932]	[0.915]
wild cluster *p*-value	[0.457]	[0.396]	[0.917]	[0.885]
ATE (endline versus midline)	1114.893	5018.534	883.636	3240.326
robust *p*-value	[0.023] **	[0.012] **	[0.049] **	[0.045] **
wild cluster *p*-value	[0.027] **	[0.021] **	[0.056] *	[0.120]
ATE (endline versus baseline)	925.973	4023.112	852.331	3082.084
robust *p*-value	[0.059] *	[0.074] *	[0.061] *	[0.074] *
wild cluster *p*-value	[0.164]	[0.204]	[0.149]	[0.232]
Observations	903	903	903	903
Households	301	301	301	301
R-squared	0.064	0.043	0.031	0.021
Treatment mean in baseline	1946.01	8455.99	1548.99	6515.77
Control mean in baseline	2013.68	9008.82	1721.84	7559.93

Notes: ** $p < 0.05$, * $p < 0.1$.

References

1. Bataglin, J.C.; Kruglianskas, I. Social Innovation: Field Analysis and Gaps for Future Research. *Sustainability* **2022**, *14*, 1153. https://doi.org/10.3390/su14031153.
2. Martens, K.; Wolff, A.; Hanisch, M. Understanding Social Innovation Processes in Rural Areas: Empirical Evidence from Social Enterprises in Germany. *Soc. Enterp. J.* **2020**, *17*, 220–239. https://doi.org/10.1108/sej-12-2019-0093.
3. O'Shaughnessy, M.; Christmann, G.; Richter, R. Introduction. Dynamics of social innovations in rural communities. *J. Rural. Stud.* **2023**, *99*, 187–192. https://doi.org/10.1016/j.jrurstud.2022.09.010.
4. Steiner, A.; Farmer, J.; Bosworth, G. Rural social enterprise—Evidence to date, and a research agenda. *J. Rural. Stud.* **2019**, *70*, 139–143. https://doi.org/10.1016/j.jrurstud.2019.08.008.
5. Kratzer, A.; Ammering, U. Rural Innovations in Biosphere Reserves—A Social Network Approach. *J. Rural. Stud.* **2019**, *71*, 144–155. https://doi.org/10.1016/j.jrurstud.2019.01.001.
6. Micelli, E.; Ostanel, E.; Lazzarini, L. The who, the what, and the how of social innovation in inner peripheries:A systematic literature review. *Cities* **2023**, *140*, 104454. https://doi.org/10.1016/j.cities.2023.104454.
7. Moulaert, F.; MacCallum, D.; Mehmood, A.; Hamdouch, A. (Eds.) *The International Handbook of Social Innovation Collective Action, Social Learning and Transdisciplinary Research*; Edward Elgar: Cheltenham, UK, 2013; ISBN: 978 1 78254 559 0.
8. Vercher, N. The role of actors in social innovation in rural areas. *Land* **2022**, *11*, 710. https://doi.org/10.3390/land11050710.
9. Janik, A.; Ryszko, A.; Szafraniec, M. Exploring the Social Innovation Research Field Based on a Comprehensive Bibliometric Analysis. *J. Open Innov. Technol. Mark. Complex.* **2021**, *7*, 226. https://doi.org/10.3390/joitmc7040226.
10. Moulaert, F.; Martinelli, F.; Swyngedouw, E.; Gonzalez, S. Towards alternative model(s) of local innovation. *Urban Stud.* **2005**, *42*, 1969–1990. https://doi.org/10.1080/00420980500279893.
11. Kluvankova, T.; Nijnik, M.; Spacek, M.; Sarkki, S.; Perlik, M.; Lukesch, R.; Melnykovych, M.; Valero, D.; Brnkalakova, S. Social Innovation for Sustainability Transformation and its Diverging Development Paths in Marginalised Rural Areas. *Sociol. Rural.* **2021**, *61*, 344–371. https://doi.org/10.1111/soru.12337.
12. Kluvánková, T.; Brnkaľáková, S.; Špaček, M.; Slee, B.; Nijnik, M.; Valero, D.; Miller, D.; Bryce, R.; Kozová, M.; Polman, N.; et al. Understanding social innovation for the well-being of forest-dependent communities: A preliminary theoretical framework. *For. Policy Econ.* **2018**, *97*, 163–174. https://doi.org/10.1016/j.forpol.2018.09.016.
13. Neumeier, S. Social innovation in rural development: Identifying the key factors of success. *Geogr. J.* **2016**, *183*, 34–46. https://doi.org/10.1111/geoj.12180.
14. Neumeier, S. Why do Social Innovations in Rural Development Matter and Should They be Considered More Seriously in Rural Development Research?—Proposal for a Stronger Focus on Social Innovations in Rural Development Research. *Sociol. Rural.* **2011**, *52*, 48–69. https://doi.org/10.1111/j.1467-9523.2011.00553.x.
15. Godin, B. Social Innovation: Utopias of Innovation from circa-1830 to the Present, Project on the Intellectual History of Innovation. Working Paper No. 11, Montréal, Quebec, Canada, 2012. Available online: http://www.csiic.ca/PDF/SocialInnovation_2012.pdf (accessed on 12 April 2024).

16. Guerrero-Ocampo, S.B.; Díaz-Puente, J.M. Social Network Analysis Uses and Contributions to Innovation Initiatives in Rural Areas: A Review. *Sustainability* **2023**, *15*, 14018. https://doi.org/10.3390/su151814018.
17. Adro, F.D.; Fernandes, C. Social entrepreneurship and social innovation: Looking inside the box and moving out of it. *Innov. Eur. J. Soc. Sci. Res.* **2021**, *35*, 704–730. https://doi.org/10.1080/13511610.2020.1870441.
18. Shearmur, R.G.; Carrincazeaux, C.; Doloreux, D. *Handbook on the Geographies of Innovation, Paperback Edition*; Edward Elgar Publishing: Cheltenham, UK; Northampton, MA, USA, 2018; ISBN: 978 1 78471 076 7.
19. Bock, B.B. Rural Marginalisation and the Role of Social Innovation; A Turn Towards Nexogenous Development and Rural Reconnection. *Sociol. Rural.* **2016**, *56*, 552–573. https://doi.org/10.1111/soru.12119.
20. Bock, B. Social innovation and sustainability; how to disentangle the buzzword and its application in the field of agriculture and rural development. *Stud. Agric. Econ.* **2012**, *114*, 57–63. https://doi.org/10.7896/j.1209.
21. Chen, H.-C.; Knierim, A.; Bock, B.B. The emergence of social innovation in rural revitalisation practices: A comparative case study from Taiwan. *J. Rural. Stud.* **2022**, *90*, 134–146. https://doi.org/10.1016/j.jrurstud.2022.02.003.
22. Bosworth, G.; Rizzo, F.; Marquardt, D.; Strijker, D.; Haartsen, T.; Thuesen, A.A. Identifying social innovations in European local rural development initiatives. *Innov. Eur. J. Soc. Sci. Res.* **2016**, *29*, 442–461. https://doi.org/10.1080/13511610.2016.1176555.
23. Gabriel-Campos, E.; Werner-Masters, K.; Cordova-Buiza, F.; Paucar-Caceres, A. Community eco-tourism in rural Peru: Resilience and adaptive capacities to the Covid-19 pandemic and climate change. *J. Hosp. Tour. Manag.* **2021**, *48*, 416–427. https://doi.org/10.1016/j.jhtm.2021.07.016.
24. van der Have, R.P.; Rubalcaba, L. Social innovation research: An emerging area of innovation studies? *Res. Policy* **2016**, *45*, 1923–1935. https://doi.org/10.1016/j.respol.2016.06.010.
25. Pol, E.; Ville, S. Social innovation: Buzz word or enduring term? *J. Socio-Econ.* **2009**, *38*, 878–885. https://doi.org/10.1016/j.socec.2009.02.011.
26. Mihci, H. Is measuring social innovation a mission impossible? *Innov. Eur. J. Soc. Sci. Res.* **2019**, *33*, 337–367. https://doi.org/10.1080/13511610.2019.1705149.
27. Antadze, N.; Westley, F.R. Impact Metrics for Social Innovation: Barriers or Bridges to Radical Change? *J. Soc. Entrep.* **2012**, *3*, 133–150. https://doi.org/10.1080/19420676.2012.726005.
28. Maree, M.; Mertens, S. The limits of economic value in measuring the performance of so-cial innovation. In *Social Innovation: Blurring Boundaries to Reconfigure Markets*; Nicholls, A.; Murdock, A., Eds; Palgrave Macmillan: New York, NY, USA, 2012; pp. 114–136.
29. Hervieux, C.; Voltan, A. Toward a systems approach to social impact assessment. *Soc. Enterp. J.* **2019**, *15*, 264–286. https://doi.org/10.1108/sej-09-2018-0060.
30. Salmoral, G.; Zegarra, E.; Vázquez-Rowe, I.; González, F.; del Castillo, L.; Saravia, G.R.; Graves, A.; Rey, D.; Knox, J.W. Water-related challenges in nexus governance for sustainable development: Insights from the city of Arequipa, Peru. *Sci. Total. Environ.* **2020**, *747*, 141114. https://doi.org/10.1016/j.scitotenv.2020.141114.
31. OECD. Multi-Dimensional Review of Peru Volume 3. From Analysis to Action. OECD Development Pathways. 2019. Available online: https://www.oecd-ilibrary.org/sites/c6c23d2c-en/index.html?itemId=/content/publication/c6c23d2c-en (accessed on 12 April 2024).
32. OECD. OECD Environmental Performance Reviews: Peru 2017. OECD and United Nations Eco-nomic Commission for Latin America and the Caribbean. 2017. Available online: https://www.oecd.org/countries/peru/oecd-environmental-performance-reviews-peru-9789264283138-en.htm (accessed on 12 April 2024).
33. Gałaś, A.; Haghighat-Khah, R.E.; Cuber, P.; Benavente, M.; Gorfinkiel, D.; Gałaś, S. The Impact of COVID-19 Pandemic on Halting Sustainable Development in the Colca y Volcanes de Andagua UNESCO Global Geopark in Peru—Prospects and Future. *Sustainability* **2022**, *14*, 4043. https://doi.org/10.3390/su14074043.
34. Maquera, G.; da Costa, B.B.F.; Mendoza, .; Salinas, R.A.; Haddad, A.N. Intelligent Digital Platform for Community-Based Rural Tourism—A Novel Concept Development in Peru. *Sustainability* **2022**, *14*, 7907. https://doi.org/10.3390/su14137907.
35. Vásquez-Pauca, M.A.; Vasquez, M.E.Z.; Castillo-Acobo, R.Y.; Gonzáles, J.L.A. Factors that influence the decision of Peruvian women to become entrepreneurs. *Rev. Venez. De Gerenc.* **2022**, *27*, 1036–1047. https://doi.org/10.52080/rvgluz.27.8.20.
36. Lazo-Porras, M.; Perez-Leon, S.; Cardenas, M.K.; Pesantes, M.A.; Miranda, J.J.; Suggs, L.S.; Chappuis, F.; Perel, P.; Beran, D. Lessons learned about co-creation: Developing a complex intervention in rural Peru. *Glob. Heal. Action* **2020**, *13*, 1754016. https://doi.org/10.1080/16549716.2020.1754016.
37. Petchey, O.L.; Gaston, K.J. Functional diversity: Back to basics and looking forward. *Ecol. Lett.* **2006**, *9*, 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x.
38. McKenzie, D. Beyond baseline and follow-up: The case for more T in experiments. *J. Dev. Econ.* **2012**, *99*, 210–221. https://doi.org/10.1016/j.jdeveco.2012.01.002.
39. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Bootstrap-based improvements for inference with clustered errors. *Rev. Econ. Stat.* **2008**, *90*, 414–427. https://doi.org/10.1162/rest.90.3.414.
40. Van Breukelen, G.J. ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *J. Clin. Epidemiology* **2006**, *59*, 920–925. https://doi.org/10.1016/j.jclinepi.2006.02.007.
41. Lin, W. Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *Ann. Appl. Stat.* **2013**, *7*, 295–318. https://doi.org/10.1214/12-aoas583.

42. E. Egbewale, B.; Lewis, M.; Sim, J. Bias, precision and statistical power of analysis of covariance in the analysis of randomized trials with baseline imbalance: A simulation study. *BMC Med. Res. Methodol.* **2014**, *14*, 49. https://doi.org/10.1186/1471-2288-14-49.
43. Duflo, E.; Glennerster, R.; Kremer, M. Using randomization in development economics research: A toolkit. In *Handbook of Development Economics*; Schultz, T.P., Strauss, J.A., Eds.; Elsevier: Amsterdam, The Netherlands, 2007; Volume 4, pp. 3895–3962.
44. Cozzubo, A.; Herrera, J.; Razafindrakoto, M.; Roubaud, F. *El Impacto de Políticas Diferenciadas de Cuarentena Sobre la Mortalidad Por COVID-19: El Caso de Brasil y Perú*; Documento de trabajo n º 501; PUCP: Lima, Peru, 2021; http://doi.org/10.18800/2079-8474.0501.
45. Mariann, V.S.; Krisztina, V. How Does Social Innovation Contribute to Solving Social Problems?- A process-oriented framework for measuring social innovation. *Eur. J. Soc. Sci. Educ. Res.* **2018**, *12*, 68–83. https://doi.org/10.26417/ejser.v12i1.p68-83.
46. Viana, L.F.C.; Hoffmann, V.E.; Junior, N.d.S.M. Regional resilience and innovation: Paper profiles and research agenda. *Innov. Manag. Rev.* **2023**, *20*, 119–131. https://doi.org/10.1108/inmr-11-2021-0225.