IFAD
Investing in rural people

Incorporating the Impact of Climate and Weather Variables into Impact Assessments: An Application to an IFAD Production Project in Rwanda

by
Nancy McCarthy
Josh Brubaker
Athur Mabiso
Romina Cavatassi

86 IFAD RESEARCH SERIES

The IFAD Research Series has been initiated by the Strategy and Knowledge Department in order to bring together cutting-edge thinking and research on smallholder agriculture, rural development and related themes. As a global organization with an exclusive mandate to promote rural smallholder development, IFAD seeks to present diverse viewpoints from across the development arena in order to stimulate knowledge exchange, innovation, and commitment to investing in rural people.

Authors:

Nancy McCarthy, Josh Brubaker, Athur Mabiso and Romina Cavatassi

ISBN 978-92-9266-291-2
Printed November 2022

IFAD
Investing in rural people

Incorporating the Impact of Climate and Weather Variables into Impact Assessments: An Application to an IFAD Production Project in Rwanda

by
Nancy McCarthy
Josh Brubaker
Athur Mabiso
Romina Cavatassi

86 IFAD RESEARCH SERIES

Acknowledgements

Funding for this research was provided by the International Fund for Agricultural Development (IFAD) under an initiative led by the Research and Impact Assessment division. The authors acknowledge funding from the Adaptation for Smallholder Agriculture Programme (ASAP), which is IFAD's flagship programme for channelling climate and environmental finance to smallholder farmers. The programme is incorporated into IFAD's regular investment processes and benefits from rigorous quality control and supervision systems. The authors would like to acknowledge various efforts that went into conducting the impact assessment, which this paper builds on. In particular, to people that have contributed to it: namely Alessandra Garbero, Mohamed Abouaziza, Tim Balint and Benjamin Wood. Special thanks go to the GIS specialist Gianluca Franceschini for his great contribution to the work and to Giuseppe Maggio who provided comments and editorial support. All omissions and errors are the authors' responsibility.

About the authors

Nancy McCarthy earned a PhD in Agriculture and Resource Economics from UC Berkeley in 1996, and a JD from the George Mason University School of Law in 2009. In 2010, McCarthy founded LEAD Analytics, a consulting firm that specializes in agricultural economics and legal analyses to effectively address growth and development objectives in developing countries, with a particular emphasis on agricultural development; natural resource management; governance, institutions and collective action; property rights and land tenure systems; and responses to climate change. McCarthy has significant experience implementing rigorous research and impact assessments while at the same time ensuring that policy-focused outputs meet the needs of a wide range of non-technical stakeholders. Prior to founding LEAD Analytics, McCarthy worked at the International Food Policy Research Institute (IFPRI) in Washington, DC, and held a joint position with IFPRI and the International Livestock Research Institute based in Nairobi, Kenya, for four years.

Josh Brubaker is a research and evaluation professional with nine years of experience in all phases of international development research. Research interests include climate change, smallholder agriculture, food security, poverty and impact evaluation. Skills include data management and analysis in particular, as well as research design and survey implementation. He holds a master's degree in International Development, with a focus in Development Economics from American University's School of International Service. Currently based in the greater Washington, DC area, he has spent time living in Uganda and working across various contexts in Eastern and Southern Africa. When not working, he enjoys spending time with his family, music and visiting the mountains.

Athur Mabiso is a senior technical specialist (economist) at the Research and Impact Division (RIA), IFAD. He works on impact assessments of IFAD-supported projects and conducts research on a variety of topics in development economics. Mabiso has extensive research experience in the areas of agricultural policy, rural investments, food security, health and nutrition. Prior to joining IFAD in 2018, Mabiso worked at the International Food Policy Research Institute, conducting research and engaging with policymakers and development partners in several developing countries in Africa and Asia. He holds a PhD in agricultural economics from Michigan State University, a master's in food and resource economics from the University of Florida and a bachelor's degree with honours in agriculture, from the University of Zimbabwe.

Romina Cavatassi is a lead natural resource economist in the Research and Impact Assessment Division of the Strategy and Knowledge Department of IFAD where she is in charge of leading the Impact Assessment and Research cluster in addition to leading a number of impact evaluation and research initiatives. She has extensive field experience having conducted applied research and implemented several projects in various countries.

Her expertise ranges from impact assessment to evidence-based analysis for decision and policymaking particularly in the field of climate change, natural resource economics, poverty alleviation, survey design, training, data collection, database management, data analysis and management. Prior to joining IFAD, Romina worked for FAO, where she focused on development and natural resource economics. She holds a PhD in natural resources and development economics from Wageningen University in the Netherlands, an MSc in environmental assessment and evaluation from the London School of Economics in the UK and a master's-level degree in economics from the University of Bologna, Italy.

Table of contents

Abstract

Along with companion papers (McCarthy et al., 2022a,c,c), this paper applies a methodological framework for incorporating current period weather and long-term climate conditions into impact assessments. More specifically, the framework applies to non-experimental impact assessments that rely on ex post data collected from both households that were beneficiaries of the project (treated households) and those that did not benefit (control households). Here, we apply the methodological framework to an IFAD project that aimed to increase high quality coffee and the performance of coffee cooperatives in Rwanda as a case study. Noting that currently there are no agreed climate and weather metrics to use in crop production estimations, we first explore the explanatory power of a wide range of weather and climate variables from a number of different sources. The exploratory search was delimited by variables consistent with economic theory and agronomic evidence, and by further evaluating statistical properties to ensure that variables were not significant by chance. Once we determine the best indicators of weather conditions and corresponding historical climate conditions, we include these variables in the first stage of the impact assessment, when a matching procedure is applied to treatment and control households; and then in the second stage regression, when the variables are included as regressors. Results show that there is some evidence of biased treatment impacts when climatic variables are not included, but more importantly, show that coffee producers are highly vulnerable to weather shocks. To generate more climate-change relevant evidence more rapidly, there is ample opportunity to more fully exploit impact assessment datasets than is commonly done.

1. Introduction

Increasing the resilience of rural households is a key part of IFAD's overarching goals. IFAD's third strategic objective deals explicitly with climate resilience: "Strengthen the environmental sustainability and climate resilience of poor rural people's economic activities" (IFAD, 2016). To reach that objective over the period 2016–2025, IFAD proposed to mainstream climate change throughout the entirety of its portfolio. IFAD's strategic framework document highlights the need to generate more evidence on climate risks and vulnerabilities to mainstream climate change by determining which specific activities are best suited to increase farm production while also increasing resilience to climate change.

In this paper, we look at the impacts of weather shocks and longer-term climate conditions on coffee production in Rwanda, in the context of the IFAD-funded "Project for Rural Income through Exports (PRICE)". Mabiso et al. (2018) (hereafter referred to as MAWB) present results for an ex post impact assessment of PRICE, and we build on this work by incorporating weather shocks and climate conditions into the analysis. The main objective of the PRICE project was to achieve sustainable increases in farmers' net returns by increasing participation in, and revenue from, export-driven value chains.[1] Project activities aimed at increasing the effective management of coffee cooperatives, increasing the availability of planting materials and providing training to farmers through farmer field schools. Results showed that PRICE had positive impacts on coffee production, led to higher prices received by farmers and to greater sales income.

To start the analysis, we begin by outlining the theoretical framework that will guide the choice of climatic variables, which is based on expected utility theory. Second, we review the econometric and agronomic literature that identifies weather and climate variables that have impacts on crop production. Both of these steps are important to help us delimit our exploration of specific weather and climate variables in the production analysis. In the third step, we systematically assess the predictive power of alternative sets of climate and weather variables in estimating crop yields. From this step, we arrive at a reduced set of such predictors to introduce into the impact assessment analysis.

In the fourth step, we include the climate variables into the propensity score matching procedure implemented by MAWB. In the final step, we introduce weather and climate variables into the regression analysis. We note here that it is important to include both long-term measures of climate conditions and current period shocks in the regressions, to ensure that current period weather variables are conditionally exogenous. This is particularly important for cross-section analyses that cannot employ household fixed effects.

The paper proceeds as follows. In section 2, we briefly provide details on key aspects of the PRICE project including the impact assessment design, and summarize results of the MAWB analysis. In section 3, we present the expected utility framework used to guide selection of variables for inclusion in our exploratory analysis and in section 4, we provide a review of relevant econometric and agronomic studies that identify specific climate and weather variables. In section 5, we present our systematic analysis of the impacts of climate and weather variables on coffee production and motivate the selection of two sets of variables to use in the remaining analysis. In section 6, we present results of the matching exercise and impact assessment. We conclude in section 7.

2. Background on the PRICE project

The PRICE project began operating in 2012 and covered a range of value chains. In this study, we focus on the project's second-stage intervention in the coffee value chain, namely the Turnaround

[1] Although the PRICE project focused on several crops in addition to coffee (tea, silk, horticulture), in this paper, we use the impact assessment data collected on household beneficiaries of the coffee production and marketing component. The PRICE coffee interventions occurred in two stages, and following reasons provided in MAWB, we focus on household data for those who benefited from the second stage activities, over the period 2016–2017.

Programme phase 2 (TAP2). An earlier phase, TAP1, was implemented from 2014 to 2015, while TAP2 was implemented from 2016 to 2017. The goal of the project was to strengthen the position of smallholder farmers in the coffee value chain. Activities mainly aimed at improving the performance of coffee cooperatives, increasing access to new coffee plants and training farmers in ensuring high-quality coffee production.

The project initially selected coffee cooperatives according to specific criteria in terms of their governance structure, financial profile and technical potential. The goal was to help underperforming cooperatives that nonetheless had committed management and transparent governance structures as well as technical potential. Ultimately, 25 cooperatives were selected under TAP1 and TAP2 each, for 50 in total. To construct a counterfactual, the impact assessment team used data collected to select coffee cooperatives and performed an ex ante propensity score matching exercise to identify suitable control cooperatives.

There are additional methodological issues to address when combining observations from TAP1 and TAP2 into the same sample, and MAWB performed the impact assessment separately for each wave. To illustrate the incorporation of climatic variables into the assessment, we use the TAP2 sample, which includes survey data on 2 048 households.

3. Theoretical framework

We start by assuming that rural people are risk-averse, and use the mean-variance approximation to expected utility of income, where farm households obtain income from risky agricultural production. Although simple, the model generates two key hypotheses: 1) Farmers will choose to allocate assets and inputs as a function of expected weather and weather variability, and 2) Realized outcomes will be a function of deviations between actual and expected weather realizations.

Consider the optimization problem below, where income is produced by only crop income. (We discuss extensions to off-farm labour and other types of rural project outcomes at the end of this section.) Income is equal to the value of output produced minus input costs, $Y = pf\left(X_i; Z, Cl\right) - \sum c_i X_i$. We posit a composite output function, $Q = f\left(X_i; Z, Cl\right)$, which is multiplied by a composite price, p. Outputs are a function of inputs chosen, X_i, as well as exogenous household and location characteristics, Z, and expected weather conditions, Cl. Relevant expected weather conditions will depend on the specific context, and may include expected seasonal rainfall, expected rainfall in a critical growing period (e.g. flowering period), expected onset of rainfall and expected day- and night-time temperatures.

Assuming that weather risk is modelled as multiplicative, we can write the expected utility maximization problem as follows:

$$\max EU\left(Y\right) = pf\left(X_i; Cl, Z\right) - \sum c_i X_i - \tfrac{1}{2}\left[pf\left(X_i; Cl, Z\right)\right]\phi_R \sigma_W^2 \qquad (1)$$

where ϕ_R is the coefficient of relative risk aversion and σ_W is the variance of expected weather. We can rewrite equation (1) as:

$$\max EU\left(Y\right) = pf\left(X_i; Cl, Z\right)\left[1 - \tfrac{1}{2}\phi_R \sigma_W^2\right] - \sum c_i X_i \qquad (2)$$

Maximization will lead to optimal input choices that are a function of expected weather, Cl, as well as relative risk aversion, weather variance, input costs and other relevant exogenous characteristics:

$$X_i^* = g\left(Cl, \sigma_W^2, \phi_R, c_i, Z\right) \qquad (3)$$

However, actual outputs can differ from expected outputs, depending on the actual realization of weather during the relevant cropping season. Thus, we can write actual production as follows:

$$Q_i^A = f\left(X_i^*\left(Cl, \sigma_W^2, \phi_R, c_i, Z\right), M\left(\frac{W^A}{Cl}\right)\right) \tag{4}$$

where the function $M\left(\frac{W^A}{Cl}\right)$ captures metrics of the difference between actual weather realized and expected weather. In general, input use will be higher and more assets will be allocated to agricultural production when the underlying climate conditions are more favourable – such as higher average and less variable rainfall, and lower likelihood of receiving high temperature shocks. Production will be higher in more favourable climate regimes, and when actual weather is more favourable to crop production than expected weather.

The above model results help determine the core variables on which data should be collected for use in the empirical analysis. First, we need to capture the impacts of expected weather on input choices and asset allocation. Next, we need to include some measure of weather variability. Finally, we need to include variables that capture how actual weather deviates from expected weather in the relevant production season.

Not all rural-based projects aim to increase crop production alone. In the present case study, the project invested in infrastructure and management capacity at the cooperative level, as well as working with coffee producers to enhance product quality and build resilience to weather shocks. We would thus expect that the project had a positive impact on average production and lowered downside risks to production at the household level, but also increased incomes and reduced income variance through higher prices and greater quality realized at the cooperative level. To capture these potential impacts, we can rewrite equation (2) as follows:

$$\max EU\left(Y\right) = p\left(P\right) f\left(X_i; P, Cl, Z\right)\left[1 - \tfrac{1}{2}\phi_R \sigma_Y^2\right] - \sum c_i X_i \tag{5}$$

where P represents project activities that have a direct impact on prices received at the cooperative level and on household coffee production. Instead of σ_W^2 found in equation (4) that captures only weather variance, we use σ_Y^2 , which captures both weather risks and non-weather income risks. More specifically, we specify $\sigma_Y^2 = f\left(\sigma_W^2, \sigma_{NW}^2\left(P\right)\right)$, where non-weather risks, $\sigma_{NW}^2\left(P\right)$ such as price fluctuations are themselves a function of project activities.

4. Literature review

While one certainly expects to find significant negative impacts of weather shocks on crop production – particularly under rainfed conditions – the literature documenting these impacts remained limited until the past decade or so. Until relatively recently, it was difficult to obtain rainfall station data and when possible, stations were often so sparse they provided limited information on how much rainfall a particular plot received. Researchers had to rely on self-reported rainfall shocks, which was often both coarse and noisy, especially when surveys covered wide geographic regions. Satellite-based products that produce rainfall estimates and indicators of “greenness” became more widely available, making it easier to control for climate and weather conditions on farm. Although economic theory does provide broad guidance on variables to use – expected weather, weather variability, current period deviations from expected weather – the theory does not cover which specific variables to use, and from which data sources, in agriculture production analyses. There is also relatively limited published agronomic research implemented specifically to determine the impacts of specific weather conditions for non-grain crops such as coffee. At the same time, the number of different data sources on which to construct weather shocks and climate conditions has proliferated. Below, we summarize recent literature that includes econometric analyses using satellite-based weather shocks and climate conditions from a number of different data sources. We focus on empirical results from studies in sub-Saharan Africa, and do not review studies that used either self-reported shocks or rainfall station data. This is followed by a review of agronomic studies that evaluate the impacts of weather on coffee.

Michler et al. (2019) use UC Santa Barbara's Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) daily rainfall estimate data, taking average daily rainfall aggregated to the ward level to match households at this level. Matching at ward level is necessary as household locations are not geo-referenced. In their analysis, they use a rainfall anomaly measure, defined as the difference between current period and mean rainfall divided by the standard deviation of rainfall over the total growing season. They also create two shock measures by dividing shocks between low and high rainfall anomalies. They also run robustness checks using dummy variables to capture more extreme shocks based on standard deviations. Results are consistent across the specifications, with low rainfall shocks having consistent negative impacts on crop outcomes. As the authors use a fixed effects specification, they do not include measures of long-term climate conditions.

Wineman et al. (2017) use CHIRPS rainfall estimate data matched to village centres, using panel data fixed effects that preclude the need to control for historical climate conditions. Weather shocks are defined as the number of dekads during the total rainy season with rainfall greater than 75 mm to capture high rainfall shocks, and the number of dekads during the rainy season with less than 15 mm of rainfall to capture low rainfall shocks. These thresholds are based on research done by Guerrero Compean (2013) in Mexico, and are designed to capture absolute thresholds. The authors do not report whether robustness checks using different thresholds, including relative thresholds, were performed. In general, they find that high rainfall shocks have positive impacts when significant across a range of household welfare indicators, while low rainfall shocks have negative impacts when significant.

Amare et al. (2018) examine the impacts of rainfall shocks on production and consumption outcomes in rural Nigeria using the National Oceanic and Atmospheric Administration's Climate Prediction Center (NOAA-CPC) African Rainfall Climatology version 2 (ARC2) dekadal rainfall estimate data. They use the natural log of the rainfall anomaly (defined as the difference between mean rainfall and previous period rainfall, divided by the standard deviation (SD)). Similar to Michler et al. (2019), they then create dummy variables to capture low and high rainfall shocks based on whether the rainfall anomaly was less or greater than 1 SD from the mean, respectively. They do not cite any sources to justify threshold choices nor report robustness checks to motivate those threshold choices. They find negative impacts of low rainfall shocks, but positive impacts of high rainfall shocks.

Arslan et al. (2017) use the ARC2 dekadal data, and dekadal temperature data from the European Centre for Medium-Range Weather Forecasts (ECMWF) in an analysis of the impacts of climatic variables on maize yields in Tanzania. For weather variables, the authors include the total season rainfall, a dummy variable capturing whether current period within-season rainfall variability exceeds long-term average within-season variability and a dummy variable capturing whether any temperatures exceeding 28°Celsius occurred during the season. In panel regressions not using fixed effects, the authors use either the coefficient of variation of total season rainfall or the average long-term total season rainfall shortfall covering years when rainfall is below its long-term average. The authors find that high intraseason variability and high temperature shocks both reduce maize yields by 16 per cent and 29 per cent, respectively. Arslan et al. (2015) also use the ARC2 dekadal rainfall data and the ECMWF dekadal temperature data to analyse maize yields in Zambia using a household panel dataset. The authors use total season rainfall, a dummy variable capturing late onset of the rains, average maximal daily temperatures through the season and, in correlated random effects models, the historical coefficient of variation of seasonal rainfall. The authors find a significant positive effect of total season rainfall, but also an unexpected positive effect of delayed rainfall onset on maize yields.

Asfaw et al. (2016) use ARC2 and ECMWF variables similar to Arslan et al. (2015), although not a dummy for delayed onset. They find negative impacts of high temperature shocks on maize yields, while the long-term coefficient of variation (CoV) of rainfall has positive impacts on the adoption of various sustainable land management practices. Alfani et al. (2021) also use ARC2 data to evaluate impacts of a drought on maize yields in Zambia, using a two-period panel dataset in which many households suffered from a severe drought shock in the second year, 2016. As household locations are not geo-referenced, they use ward-level averages to generate weather and climate variables. The authors run a correlated random effects model of maize yield, and include the absolute per cent deviation of total season rainfall, the CoV of total season rainfall and a dummy for a drought shock to capture non-linear impacts of large deviations from average. The drought shock takes a value of one if, during the total growing season, rainfall was below the minimum of rainfall received over the period 1983–2015. Results indicate that the drought reduced yields between 29 per cent and 41 per cent. The CoV is also negative and significant, consistent with both theoretical and empirical results that farmers in high rainfall variability environments are less likely to invest in crop productivity.

Pape and Wollburg (2019) use the United States Geological Survey's EROS Moderate Resolution Imaging Spectroradiometer (NDVI eMODIS) data to generate percentage deviation of NDVI in two critical rainy seasons from mean NDVI in three "normal" years preceding the drought (Somalia drought in the second 2016/first 2017 seasons). They also include the average deviation from mean NDVI over the period 2002–2013 to control for propensity to experience a drought. Results show generally negative impacts, although these are difficult to interpret in terms of impacts of current period shocks as the authors use the preceding period NDVI values. Mejia-Mantilla and Hill (2017) use the crop Water Requirement Satisfaction Index (WRSI) matched to households, but it is unclear what time period is covered. The authors use fixed effect models, so do not directly control for long-term probability of water stress. They find positive effects of higher water satisfaction on agricultural incomes.

Most of the above studies look at impacts of GIS-based measures of weather shocks on grain production, although some include a wider range of crop and livestock outcomes (e.g. Michler et al. (2019) and Wineman et al. (2017)). There is much less evidence on the impacts of weather shocks on coffee production (Hakorimana and Akçaöz,, 2017). In Rwanda, there is one coffee crop per year, and the flowering period tends to start in September (NAEB, n.d.). Coffee begins to flower following the onset of rains, and then transitions to the fruiting period in which the coffee cherry rapidly develops and forms a coffee bean (Ashine, 2019; UCDA, 2019). About five weeks after beans are fully formed, the cherry will change colour and be ready to harvest (UCDA, 2019). Weather extremes during the fruiting period when cherries are expanding rapidly and beans form are particularly important for determining final yields (UCDA, 2019; DaMatta et al., 2007; Cannell, 1974). Others have noted that it is important that the period preceding rainfall onset be relatively dry to stimulate flowering once rain does start (Haarer, 1958; DaMatta & Ramalho, 2006). Higher temperatures during the flowering period, and particularly in the fruiting period, also reduce coffee yields (DaMatta et al., 2007).

In addition to the household survey-based results and agronomic evidence, there are also studies that attempt to compare the performance of different rainfall products, generally by comparing the different rainfall estimates with rainfall gauge measurements (Dinku et al., 2018; Joseph et al., 2020; Logah et al., 2021). Many studies find that CHIRPS outperforms products such as ARC2, while other studies find that CHIRPS performs more similarly to the Tropical Rainfall Measuring Mission (TRMM) and the Tropical Applications of Meteorology using Satellite data and ground-based observations TAMSAT (Dinku et al., 2018; Joseph et al., 2020; Macharia et al., 2020). However, the above results base performance against rain gauge data, which by definition cannot test how well such data sources (and variables created from them) perform for households located far away from such stations. A wide range of studies also determine that dekadal data is more closely correlated with rainfall gauge data than daily data (Ouma et al., 2012; Dembélé and Zwart; 2016; Zwart et al., 2018; Coz and van de Giesen, 2020; Logah et al., 2020).

In summary, the empirical evidence suggests that households subject to extreme weather events often suffer losses in crop yields and agricultural income. However, it is difficult to compare results as the studies use a wide range of different definitions of weather shocks using different rainfall and temperature data sources. Although the literature on the topic is gaining ground – especially the types of analysis found in Michler et al. (2021) that use household data combined with multiple satellite data sources and weather variables – at present, there is not sufficient evidence to use a single data source and specific variables. Best practice would argue for collecting data from a range of sources, while choice of which specific variables to use should be guided by both agronomic evidence as well as economic theory, and using dekadal observations.

5. Analysis of impacts of weather and climate variables on coffee production

For the purposes of this paper, we will use the term "climate variables" to refer to measures of long-term averages and variability in rainfall and temperature, while we use the term "weather variables" to refer to rainfall and temperature variables that are measured during the season in which agriculture production data were collected. We use "climatic variables" to refer to both weather and climate variables.

For coffee production, MAWB used two measures, harvested coffee per hectare and coffee per tree. The two perform similarly, although explanatory power is generally greater for coffee harvested per hectare, and so we run our tests of the impacts of climatic variables on coffee production using this variable. We start our iterative exploration by running regressions on coffee yields using observations on households involved in the second phase of the TAP programme (TAP2) and control households.

5.1. Climatic data sources and climate and weather variable construction

To begin the analysis, we must first determine which climate and weather variables to use in our regressions and also determine which climate data source provides the best predictive power in our analysis. For each of our climate datasets, we create a number of variables identified in the previous literature, and systematically test which variables, constructed from which climate dataset, perform best in predicting relevant agricultural production outcomes. Results of systematically testing both variables and data sources will contribute to a sparse literature on identifying the best climate and weather variables to include in production analyses, which should help in generating comparable empirical results across studies.

For rainfall estimate data, we use the ARC2 and CHIRPS data sources. This choice to limit the analysis to these two data sources was motivated by the fact that many of the extant studies use either CHIRPS or ARC2, and so this choice facilitates comparison. Second, CHIRPS has been found to perform similarly to other products, such as TAMSAT and TRMM (now IMERG), in part because those products use similar inputs and processes for integrating gauge data versus ARC2, which uses fewer inputs and a simpler process for integrating gauge data (Dinku et al., 2018; Coz and van de Giesen, appendix B, 2020). Additionally, although fewer household-level econometric studies have used the NDVI, SPI and SPEI datasets, there is significant interest in being able to use these datasets to predict crop production outcomes and thus identify geographic locations suffering from specific weather shocks (Klisch and Atzberger, 2016; Sruthi and Aslam, 2015). NDVI is used as part of FEWSNET's Early Warning eXplorer (EWX software) and FAO's Agricultural Stress Index System (ASIS), the U.S. National Integrated Drought Information System (NIDIS) provides data on SPI as well as NDVI, FEWSNET regularly provides SPI data in its country-level food security outlooks and Vicente-Serrano et al. (2010) developed the SPEI Global Drought Monitor.

More specifically, we use NOAA's ARC2 dataset covering the period of 1983–present, and the CHIRPS dataset covering the period 1981–present. For CHIRPS, we have both daily and dekadal (10-day) observations, which we hereafter refer to as CHIRPS Day and CHIRPS Dek, respectively. For temperature, we use data from the European Centre for Medium-Range Weather

(ECMWF) ERA INTERIM reanalysis model data. The SPI and SPEI indices were created using ARC2 and CHIRPS dekadal data combined with the ECMWF temperature data. Because the SPI and SPEI created with ARC2 and with CHIRPS Dek were highly correlated, we proceeded with just SPI and SPEI created using ARC2 data. We also used two NDVI datasets, NOAA's Climate Data Record of Advanced Very High Resolution Radiometer (AVHRR) Surface Reflectance and the United States Geological Survey's EROS Moderate Resolution Imaging Spectroradiometer (hereafter referred to as NOAA-E. All GIS data were matched at the community level, as household coordinates were not available.

5.2. Climate and weather variables used in the production analysis

Although the econometric and agronomic evidence helps guide the variables to create, it does not provide us with definitive specific variables. For instance, there are different time periods over which cumulative rainfall can be defined; here, we consider three different periods. The first period is the total rainy season, which is constructed based on the onset and cessation of rainfall. Onset of rainfall for coffee production is defined as any dekad starting from the beginning of August for which at least 25 mm of rain falls during a dekad and is then followed by a dekad with at least 20 mm of rain (Tadross et al., 2009). Cessation occurs when three consecutive dekads experience less than 20 mm of rainfall after February (Tadross et al., 2009). We refer to this as the rainy period. The second is the "fruiting" period, which is defined as the eighth to fourteenth dekads following the onset of rains. The fruiting period definition captures rainfall during that critical period, but does not capture potential negative impacts of delayed onset that accrue irrespective of the amount of rainfall during a specific time after onset. The third time period covers the third dekad of November through the first dekad of January; hereafter referred to as the NJ period. This period covers the expected fruiting period, and can thus capture negative impacts of delayed onset as well as deviations from expected rainfall.

Next, we consider different measures of current period weather shocks. To facilitate comparability with extant studies, we evaluate three different deviation measures:[2]

1. The absolute per cent difference of current period weather observations from mean weather.

2. The absolute per cent difference for current period observations that are below the mean, and per cent difference for current period observations that are above the mean. This specification allows for different impacts of dry versus wet conditions.

3. Threshold dummy variables for both low and high weather differences. Relatively small deviations from expected weather may have limited impacts on production, and the per cent difference measures may mask non-linear impacts. For ARC2, CHIRPS Dek, CHIRPS Day, NDVI-A and NDVI-E per cent difference variables, we start the thresholds at 16 per cent and increase the threshold, by 2 per cent, up to the highest observed per cent difference experienced by at least 5 per cent of locations. Because SPI and SPEI are expressed in SD, we start the threshold at a .1 SD, increasing by .2 SD up to the highest threshold experienced by at least 5 per cent of locations.

To control for the preceding dry season conditions, we include a dummy variable for whether any dekad two months preceding the onset of the rainy season received rainfall greater than 25 mm. As noted in the literature review, coffee bushes are most productive when the period preceding the growing season is dry (DaMatta et al., 2007). As a fixed threshold is not provided in the literature, we use the threshold corresponding to the first-dekad minimum required for the onset of the rainy season. Even with this fairly high threshold, about 45 per cent experienced dry period rain. We did not systematically assess other dry season rainfall thresholds. With respect to

[2] Given the relatively limited geographic coverage of the project, we note that the rainfall anomalies were highly correlated with per cent differences. However, as there is no theoretical reason to prefer anomalies (and theory better supports use of per cent differences), and because anomalies can be misleading in different applications (e.g. where households are experience very different rainfall regimes), here we present results for the per cent differences only.

temperatures, while the literature suggests that high temperature shocks can have negative impacts on coffee yields, we note that seasonal temperatures covering the relevant growing season never exceeded critical values. For instance, there were no observations of noon temperatures exceeding 35°C or midnight temperatures exceeding 28°C. Average temperatures do not vary substantially across the area covered, and were never significant as would be expected. We do not report further on temperature variables in any of the analyses reported below, although we note that temperature is used in creation of the SPI and SPEI variables. It is worth stressing that the ability of researchers to uncover the impacts of different weather shocks will be a function of the weather conditions prevailing during the production season observed in the dataset. In this particular case, we cannot uncover the impact of high temperature shocks because such shocks are not observed in the data. Of course, that does not mean that high temperature shocks do not have impacts on coffee production.

We next consider variables that control for underlying climate conditions, including expected rainfall (mean rainfall) and the CoV of rainfall corresponding to the weather variables used (e.g. ARC2 mean flowering period rainfall when the weather variables were constructed using ARC2 and covering the flowering period). However, we note that the correlation between expected rainfall and the CoV ranges from moderate to very high (and negative as we expect), from −.38 to above −.95. Correlation with temperature, and the SPI and SPEI were quite high as well. This led us to create a climate index. Similar to wealth and other indices, the climate index allows us to capture a wider range of underlying climate conditions than possible when using each component as a regressor.

To create the index, we include the mean ARC2 and CHIRPS Dek rainfall covering the relevant time period and the probability of receiving moderately high levels for SPI and SPEI (values ranging between 0 and 1) over the relevant time period. Higher levels of these four variables captures favourable climatic conditions. To capture the potential for receiving damagingly low rainfall, we include the CoV for rainfall realizations below the mean using both ARC2 and CHIRPS Dek sources covering the relevant period. To capture the potential for receiving damagingly high rainfall, we include the CoV for rainfall realizations above the mean using both ARC2 and CHIRPS Dek sources covering the relevant time period. We run a principal component factor analysis using these eight variables. The scoring coefficients on the first factor for the different indices were quite similar, and lead us to interpret the index as capturing unfavourable climate conditions. In particular, ARC2 and CHIRPS mean rainfall and the SPI and SPEI dummies have negative scoring coefficients, while the low and high rainfall CoVs for both ARC2 and CHIRPS have positive coefficients.

Although we ran the production analyses using all of the data described above, we do not present results for the NDVI-A, NDVI-E or CHIRPS Day. The two NDVI measures were never significant. This may be because NDVI captures a wider range of factors affecting local “greenness” – just as crop yields are a function of many variables in addition to weather. While NDVI may pick up important changes over time, given that non-weather factors can also explain NDVI, it may prove less useful for distinguishing spatial differences, as is the case in our cross-section data analyses. With respect to CHIRPS Day, we note that a number of studies have found that dekadal data performs better in terms of matching rain gauge data than daily data (Ouma et al., 2012; Dembele and Zwart; 2016; Zwart et al., 2018; Coz and van de Giesen, 2020; Logah et al., 2020). In our case, daily data variables were either not significant, or when significant, had larger standard errors than the dekadal variables, so these results are consistent with the extant literature.

5.3. Climate and weather variables: descriptive statistics

Table 1 lists descriptive statistics for the weather variables. The second column includes the per cent of households in which the weather variables are below the mean, and the third column includes the absolute average per cent difference from the mean for households that experience below-mean rainfall. The fourth column includes the per cent of households that experienced above-mean conditions and the fifth column includes the average per cent difference from the mean for households that experienced above-mean rain.

Table 1. Descriptive statistics, weather variables.

	HH Below, %	|Diff Below|, %	HH Above, %	Diff Above, %
ARC2				
Rainy season	32	32	68	54
Fruiting period	63	14	37	9
Nov 20−Jan 10	71	14	29	18
CHIRPS Dek				
Rainy season	5	22	95	33
Fruiting period	40	15	60	10
Nov 20−Jan 10	87	10	13	5
SPI				
Rainy season	0		100	70
Fruiting period	12	24	88	49
Nov 20−Jan 10	37	17	63	29
SPEI				
Rainy season	1	13	99	69
Fruiting period	7	60	93	40
Nov 20−Jan 10	83	37	17	32

While all perform somewhat similarly across the seasons, there are still marked differences in the percentage of households falling below-mean and above-mean across the weather categories. Looking at below-mean deviations, we note that most sources show increasing incidence of below-mean weather across the rainy, fruiting and NJ seasons, respectively. Most sources show relatively modest average deviations for those locations receiving below-mean rainfall, except for the SPEI fruiting period observations. Looking at above-mean observations, all variables suggest that potentially damaging high rainfall occurred over the rainy season. Finally, most variables show moderate above-mean deviations in the fruiting and NJ period. Moderate above-mean deviations in these seasons may actually have positive impacts on coffee production.

5.4. Additional explanatory variables

In addition to the weather and climate variables, we include a number of standard household and location characteristics in our production function. Production inputs include the hectares in coffee production (in natural logs), household labour (in natural logs) and the number of coffee bushes (in natural logs). We use dummy variables for organic and inorganic fertilizer use, pesticides and hired labour, and an index of productive assets.[3] In terms of factors that may affect factor productivity, we include two GIS variables capturing topographic characteristics, elevation and slope. Household characteristics include the number of years the household has cultivated coffee (in natural logs), household heads' years of schooling, a dummy for whether the household head is female, a dummy variable for whether the household had previously received training from the coffee cooperative to which they belong and a wealth index that captures consumer durables and housing characteristics.[4] Following MAWB, we also include variables constructed from data collected from the coffee cooperatives to which farmers belong, specifically, the average tons of coffee cherries collected in logs and the coffee-washing station's capacity utilization rate.

[3] The productive asset index is created from the scores of the first principle factor from a principal components factor analysis, which includes the following productive assets: number of water cans, sickles/machetes, water pumps, hand carts/wheelbarrows, ox carts, ox ploughs, tractors, tractor ploughs, motorized pumps, mechanical dryers, solar dryers, grain mills, poultry houses, livestock enclosures, storage houses, livestock barns, pig sties and other agricultural assets.
[4] The housing index is constructed using multiple correspondence analysis on categorical variables that capture the building materials used in the dwelling walls, roof and floor, as well as a categorical variable for the type of toilet facility the household uses, the number of rooms occupied in the dwelling and a dummy for dwelling electrification.

5.5. Results of production analysis

We start by running our coffee yield regressions using the absolute per cent difference of current period rainfall weather from long-term average. Table 2 lists results for our 15 weather season categories, controlling for long-term climate with the climate index. The second and third columns of Table 2 include results for variables created using rainy season data for two specifications: the first using the absolute difference and the second using low and high absolute differences. The fourth and fifth columns include results for the two specifications in the fruit period, and the sixth and seventh columns results for the NJ period.

Table 2. Impacts of weather difference on coffee yield over three periods.

	Coffee yield (kg/ha)											
	Rainy				Fruit				NJ			
	I		II		I		II		I		II	
ARC2												
|% Diff|	−0.474				−0.261			−	0.604			
|% Diff| if < 0			−0.316				−1.667				0.01	
|% Diff| if > 0			−0.683	**			1.466				1.434	
Climate index	−0.029		−0.038		−0.105		−0.086		−0.127		−0.116	
Constant	0.383		0.368		0.34		0.342		0.165		−0.247	
# Obs.	2048		2048		2048		2048		2048		2048	
Adj. R^2	0.517		0.518		0.517		0.519		0.517		0.517	
CHIRPS Dek												
|% Diff|	−0.868				−1.108				−0.124			
|% Diff| if < 0			−0.003				−0.965				0.199	
|% Diff| if > 0			−0.697				−1.173				−3.331	
Climate index	−0.019		−0.056		−0.052		−0.047		−0.047		0.01	
Constant	0.505		0.313		0.56		0.567		0.18		0.323	
# Obs.	2048		2048		2048		2048		2048		2048	
Adj. R^2	0.517		0.517		0.517		0.517		0.517		0.517	
SPI3												
|% Diff|	1.834	***			1.197	***			0.511			
|% Diff| if < 0			0.000				1.034				−1.033	
|% Diff| if > 0			1.834	***			1.194	***			1.125	**
Climate index	−0.187		−0.187		−0.025		−0.003		−0.221		−0.097	
Constant	−0.789		−0.789		−0.908		−0.924		0.167		−0.597	
# Obs.	2048		2048		2048		2048		2048		2048	
Adj. R^2	0.522		0.522		0.521		0.521		0.517		0.519	
SPEI3												
|% Diff|	0.682	**			0.483	*			−0.271			
|% Diff| if < 0			2.014				−0.189				−0.351	
|% Diff| if > 0			0.691	**			0.744	**			0.677	
Climate index	−0.067		−0.085		0.019		0.161		−0.086		−0.114	
Constant	0.036		0.047		−0.323		−0.536		0.485		−0.184	
# Obs.	2048		2048		2048		2048		2048		2048	
Adj. R^2	0.519		0.519		0.518		0.519		0.517		0.517	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

The results in Table 2 show that the weather and climate variables have limited impacts on coffee yields. However, there is a significant negative impact of above-mean rain using the ARC2 rainy season variable. The SPI and SPEI variables perform similarly, and appear to be picking up a positive impact of above-mean rainfall on coffee yields.

As highlighted in the descriptive statistics, there is a wide range of realizations for many variables, and so we next evaluate whether we can identify thresholds below and above which rainfall deviations have significant negative impacts on grain yields. Table 3a presents a subset of results generated at different thresholds for below-mean deviations, using the climate index to control for climate conditions. Results in Table 3a are given for ARC2 fruit and NJ and for SPIE NJ. All other weather season variables either gave insignificant results across the board or did not have realizations above the minimum threshold.

The first, third and fifth columns of Table 3a list the deviation (per cent difference or SD) used in the regression, while the second, fourth and sixth columns list the estimated coefficient. The evidence suggests that there are negative impacts of rain shortfalls using the ARC2 and SPEI variables. Impacts are consistent across specifications, although the range of negative impacts is greater for the ARC2 NJ and SPEI NJ specifications.

Table 3a. Below-mean threshold effects.

% below	ARC2 fruit		% below	ARC2 NJ		SD below	SPEI NJ	
20	−0.272	*	20	−0.319	**	0.60	−0.281	*
22	−0.234	**	22	−0.243		0.65	−0.321	**
			24	−0.340	***	0.70	−0.340	
			26	−0.341	***	0.75	−0.485	**

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

Table 3b presents a subset of results generated at different above-mean per cent differences. For the most part very high rainfall realizations are observed only over the entire rainy period. Both ARC2 rainy and CHIRPS Dek rainy show a wide range of negative impacts for rainfall realizations above 60 per cent for ARC2 rainy and above 50 per cent for CHIRP Dek rainy. None of the deviation thresholds for SPI or SPEI picked up negative impacts of high rainfall shocks, either because high deviations were not observed or were not significant.

Table 3b. Above-mean threshold effects.

% above	ARC2 RS		% above	CHIRPS Dek RS	
62	−0.434	***	54	−0.487	***
66	−0.388	***	56	−0.493	***
70	−0.377	**	58	−0.493	***
74	−0.993	***	60	−0.544	***
78	−0.660	**	62	−0.643	***

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

Our sixth specification was to evaluate using both the per cent difference below-mean and above-mean with thresholds. This specification allows for both linear and non-linear impacts of weather shocks on coffee production. We ran regressions for the subset of negative and significant impacts reported in Tables 3a and 3b. The linear impacts largely remain insignificant while threshold effects remain significant, so we do not report results here.

From the above analysis, we select two sets of weather variables to use in the full impact assessment analysis. The first set includes: 1a) a dummy variable for the ARC2 NJ 24 per cent below-mean threshold, to capture negative impacts of rainfall deficits in a critical part of the growing season, 1b) a dummy variable for the ARC2 rainy season 62 per cent above-mean rainfall threshold, to capture a negative shock associated with damagingly high above-mean rainfall and 1c) a dummy variable for SPI NJ values falling in the range [.55, 1], to capture positive impacts of moderately high rainfall shocks. The second set includes: 2a) a dummy variable for the CHIRPS rainy season 60 per cent threshold, and 2b) a dummy variable for SPI NJ values falling in the range [.55, 1]. In all cases, we use the climate index to control for climate conditions.

5.6. Summary and discussion of weather and climate variable impacts

Looking first at Table 1, we note that all variables suggested that a significant portion of locations experienced high rainfall shocks during the rainy season. However, only ARC2 and SPEI suggest that some locations received very low rainfall realizations. Looking at the results in Tables 2 and 3, we draw six lessons. First, linear shocks, captured by absolute per cent differences, are generally not significant predictors of negative weather shocks. Instead, only non-linear thresholds of low and high rainfall realizations were significant predictors. About half of the studies cited in the literature review find significant linear impacts, while more than half find negative non-linear threshold effects. One major difference is that all econometric studies reviewed focused on annual grain crops, and none focused on a perennial crop such as coffee. Perennial crops are often less vulnerable to modest weather shocks than annuals (Snapp et al., 2018 and references contained therein), and our results are consistent with that empirical observation.

Second, the time period covered by the weather variables matters. For instance, we find no impacts of low weather shocks when thresholds are measured over the entire rainy season, but do find negative threshold impacts for ARC2 and SPEI over the NJ period. The opposite is true for high rainfall shocks; no high rainfall shock thresholds are significant when constructed over the fruit or NJ periods, but ARC2 and CHIRPS Dek thresholds are significantly negative over a range of threshold values covering the entire rainy season. In most of the published studies, it is unclear over what time period the weather variables were constructed.

Third, while the ARC2 and CHIRPS Dek rainy variables are significant negative predictors of high rainfall shocks, the SPI and SPEI variables did not pick up negative high rainfall shocks. Fourth, while SPI and SPEI did not pick up the damagingly high rainfall shock, they did pick up positive impacts of moderately high rainfall. Fifth, although ARC2 and CHIRPS Dek variables exhibit fairly low correlations, threshold variables from both sources were able to pick up negative impacts of high rainfall shocks.

Overall, our evidence suggests that threshold dummies performed best at capturing negative weather shocks. SPI and SPEI, on the other hand, only captured positive impacts of moderately high rainfall. Combined with results from previous empirical studies, the evidence presented here suggests that researchers still require to systematically evaluate different variables from different sources to determine which weather variables best capture weather shocks for specific crops. Presenting results from systematic evaluations would increase our understanding of the performance of climatic different variables across different crops and agro-ecological contexts.

6. Impact assessment results

In this section, we start by largely reproducing the results found in MAWB, which starts by matching households using the following variables: length of membership in a coffee cooperative; length of time as a coffee producer; number of rooms in home, time taken to collect water and whether or not had electricity, all five years ago; age of household head; household size; distance from home to cooperative and from the cooperative to the capital, Kigali; average level of schooling of household adults; and number of agricultural activities undertaken by household members. The authors then analyse the impacts of the TAP2 programme on various outcomes, although here we focus exclusively on coffee production outcomes. MAWB report results on total coffee production in kg and net coffee income. Here, we also evaluate TAP2 impacts on coffee yields (kg/ha).

We did make some changes to the original specification and thus there are slight differences between reported results for MAWB and our results. These changes facilitate comparison across our additional specifications. First, we control for standard errors clustered at the village level. We also drop observations that were missing any production data rather than imputing those values, which leads to a sample size of 2 052 versus the sample size of 2 094 used in the MAWB analysis.

Explanatory variables used in the MAWB regression analysis included: the cost to join the cooperative (in logs); total landholdings (in logs); gender, age and years of schooling, age of the household head; number of household members engaged in wage work; dummies for receipt of extension advice on five agricultural activities; a dummy equal to 1 if any household member belongs to a non-TAP2 coffee cooperative; number of shocks households' experienced in the previous year; and a dummy for whether any household member had received any training from the coffee cooperatives. They also included information collected from the coffee cooperatives to which farmers belong, specifically, the average tons of coffee cherries collected in logs and the coffee-washing station's capacity utilization rate.

As shown in Table 4, there are significant positive impacts of the TAP2 project on coffee harvest and coffee income, consistent with the estimates reported in table 1.1 in MAWB.

Table 4. Coffee outcomes, impact of TAP following MAWB specification.

	Coffee yield (kg/ha)		**Coffee harvest (kg)**		**Net coffee income**	
TAP2	0.379	*	0.611	***	0.347	***
Constant	7.851	***	3.11	***	10.374	***
# Obs.	2052		2052		2052	
Adj. R^2	0.212		0.198		0.068	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

6.1. MAWB matching and regression specification, but including weather and climate

In the next specification, we use the same matching and production variables as above but include our two sets of weather variables and the climate index. A priori, we expect that the TAP2 coefficient on coffee yields would be biased downward in the absence of controlling for climate conditions, as TAP2 households are located in less favourable climate environments. The bias on total coffee harvested is more difficult to sign, however. Under certain economic conditions, we would expect greater inputs and investments per unit land under more favourable climate conditions, while land size cultivated may be larger under less favourable conditions. From the descriptive statistics found in the Appendix, Table A2, we know that TAP2 households indeed have more land in coffee production.

From Table 5a which provides results using the first weather set variables, we note that TAP2 only remains significant on net coffee income, while the weather shock coefficients are generally insignificant. From Table 5b, we note that the TAP2 impacts regain significance and are more in line with what we expect, particularly for coffee yields. Project impacts on coffee yields in Table 5b are greater than in Table 4 and coffee harvest is lower; additionally, the climate index is positive and significant in coffee harvest, consistent with more extensive coffee production in areas with less favourable climate conditions. Of the weather variables, only the high rainfall shock has a significant negative impact and only on coffee yields.

Results suggest that simply adding the weather and climate variables as regressors may not adequately capture the impacts of these variables where they differ across treatment and control households. Additionally, while the matching and regression procedure may adequately control for confounders related to treatment, they may not adequately control for confounders that affect the weather and climate variables. For instance, the R2 are relatively low, at around .2 for production and .07 for income. While the primary goal of impact assessments is to recover project impacts, without a theoretically grounded and rich set of regressors, the impacts of other contextual variables – such as weather shocks – are likely to be biased. In many cases, the dominant focus on project impacts means that these expensive datasets are not exploited to their fullest, and miss opportunities for generating evidence on climate resilience.

Table 5a. Coffee outcomes, MAWB specification; S1 weather and climate variables.

	Coffee yield (kg/ha)		Coffee harvest (kg)		Net coffee income	
TAP2	0.357		0.404		0.341	**
Dummy, ARC2 low NJ	0.489	*	−0.046		−0.12	
Dummy, SPI high NJ	−0.091		0.299		−0.138	
Dummy, ARC2 high rainy	−0.042		0.016		0.047	
Climate index	0.04		0.218		0.117	
Constant	7.718	***	3.698	***	10.479	***
# Obs.	2048		2048		2048	
Adj. R^2	0.216		0.212		0.071	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

Table 5b. Coffee outcomes, MAWB specification; S2 weather and climate variables

	Coffee yield (kg/ha)		Coffee harvest (kg)		Net coffee income	
TAP2	0.512	**	0.462	*	0.316	**
Dummy, SPI high NJ	−0.249		0.248		−0.108	
Dummy, CHIRPS high rainy	−0.482	**	−0.245		0.082	
Climate index	0.111		0.243	*	0.11	
Constant	7.622	***	3.588	***	10.501	***
# Obs.	2048		2048		2048	
Adj. R^2	0.216		0.214		0.071	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

6.2. Updated matching and regressions to include climatic variables

In this section, we first include the climate index in the matching exercise, and then include a richer set of regressors into the regression analysis. Table 6 gives the pre- and post-match summary statistics for the climate index, first using the MAWB matching variables, and then using the MAWB variables and the climate index. When using just the MAWB matching variables, the climate index is significantly different, so that matched controls are still located in areas with more favourable climate conditions. However, when matching on the MAWB variables and the climate index, the difference between treatment and controls after the match is not significant.

Table 6. MAWB matching versus MAWB+climate index matching.

	Treated	Control	p-value	Treated	Control	p-value
MAWB matching						
Climate index	0.349	−0.203	0.000	0.349	0.019	0.000
MAWB+climate index matching						
Climate index	0.349	−0.203	0.000	0.349	0.315	0.423

It is instructive to also look at how well the component variables of the climate index are matched using the climate index in the matching. In many empirical applications, researchers will need to choose which climate variables to use in matching before knowing which climate variables will be the most important to control for specific project outcomes. Although one can always match ex post as we are doing here, being able to match on the “right” climate variables should increase the ability to choose locations to select well-matched control households. Given the weather-related descriptive statistics in Table 1, one might suspect that matching on variables covering one seasonal time period from one data source may not help to match climate variables covering different seasonal time periods and from different sources. Using the climate index, however, may improve matching on a wider range of variables. To facilitate comparison between pre- and post-matching results, in Table 7, we present results for standardized differences in the variables between treatment and controls, with unmatched differences and corresponding p-values in the second and third columns, and differences and p-values after matching on the climate index in the fourth and fifth columns. As shown in Table 7, most of the climate variables remain significantly different after matching, although most differences are smaller than pre-match differences.

Table 7. Standardized differences of climate variables, pre- and post-matching on the climate index.

	Pre-match treat-control std. dif.	Difference, p-value	Post-match treat-control std. dif.	Difference, p-value
Climate index	0.409	0.000	0.025	0.423
ARC2 mean rain NJ	−0.474	0.000	−0.075	0.007
ARC2 CoV low rain NJ	0.589	0.000	0.288	0.000
ARC2 CoV high rain RS	0.520	0.000	0.197	0.000
CHIRPS mean rain NJ	−0.124	0.001	0.208	0.000
CHIRPS CoV low rain NJ	0.494	0.000	0.204	0.000
CHIRPS CoV high rain RS	0.329	0.000	−0.004	0.889
SPI moderate high	−0.045	0.176	0.129	0.001
SPEI moderate high	0.315	0.000	0.589	0.000

Overall, our matching results suggest that while it may be difficult to determine the optimal climate variables on which to match ex ante, an index of variables may be the best – if imperfect – option, particularly if the goal is to control for exposure to damaging weather shocks.

6.3. Updated matching and updated regression specification

Results for the updated matching and regression specifications are shown in Tables 8a and 8b. The tables show that adjusted R2 are more than double those using the original regression variables, now ranging from .46 to .5 for production and .18 for net coffee income.

The TAP2 coefficients are positive and significant across coffee outcomes and weather specifications, similar to results in Table 4. While TAP2 impacts in Tables 8a and 8b are similar for coffee yields and for net coffee income, the impacts are much lower for coffee harvest. The latter suggests that omitting climatic variables leads to an upward bias in net harvest consistent with more extensive coffee production captured in Tables 5a and 5b.

With respect to the climatic variables, Tables 8a and 8b show that modest positive rainfall shocks have positive impacts on coffee yields and harvest, but no significant impact on incomes. High rainfall shocks have significant negative impacts on yields and harvest, although only the ARC2 high rainfall shock has a significant impact on net coffee income. The magnitude of the impact of these high rainfall shocks is quite high. For instance, ARC2 high rainfall shock leads to a 35 per cent reduction in coffee yields (confidence interval [30,39], a 32 per cent reduction in coffee harvested (CI [28,35]) and an 18 per cent reduction in coffee income (CI [15,22]). The CHIRPS high rainfall shock leads to a 37% reduction in coffee yields (CI[34,40]) and a 29 per cent reduction in coffee harvest (CI [26,32]). The climate index only has a significant impact on net coffee income, which is negative. The more unfavourable the climate conditions, the lower the net coffee income. The latter would be consistent with greater costs and with lower prices, and lower prices may reflect lower quality beans in such less favourable areas.

Table 8a. Coffee outcomes, updated matching and production variables; S1 weather and climate variables.

	Coffee yield (kg/ha)		Coffee harvest (kg)		Net coffee income	
TAP2	0.374	*	0.293	*	0.367	**
Dummy, ARC2 Low Rain NJ	−0.211		−0.185	*	0.022	
Dummy, High SPI NJ	0.555	***	0.573	***	−0.135	
Dummy, ARC2 High Rain RS	−0.442	***	−0.382	***	−0.216	**
Climate Index	−0.446		−0.405		−0.425	*
Constant	0.124		0.428		9.164	***
# Obs.	2 048		2 048		2 048	
Adj. R^2	0.495		0.461		0.181	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

Table 8b. Coffee outcomes, updated matching and production variables; S2 weather and climate variables

	Coffee yield (kg/ha)		Coffee harvest (kg)		Net coffee income	
TAP2	0.401	*	0.316	*	0.365	**
Dummy, high SPI NJ	0.340	**	0.404	***	−0.163	
Dummy, CHIRPS high rain RS	−0.412	**	−0.326	**	−0.050	
Climate index	−0.530		−0.475		−0.434	**
Constant	0.341		0.594		9.192	***
# Obs.	2 048		2 048		2 048	
Adj. R^2	0.499		0.464		0.180	

*Asterisks denote significance; * p<0.10, ** p<0.05, *** p<0.01.*

To summarize, results suggest that including weather and climate shocks without matching on underlying climate variables can lead to biased treatment coefficients, although in the above case, this is observed only for net coffee harvest. More importantly, the impacts of weather shocks are sensitive to the control variables included. The most robust result is the negative impact of extremely high rainfall shock on coffee production over the growing season. Positive impacts of moderately high rainfall in the critical period are also fairly robust across specifications, while the impacts of low rainfall shocks are the most sensitive to the production function specification.

The results also suggest that coffee farmers in Rwanda are vulnerable to weather shocks, with large negative impacts on yields and coffee harvested. It is also interesting to note that the impacts on total coffee harvested are less than the impacts on coffee yields, which suggests that the land extensification strategy can ameliorate some of the damage due to high rainfall shocks. Without this richer analysis, however, these important insights would be lost.

6.4. Resilience

The above regressions capture the impact of TAP2 treatment on average coffee production outcomes. To investigate whether TAP2 also reduced the likelihood of experiencing very poor production outcomes, we created a dummy variable equal to 1 if the households' coffee yields were less than or equal to 20 per cent of district-level average yields. First, we note that the marginal effects of TAP2 on coffee yields at observed values and at means are negative and significant using both of the weather sets, which means that TAP2 was effective at reducing the probability of realizing very low coffee production outcomes. To highlight the impacts of TAP2 under weather shocks, in Table 9, we show the predicted probabilities of very low yields at observed values for TAP2 and control households, and under the different rainfall shocks. The first two rows contain results for those who did and did not experience a low rain shock, respectively; while the columns indicate the predicted probabilities of low yield outcomes and respective standard errors for TAP2 and controls, respectively. The third row contains the p-value corresponding to a test of differences between those who did and did not experience a shock, while the final column contains p-values corresponding to a test of differences between TAP2 and controls.

Looking first at the low rainfall shock outcomes, we see that the difference between TAP2 and controls under a low rainfall shock is significantly lower, while the difference for those not experiencing a shock is not significant. For those facing a high rainfall shock using Weather Set 1, TAP2 households faced significantly lower probabilities than control households under both shock and no shock. If instead we compare TAP2 under low and high rainfall shocks, we note that there is no statistical difference in predicted probabilities, while control households under both low and high shocks have higher probabilities of low yields versus controls not experiencing shocks. Turning next to the high rainfall shocks from Weather Set 2, we note again that TAP2 households have better outcomes under both shocks and no shocks vis-à-vis the controls. However, TAP2 households experiencing a high rainfall shock have higher probabilities of low yields than those who do not experience a shock.

Table 9. Predicted probabilities of low yields, TAP2 and controls, under shock and no shock.

			Treated		Control		
			Prob.	SE	Prob.	SE	p-value
Weather Set 1	Low rain	Shock	0.229	0.047	0.369	0.065	0.000
		No shock	0.164	0.014	0.213	0.023	0.120
	p-value		0.234		0.036		
	High rain	Shock	0.202	0.037	0.291	0.036	0.030
		No shock	0.166	0.018	0.222	0.021	0.089
	p-value		0.463		0.074		
Weather Set 2	High rain	Shock	0.325	0.052	0.483	0.084	0.045
		No shock	0.159	0.011	0.228	0.021	0.014
	p-value		0.002		0.002		

To summarize, TAP2 households are less likely than controls to suffer low yields due to weather extremes, and results are robust across both weather sets. Using the Weather Set 1 specification, results indicate that TAP2 has an additional resilience benefit as predicted probabilities are not statistically different between those experiencing or not experiencing a high rainfall shock. However, using the Weather Set 2 specification, we do not see this additional benefit. Overall, evidence suggests that TAP2 has positive impacts on this dimension of resilience, with some evidence that TAP2-enabled households suffering shocks do as well as those that do not suffer shocks.

7. Summary and concluding comments

Guided by economic theory and agronomic evidence, we started by evaluating the performance of a wide range of weather and climate variables from a number of different climatic data sources in terms of predicting coffee production. Descriptive statistics suggest that the same variables constructed with different sources give different percentages of households estimated to have received either below-mean or above-mean rainfall. While ARC2 and CHIRPS variables suggested that some locations received damagingly high rainfall, and ARC2 variables suggested that some locations received damagingly low rainfall, the SPI and SPEI variables did not suggest either damagingly low or high rainfall. SPI and SPEI variables did suggest that some locations received moderately high rainfall that can increase production.

Looking next at impacts of weather shocks and climate conditions on coffee yields, the analysis showed that variables capturing linear impacts of weather shocks generally were not significant predictors of coffee yields. Separating per cent differences for below-mean and above-mean differences also yielded limited significant results, but suggested that the SPI in particular was picking up positive impacts of above-mean differences across all three time periods. Next, we evaluated whether identifying threshold values may better pick up non-linear impacts of shocks on yields. We found negative impacts of ARC2 in the critical growth months of November to the first dekad of January, while both ARC2- and CHIRPS-based thresholds captured negative impacts of high rainfall events accumulated through the entire growing season. Altogether, results suggest that drier than expected conditions can have significant negative effects when such conditions characterize the critical growth months. However, high rainfall shocks only have negative impacts when looking at rainfall totals through the full growing season. Overall, the results are sensitive to the data source used to construct variables, the time period covered and linear versus non-linear impacts.

We then re-ran the MAWB impact assessment analysis, employing inverse probability weighted regression analyses for four coffee production outcomes, largely reproducing the MAWB results that show positive impacts of treatment on coffee production and net coffee incomes. We then included two sets of weather and climate variables into the MAWB specifications, and found that treatment effects disappeared, while weather and climate variables also had no significant impact. The MAWB matching procedure did not include climate variables, while the descriptive statistics in Table A2 and the matching analysis in Table 6 suggest that treatment households were located in areas with less favourable climate conditions but also had a larger area of land in coffee production. Theoretically, one would expect that failure to match on the climate variables would bias the TAP2 impact on yields downward, with ambiguous impacts on total coffee harvested. When we instead matched on the climate index, the treatment variables re-gained positive and significant coefficients, although results suggest that only the TAP2 coefficient on coffee harvested was biased upward when climatic variables were not included.

Results also show that weather shocks had significant and large impacts on coffee production, and the probability of experiencing very low yields, a measure of resilience. Results show that TAP2 had significant impacts on lowering that probability vis-à-vis controls.

Overall, the analysis shows that including weather and climate variables into project impact assessments can provide information on how these variables affect the overall outcomes of interest, generating valuable insights into future project design. TAP2 beneficiaries were able to better withstand weather shocks than control households, but those shocks still had negative impacts. The latter indicates the need to continue refining the project approach (and future project designs) to handle the increasing frequency of weather shocks.

Our results also suggest that choosing what data sources to use and which exact variables to create will continue to be a difficult task confronting researchers. While economic theory and agronomic evidence can provide guidance, there remain many choices on constructing specific variables from specific sources. In this case study, the same variables created from different sources generate quite different descriptive statistics on per cent of households experiencing either below-mean or above-mean weather conditions, as well as the average size of that difference. Our analysis also suggests that non-linear thresholds did a better job of explaining impacts on coffee production rather than the simple per cent difference, either alone or split into below and above differences. But, much more work remains to be done to corroborate this result. Finally, where the variables are more consistent across data sources, impacts on coffee production are also more robust across specifications, for example the negative impacts of above-mean rainfall differences across the total rainy season. To gain a wider understanding of the predictive power of different climatic variables constructed using different data sources, researchers should report results from systematically testing the GIS-based data they collect and analyse.

References

Alfani, F., Arslan, A., McCarthy, N., Cavatassi, R. and Sitko, N. (2021). Climate resilience in rural Zambia: evaluating farmers' response to El Niño-induced drought. *Environment and Development Economics*, *26*(5-6): 582-604.

Amare, M., Jensen, N., Shiferaw, B., & Cissé, J. (2018). Rainfall shocks and agricultural productivity: Implication for rural household consumption. *Agricultural Systems, 166*, 79-89.

Arslan, A., McCarthy, N., Lipper, L., Asfaw, S., Cattaneo, A., & Kokwe, M. (2015). Climate Smart Agriculture?: Assessing the adaptation implications in Zambia. *Journal of Agricultural Economics, 66*(3), 753-780.

Arslan, A., Belotti, F. and Lipper, L. (2017). Smallholder productivity and weather shocks: Adoption and impact of widely promoted agricultural practices in Tanzania. *Food Policy*, *69*: 68-81.

Asfaw, S., McCarthy, N., Lipper, L., Arslan, A., & Cattaneo, A. (2016). What determines farmers' adaptive capacity?: Empirical evidence from Malawi. *Food Security, 8*(3), 643-664.

Ashine, T.A. 2019. Determination of Crop Water Requirement and Irrigation Water Requirement for Coffee Arabica (*Coffea Arabica* L.): Case Study of Somodo Watershed. *International Journal of Research Studies in Agricultural Sciences*, 5(6): 1–8.

Cannell, M. (1974). Factors affecting Arabica coffee bean size in Kenya. *Journal of Horticultural Science, 49*(1), 65–76.

Coz, C. and van de Giesen, N. (2020). Comparison of Rainfall Products over sub-Saharan Africa. *Journal of Hydrometeorology*. 21(4): 553-596.

DaMatta, F., & Ramalho, J. (2006). Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology, 18*(1), 55–81.

DaMatta, F., Ronchi, C., Maestri, M., & Barros, R. (2007). Ecophysiology of coffee growth and production. *Brazilian Journal of Plant Physiology, 19*(4), 485–510.

Dembélé, M. and Zwart, S.J. (2016). Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *International Journal of Remote Sensing*, *37*(17), 3995-4014.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H. and Ceccato, P., (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*. 144: 292-312.

Giertz, Á., Mudahar, M., Gray, G., Rubaiza, R., Galperin, D., & Suit, K. (2015). *Rwanda Agricultural Sector Risk Assessment.* Washington, DC: World Bank.

Guerrero Compeán, R. (2013). Weather and Welfare: Health and Agricultural Impacts of Climate Extremes, Evidence from Mexico. IDB Working Paper No. 4239. Inter-American Development Bank, Washington DC.

Hakorimana, F., and Akçaöz, H. (2017). The climate change and Rwandan coffee sector." *Turkish Journal of Agriculture-Food Science and Technology,* 5(10): 1206-1215.

Haarer, A. (1958). *Modern Coffee Production.* London: Leonard Hill.

Hatfield, J., & Prueger, J. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes, 10*, 4–10.

Hatfield, J., Boote, K., Kimball, A., Ziska, L., Izaurralde, R., Ort, D., . . . Wolfe, D. (2011). Climate impacts on agriculture: Implications for crop production. *Agronomy Journal, 103*(2), 351–370.

IFAD. (2016). IFAD Strategic Framework 2016-2025: Enabling Inclusive and Sustainable Rural Transformation. IFAD, Rome, Italy.

Joseph, J.E, Akinrotimi, O.O., Rao, K.P.C., Ramaraj, A.P., Traore, P.S.C., Sujatha, P., Whitbread, A.M. (2020). The usefulness of gridded climate data products in characterizing climate variability and assessing crop production. CCAFS Working Paper No. 322 Wageningen, the Netherlands: CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS).

Klisch, A. and Atzberger, C. (2016). Operational drought monitoring in Kenya using MODIS NDVI time series. *Remote Sensing*, 8(4): 267.

Laderach, P., Lundy, M., Jarvis, A., Ramirez, J., Portilla, E. P., Schepp, K., & Eitzinger, A. (2011). Predicted Impact of Climate Change on Coffee Supply Chains. In W. L. Filho (Ed.), *The Economic, Social and Political Elements of Climate Change* (pp. 703–723). Climate Change Management.

Logah, F.Y., Adjei, K.A., Obuobie, E., Gyamfi, C. and Odai, S.N. (2021). Evaluation and Comparison of Satellite Rainfall Products in the Black Volta Basin. *Environmental Processes*, 8(1): 119-137.

Mabiso, A., Abouaziza, M., Wood, B.D.K., and Balint, T. (2018). Project for the rural income through exports(PRICE), Rwanda: Impact Assessment Report. IFAD, Rome, Italy.

Macharia, J.M., Ngetich, F.K. and Shisanya, C.A. (2020). Comparison of satellite remote sensing derived precipitation estimates and observed data in Kenya. *Agricultural and Forest Meteorology*, 284: 107875.

McCarthy, N., Brubaker, J., Mabiso, A. and Cavatassi, R. In press, a. Incorporating the Impact of Climate and Weather Variables into Impact Assessments: An Application to an IFAD Grain Storage Project Implemented in Chad. IFAD Research Series. IFAD, Rome, Italy.

McCarthy, N., Cavatassi, R., and Mabiso, A. In press, b. Incorporating the Impact of Climate and Weather Variables into Impact Assessments: An Application to an IFAD Resilience Building Project in the Mekong Delta, Vietnam. IFAD Research Series. IFAD, Rome, Italy.

McCarthy, N., Cavatassi, R., and Mabiso, A. In press, c. Incorporating the Impact of Climate and Weather Variables into Impact Assessments: An Application to an IFAD Livestock Project in Mozambique. IFAD Research Series. IFAD, Rome, Italy.

Mejia-Mantilla, C. and Hill, R.V. (2017). With a Little Help: Shocks, Agricultural Income, and Welfare in Uganda. *World Bank Policy Research Working Paper* No. 7935. World Bank, Washington, DC.

Michler, J., Baylis, K., Arends-Kuenning, M., & Mazvimavi, K. (2019). Conservation agriculture and climate resilience. *Journal of Environmental Economics and Management, 93*, 148-169.

Michler, J.D., Josephson, A., Kilic, T. and Murray, S. (2021). Estimating the Impact of Weather on Agriculture. Policy Research Working Paper No. 9867. Washington, DC: World Bank.

NAEB (National Agricultural Export Development Board). (n.d.) Rwanda Coffee: A Second Sunrise. Available at: https://naeb.gov.rw/index.php?id=48

Ouma, Y.O., Owiti, T., Kipkorir, E., Kibiiy, J. and Tateishi, R., 2012. Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: daily, decadal and monthly evaluations. *International Journal of Remote Sensing*, *33*(24), 7662-7684.

Pape, U.J. and Wollburg, P.R., 2019. Impact of Drought on Poverty in Somalia. *World Bank Policy Research Working Paper* No. 8698. World Bank, Washington DC.

Snapp, S., Roge, P., Okori, P., Chikowo, R., Peter, B. and Messina, J. (2019). Perennial grains for Africa: Possibility or pipedream?. *Experimental Agriculture*, *55*(2): 251-272.

Sruthi, S. and Aslam, M.M. (2015). Agricultural drought analysis using the NDVI and land surface temperature data: A case study of Raichur district. *Aquatic Procedia*, *4*: 1258-1264.

Tadross, M., Suarez, P., Lotsch, A., Hachigonta, S., Mdoka, M., Unganai, L., . . . Muchinda, M. (2009). Growing-season rainfall and scenarios of future change in southeast Africa: Implications for cultivating maize. *Climate Research, 40*, 147-161.

UCDA (Uganda Coffee Development Authority). 2019. Arabica Coffee Handbook. Kampala: UCDA.

Vicente-Serrano, S.M., Van der Schrier, G., Beguería, S., Azorin-Molina, C. and Lopez-Moreno, J.I. (2015). Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *Journal of Hydrology*, 526: 42-54.

Wineman, A., Mason, N., Ochieng, J., & Kirimi, L. (2017). Weather extremes and household welfare in rural Kenya. *Food Security, 9*(2), 281-300.

Zwart, S.J., Mishra, B. and Dembélé, M. (2018). Satellite rainfall for food security on the African continent: performance and accuracy of seven rainfall products between 2001 and 2016. Abstract from Remote Sensing and Hydrology Symposium, Córdoba, Spain.

Data appendix.

Table A1. Sample size

	Treated	Control
Observations	691	1 361

Table A2. Descriptive statistics (1 of 2)

	Treated		Control	
	Mean	SD	Mean	SD
Selected dependent variables				
Coffee yield (kg/tree)	1.411	2.221	1.799	2.008
Coffee yield (kg/ha)	4 484	7 052	4 735	8 524
Coffee harvest (kg)	1 082	2 058	1 013	1 272
Coffee income	273 854	261 420	291 873	239 980
Dummy, low yield (kg/tree)	0.213	0.410	0.193	0.395
Dummy, low yield (kg/ha)	0.191	0.393	0.194	0.396
IA matching vars				
Years of co-op membership	9.848	4.914	9.814	5.520
Years cultivating coffee	27.9	14.9	29.3	14.5
# rooms occupied 5 yrs ago	3.443	1.202	3.525	1.254
Time to reach water 5 yrs ago	22.6	20.5	17.3	15.2
Dummy, electricity 5 yrs ago	0.093	0.290	0.133	0.340
HH head age	55.7	13.1	54.8	13.5
HH members	5.064	2.199	4.864	2.081
Farmer distance to co-op	4.477	3.995	3.115	3.014
Co-op distance to Kigali	65.2	27.3	71.9	38.9
Avg years education	5.875	2.506	6.089	2.376
HH members in agriculture	2.488	1.184	2.619	1.333
IA control vars				
HH fee to join co-op	19 233	33 836	21 658	41 095
Cherries processed by co-op	233.2	291.3	552.0	406.9
Co-op washing station utilization	62.6	53.1	80.9	33.5
HH total agricultural land area	8.3	31.2	20.1	49.8
Dummy, HH head female	0.200	0.400	0.198	0.399
HH head age	55.7	13.1	54.8	13.5
HH head years of schooling	4.562	3.342	4.744	3.299
# HH members wage employed	0.094	0.329	0.122	0.383
Dummy, crop advice	0.505	0.500	0.491	0.500
Dummy, ag inputs advice	0.469	0.499	0.438	0.496
Dummy, harvest advice	0.274	0.446	0.271	0.445
Dummy, post-harvest advice	0.211	0.409	0.171	0.377
Dummy, sales advice	0.013	0.113	0.032	0.175
Dummy, member of other co-op	1.818	0.386	1.936	0.245

Table A2. Descriptive statistics (2 of 2)

	Treated		Control	
	Mean	SD	Mean	SD
IA control vars (continued)				
# HH shocks experienced	1.349	1.471	1.159	1.368
# HH income sources	1.654	0.810	1.855	0.930
TAP_OTHER	0.208	0.406	0.159	0.366
Dummy, training from co-op in 5	0.537	0.499	0.494	0.500
Climate Set 1				
Mean rainfall NJ	141	23	160	34
CoV rainfall NJ	0.377	0.042	0.362	0.038
Climate Set 2				
Climate index†	0.349	0.868	−0.203	1.033
Weather Set 1				
Dummy, ARC2 low rain NJ	0.171	0.377	0.076	0.265
Dummy, High SPI NJ	0.245	0.430	0.016	0.126
Dummy, ARC2 high rain RS	0.398	0.490	0.137	0.344
Dummy, ARC2 dry pd rain	0.304	0.460	0.517	0.500
Weather Set 2				
Dummy, high SPI NJ	0.245	0.430	0.016	0.126
Dummy, CHIRPS high rain RS†	0.142	0.349	0.052	0.223
Dummy, ARC2 dry pd rain	0.304	0.460	0.517	0.500
New production function				
Elevation (km)	1.705	0.179	1.722	0.177
Slope	12.1	8.4	15.0	7.9
Average age of coffee trees	23.8	15.1	24.1	13.9
Any labour hired on coffee plot	0.605	0.489	0.620	0.486
Hectares under coffee	0.965	3.934	5.422	52.113
# coffee trees owned	1034	1231	683	716
Value of labour	19000	40881	29311	79092
Dummy, organic fertilizer	0.048	0.213	0.050	0.218
Dummy, inorganic fertilizer	0.151	0.358	0.112	0.316
Dummy, pesticide	0.130	0.337	0.100	0.300
Years cultivating coffee	27.909	14.871	29.345	14.533
Dummy, HH head female	0.200	0.400	0.198	0.399
HH head years of schooling	4.562	3.342	4.744	3.299
Housing index	0.361	0.179	0.379	0.180
Productive assets index	0.986	0.782	1.171	0.936
Dummy, training from co-op in 5	0.537	0.499	0.494	0.500
HH fee to join co-op	19233	33836	21658	41095
Cherries processed by co-op	233.2	291.3	552.0	406.9
Co-op washing station utilization	62.6	53.1	80.9	33.5
Distance to road (km)	4.173	3.346	3.816	2.850

† *indicates four missing observations due to limitations of the climate dataset.*

International Fund for Agricultural Development
Via Paolo di Dono, 44 - 00142 Rome, Italy
Tel: +39 06 54591 - Fax: +39 06 5043463
Email: ifad@ifad.org
www.ifad.org

facebook.com/ifad
instagram.com/ifadnews
linkedin.com/company/ifad
twitter.com/ifad
youtube.com/user/ifadTV

ISBN 978-92-9266-291-2