SCUOLA DI DOTTORATO

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Economics, Management and Statistics

PhD program in Economics, Statistics and Data Science - Cycle XXXVII

Curriculum in Economics

Essays in Development Economics

Matteo Ferraro

Registration number 883957

Supervisor: Prof. Andrea Guariso

Co-supervisor: Prof. Selene Ghisolfi

Coordinator: Prof. Matteo Manera

ACADEMIC YEAR 2024/2025

Introduction

This PhD thesis investigates two central problems for households living in rural areas of Uganda. The first chapter examines how short-run agricultural income shocks affect household fertility choices; the second one studies the effects of asymmetrical information in the coffee market, presenting a technology-driven solution to improve the position of smallholder producers in the coffee value chain.

Chapter 1 develops an index that interacts global cash-crop price variation with local agro-climatic suitability to generate plausibly exogenous income shocks at household level. Using birth-history and panel data of Ugandan households (DHS/UNPS), the analysis shows that positive cash-crop shocks increase both the preferred waiting time for the next birth and realized birth intervals. Results from survey data and a complementary vignette experiment run among Ugandan coffee farmers support the hypothesis that this effect is driven by the higher marginal cost of time devoted to childbearing.

Chapter 2, written with professors Tessa Bold, Jakob Svensson and Selene Ghisolfi, documents a market-for-lemons problem in the Ugandan coffee market. Drawing on a detailed field survey collecting sales-level transaction data in Ibanda (Uganda), the chapter shows that the information asymmetry in the value chain (driven by the loss of observable signals of quality during processing) reduces farmers' return to quality in the coffee beans market. The chapter then presents a market-based solution to this problem, based on a smartphone app that produces bag-level quality metrics at the point of sale.

Contents

Chapter 1

Cash crops and birth spacing in rural Uganda

Introduction

The relationship between income and fertility has been widely studied in economics, starting from the seminal work of Becker (Becker 1960). Data consistently show that richer households in developed countries tend to have fewer children (Doepke, Hannusch, et al. 2023) and smaller family size is often associated to better development outcomes for both mothers and their children, at least in developing countries (Bailey, Hatton, and Inwood 2016; Baranowska-Rataj, Barclay, and Kolk 2017; Öberg 2015; Ponczek and Souza 2012). However, not only the number of births matters for maternal and child health: birth spacing plays a relevant role, as empirical evidence shows that markedly short birth intervals are associated with adverse outcomes (Conde-Agudelo, Rosas-Bermudez, and Kafury-Goeta 2012; Swaminathan et al. 2020). To mitigate health risks to children and mothers, the World Health Organization thus recommends birth-to-pregnancy intervals of at least two years, which corresponds approximately to three years between births (WHO 2007).

Despite these clear health benefits, achieving optimal birth spacing remains a challenge, particularly in low-income settings where economic conditions profoundly influence household reproductive decisions (Arega et al. 2024; Shifti et al. 2020). Developing countries are also the ones where adverse outcomes of short birth intervals are stronger (Beyene et al. 2025; Islam et al. 2022; R. Miller and Karra 2020), increasing probability of infant mortality and malnutrition. Molitoris, Barclay, and Kolk (2019), for example, estimate that increasing birth-to-birth intervals from 12 to 24 months is associated with at least 30% reduction in infant mortality risks in developing countries.

This paper investigates how short-term income variations influence both desired and actual birth spacing among rural Ugandan households. Understanding how farmers' decisions regarding birth spacing respond to these temporary income shifts is crucial to design effective public health and development policies. On one hand, a positive income shock could lead households to have their next child sooner (an *income effect*), as children are often treated as normal goods and greater resources can ease the financial burden of child-rearing. On the other hand, if this shock is due to higher returns of farm labour, also the opportunity cost of time devoted to pregnancy increases. This would incentivize households, particularly women, to reallocate time toward agricultural work and postpone the next birth to capitalize on the favourable economic conditions (a *substitution effect*). The observed net effect on birth spacing would thus depend on which of these two forces dominates.

To explore this relationship, I employ a quasi-experimental design. I build a proxy for exogenous variation in cash crop income by interacting plausibly exogenous year-to-year fluctuations in international commodity prices (the "shift") with pre-determined local agro-climatic suitability for specific cash crops (the "share"). This shift-share strategy allows me to estimate the causal impact of income shocks on fertility timing, overcoming the endogeneity inherent in directly regressing income on reproductive choices.

The analysis proceeds in three steps. First, I validate the income proxy using panel data from the Uganda National Panel Survey (UNPS), showing that a positive shock in the price index leads to a significant increase in the value of crop sales and in the number of days worked on family plots. Second, using data from the Demographic and Health Surveys (DHS), I examine the effect of the income proxy on fertility *preferences*. I find that while it does not affect the desire to have another child, it significantly increases the preferred waiting time until the next birth for both husbands and wives. Third, I analyze fertility *outcomes* using DHS birth histories: the results show that a positive income shock in the year of a birth increases the interval to the subsequent one. Finally, evidence from a complementary survey experiment reveals that a pure income effect would decrease birth spacing, supporting the interpretation that main results are driven by a dominating substitution effect.

This paper contributes to the literature on the determinants of birth spacing and family planning behaviour. Household economic status consistently emerges as one of the strongest correlates of short birth intervals both in low- and middle-income countries (Pimentel et al. 2020; Schultz 1997) and high income ones (Balbo, Billari, and Mills 2013). This paper adds causal evidence of a relationship between income and birth spacing, moving beyond the mainly correlational results of epidemiological studies and documenting a mechanism through which income volatility directly impacts health outcomes, as short birth intervals are a primary driver of maternal and infant mortality in Sub-Saharan Africa. A growing body of work interprets fertility and family planning as part of households' response to income risk and economic shocks. Kosec and Song (2021), for example, use panel data from Kyrgyzstan to show that households facing negative income shocks increase contraceptive use, effectively using family planning as a buffer against economic uncertainty; similar patterns are documented in low income countries and among agricultural households (Adsera and Menendez 2011; Chatterjee and Vogl 2018). Since farm incomes are highly sensitive to climate variability and seasonal cycles, fertility decisions often display seasonal patterns and respond to agricultural shocks (Dorélien 2016). Climate induced disruptions and droughts (Corno, Hildebrandt, and Voena 2020; Jung 2023) or accidental crop losses (Alam and Pörtner 2018) can alter family planning intentions and may push couples to shift the timing of births, with effects that depend on the degree of household reliance on farming and on local rural conditions. The responsiveness of birth spacing to economic conditions is reinforced by historical evidence (Boberg-Fazlić et al. 2021; Cinnirella, Klemp, and Weisdorf 2017), where in some cases shocks can impact couples behaviour also beyond the short-term, leading to overall fertility decline (Kitchens and Rodgers 2023). With respect to this literature, this paper focuses on more ordinary short-term variation in agricultural income rather than on large adverse events such as droughts or crop failures, providing evidence on how reproductive decisions are shaped by economic factors even in the absence of severe subsistence constraints.

In the second place, the analysis is also related to economic models in which fertility responds to both income and substitution effects, where children are modeled as normal goods but childrearing requires parental time that could instead be supplied to the labour market (Becker 1960; Becker and Barro 1988; Becker and Lewis 1973; Doepke, Hazan, and Maoz 2015; Willis 1973). In these frameworks higher income relaxes the budget constraint and can encourage earlier childbearing, while higher wages or returns to time intensive work increase the opportunity cost of childbearing and can induce postponement. By exploiting plausibly exogenous variation in cash crop prices that raises the return to farm labour, this paper provides empirical evidence on this substitution channel also among Ugandan farmers, validating the relevance of these household decision-making models for the developing world.

This paper is organized as follows: section 1 introduces the data and the identification strategy, section 2 presents the results and section 3 explains setting and evidence coming from the survey.

1 Identification strategy and data

1.1 Data

To empirically test the relationship between income shocks and birth spacing, this study relies on a combination of agricultural, administrative, and survey data. I employ data on the suitability for cultivation of cash crops and their international market prices to construct a proxy for household income variation.Household survey data are used to validate this proxy and to measure fertility preferences and outcomes. Finally, I complement these observational data with a primary survey experiment designed to isolate the income effect mechanism. Table 1 provides a summary of all data sources used.

Table 1. Summary of Data Sources and Variables

Dataset	**Description & Sample**	**Key Variables**	**Use in Analysis**
FAO GAEZ	Spatial raster data (0.5° × 0.5°) of agro-climatic land suitability (1961-1990 baseline).	Crop suitability index (0-100) for coffee, cotton, tea, tobacco.	Construction of the shift-share instrument ("share").
World Bank & FAOSTAT	Annual international commodity price and production time series (1995-2015).	Real international prices (USD); export values.	Construction of the shift-share instrument ("shift").
UNPS (2009-2011)	Balanced panel of 1051 agricultural households harmonized by FAO RULIS.	Value of crops sold; agricultural and total income; days worked on own plots.	Validation of the income proxy.
DHS (2006, 2011, 2016)	Pooled cross-section of 1301 agricultural couples and births panel.	Desire for children; preferred waiting time; birth intervals (years); controls.	Analysis of fertility preferences and actual birth spacing outcomes.
Survey (2025)	Cross-sectional survey of 626 coffee-growing households in Ibanda district.	Responses to income vignettes; knowledge of spacing.	Testing the presence of an income effect.

1.1.1 Agricultural and price data

To characterize the local potential for cash crop production, I utilize the agro-climatic suitability index from the FAO Global Agro-Ecological Zones (GAEZ) database. This index expresses how suitable every 0.5° × 0.5° raster cell (approximately 55 km × 55 km at the equator, see Figure 1 for an example) is for the cultivation of specific crops, with values ranging from 0 (not suitable) to 100 (highly suitable). The index models crop suitability as a function of agro-edaphic variables (soil nutrient availability, retention capacity, rooting conditions and oxygen availability) using their averages over long time periods (1961-1990 in this case)[1].

[1]Since different levels of inputs (irrigation and fertilizers) may differentially affect the suitability of soils with different characteristics, I consider the index calculated by the FAO in the "intermediate input, rain-fed" scenario. The data can be obtained from `http://gaez.fao.org/`.

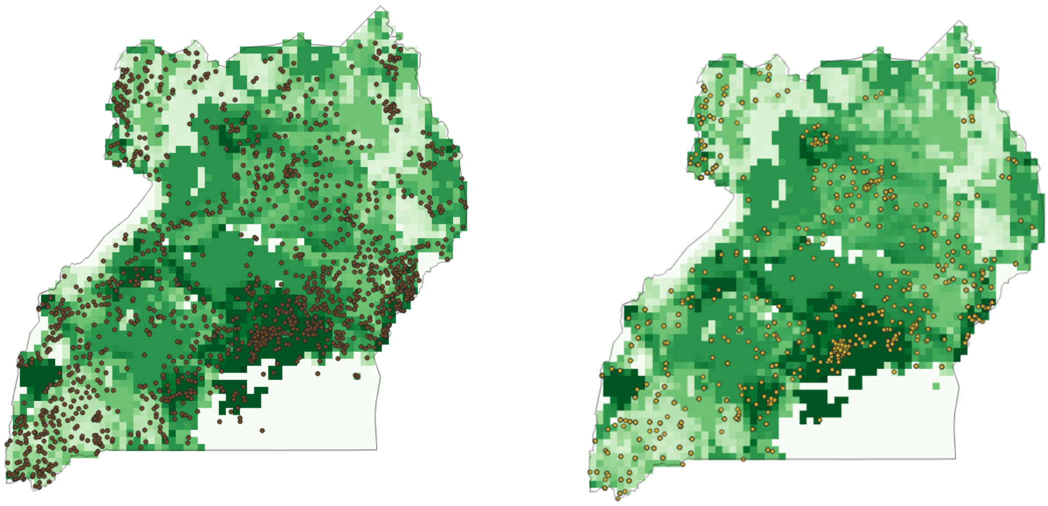


Figure 1. An example of suitability for coffee. The white cells have no data, while those with higher suitability for coffee cultivation are represented with a darker green. Dots represent DHS clusters (left) or UNPS enumeration areas (right). Source: FAO GAEZ v4.

These spatial data are combined with time-series economic data. To identify the relevant cash crops, I rely on production and export values from the FAOSTAT database, specifically examining the composition of Ugandan agricultural exports from 1995 to 2005[2]. Based on these data, the analysis focuses on the four primary cash crops: coffee, cotton, tea, and tobacco. The "shift" component of the income proxy is derived from the annual averages of international real prices for these crops, retrieved from the World Bank Commodity Markets Pink Sheet[3].

1.1.2 Survey Data

Uganda National Panel Survey To validate that the constructed price index acts as a genuine proxy for agricultural income, I employ the Uganda National Panel Survey (UNPS). Specifically, I utilize the agricultural modules from the 2009, 2010, and 2011 waves, which have been harmonized by the FAO Rural Livelihoods Information System (RULIS)[4]. The analysis focuses on a balanced panel of 1051 households ($N = 3153$ observations) that reported agricultural activities across all three waves. Key variables elaborated from this dataset include the total value of crops sold in the last 12 months, household income derived specifically from agricultural activities (including crop sales, wages, and livestock), and total household income including transfers and non-farm wages. To assess labour responses, I also utilize data on the number of household members working on family plots and the total days of labour supplied by household members.

[2] Data retrieved from `https://www.fao.org/faostat/en/#data/TI`.

[3] Data retrieved from `https://www.worldbank.org/en/research/commodity-markets`.

[4] Data retrieved from `https://microdata.worldbank.org/index.php/catalog/lsms/` and `https://www.fao.org/in-action/rural-livelihoods-dataset-rulis/en/`.

Demographic and Health Surveys (DHS) The core analysis of fertility behaviour relies on the Uganda Demographic and Health Surveys (DHS). I pool data from the 2006, 2011, and 2016 waves to maximize statistical power. To ensure representativeness across the pooled sample, individual weights were normalized within each survey year and re-scaled so that each wave contributes equally to the analysis. The data is utilized in three distinct forms. First, to analyze stated fertility preferences, I construct a cross-sectional sample of 1301 couples ($N = 1301$) where both the husband and wife are self-employed in agriculture. For these couples, I extract variables on the desire for an additional child and the preferred waiting time for the next birth. Second, to analyze actual birth spacing intervals of these couples, I construct a panel of realized birth-to-birth intervals ($N = 5661$). Third, to analyze the annual probability of birth, I reshape these birth histories into a retrospective woman-year panel ($N = 17680$) covering the period 1995-2015. To strictly measure birth spacing rather than the onset of fertility, this panel is left-censored to start the year following each woman's first birth. Control variables extracted for the analysis include respondent age, education level, religion, and district of residence.

Survey Experiment To complement the observational analysis and isolate the pure income effect, I utilize primary data collected through a field survey in the Ibanda district of Uganda during July 2025. The sample consists of one respondent for each of the 626 interviewed households, who are actively cultivating Arabica coffee. This dataset includes responses to a randomized vignette experiment regarding fertility timing under different coffee price scenarios. The key treatment variables are dummy indicators for the *High Income* and *Low Income* vignettes (relative to a *Normal Income* baseline).

1.2 Identification strategy

This section discusses the identification strategy: the choice of the relevant cash crops, the construction of the proxy and the implications of the recent literature on shift-share instruments[5] for the robustness, validity and interpretability of the proxy with respect to this setting.

1.2.1 Which cash crops?

The identification strategy relies on the effect of international prices of cash crops to obtain a proxy of yearly agricultural income variation for households. As reported by

[5] Among the others: Adao, Kolesár, and Morales (2019), Borusyak, Hull, and Jaravel (2022) and Goldsmith-Pinkham, Sorkin, and Swift (2020); I followed the applied checklist provided by Borusyak, Hull, and Jaravel (2025).

Byrnes (1992) the traditional cash crops in Uganda are coffee[6], tea, cotton and tobacco. The role of these crops is confirmed by FAO data (FAO 2025). In the period of analysis (1995-2015 for the birth history analysis of the DHS sample and 2005, 2006 and 2011 for the questions about fertility preferences) each of these four crops represent at least 1% of the agricultural exports (median share between 1995 and 2015: 57% for coffee, 5% for tea, 2% for cotton and tobacco) and the prices along the value chain is plausibly influenced by international prices since a relevant share of the production is then exported (median share between 1995 and 2015: 92% for coffee, 75% for tobacco, 20% cotton and 13% for tea). Note also that Uganda is not one of the major producers of any of these crops, so that it is unlikely that international price changes is driven by Ugandan supply variation[7].

1.2.2 An index measuring the suitability-weighted exposure to the international prices

To capture the effect of short-term income variation on desired fertility, I build a shift-share proxy for local income variation. The core idea (borrowed from Berman et al. (2023)) is that international prices for specific crops (*shift*) unequally affects the crop-related income of households living in different areas of Uganda because of their different relative suitability for the cultivations of those crops (*shares*). In other words, farmers living in more coffee-suitable land will probably cultivate coffee and therefore be affected by coffee prices than farmers living in less coffee-suitable land.

For each household, based on its location, I retrieve the suitability for the cultivation of cotton, tea, coffee and tobacco and compute its suitability-weighted exposure to the international price of cash crop $i \in \{\text{coffee, tobacco, cotton, tea}\}$ in year t as

$$\text{price\_index}_t = \frac{\text{suitability}_i}{\sum\limits_{j \in \{\text{cash crops}\}} \text{suitability}_j} * \mathrm{P}_{i,t} \tag{1.1}$$

and define the *yearly variation* of this index by plugging $\Delta \mathrm{P}_{i,t} = log(\mathrm{P}_{i,t}) - log(\mathrm{P}_{i,t-1})$ in the formula (while suitability remains constant), since $\Delta \mathrm{P}_{i,t}$ will be percent price variation of crop i between years $t-1$ and t in the international markets. Figure 2 depicts some predicted changes based on different price variations.

[6] Arabica coffee has been by far the most important internationally-traded cash crop in Uganda, while recently also robusta coffee has increased its market share (USDA Foreign Agriculture Service 2025). The analysis in the following subsection uses the suitability for the cultivation of arabica coffee and its international prices variation; being highly correlated, almost nothing would change if we used a weighted average considering also the suitability for cultivation of robusta and its international prices variations.

[7] The production of most important agricultural product coffee, has never exceeded the 10% of world production (International Coffee Organization, `http://dev.ico.org/documents/review8e.pdf` and `https://www.ico.org/documents/cy2017-18/cmr-1017-e.pdf`)

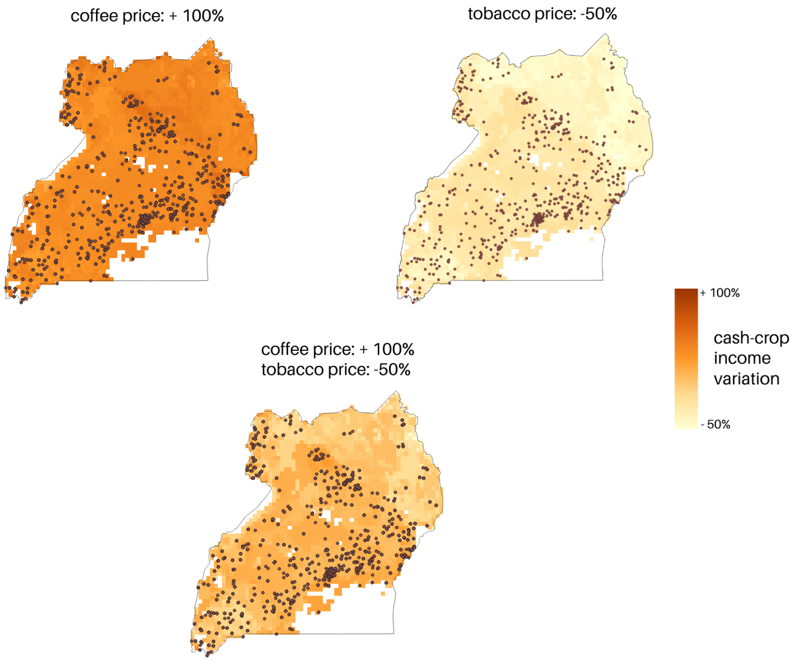


Figure 2. Examples of the proxy for yearly cash crop income variation computed on different price movements. Dots represent the location of DHS clusters. Within each map, darker areas are predicted to have a higher income variation than lighter ones.

1.2.3 The idealized experiment and the identification assumptions

Analysing the interplay between cash crop prices and birth spacing is difficult because we do not observe actual income variation of couples interviewed by the DHS, since this variable is not captured by the survey. Even if we observed this variable, we could not regress it on fertility decisions since those are contemporaneously determined with labour decisions (i.e. how much labour effort to provide in agriculture).

The idealized experiment in this setting would be a randomized control trial on households involved in agricultural economic activities that want an additional child. Farmers in one treatment arm should be offered to sell their cash crops to higher prices and farmers in another treatment arm should receive an unconditional cash transfer. Those households that can sell at higher prices face a higher opportunity costs of time devoted to parental care and may decide to have the child later with respect to the control group (if the substitution effect prevails); those household that receive the cash transfer should instead supply less agricultural labour while remaining at the same income level and decide to have the child sooner. In this paper, the second arm of the idealized experiment is proxied by the survey experiment in Section 3, aimed at understanding if an income effect exists and which impact it has on birth spacing preferences, while the first arm is proxied by employing quasi-random shifts in international prices which determine income variation based on the suitability of the soil. The recent literature on shift-share designs argues that this analysis is valid only if certain methodological precautions are taken. To meet these requirements, I (*i*) address the incomplete-share problem by normalizing crop suitability so that shares sum to one (Equation 1.1), thereby ruling out correlation between the proxy and the error term; (*ii*) lag the shares to the start of the natural experiment by using FAO GAEZ indices from the 1961–1990 baseline, which predates the collection of both UNPS and DHS outcomes; (*iii*) implement robustness checks standard in the literature, including falsification exercises and standard errors robust to share-level correlation (Section 2.4).

1.2.4 Weighting

To ensure comparability across the pooled 2006, 2011, and 2016 DHS waves, I re-normalize sampling weights. Individual weights are first standardized within each survey year and then re-scaled such that the sum of weights is identical across waves. This ensures that each survey round contributes equally to the aggregate analysis, preventing years with larger sample sizes from driving the pooled estimates (Rutstein and Staveteig 2014). The main findings are robust to the use of original, unadjusted weights.

1.3 Descriptives

1.3.1 Shifts and shares

The key ingredients of the proxy variable are the international prices of the four cash crops and the suitability for their cultivation at household cluster coordinates. Figure 3 shows how prices of selected cash crops vary in the period from 1995 to 2015 (maximum range of outcome variables); since the main specifications include a standardized version of the the shift-share proxy the figure depicts both the price series indexed to 1995 and the one represented by z-scores.

Figure 4 depicts the distribution of suitability for cultivation of the crops at the household coordinates of the 4740 household interviewed by DHS during the 2005, 2011 and 2016 waves. Coffee and tea suitability distribution is widespread, while tobacco and cotton have a peak at lower levels of suitability.

1.3.2 Income

The first part of the result section investigates whether the variable index captures the variation in households agricultural income. The analysis is carried out on a sub-sample of households in the Uganda National Panel Survey that was interviewed in three subsequent years about their agricultural activity; table 2 shows the descriptive statistics of relevant

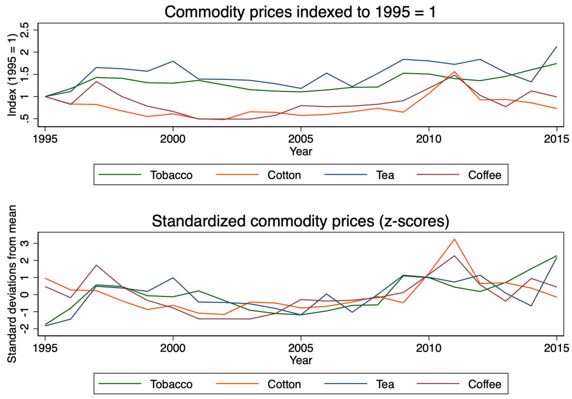


Figure 3. International real prices of cotton, tea, tobacco and coffee. The upper panel shows the price indexed to 1995 = 1, the lower standardized prices (z-scores). Source: FAO (2025)

.

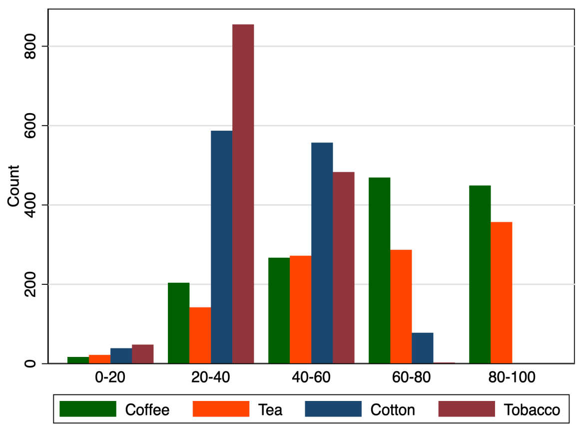


Figure 4. Histogram of suitability for cultivation of coffee, cotton, tea and tobacco for the households of the DHS sample. Source: FAO and IIASA (2020)

.

	N	Mean	Std. Dev.	Min	Max
Value of crop sold (real USD)	3153	182.33	322.92	0.00	6400.96
Tot. income (real USD)	3153	932.32	1104.95	0.00	16175.25
Agr. income (real USD)	3153	454.87	384.29	0.00	4762.95
# of workers in own plots	3153	2.74	1.57	0.00	49.00
# of working days in own plots	3153	258.74	197.02	0.00	1979.00
Days per worker	3151	106.39	82.02	0.00	883.80
Hectars of land	3099	1.33	1.13	0.02	14.09

Table 2. Summary statistics for the balanced panel of 1051 agricultural households from the Uganda National Panel Survey (UNPS) for survey years 2009-2011. The table reports means, standard deviations, and ranges for key household-level agricultural production, income, and labour market outcomes. All monetary values are reported in real 2010 US dollars.

	N	Mean	Std. Dev.	Min	Max
Husband wants another child	1301	0.58	0.49	0.00	1.00
Wife wants another child	1301	0.54	0.50	0.00	1.00
Both want more children	1301	0.43	0.50	0.00	1.00
Waiting time for birth (wife, years)	550	2.51	1.66	0.00	6.00
Waiting time for birth (husband, years)	552	2.24	1.53	0.00	6.00
Age husband	1301	35.76	8.91	18.00	54.00
Age wife	1301	30.34	8.23	15.00	49.00
Number of children	1301	4.12	2.51	0.00	12.00

Table 3. Descriptive statistics for the primary analysis sample of 1,301 agricultural self-employed couples from the pooled 2006, 2011, and 2016 waves of the Demographic and Health Surveys (DHS). The table presents summary statistics for stated fertility preferences and key demographic characteristics.

variables, for the 1051 household $\times$ 3 waves. The value of crop sold, the total and the agricultural income variables are in real 2010 dollars (as the international crop prices).

1.3.3 Fertility preferences

The analysis on fertility behaviour uses a sample of couples interviewed in the Demographic and Health Surveys (DHS) in 2006, 2011 and 2016; pooling the three waves yields a sample of 4740 couples. Since the DHS does not collect direct measures of household or individual income, I rely on the occupation module to subset the analysis on those couples where both of husband and wife declare to be agricultural self-employed (1301 couples across the three waves, 36% of the overall sample).

Table 3 shows descriptive statistics of relevant variables: in more than 40% of the couples both husband and wife want an additional child, with an average preferred waiting time of 2.2 years. If we take into account the year of the last of birth of the couple and calculate the spacing implied by husbands' preferences, in 88 out of 550 (22%) cases the waiting time would not be adequate to ensure the safety of the mother and the coming children. A smaller share (16%) of waiting times implied by wives preferences is potentially harmful, according to the WHO guidelines, for children and mothers health.

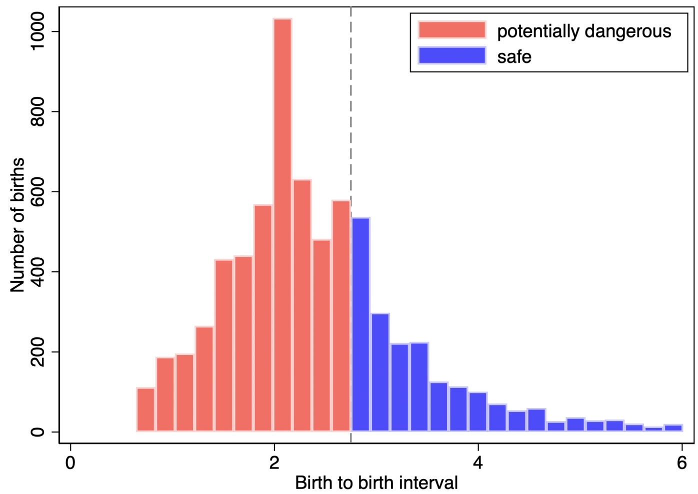


Figure 5. Histogram of the distribution of birth-to-birth intervals, measured in years, from the pooled DHS birth history data for the sample of agricultural self-employed couples. The vertical dashed line at 2.75 years shows the minimum birth-to-birth interval recommended by the World Health Organization (WHO). Shorter intervals, associated with higher health risks, are shaded in red.

1.3.4 Birth histories

To test whether stated fertility preferences translate into actual behaviour, I move from intentions to revealed outcomes by analyzing detailed birth histories from the DHS. In this case I look at birth intervals, corresponding to the time elapsed between birth $k-1$ and birth k. The distribution of these intervals is depicted in figure 5, which shows that a considerable number of births occur outside the safe window prescribed by the WHO.

2 Results

2.1 Income proxy validation

A crucial first step is to establish that the price index is a valid proxy for agricultural income shocks. The next sections the analysis confirms that international prices transmit to local prices and that the index correlates with household-level economic outcomes. I also investigate the relationship between the index and the hours worked by the household members in their own plot.

2.1.1 International versus local prices

A crucial assumption for the identification is the existence of a correlation between local and international prices. To verify this, records of tobacco, cotton, tea, and coffee sales were extracted from the full set of Ugandan national panel surveys (2005 to 2020). Unit

	Local prices			
	Coffee	Tea	Tobacco	Cotton
Crop international price	0.013**	0.027*	0.012**	0.019***
	(0.005)	(0.013)	(0.005)	(0.002)
R-squared	0.005	0.263	0.065	0.195
Observations	1087	14	75	261

s.e. in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Evidence of international-to-local price transmission. The table reports coefficients from OLS regressions of inflation-adjusted local farm-gate prices (in USD) on international real prices for the four main cash crops. Each column represents a separate regression for the specified crop.

prices per kilogram were calculated by dividing the transaction value by the quantity sold, adjusted for inflation, and converted to US Dollars. To mitigate the influence of outliers and reporting errors, the analysis retains only prices falling between the 25th and 75th percentiles.

Table 4 shows the results of regressing prices received by farmers against international real prices (in USD), confirming that higher international prices are positively correlated with higher local prices.

2.1.2 Household income variation

The three consecutive Ugandan National Panel Survey containing the agricultural module give us the opportunity to see if the proxy is correlated to variation in revenues from crop selling and days worked in the household plots. I exploit the FAO RuLIS database, which calculates an household-level indicator of income derived from crop sales based on the data collected by the UNPS. In particular, I consider at yearly data collected in 2009, 2010 and 2011[8]. Based on income data, I define the ratio between income from crop selling and agricultural income, $\frac{\text{cropsold}}{\text{agrincome}}$ and the ratio between income from crop selling and total income $\frac{\text{cropsold}}{\text{totincome}}$. Based on work data, I create the variable $\frac{\text{days}}{\text{worker}}$, the number of days worked on household plots per household worker.

I consider only households i that are interviewed in each of the three years and calculated year on year variations of the relevant variables in the year $t \in \{2010, 2011\}$ as

$$\Delta \mathrm{Y}_{i,t} = \frac{\mathrm{Y}_{i,t} - \mathrm{Y}_{i,t-1}}{\mathrm{Y}_{i,t-1}}.$$

I then estimate the specification

$$\Delta \mathrm{Y}_{i,t} = \beta_0 + \beta_1 std(\text{price\_index})_{i,t} + \delta_i + \gamma_t + \varepsilon_{i,t} \tag{1.2}$$

[8]For example, the 2009 income is retrieved in the from the 2009-2010 UNPS that asks about crop sold in the first (March-June) and second (October-December) seasons of 2009.

where $std(\text{price\_index})_{i,t}$ is the standardized value of *price_ index* (the suitability-weighted exposure to variation of the international prices of cash crops between year $t-1$ and year t), δ_i and γ_t are household and time fixed effects. Standard errors are clustered at enumeration area level (the primary sampling unit, or psu). Results are in table 5.

A one-standard deviation increase in the suitability-weighted exposure to the international prices is associated with a 50% increase of the value of the crop sold; crop selling becomes a higher share of agricultural income (going from 42% to 70% on average) and of total income. Both total number of days worked and number of day per worker increase, by 23% and 32% respectively, while the number of workers (not shown in the table) is not affected. These results suggest that increase in international prices of cash crops translates into a meaningful increase in the return of crop selling household and it is associated to an increase in agricultural working days along the intensive margin; based on the theoretical framework sketched in the subsection 1.2.3 couples work more and defer the next birth thus increasing birth spacing.

Unfortunately, UNPS data do not report working days disaggrated by household member; it is therefore not possible to precisely measure the intensive or extensive margin of increase in wives labour. At the same time, the validity of this effect as a mechanism for the results in the following sections relies on the assumption that female labour supply is responsive to cash crop price fluctuations. I argue this is the case in the setting of this paper. While cash crops in Uganda are often marketed by male household heads, the labour required to produce them is heavily dependent on women. For example, literature on Ugandan agriculture documents that women provide approximately 70% of the labour in coffee production and are actively engaged in labor-intensive tasks such as weeding, harvesting, and post-harvest processing (Hill and Vigneri 2014). In the sample of UNPS panel data, I actually find that women are integral to crop production: in 72% of households the wife is listed among the top three labourers on the household plots.

2.2 Reduced-form effect on fertility preferences

This subsection investigates the relationship of short-term income variation on desired fertility. As mentioned in section 1.3.3, the Demographic Health Survey surveyed, in Uganda, a total of 4740 couples split in the three waves of 2006, 2011 and 2016 asking fertility preferences (whether they want an additional child an when) to both husbands and wives. I subset this sample to those who agricultural self-employed (i.e. those who operate their own farm or agricultural enterprise independently), which leaves us with roughly 1300 observations[9]. For each couple, based on its location and year of interview, I compute the *price_ index* variable as discussed in section 1.2.

[9]Since `reghdfe` removes singletons to have a correct estimation, this number lowers in the regressions table.

	Δ cropsold	Δ $\frac{\text{cropsold}}{\text{agr. income}}$	Δ $\frac{\text{cropsold}}{\text{tot. income}}$	Δ days	Δ $\frac{\text{days}}{\text{workers}}$
std(price_index)	0.544***	0.718***	0.847***	0.244**	0.347***
	(0.155)	(0.225)	(0.281)	(0.101)	(0.112)
Year FE	yes	yes	yes	yes	yes
Household FE	yes	yes	yes	yes	yes
Dep. var. mean	140.839	0.402	0.233	244.094	97.224
Price index mean	-0.013	-0.013	-0.013	-0.013	-0.013
Price index SD	0.165	0.165	0.165	0.165	0.165
R-squared	0.394	0.452	0.484	0.467	0.439
Observations	2102	2102	2102	2096	2096

s. e. in parentheses, clustered at psu level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Validation of the cash crop price index. The table presents estimates of the effect of a one-standard-deviation change in the `price_index` on year-on-year changes in household income and labour outcomes from the UNPS panel. Dependent variable mean is calculated on the variable x, not on the Δx. All specifications include household and year fixed effects.

I then estimate the specification

$$\text{Y}_i = \beta_0 + \beta_1 std(\text{price\_index})_{i,t} + \alpha X_i + \varepsilon_i \tag{1.3}$$

separately for husbands and wives. The outcome Y_i is either a dummy indicating if the respondent wants an additional child or, conditional on willingness to have another child, the preferred waiting time for the next child; the variable $\texttt{price\_index}_{i,t}$ proxies the standardized variation year $t-1$ to year t of cash crop income and X_i contains age of the husband and age of the wife (as control variables), year of last birth, district[10] and time fixed effects. Standard error are clustered at district level.

Variation in international prices does not have an effect on the willingness to have an additional child. Conditional on wanting it, it increases the waiting time by 0.46 years for husbands and 0.30 years for wives (18% and 12% of the mean of the dependent variable, respectively).

2.3 Reduced-form effect on birth spacing

I then exploit detailed information on birth histories of those couples interviewed in the DHS to assess the extent to which intentions to have a deferred birth because of a positive income variation are then put to practice.

[10]Since each couple is surveyed only once, I control for geographical heterogeneity by adding a fixed effect for each of the 56 Ugandan districts as defined in the 2002 Ugandan Census considered in the survey. Each district contains 5 to 15 clusters for each wave.

	Additional child		Waiting time	
	Hus	Wife	Hus	Wife
std(price_index)	0.038	-0.011	0.464***	0.302**
	(0.025)	(0.018)	(0.131)	(0.122)
Age husband	-0.013***	-0.006**	-0.016	-0.038**
	(0.003)	(0.003)	(0.015)	(0.017)
Age wife	-0.005	-0.015***	0.025	-0.016
	(0.004)	(0.004)	(0.035)	(0.037)
District FE	yes	yes	yes	yes
Wave FE	yes	yes	yes	yes
Parity (No. Children) FE	yes	yes	yes	yes
Last Birth Year FE	yes	yes	yes	yes
# Districts	54	54	51	51
# Waves	3	3	3	3
# Parity FE	14	14	9	9
# Last Birth Year FE	29	29	17	17
Dep. var. mean	0.564	0.517	2.671	2.386
Price index mean	0.056	0.056	0.053	0.052
Price index SD	0.067	0.067	0.067	0.067
R-squared	0.284	0.430	0.231	0.139
Observations	1189	1189	462	463

s. e. in parentheses, clustered at district level
* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 6. Reduced-form estimates of the effect of price shocks on stated fertility preferences. The table reports OLS estimates of the impact of the standardized `price_index` on the desire for an additional child (columns 1-2) and the preferred waiting time in years (columns 3-4), conditional on both members wanting another child.

2.3.1 Effect on birth-spacing

To test whether fertility preferences translate into actual behaviour, I move from stated intentions to revealed outcomes by analysing detailed birth histories from the DHS. For each couple considered in the analysis of the previous section, I recover the length of birth-to-birth spacing intervals from the second birth onward, which will be the outcome variable of interest. The key explanatory variable is the lagged price index, capturing the suitability-weighted variation in international prices in the year of the previous birth (i.e. the variation in income from $t_{k-1} - 1$ to t_{k-1}), which are the conditions relevant to the decision to have the next child. This identification strategy hinges on using the quasi-random price shock realized at the time of the previous birth as a pre-determined proxy that is plausibly exogenous to the forward-looking decision of when to have the next child.

In particular, I estimate the specification

$$\text{spacing}_{i,k} = \beta_0 + \beta_1 std(\text{price\_index})_{i,t_{k-1}} + \gamma_{t_{k-1}} + \alpha X_i + \epsilon_{i,k}$$

where $\text{spacing}_{i,k}$ is the time in years between birth $k-1$ and birth k for woman i, the variable $std(\text{price\_index})_{i,t_{k-1}}$ is the standardized cash-crop price shock experienced by woman i during the year of the previous birth, $\gamma_{t_{k-1}}$ controls for the year of the previous birth and the vector X_i contains a set of fixed effects: either district, ages of husband

and wife and parities (to fully mirror specification 1.3) or woman fixed effects.

Table 7 presents the results of this analysis. In column 1, I only consider each mother's second birth. One standard deviation increase in the suitability-weighted exposure to international prices at the time of the first birth is associated to an increase of 0.3 years in the birth spacing of the second birth, similar to the coefficients estimated in table 6. If we consider all the births and account for parity fixed effects (column 4) the magnitude of this coefficient halves, despite remaining significant at 95% level, suggesting that the effect is more pronounced at lower parities. This coefficient is comparable to that obtained by controlling for mother fixed effect, presented in column 5. In the sample, the average birth-to-birth interval is approximately 29 months, significantly below the recommended threshold. A delay of 4.2 months represents a 15% increase in the average waiting time, effectively bridging nearly half the gap between the observed mean and the 3-year safety target. More over, since the relationship between birth spacing and adverse health outcomes is non-linear, by shifting the distribution of births to the right even a *small* average delay moves a significant mass of births out of the highest-risk zone (under 24 months) and into a safer temporal window. Taken together, these results suggest that the positive variation in agricultural income translate into longer birth-to-birth intervals, which could be beneficial to both mothers and children health.

2.3.2 Effect on the yearly probability to have birth

For households affected by an increase in the suitability-weighted price variation, higher birth spacing should also translate into a lower probability to give birth in a given year. To test this, following Kitchens and Rodgers (2023), I build an yearly panel of births for each woman who belongs to a couple of both agricultural self-employed; in order to focus on birth spacing I then set as starting point of each birth history the year after the first birth, thus dropping any observations before that year and exclude observation before 1995 (the year Uganda entered in the WTO). This results in 17680 observations for 1519 women ranging from 1995 to 2015.

I then estimate the baseline specification

$$\text{had\_birth}_{i,t} = \beta_0 + \beta_1 std(\text{price\_index})_{i,t-1} + \delta_i + \gamma_t \tag{1.4}$$

where *had_birth* is a dummy equal to one if woman i had birth in year t and zero otherwise, $price\_index_{i,t-1}$ the standardized proxy for the variation of cash crop income from year $t-2$ to year $t-1$, δ_i are either districts (and in one specification I add religion, cohorts fixed effects) or woman fixed effects and γ_t and year fixed effects. Standard errors are clustered at primary sampling unit level. Table 8 shows the results of the estimation: in column 1 and 2 I employ a model with years and districts fixed effects. Column 3 switches to individual fixed effects, while column 4 repeats the estimation on a smaller

	Dependent variable: birth spacing (in years)				
	(1)	(2)	(3)	(4)	(5)
std(price_index) at previous birth	0.354***	0.124**	0.120**	0.137**	0.111*
	(0.126)	(0.060)	(0.060)	(0.056)	(0.061)
Age Husband	0.011			0.015***	
	(0.007)			(0.004)	
Age Wife	0.068***			0.063***	
	(0.011)			(0.007)	
Mother FE					Yes
District FE	Yes	Yes	Yes	Yes	
Previous Birth Year FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes
Parity FE			Yes	Yes	
Sample: only first birth interval	Yes	All	All	All	All
# Districts	54	54	54	54	
# Waves	3	3	3	3	3
# Couples					1,219
# Previous Birth Years	25	25	25	25	25
# Parity Orders			12	12	
Dep. var. mean	2.420	2.469	2.469	2.469	2.463
Price index mean	0.018	0.029	0.029	0.029	0.030
Price index SD	0.169	0.167	0.167	0.167	0.168
Adjusted R-squared	0.122	0.031	0.034	0.107	0.076
Observations	1,420	5,661	5,661	5,661	5,446

Standard errors in parentheses, clustered at the mother level.

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}p < 0.01$

Table 7. Reduced-form estimates of the effect of price shocks on actual birth-to-birth spacing. The dependent variable is the birth interval in years. The key independent variable is the standardized `price_index` in the year of the previous birth. Column (1) restricts the sample to the first birth interval. Columns (2)-(5) use the full sample of all birth intervals, progressively adding controls and employing mother fixed effects in column (5). Standard errors clustered at the mother level in parentheses.

	(1) had_birth	(2) had_birth	(3) had_birth	(4) had_birth
std(price_index)	-0.027**	-0.028**	-0.027**	-0.030**
	(0.012)	(0.012)	(0.012)	(0.015)
Mother FE	yes	yes	yes	yes
District FE	yes	yes		
Year FE			yes	yes
Wife coohort FE		yes		
Hus. coohort FE		yes		
Religion FE		yes		
# Years	24	24	24	24
# Districts	54	54		
# Women			1519	1493
# Religions		9		
# Wife cohorts		43		
# Hus. cohorts		46		
Dep. var. mean	0.311	0.311	0.312	0.298
Price index mean	0.015	0.015	0.015	0.010
Price index SD	0.141	0.141	0.141	0.147
Observations	17680	17680	17657	12665

s. e. in parentheses, clustered at psu level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8. Reduced-form estimates of the effect of international cash crop price shocks on the annual probability of giving birth, using a woman-year panel from DHS birth histories. The dependent variable is a dummy equal to one if a birth occurred in year t. The key independent variable is the standardized `price_index` from year t-1. Column (1) includes district and year fixed effects, column (2) adds demographic controls, column (3) replaces district with woman fixed effects, and column (4) restricts the sample based on the woman's ideal number of children. Standard errors are clustered at the PSU level.

sample obtained by removing, for each woman, the observations representing the years after she achieved her ideal number of children (an information that is collected by the DHS).

Estimates are consistent through all the specifications and report a negative effect of an increase in the independent variable on woman's probability to give birth. By considering column 3 as the preferred specification, we can observe that an increase from year $t-2$ to year $t-1$ of one standard deviation in the suitability-weighted exposure to international prices leads to a 0.027 decrease in the probability of a woman to give birth in year t , roughly 9% of the mean of the dependent variable (0.312), which implies an effect on birth spacing similar to that estimated in tables 6 and 7.

2.4 Robustness tests

2.4.1 Randomization inference

To assess the robustness of the findings, I implement randomization inference following the framework illustrated by Heß (2017). Randomization inference provides a distribution-free approach to hypothesis testing that is particularly well-suited to shift-share designs, where standard asymptotic inference may be unreliable in finite samples (Borusyak, Hull, and Jaravel 2022). The core idea is to construct the exact null distribution of the test statistic by repeatedly permuting the treatment variable (in this case, the shift-share price index) across observations, while holding fixed all other aspects of the data and model specification.

Specifically, I use the `ritest` command in Stata to conduct 200 random permutations of the `price_index` variable. For each permutation, I re-estimate our baseline specification and compute the coefficient estimate. This generates an empirical distribution of the treatment effect under the sharp null hypothesis that the price index has no effect on birth spacing. The p-value is then calculated as the proportion of permuted test statistics that are as extreme or more extreme than the observed test statistic. Table 9 presents the results, which confirm the robustness of the analysis to the randomization inference procedure.

2.4.2 Standard errors robust to spatial correlation

The baseline analysis accounts for the non-independence of observations by clustering standard errors at the district level. However, this approach assumes that the error term ϵ is correlated for households within the same district but is strictly uncorrelated across district boundaries. This assumption may be overly restrictive, as unobserved factors affecting fertility decisions may be spatially correlated and transcend administrative borders (Cameron and D. L. Miller 2015). If such spatial autocorrelation exists, standard clustering can be an imprecise approximation of the true error covariance structure, potentially leading to biased standard errors.

To ensure the robustness of the findings to more flexible patterns of spatial dependence, the main specifications are re-estimated using spatially robust standard errors. I correct for spatial autocorrelation using a Bartlett (triangular) kernel, which assumes that the correlation between observations declines linearly with distance, reaching zero at a specified cutoff. A cutoff distance of 150 kilometers is chosen to allow for correlation across a reasonably large geographic area. The results of this robustness check are presented in Table 10. The estimated coefficients for the price index shock on both fertility preferences and birth outcomes remain mostly statistically significant. This confirms that the main findings are not an artifact of unmodeled spatial correlation.

Outcome	Coefficient	Permutations	RI p-value
Panel A: Income and Labor Market Outcomes (UNPS)			
Value of Crop Sales	0.544	200	0.000
Crop Sales / Agr. Income	0.718	200	0.000
Crop Sales / Total Income	0.847	200	0.000
Days Worked	0.244	200	0.000
Days Worked per Worker	0.347	200	0.000
Panel B: Stated Fertility Preferences (DHS)			
Wife's Desired Wait (Years)	0.464	200	0.000
Husband's Desired Wait (Years)	0.302	200	0.005
Panel C: Actual Birth Spacing (Years, DHS)			
Model 1: 2nd Birth, Ages FE	0.335	200	0.000
Model 2: District FE	0.123	200	0.000
Model 3: + Parity FE	0.120	200	0.000
Model 4: + Ages FE	0.141	200	0.000
Model 5: Couple FE	0.111	200	0.000
Panel D: Probability of Birth in a Given Year (DHS)			
Model 1: District, Year FE	-0.027	200	0.000
Model 2: + Controls	-0.028	200	0.000
Model 3: Couple, Year FE	-0.027	200	0.000
Model 4: M3, below ideal N	-0.030	200	0.000

Notes: The table reports results from randomization inference tests based on 200 permutations of the price index shock. The RI p-value is the two-sided share of permutations yielding a coefficient at least as large in absolute value as the observed one. Panel A uses the UNPS panel. Panels B, C, and D use DHS data.

Table 9. Randomization inference tests for main specifications. The table reports the estimated coefficients and randomization inference p-values from 200 permutations of the price index shock. The p-value represents the share of permutations yielding a coefficient with an absolute value greater than or equal to the observed coefficient. Panels A, B, C, and D correspond to the main specifications for income outcomes (Table 4), stated preferences (Table 5), actual birth spacing (Table 6), and the probability of birth (Table 7), respectively.

	Stated Prefs.		Birth Spacing					P(had birth)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
std(price_index)	0.464***	0.302***	0.347***	0.124**	0.122**	0.120**	0.111	-0.027*
	(0.131)	(0.093)	(0.113)	(0.059)	(0.060)	(0.056)	(0.075)	(0.016)
Age husband	-0.016	-0.038**	0.011			0.011**		
	(0.017)	(0.015)	(0.007)			(0.005)		
Age wife	0.025	-0.016	0.068***			0.032***		
	(0.032)	(0.030)	(0.011)			(0.004)		
Household FE	Yes	Yes						
Year FE	Yes	Yes						Yes
District FE			Yes	Yes	Yes	Yes		
Wave FE			Yes	Yes	Yes	Yes	Yes	
Parity FE			Yes	Yes		Yes		
Last Birth Year FE			Yes	Yes				
Previous Birth Year FE					Yes	Yes	Yes	
Mother FE							Yes	Yes
Dep. Var. Mean	2.671	2.386	2.421	2.470	2.470	2.470	2.464	0.312
Observations	462	463	1,416	5,661	5,661	5,661	5,446	17,657

Notes: Standard errors corrected for spatial correlation up to 200km using a Bartlett kernel.
* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 10. Robustness of main estimates to spatial autocorrelation. This table re-estimates the main specifications for stated preferences, birth spacing, and birth probability using standard errors that are robust to spatial correlation. A Bartlett kernel with a 150 km cut-off is used to allow for spatial dependence that declines linearly with distance. Point estimates are identical to those in Tables 6, 7, and 8.

2.4.3 Estimation on a placebo sample

To further strengthen the validity of the research design, I re-estimate the main specifications on a placebo sample comprising couples where neither the husband nor the wife is engaged in agricultural self-employment. Since these households' incomes are not directly linked to crop cultivation, the proxy constructed from agricultural price shocks should have no effect on their reproductive behaviour. As shown in Table 10, the estimated coefficient of the price index is statistically indistinguishable from zero across all fertility outcomes for this non-agricultural group. This null result provides evidence against confounding factors that might be correlated with the proxy and affect all households' fertility choices, thus bolstering a causal interpretation of the main findings.

3 Is there an income effect? Hints from a survey experiment

To complement the analysis of the effects of income variation on birth spacing I run a small survey experiment on coffee-producing households in the district of Ibanda, Uganda. Survey experiment are an increasingly popular method to explore human behaviour in conterfactual scenarios and extract stylized facts (Stantcheva 2023).

To investigate the perceived link between income volatility and fertility timing, I run a small survey module comprising direct questioning and a randomized vignette experiment. The questions were administered to an household member (either the household head or her/his spouse) of 626 households growing Arabica coffee in the district of Ibanda,

	Stated Prefs.		Birth Spacing			P(had birth)
	(1) Wife Wait	(2) Husb Wait	(3)	(4)	(5)	(6)
std(price_index)	-0.272	0.230	0.056	-0.033	0.075	-0.012
	(0.333)	(0.315)	(0.185)	(0.097)	(0.126)	(0.019)
Age Husband	-0.021	-0.024	0.013	0.015*		
	(0.022)	(0.021)	(0.013)	(0.009)		
Age Wife	-0.002	0.036	0.159***	0.166***		
	(0.032)	(0.042)	(0.022)	(0.019)		
District FE	Yes	Yes	Yes	Yes		
Wave FE	Yes	Yes	Yes	Yes	Yes	
Parity FE	Yes	Yes		Yes		
Last Birth Year FE	Yes	Yes				
Previous Birth Year FE			Yes	Yes	Yes	
Mother FE					Yes	Yes
Year FE						Yes
# Districts	49	49	52	54		
# Waves	3	3	3	3	3	
# Parity FEs						
# Last Birth Year FEs	15	16	25	25	25	
# Mother FEs					586	911
Dep. Var. Mean	2.857	2.678	2.794	2.830	2.789	0.246
Observations	388	395	795	2,352	2,125	9,375

Notes: Placebo tests on non-agricultural households. Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11. Placebo test using non-agricultural households. This table re-estimates the main specifications for stated preferences (columns 1-2), actual birth spacing (columns 3-5), and the annual probability of birth (column 6) on a sample of couples interviewed by the DHS where neither of the two members is engaged in agricultural self-employment. The consistently statistically insignificant coefficients support the exogeneity assumption of the research design.

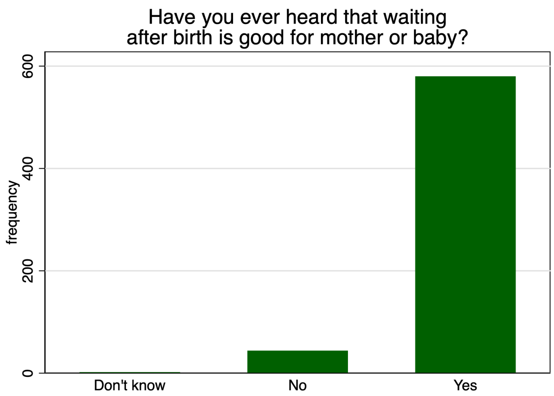
Uganda, during the month of July 2025. Initially, respondents were asked a series of questions about their general knowledge and beliefs about family planning and the influence of economic factors on fertility. Specifically, participants were asked three closed-answer questions:

1. *Have you ever heard that waiting some time after giving birth before getting pregnant again is good for the mother or baby?*

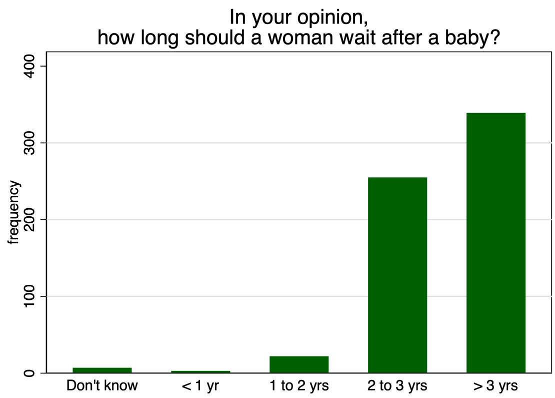
2. *In your opinion, how long should a woman wait after having a baby before getting pregnant again?*

3. *Think about all the coffee farmers you know. Do you think how much money they make affects when they decide to have another child?*

The answers to the first two questions are summarized in Figure 6. The majority of the respondents heard about spacing and know what is the recommended timing between births.



(a) Heard about spacing



(b) Ideal birth spacing

Figure 6. Awareness about birth spacing

Moving on to the third question, a lot of the respondent don't think that for coffee farmers how much money they make influences the timing of births (figure 7). This could be a threat for the validity vignette experiment as we need to check that its results hold in the full sample and in the subset of those who *ex-ante* thought that a relationship between income and spacing existed.

Following these direct questions, participants were randomly assigned to one of three vignette conditions to explore the perceived causal effect of income shocks on fertility timing. Each vignette introduced by the same introduction

- *Denis and Beatrice are married. They grow coffee. They have two children. They want another child, but they have not yet decided when.*

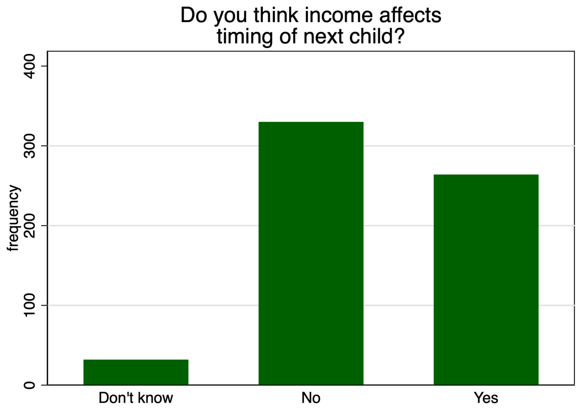


Figure 7. Share of respondents who think that income affects birth timing among farmers.

and then varied in one of the three income scenario:

- in the *High Income* version, respondents were told: *"Let's imagine that the price of green coffee beans is very high. They sell their coffee and get more money than usual."*
- in the *Low Income* version, the scenario presented was: *"Let's imagine that the price of green coffee beans is very low. They sell their coffee and get less money than usual."*
- Finally, the *Normal Income* version served as a control, stating: *"Let's imagine that the price of green coffee beans is normal. They sell their coffee and get about the same money as they usually do."*

After each vignette, respondents were asked to consider the specific income scenario and answer:

- *Do you think the fact they have [more/less/the same amount of] money would make them want the next child...* with response options including *"Sooner," "Later," "No difference," "No one decides - it just happens", and "Don't know."*

This experimental design allows for a direct assessment of how perceived income changes influence decisions about the timing of subsequent births. I then estimate specification similar to equation 1.3:

$$\text{answer}_i = \beta_0 + \beta_1 \text{high\_vignette}_i + \beta_2 \text{low\_vignette}_i + \alpha X_i + \epsilon_i$$

where answer_i is a series of dummy (one for each possible answer: *sooner*. *later*, *no difference*, *no one decides - it just happens* and *don't know*), *high_vignette* and *low_vignette*

	DK		Fate		Later		Same		Sooner	
	Baseline	FE	Baseline	FE	Baseline	FE	Baseline	FE	Baseline	FE
High vignette	0.002		0.017		-0.216***		-0.020		0.216***	
	(0.011)		(0.027)		(0.043)		(0.049)		(0.035)	
		-0.002		0.028		-0.199***		-0.018		0.191***
		(0.011)		(0.030)		(0.045)		(0.050)		(0.035)
Low vignette	-0.001		0.019		0.105**		-0.063		-0.060***	
	(0.010)		(0.027)		(0.050)		(0.050)		(0.019)	
		-0.002		0.032		0.102*		-0.069		-0.062**
		(0.008)		(0.032)		(0.054)		(0.064)		(0.023)
Age		0.001**		0.000		-0.002		0.000		0.001
		(0.000)		(0.001)		(0.001)		(0.002)		(0.001)
Female		0.001		0.004		-0.015		0.032		-0.022
		(0.010)		(0.023)		(0.036)		(0.048)		(0.028)
Last birth FE		Yes		Yes		Yes		Yes		Yes
EA FE		Yes		Yes		Yes		Yes		Yes
R-squared	0.000	0.012	0.001	0.003	0.085	0.083	0.003	0.004	0.137	0.119
Observations	626	605	626	605	626	605	626	605	626	605
Dep. var. mean	0.011	0.012	0.083	0.083	0.335	0.337	0.447	0.450	0.123	0.119

s.e. in parentheses, clustered at village level

* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 12. Effect of hypothetical income shocks on fertility timing preferences from a vignette experiment. The table reports coefficients from OLS regressions where the dependent variables are dummies for the respondent's choice about a fictional couple's birth timing. The independent variables are indicators for assignment to the 'High' or 'Low' income scenarios, with the 'Normal income' scenario as the reference category. For each outcome, the first column is a baseline estimate and the second includes age and sex fixed effects.

are the two high and low income vignette (so that the normal scenario is the reference category) and X_i are individual controls: age, sex and fixed effects for last year of births and village. standard errors are clustered at village level. Table 12 presents the results of the two specifications with and without fixed effects for each outcome. The results consistently point towards a positive relationship between income and birth spacing: in the *High (Low) income* scenario respondent were less (more) likely to answer that *Denis and Beatrice*, the fictional couple, wanted a children later (sooner) with respect to the *Normal income* scenario. Reassuringly, the probabilities of answering *Don't know*, *No one decides - it just happens* and *Same* were instead not altered by the scenario each respondent was randomly assigned to.

As a robustness check, I re-estimated the same specifications on the sub-sample of respondent who answered that *How much money (coffee farmers) make affects when they decide to have another child*. The results (table 13) of the in the high income setting are qualitatively the same, with some coefficients becoming stronger in magnitude.

	DK		Fate		Later		Same		Sooner	
	Baseline	FE	Baseline	FE	Baseline	FE	Baseline	FE	Baseline	FE
High vignette	-0.011		-0.021		-0.445***		-0.077		0.555***	
	(0.011)		(0.022)		(0.065)		(0.062)		(0.061)	
		-0.009		0.001		-0.489***		-0.086		0.583***
		(0.021)		(0.013)		(0.097)		(0.079)		(0.076)
Low vignette	0.000		0.011		0.044		0.033		-0.089***	
	(0.016)		(0.029)		(0.073)		(0.067)		(0.034)	
		0.005		0.039		0.009		0.034		-0.086*
		(0.014)		(0.033)		(0.103)		(0.090)		(0.047)
Age		0.001		0.002*		-0.003		-0.002		0.002
		(0.001)		(0.001)		(0.002)		(0.002)		(0.002)
Female		0.020		0.017		0.008		-0.004		-0.041
		(0.020)		(0.035)		(0.077)		(0.061)		(0.053)
Last birth FE		Yes		Yes		Yes		Yes		Yes
EA FE		Yes		Yes		Yes		Yes		Yes
R-squared	0.004	0.027	0.006	0.037	0.192	0.194	0.011	0.014	0.427	0.469
Observations	264	248	264	248	264	248	264	248	264	248
Dep. var. mean	0.008	0.008	0.030	0.028	0.473	0.492	0.242	0.234	0.246	0.238

s.e. in parentheses, clustered at village level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13. Robustness check of the vignette experiment on a restricted sample. This table replicates the analysis from Table 11, restricting the sample to respondents who stated ex-ante that income affects fertility timing. The results are qualitatively similar to the full sample, confirming the robustness of the experimental findings.

Conclusions

This paper investigates the causal effect of short-term income shocks on birth spacing decisions in rural Uganda. I considered two competing hypotheses: a positive income shock could either shorten birth intervals by easing the financial costs of child-rearing (the income effect) or lengthen them by increasing the opportunity cost of time spent on childcare instead of agricultural work (the substitution effect).

To disentangle these mechanisms, I employ a quasi-experimental shift-share design. I construct an index of exogenous income shocks by interacting plausibly exogenous year-to-year fluctuations in international commodity prices with pre-determined, local agro-climatic suitability for four main cash crops: coffee, tea, cotton, and tobacco.

The analysis proceeded in three stages. First, I validated the proxy using data from the Uganda National Panel Survey, showing that a positive shock to the price index leads to significant increases in the value of crop sales and days worked on the family farm. Second, using data on agricultural couples from the Demographic and Health Surveys (DHS), I found that these positive income shocks do not change the desire for more children but significantly increase the preferred waiting time until the next birth for both spouses. Third, I confirmed these preferences translate into behaviour by analyzing DHS birth histories. A positive price shock is associated with a longer subsequent birth-to-birth interval and a lower probability of giving birth in a given year. The magnitude of this effect is clinically relevant: one-standard deviation increase in the suitability-weighted exposure to international price of cash crops delays birth by 0.35 years, a 15% increase over the average birth interval observed in the sample (2.4 years) moving a significant mass of households out of the highest-risk window (under 2 years) into a safer temporal range.

These findings are robust to a series of tests, including randomization inference, controls for spatial correlation, and a placebo test on non-agricultural households. The results consistently suggest that for temporary income shocks, the substitution effect dominates the income effect. This analysis suggests that when agricultural prices are high, women in farming households reallocate their time toward income-generating activities, thereby postponing the next birth. A complementary survey experiment isolates a pure income effect, finding that hypothetical high-income scenarios are associated with a desire for shorter birth intervals, which further supports the interpretation that the substitution effect drives the main quasi-experimental results.

From a policy perspective, these findings suggest that agricultural development policies which boost farmers' incomes may have the beneficial secondary short-term effect of promoting healthier birth spacing. By increasing the return to female labour in agriculture, such policies can align economic objectives with public health goals for improved maternal and child well-being.

References for Chapter 1

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales (2019). "Shift-share designs: Theory and inference". In: *The Quarterly Journal of Economics* 134.4, pp. 1949–2010.

Adsera, Alicia and Alicia Menendez (2011). "Fertility changes in Latin America in periods of economic uncertainty". In: *Population studies* 65.1, pp. 37–56.

Alam, Shamma Adeeb and Claus C Pörtner (2018). "Income shocks, contraceptive use, and timing of fertility". In: *Journal of Development Economics* 131, pp. 96–103.

Arega, Gezachew Gebeyehu et al. (2024). "Spatial variation of short birth intervals and their determinant factors among reproductive women in Ethiopia using a geographically weighted regression model". In: *Frontiers in Medicine* 11, p. 1363844.

Bailey, Roy E, Timothy J Hatton, and Kris Inwood (2016). "Health, height, and the household at the turn of the twentieth century". In: *The Economic History Review* 69.1, pp. 35–53.

Balbo, Nicoletta, Francesco C Billari, and Melinda Mills (2013). "Fertility in Advanced Societies: A Review of Research: La fécondité dans les sociétés avancées: un examen des recherches". In: *European Journal of Population/Revue europeenne de demographie* 29.1, pp. 1–38.

Baranowska-Rataj, Anna, Kieron Barclay, and Martin Kolk (2017). "The effect of number of siblings on adult mortality: Evidence from Swedish registers for cohorts born between 1938 and 1972". In: *Population studies* 71.1, pp. 43–63.

Becker, Gary (1960). "An economic analysis of fertility". In: *Demographic and economic change in developed countries*. Columbia University Press, pp. 209–240.

Becker, Gary and Robert J. Barro (1988). "A Reformulation of the Economic Theory of Fertility". In: *The Quarterly Journal of Economics* 103.1, pp. 1–25. DOI: 10.2307/1882640.

Becker, Gary and H Gregg Lewis (1973). "On the interaction between the quantity and quality of children". In: *Journal of political Economy* 81.2, Part 2, S279–S288.

Berman, Nicolas et al. (2023). "Crop prices and deforestation in the tropics". In: *Journal of Environmental Economics and Management* 119, p. 102819.

Beyene, Fentahun Yenealem et al. (2025). "Effect of short inter-pregnancy interval on perinatal and maternal outcomes among pregnant women in SSA 2023: Systematic review and meta-analysis". In: *PLoS One* 20.1, e0294747.

Boberg-Fazlić, Nina et al. (2021). "Disease and fertility: Evidence from the 1918–19 influenza pandemic in Sweden". In: *Economics and Human Biology* 43. Originally published as IZA Discussion Paper No. 10834, June 2017, p. 101020. DOI: 10.1016/j.ehb.2021.101020.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2022). "Quasi-experimental shift-share research designs". In: *The Review of economic studies* 89.1, pp. 181–213.

— (2025). “A practical guide to shift-share instruments”. In: *Journal of Economic Perspectives* 39.1, pp. 181–204.

Byrnes, Rita M (1992). “Uganda: A country study”. In: Library of Congress - Country Studies.

Cameron, A Colin and Douglas L Miller (2015). “A practitioner’s guide to cluster-robust inference”. In: *Journal of human resources* 50.2, pp. 317–372.

Chatterjee, Shoumitro and Tom Vogl (2018). “Escaping Malthus: economic growth and fertility change in the developing world”. In: *American Economic Review* 108.6, pp. 1440–1467.

Cinnirella, Francesco, Marc Klemp, and Jacob Weisdorf (2017). “Malthus in the bedroom: Birth spacing as birth control in pre-transition England”. In: *Demography* 54.2, pp. 413–436.

Conde-Agudelo, Agustin, Andres Rosas-Bermudez, and Ana C. Kafury-Goeta (2012). “Birth Spacing and Risk of Adverse Perinatal Outcomes: A Meta-analysis”. In: *JAMA* 295.15, pp. 1809–1823.

Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena (2020). “Age of marriage, weather shocks, and the direction of marriage payments”. In: *Econometrica* 88.3, pp. 879–915.

Doepke, Matthias, Anne Hannusch, et al. (2023). “The economics of fertility: A new era”. In: *Handbook of the Economics of the Family*. Vol. 1. 1. Elsevier, pp. 151–254.

Doepke, Matthias, Moshe Hazan, and Yishay D. Maoz (2015). “The Baby Boom and World War II: A Macroeconomic Analysis”. In: *The Review of Economic Studies* 82.3, pp. 1031–1073. DOI: `10.1093/restud/rdv001`.

Dorélien, Audrey M (2016). “Birth seasonality in sub-Saharan Africa”. In: *Demographic Research* 34, pp. 761–796.

FAO (2025). *FAOSTAT: Crops and livestock products*. `URL`. Accessed on June 2025. Licence: CC-BY-4.0.

FAO and IIASA (2020). *Global Agro-Ecological Zones (GAEZ v4) – Data Portal User’s Guide*. Food, Agriculture Organization of the United Nations (FAO), and International Institute for Applied Systems Analysis (IIASA). Rome.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). “Bartik instruments: What, when, why, and how”. In: *American Economic Review* 110.8, pp. 2586–2624.

Heß, Simon (2017). “Randomization inference with Stata: A guide and software”. In: *The Stata Journal* 17.3, pp. 630–651.

Hill, Ruth V. and Marcella Vigneri (2014). “Mainstreaming gender sensitivity in cash crop market supply chains”. In: *Gender in agriculture: Closing the knowledge gap*. Ed. by Agnes R. Quisumbing et al. Springer.

Islam, Mohammad Zahidul et al. (2022). “Negative effects of short birth interval on child mortality in low-and middle-income countries: a systematic review and meta-analysis”. In: *Journal of global health* 12, p. 04070.

Jung, Jaehyun (2023). “Can Abortion Mitigate Transitory Shocks? Demographic Consequences under Son Preference”. In: *Journal of Human Resources*.

Kitchens, Carl T and Luke P Rodgers (2023). “The impact of the WWI agricultural boom and bust on female opportunity cost and fertility”. In: *The Economic Journal* 133.656, pp. 2978–3006.

Kosec, Katrina and Jie Song (2021). “The effects of income fluctuations on undernutrition and overnutrition across the lifecycle”. In: *Health economics* 30.10, pp. 2487–2509.

Miller, Ray and Mahesh Karra (2020). “Birth spacing and child health trajectories”. In: *Population and Development Review* 46.2, pp. 347–371.

Molitoris, Joseph, Kieron Barclay, and Martin Kolk (2019). “When and Where Birth Spacing Matters for Child Survival: An International Comparison Using the DHS”. In: *Demography* 56.4, pp. 1349–1370.

Öberg, Stefan (2015). “Sibship size and height before, during, and after the fertility decline: a test of the resource dilution hypothesis”. In: *Demographic Research* 32, pp. 29–74.

Pimentel, Juan et al. (2020). “Factors associated with short birth interval in low-and middle-income countries: a systematic review”. In: *BMC pregnancy and childbirth* 20.1, p. 156.

Ponczek, Vladimir and Andre Portela Souza (2012). “New evidence of the causal effect of family size on child quality in a developing country”. In: *Journal of Human Resources* 47.1, pp. 64–106.

Rutstein, Shea O. and Sarah Staveteig (2014). *Making the Demographic and Health Surveys Wealth Index Comparable*. Tech. rep. DHS Methodological Reports No. 9. Rockville, Maryland, USA: ICF International. URL: `https://dhsprogram.com/pubs/pdf/MR9/MR9.pdf`.

Schultz, T Paul (1997). “Demand for children in low income countries”. In: *Handbook of population and family economics* 1, pp. 349–430.

Shifti, Desalegn Markos et al. (2020). “Individual and community level determinants of short birth interval in Ethiopia: a multilevel analysis”. In: *PloS one* 15.1, e0227798.

Stantcheva, Stefanie (2023). “How to run surveys: A guide to creating your own identifying variation and revealing the invisible”. In: *Annual Review of Economics* 15.1, pp. 205–234.

Swaminathan, Akshay et al. (2020). “Association between interpregnancy interval and subsequent stillbirth in 58 low-income and middle-income countries: a retrospective analysis using Demographic and Health Surveys”. In: *The Lancet Global Health* 8.1, e113–e122.

USDA Foreign Agriculture Service (May 2025). *Coffee Annual: Uganda*. Tech. rep. UG2025-0001. Documents record-high robusta production and ongoing expansion of robusta’s market share in Uganda. Nairobi: United States Department of Agriculture.

WHO (2007). "Report of a WHO technical consultation on birth spacing: Geneva, Switzerland 13-15 June 2005". In: *Report of a WHO technical consultation on birth spacing: Geneva, Switzerland 13-15 June 2005*.

Willis, Robert J. (1973). "A New Approach to the Economic Theory of Fertility Behavior". In: *Journal of Political Economy* 81.2, S14–S64.

Chapter 2

A market-based solution to overcome information asymmetries in the Ugandan coffee value chain

with Tessa Bold, Jakob Svensson and Selene Ghisolfi[1]

Introduction

Global poverty remains one of the most significant challenges worldwide, particularly affecting rural populations in Sub-Saharan Africa. In these contexts, growth in the agricultural sector is a particularly powerful force for poverty reduction (World Bank 2007), as poorer households depend directly on agriculture for their livelihoods, and improvements in the sector quickly translate into increased incomes, employment, and local food security for these populations (Christiaensen, Demery, and Kuhl 2011; Ivanic and Martin 2018). Consequently, a market-based value chain approach has become a central pillar of pro-poor development strategy (World Bank 2019), as producer's position within the value chain is a critical determinant of the economic value they can create and capture (Alfaro-Ureña, Manelici, and Vasquez 2022; Hansman et al. 2020; Rodriguez-Clare 1996).

Smallholder farmers in low-income countries, however, typically enter GVCs at the initial stage: the provision of raw, unprocessed commodities. While subsequent stages of processing and refinement add progressively more value, these activities are predominantly performed by larger domestic exporters or firms in wealthier countries (World Bank 2019). This structural pattern perpetuates a low-income equilibrium for the farmers who form the base of these chains. A primary explanation for this phenomenon lies in

[1]Tessa Bold and Jakob Svensson are professors of Economics at the Institute for International Economic Studies of the Stockholm University; Selene Ghisolfi is Assistant Professor in Economics at Catholic University of the Sacred Hearth in Milan.

a fundamental trade-off between value creation and quality control: as farmers perform post-harvest processing to transform raw agricultural products into higher-value goods the underlying quality of the product often becomes more difficult for buyers to observe. Asymmetrical information penalizes producers of high-quality goods, erodes incentives for quality-enhancing investments, and can lead to a market dominated by low-quality trade (Akerlof 1970).

This paper documents the phenomenon of the *market for lemons* in the coffee value chain using data from a field experiment with farmers in Ibanda (Uganda) and presents a market-based solution to address the asymmetrical information in the value chain. Based on a sketched conceptual framework, we develop some hypothesis about the consequences of the asymmetric information in the coffee value chain and then test them empirically on sales-level data collected on the field. Results show that higher quality is rewarded only in the market of the unprocessed good (cherries) and that batches of cherries of lower quality are more likely to be sold as processed good (beans); consequently, lower relational contracts are more frequent in the cherry market (where quality can be rewarded) and cherry sellers are less likely to feel the price they receive for their coffee is unfair. We then propose a market-based solution to overcome this asymmetric information problem.

This paper contributes to three strands of literature. First, we add to the evidence on quality upgrading in agricultural value chains. While the potential for poverty reduction through value-addition is well documented (Bold et al. 2022), we provide empirical evidence of a specific structural barrier: the trade-off between value creation and observability. Recent evidence from agricultural markets in developing countries documents this dynamic across various contexts (such as wheat markets in Ethiopia (Do Nascimento Miguel 2024), dairy products (Rao and Shenoy 2023), food safety in general (Hoffmann et al. 2023; Nindi, Bauchet, and Ricker-Gilbert 2023)) and a study in the Ugandan coffee sector specifically finds that processing can be associated with lower farm-gate returns when quality cannot be observed, producing a "value degrading" effect for farmers (Arslan, Gregg, and Wollni 2024).

In the second place, we contribute to the literature on information asymmetries and relational contracts. Vertical integration and relational contracts attempt to solve the lemons problem within value chains: in Rwanda's coffee sector, for instance, relational contracts between coffee mills and farmers allow mills to provide critical inputs and training in exchange for quality and loyalty (Vrolijk 2023). Through integration with these mills, powerful exporters can drive quality upgrading and improve prices (Atkin, Khandelwal, and Osman 2017). However foreign ownership may limit domestic value capture (Rodriguez-Clare 1996), while domestic owners face social pressures that hinder efficiency (Macchiavello and Morjaria 2022). Furthermore, since the value of these arrangements relies on relationship stability (Macchiavello 2022; Macchiavello and Morjaria 2015) to bundle premiums with extension services (Macchiavello and Miquel-Florensa 2019), restrictions on integration can inadvertently harm farmers (Leone et al. 2025). We complement this literature by documenting the limits of these mechanisms. We find that while relational contracts exist in the observable cherry market, they struggle to sustain cooperation in the opaque bean market. Similar alternative solutions such as quality certifications often fail to transmit price premiums back to the producer (Jena, Stellmacher, and Grote 2017; Minten et al. 2018), highlighting the persistence of the market failure.

Finally, we introduce a novel, market-based solution using low-cost technology to address information asymmetry by making quality verifiable at the point of transaction. While technically feasible in laboratory settings (Chen, Chiu, and Zou 2022; García, Candelo-Becerra, and Hoyos 2019), these tools must be deployable in low-infrastructure environments to bridge the gap between potential and reality. We propose and pilot a smartphone-based verification tool.

The remainder of this paper is organized as follows. Section 1 provides background on value creation and quality uncertainty and presents the research setting and data. Section 2 outlines the empirical strategy and presents the main results. Section 3 describes our proposed solution.

1 Research setting and data

1.1 The problem

Processing can obscure the information required for quality-based trade and thus lower the engagement of farmers in value-added activities. Coffee provides a clear illustration of this dilemma: a farmer can either sell fresh, unprocessed coffee cherries or undertake processing - drying, sorting, and dehulling - to produce green coffee beans, a higher-value product. The quality of unprocessed cherries is relatively easy for a buyer to assess visually: a batch can be quickly evaluated for ripeness (indicated by a uniform, deep red color) and the absence of physical defects and buyers can thus confidently pay a premium for high-quality raw material.

Once the cherry's pulp is removed and the bean is dried, crucial visual cues about the original fruit's ripeness are lost. In a typical field setting, it is nearly impossible to determine with certainty whether a given sample of green beans originated from exclusively ripe cherries or from a mix that included unripe or spoiled ones.

The information asymmetry fundamentally alters market dynamics. Rational, risk-averse buyers, unable to distinguish quality, will offer a price that reflects the average quality they expect to find in the market. Consequently, the incentive for farmers to invest the extra effort and resources required to produce high-quality processed goods is severely diminished. This creates a market characterized by adverse selection, where low-quality products are more likely to be traded, potentially driving high-quality producers

Figure 1. Arabica cherry color spectrum at different stages of ripeness. The highlighted cherries have the perfect level of ripeness.

out of the market entirely.

1.2 Setting

The project is located in the Ibanda District, in southwestern Uganda. Our study population comprises over 600 coffee-farming households distributed across 22 villages. These villages are situated at elevations between 1500 and 1800 meters above sea level, ideal for cultivating Arabica coffee. The farmers are smallholders, cultivating an average of 1.2 acres of coffee, often intercropped with banana trees. Most households rely on subsistence farming to survive, and poverty is prevalent. Previous research in Uganda has highlighted the challenges farmers face in adopting new technologies and accessing profitable markets (Kassie, Shiferaw, and Muricho 2011; Kijima, Otsuka, and Sserunkuuma 2011). Only one in five households in our sample has access to electricity, and one in ten reports consuming only one meal per day.

(a) A coffee plot in Ibanda (credits to Amos)

(b) A map of the villages included in the survey. Source: Google Earth.

Figure 2. The setting of the study.

Variable	2021	2023	2024	2025
Observations	711	330	199	200
Avg arabica acres	1.264	2.938	1.217	0.920
Avg score in quality quiz (out of 5)	2.56	3.02	2.91	2.95
Price unfair	.	.	0.73	0.55
No primary education	0.29	0.36	0.39	0.41
Primary education	0.29	0.44	0.43	0.40
Secondary education	0.13	0.18	0.17	0.18
Sold cherries at least once	0.25	0.26	0.13	0.09
Sold beans at least once	0.74	0.78	0.89	0.92

Table 1. This table presents sample means for key farmer and household characteristics, disaggregated by survey year. The sample includes baseline survey data from 2021 and 2023, and control group data from the 2024 and 2025 follow-up surveys.

	Cherries				Beans			
	2021	2023	2024	2025	2021	2023	2024	2025
Observations	455	147	74	64	756	296	301	376
Avg price (UGX/kg)	939.8	1347.6	1792.6	2219.5	5998.0	8433.2	10958.5	14623.4
Avg quantity (kg)	53.7	67.1	86.6	100.2	111.6	55.9	144.2	112.8
Avg processing costs (UGX/kg)	27.9	0.0	5.2	34.9	68.5	29.6	18.2	35.1
Avg profits (UGX/kg)	891.5	1328.6	1787.5	2550.3	5920.5	8286.2	10689.2	14928.8
Avg share of red cherries (%)	91.1	83.7	96.8	97.6	90.3	85.5	93.9	94.0
Post-harvest cleaning share	0.25	0.27	0.30	0.28	0.01	0.04	0.20	0.05
Share of sales checked	0.81	0.73	0.76	0.50	0.76	0.78	0.71	0.63
Issues found	0.21	0.43	0.32	0.55	0.10	0.39	0.34	0.42
Sold to usual buyer	0.79	0.68	0.51	0.66	0.54	0.56	0.51	0.49

Table 2. This table reports sample means for coffee sale transactions, disaggregated by product type (cherries or beans) and survey year. Prices, costs, and profits are reported in Ugandan Shillings (UGX). The share of red cherries is self-reported by farmers for each specific sale.

1.3 Data

The analysis relies on primary data collected from smallholder coffee farmers in the Ibanda District. The dataset is constructed as an unbalanced panel covering four harvest seasons: 2021, 2023, 2024, and 2025. Baseline data were collected in 2021 and 2023, while starting from 2024 a subset of farmer was part of an intervention aimed at creating a market for high quality coffee. The randomization was carried out within each village and for the purpose of this analysis we retain only the observations from the baseline waves and the control group observations from the follow-up surveys of 2024 and 2025.

1.3.1 Sampling and timeline

The study population comprises over 600 coffee-farming households distributed across 22 villages. The baseline sample was established in the 2021 and 2023 waves. For the 2024 and 2025 waves, we retained a subsample of approximately 200 households. These households constitute the pure control group of a randomized controlled trial designed to evaluate the technological solution described in section 3. As such, their production and marketing conditions reflect the status quo, unaffected by external interventions. Surveys were administered at the end of an harvest season (spring 2021, fall 2023, spring 2024 and spring 2025) to capture realized production and sales outcomes for that specific period.

1.3.2 Key variables

Sales data Each survey wave includes a fixed core questionnaire that captures households' demographics, agricultural activities, and coffee-specific outcomes. For every individual sale made during the harvest season, farmers report the processing stage of the coffee sold alongside the exact quantity sold (in kilograms) and the price received per kilogram. They also report key transaction characteristics, including the identity of the buyer (for example, whether they are the farmer's usual buyer), whether any quality inspection was performed at the point of sale, and any immediate monetary costs for that transaction, such as transport, dehulling fees, and other harvesting and post-harvest expenses (for instance, hiring costs). In addition, the survey elicits coffee quality knowledge and practices by assessing the farmer's understanding and implementation of quality-enhancing techniques (practices used to ensure high coffee quality, such as selective harvesting of red cherries and proper drying). The analysis is restricted to sales of red cherries and green beans, which together account for 99% of coffee sales in our sample. Table 3 reports summary statistics of farmer-level variables by year, while Table 4 presents summary statistics of sale-level variables by year and processing level.

Measuring coffee quality Coffee quality is a multidimensional concept shaped by both agronomic conditions and post-harvest processes. In this study, we focus on the ripeness of a coffee cherry at the time of harvest as a key observable indicator of intrinsic bean quality. Bean flavour is strongly correlated with ripeness: unripe cherries tend to produce bitter coffee due to underdeveloped sugars, whereas ripe cherries yield higher sweetness and aromatic quality. Consequently, cherry color at harvest serves as an important proxy for expected cup quality and market value.

Ripeness is also a dimension of quality over which farmers retain direct control: through increased effort and attentiveness during harvesting, farmers can selectively pick only ripe cherries. Unlike structural constraints such as soil conditions or altitude, which evolve slowly and are largely beyond farmers' control, harvesting practices are likely to be responsive to short-term economic incentives. If price premiums reward higher-quality beans, farmers can adjust labour allocation to improve ripeness selection; this responsiveness makes ripeness an ideal starting point for a market, quality-based incentive design. We measure this using a self-reported continuous variable ranging from 0 to 100, representing the percentage of fully ripe cherries in the batch harvested for a specific sale. To validate the robustness of this measure, we also collect data on farmers' agronomic knowledge via a visual quiz, assessing their ability to identify quality defects and proper processing techniques. Additionally, we construct binary indicators for "high quality" (if the share of red cherries exceeds the sample median of 90%) and for whether the farmer performed additional post-harvest cleaning (sorting) to remove defects.

Variable	2021	2023	2024	2025
Observations	711	330	199	200
Avg arabica acres	1.264	2.938	1.217	0.920
Avg score in quality quiz (out of 5)	2.56	3.02	2.91	2.95
Price unfair	.	.	0.73	0.55
No primary education	0.29	0.36	0.39	0.41
Primary education	0.29	0.44	0.43	0.40
Secondary education	0.13	0.18	0.17	0.18
Sold cherries at least once	0.25	0.26	0.13	0.09
Sold beans at least once	0.74	0.78	0.89	0.92

Table 3. This table presents sample means for key farmer and household characteristics, disaggregated by survey year. The sample includes baseline survey data from 2021 and 2023, and control group data from the 2024 and 2025 follow-up surveys.

	Cherries				Beans			
	2021	2023	2024	2025	2021	2023	2024	2025
Observations	455	147	74	64	756	296	301	376
Avg price (UGX/kg)	939.8	1347.6	1792.6	2219.5	5998.0	8433.2	10958.5	14623.4
Avg quantity (kg)	53.7	67.1	86.6	100.2	111.6	55.9	144.2	112.8
Avg processing costs (UGX/kg)	27.9	0.0	5.2	34.9	68.5	29.6	18.2	35.1
Avg profits (UGX/kg)	891.5	1328.6	1787.5	2550.3	5920.5	8286.2	10689.2	14928.8
Avg share of red cherries (%)	91.1	83.7	96.8	97.6	90.3	85.5	93.9	94.0
Post-harvest cleaning share	0.25	0.27	0.30	0.28	0.01	0.04	0.20	0.05
Share of sales checked	0.81	0.73	0.76	0.50	0.76	0.78	0.71	0.63
Issues found	0.21	0.43	0.32	0.55	0.10	0.39	0.34	0.42
Sold to usual buyer	0.79	0.68	0.51	0.66	0.54	0.56	0.51	0.49

Table 4. This table reports sample means for coffee sale transactions, disaggregated by product type (cherries or beans) and survey year. Prices, costs, and profits are reported in Ugandan Shillings (UGX). The share of red cherries is self-reported by farmers for each specific sale.

2 Conceptual framework, empirical specifications and results

2.1 Conceptual Framework

What is the impact of the trade-off between value adding and quality screening in the early steps of the coffee value chain?

1. **Disappearing quality premia in the market for the processed good.** The statistical relationship between a farmer's production quality and the price they receive is strong and positive for fresh cherries sales (where quality is observable)

but weak or non-existent for processed green beans sales (where quality becomes unobservable).

2. **Adverse selection in the market for processed goods.** Batches of cherries with lower quality and farmers who perform poorer pre-harvest practices will be more present in the market for beans with respect to that of cherries.

3. **Quality checks, relational contracts and farmers perception.** Quality checks are more present in the unprocessed good market, and, for a given level of quality of the harvest, issues are more frequently detected in the cherries. Relational contracts are more frequent where quality is verifiable and rewarded. Farmers who sell higher quality processed goods will perceive the market as unfair.

2.2 Empirical Specifications and results

The core of the analysis is a series of regressions designed to test for the presence and characteristics of a "market for lemons" based on the implications outlined in the conceptual framework sketched in section 2.

2.2.1 Disappearing quality premia in the market for the processed good

In order to find if higher quality of the product (proxied here by the share of red cherries harvested by the farmer to be sold as cherries or be processed and sold as green beans) influences the price received, we estimate the following regression separately for cherries and beans sales:

$$\text{price per kg}_{i,j,ea,t} = \beta_0 + \beta_1 \text{red cherries}_{i,j,ea,t} + \text{kg sold}_{i,j,ea,t} + \text{controls}_{j,ea,t} + \delta_{ea} + \gamma_t + \epsilon_{i,j,ea,t}$$

where the outcome is the price per kilogram of cherries or beans sold (or the profit per kg, net of other costs such as transportation and dehulling costs) in sale i of farmer j living in village ea in year t; $\textit{red cherries}_{i,j,ea,t}$ is the self-reported share of red cherries out of the total harvested for that sale; $\textit{kg sold}_{i,j,ea,t}$ is the amount of either cherries or beans sold, $\textit{controls}_{j,ea,t}$ include acres of land devoted by the farmer to cultivation of arabica coffee and level of knowledge of coffee practices (as measured by our quiz), while δ_{ea} and γ_t are enumeration area and year fixed effects. Standard errors are clustered at household level.

Higher quality of harvested cherries is associated to higher prices for farmers in the market for unprocessed goods (cherries) but not in the market of processed ones (beans). The coefficient implies that a 10 percentage point increase in the share of red cherries corresponds to an increase of 3% of the price per kg. It is important to note that this null

	Cherries			Beans		
	(1) price/kg	(2) sprice/kg	(3) profit/kg	(4) price/kg	(5) price/kg	(6) profit/kg
Share red cherries harvested (%)	2.828**	2.828**	2.903**	4.031	4.031	4.233
	(1.343)	(1.343)	(1.213)	(4.016)	(4.016)	(4.120)
Quality quiz score	-16.573	-16.573	-21.094	-69.603	-69.603	-64.961
	(30.412)	(30.412)	(29.865)	(61.402)	(61.402)	(62.017)
Arabica acres	1.470***	1.470***	1.432***	0.060	0.060	0.083
	(0.435)	(0.435)	(0.413)	(2.666)	(2.666)	(2.649)
Kg sold	0.618***	0.618***	0.551***	0.985**	0.985**	1.021**
	(0.201)	(0.201)	(0.163)	(0.491)	(0.491)	(0.500)
Enumeration area FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
# villages	18	18	18	22	22	22
# years	4	4	4	4	4	4
dep. var. mean	1092.332	1092.332	1065.408	9444.288	9444.288	9407.036
R^2	0.843	0.843	0.843	0.849	0.849	0.850
observations	473	473	467	1580	1580	1577

s. e. in parentheses, clustered at the household level

* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 5. This table reports OLS estimates of the relationship between coffee quality (self-reported % of red cherries) and the sale price/profit per kg. Columns (1)-(3) present results for unprocessed cherry sales, while columns (4)-(6) present results for processed bean sales.

result holds for both the gross price and the net profit; this rules out the hypothesis that the disappearance of the premium is simply due to margin compression (where higher input costs for processed beans offset a price premium).

2.2.2 Adverse selection in the market for processed goods

To investigate if the quality of cherries sold as processed or unprocessed good is different we estimate the equation

$$\text{red cherries}_{i,j,ea,t} = \beta_0 + \beta_1 \text{high quality}_{i,j,ea,t} + \text{kg sold}_{i,j,ea,t} + \text{controls}_{j,ea,t} + \delta_{ea} + \gamma_t + \epsilon_{i,j,ea,t}$$

where *red cherries*$_{i,j,ea,t}$ is a dummy equal to 1 if the sale i of farmer j was a cherry sale (0 if it was a beans sale), and *high quality*$_{i,j,ea,t}$ is either a dummy equal to 1 if the share of red cherries harvested by the farmer for that sale is above the median (90%) or a dummy equal to 1 if additional post-harvest cleaning (removing defective cherries, or sticks and stones) was carried out by the farmers for the coffee sold in that sale.

Batches of cherries whose quality are above the median are more 5% more likely to be sold as cherries, while those where farmers carry out additional cleaning (to remove unripe cherries or waste) are 14% more likely to be sold as cherries.

	Sold as red cherries		
	(1)	(2)	(3)
Share of red cherries above the median	0.050**		0.048**
	(0.022)		(0.022)
Additional cleaning		0.145***	0.143***
		(0.053)	(0.052)
Kg harvested	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Arabica acres	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Quality quiz score	-0.020	-0.021	-0.020
	(0.020)	(0.020)	(0.020)
Village FE	yes	yes	yes
Year FE	yes	yes	yes
# villages	22	22	22
# years	4	4	4
dep. var. mean	0.229	0.229	0.229
R^2	0.330	0.333	0.337
observations	2045	2046	2045

s. e. in parentheses, clustered at the household level
* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 6. This table presents OLS estimates testing for adverse selection. The dependent variable is a dummy equal to 1 if the sale was of unprocessed cherries and 0 if it was of processed beans. The key independent variables are indicators for high quality: whether the share of red cherries was above the sample median (90%) and whether the farmer performed additional post-harvest cleaning.

2.2.3 Quality checks, relational contracts and farmers perception

For each sale we also ask if the buyer has received any quality check of the coffee, if any issue was found, and why he or she sold to that buyer. We estimate the equation

$$\text{outcome}_{i,j,ea,t} = \beta_0 + \beta_1 \text{red cherries}_{i,j,ea,t} + \text{kg sold}_{i,j,ea,t} + \text{usual buyer}_{i,j,ea,t} + \text{controls}_{j,ea,t} + \delta_{ea} + \gamma_t + \epsilon_{i,j,ea,t}$$

where $outcome_{i,j,ea,t}$ is one of the three dummies *quality check*, i.e. whether the buyer in the sale i of farmer j checked the quality of the coffee (cherries or beans) before buying it, *issues*, i.e. whether issues were found, and *usual buyer*, i. e. whether the farmer sold to that buyer because he or she was the usual one - the latter variable is also used as a control for the first two outcomes. The variable $cherries_{i,j,ea,t}$ is a dummy equal to 1 if the sale was a cherry sale and 1 if it was a beans one.

The results in table 7 show that cherry sales are more likely to be completed with a *usual buyer*, coherently with idea that relational contracts are more easily established when quality is observable. It is not more likely to have a quality check if you sell cherries with respect to beans, but conditional on quality (share of red cherries harvested) cherry batches are more likely to be found with quality issues, consistently with cherry quality being more observable than quality of beans.

	(1) usual buyer	(2) quality check	(3) issues
Red cherries	0.154***	-0.046	0.077**
	(0.044)	(0.040)	(0.039)
Share of red cherries (%)	-0.003**	0.001	-0.004**
	(0.001)	(0.001)	(0.001)
Arabica acres	0.002***	0.001***	-0.000
	(0.000)	(0.000)	(0.001)
Quality quiz score	-0.022	0.076***	-0.110***
	(0.019)	(0.019)	(0.017)
Kg sold	-0.000	0.000**	-0.000
	(0.000)	(0.000)	(0.000)
Usual buyer		-0.052**	0.014
		(0.026)	(0.022)
Village FE	yes	yes	yes
Year FE	yes	yes	yes
# villages	22	22	22
# years	4	4	4
dep. var. mean	0.555	0.725	0.288
R^2	0.117	0.055	0.138
observations	2054	1985	2054

s. e. in parentheses, clustered at the household level
* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 7. This table reports OLS estimates on the characteristics of sale transactions. The dependent variables are dummies for whether the sale was to a usual buyer (column 1), whether a quality check was performed (column 2), and whether quality issues were found (column 3). The main independent variable is a dummy equal to 1 for cherry sales.

In 2024 and 2025 we also ask farmers if they felt the price they received in that season for their coffee was higher/lower or in line for the quality of the coffee they sold. Based on the answer we define the dummy *price unfair* which is equal to 1 if the farmer felt that the price she received was unfair for the quality of the coffee she sold. We then estimate the equation:

$$\text{price unfair}_{j,ea,t} = \beta_0 + \beta_1 \text{cherry seller}_{j,ea,t} + \text{mix seller}_{j,ea,t} + \text{monitoring}_{j,ea,t} + \delta_{ea} + \gamma_t + \epsilon_{j,ea,t}$$

where *cherry seller* and *bean seller* are two dummies equal to 1 if the farmer sells only cherries or sells a mix of cherries and beans. Farmers that sell only cherries are 25% less likely to answer that the price they receive is unfair considering the quality the quality of the coffee they sell.

2.3 Limitations

This study has some limitations. First, our primary measure of coffee quality, the percentage of red cherries in a harvested batch, is self-reported by farmers. This reliance on

	price unfair	
	(1)	(2)
Cherry seller	-0.254***	-0.234**
	(0.093)	(0.093)
Mix seller	0.023	0.031
	(0.064)	(0.062)
Monitors local prices		-0.062
		(0.059)
Monitors international prices		0.254**
		(0.123)
Village FE	yes	yes
Year FE	yes	yes
# villages	22	22
# years	2	2
dep. var. mean	0.641	0.641
R^2	0.076	0.084
observations	387	387

s. e. in parentheses, clustered at the household level
* $p < 0.10$, ** $p < 0.05$, $^{***}p < 0.01$

Table 8. This table reports OLS estimates from the 2024 and 2025 survey waves. The dependent variable is a dummy equal to 1 if the farmer reported the price received as "unfair" for the quality sold. The main independent variables are dummies indicating whether a farmer sells only cherries or a mix of cherries and beans during a season (the omitted category is bean-only sellers).

self-reported data may introduce measurement error due to recall bias or social desirability bias. While we have no *a priori* reason to believe that such reporting errors would be systematically different for sales of cherries versus beans, objective, third-party quality verification for each transaction would provide a more robust validation of our findings. Second, the empirical analysis is primarily descriptive and only correlational; the OLS regressions document patterns that are consistent with the predictions of information asymmetry theory and the goal of the analysis is to provide a diagnosis of a market failure rather than to isolate a specific causal parameter. Third, the external validity of our findings warrants careful consideration. This study is situated in the specific context of Ibanda; in regions with strong, well-managed farmer cooperatives that perform collective processing and marketing, the information asymmetry problem between the farmer and the first buyer might be less severe. Consequently, while the underlying economic principle of information asymmetry is universal, the specific empirical patterns we observe may not generalize to contexts with different market structures or institutional safeguards.

3 A market-based solution

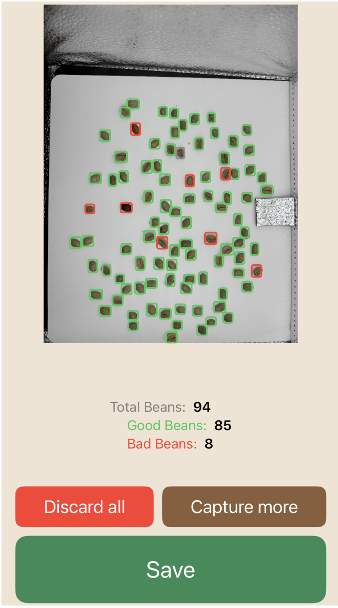
To overcome the trade-off between value creation and and quality detection we propose a solution based on (i) farmers' knowledge about best practices of coffee farmers (ii) clear understanding of the impact of farmers' harvesting choices on beans quality (iii) reliable and scalable detection of beans quality for buyers.

Knowledge about quality practices in all phases of coffee production is spread out at village level. Within each group, farmers nominate a peer to monitor compliance and distribute educational material, reinforcing incentives to adopt a quality-enhancing behavior. Along with correct cherry-picking, good post-harvest processing reduces the prevalence of defects such as sourness, fungality, and breakage (Huang et al. 2025).

Information about how good harvest and post-harvest activities translate into good quality coffee beans was acquired empirically, building a dataset that links the observable properties of green beans to the original cherry color at harvest. We procured samples of red and green cherries, mixed them in batches with various proportion of green cherries (100% red, 95% red + 5% green and so on), dried them using conventional methods to produce green beans and then drew repeated 300-gram samples (approximately 2,000 beans) from each batch. Each bean was then classified through laboratory inspection following standard coffee grading criteria and labelled as bad (i.e. deformed, broken, insect-damaged, black, partially black, brown) or good. Thanks to this analysis we defined critical thresholds of defective beans that, if reliably detected, instruct buyers on the quality of the bag the farmer is trying to sell and discriminate between bags of beans coming from 100% red cherries (to buy) and those coming from mixed cherries (to reject).

Quality detection is performed through a smartphone application called *Beaney*. The application utilizes a two-stage pipeline that occurs entirely on-device: a YOLO-based detector model identifies all beans within a photo, and a FastViT-based classifier assigns each detected bean to a quality class (good or bad). Both methods have been extensively employed in the literature of applications of computer vision to the classification of single arabica coffee beans (Adiwijaya, Sarno, and Wijaya 2024; Manansala and Paglinawan 2024; Nugroho et al. 2025). Automated classification enables the application to calculate bean quality distributions, count defective beans, output color-coded overlays for immediate interpretation by buyers in the field and ultimately decide whether to buy a coffee bag with a price-premium for high quality or not. The app is used together with a portable photo box, that helps to evaluate coffee beans in consistent light conditions (figure 3 presents a screenshot of the app and a picture of the photo box).

(a) A photo box filled up with beans to classify.



(b) Beans classification through Beaney.

Figure 3

Conclusion

This paper documents a classic example of market failure in a context where its consequences are profound: the smallholder coffee sector in Uganda. We provide evidence that the initial steps of the coffee value chain are plagued by a market for lemons: when coffee cherries are sold fresh, their quality is observable, and farmers are rewarded with price premia for higher-quality harvests. However, once farmers invest in processing to create green coffee beans, a higher-value product, the underlying quality becomes difficult for buyers to observe; the information asymmetry leads to the erosion of quality premia and creates adverse selection, where lower-quality products are more likely sold in the processed market. This dynamic traps farmers in a low-value equilibrium, penalizing their efforts to improve quality and capture more of the value they create. To address this market failure, we propose a scalable, market-based solution centered on *Beaney*, a smartphone application that enables low-cost, on-the-spot quality verification of green coffee beans. By making quality transparent at the point of sale, such a tool can restore the link between quality and price, re-establishing incentives for farmers to invest in quality-enhancing practices in the market for coffee beans, improving their position in the coffee value chain and consequently their income.

References for Chapter 2

Adiwijaya, Nelly Oktavia, Riyanarto Sarno, and Dedy Rahman Wijaya (2024). "Coffee Defects Detection Based on Green Bean Images Using YOLO Architecture". In: *2024 IEEE 2nd International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT)*. IEEE, pp. 314–319.

Akerlof, George A. (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism". In: *The Quarterly Journal of Economics* 84.3, pp. 488–500. DOI: 10.2307/1879431.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez (Aug. 2022). "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages*". In: *The Quarterly Journal of Economics* 137.3, pp. 1495–1552. ISSN: 0033-5533. DOI: 10.1093/qje/qjac006.

Arslan, Cansın, Daniel Gregg, and Meike Wollni (2024). "Paying more to make less: value degrading in the coffee value chain in eastern Uganda". en. In: *American Journal of Agricultural Economics* 106.1, pp. 96–117. DOI: 10.1111/ajae.12389.

Atkin, David, Amit K. Khandelwal, and Adam Osman (May 2017). "Exporting and Firm Performance: Evidence from a Randomized Experiment*". In: *The Quarterly Journal of Economics* 132.2, pp. 551–615. ISSN: 0033-5533. DOI: 10.1093/qje/qjx002.

Bold, Tessa et al. (2022). "Market access and quality upgrading: Evidence from four field experiments". In: *American Economic Review* 112.8, pp. 2518–52.

Chen, Shih-Yu, Ming-Feng Chiu, and Xue-Wei Zou (2022). "Real-time defect inspection of green coffee beans using NIR snapshot hyperspectral imaging". In: *Computers and Electronics in Agriculture* 197, p. 106970.

Christiaensen, Luc, Lionel Demery, and Jesper Kuhl (Nov. 2011). "The (evolving) role of agriculture in poverty reduction—An empirical perspective". In: *Journal of Development Economics* 96.2, pp. 239–254. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2010.10.006.

Do Nascimento Miguel, Jérémy (Oct. 2024). "Returns to quality in rural agricultural markets: Evidence from wheat markets in Ethiopia". In: *Journal of Development Economics* 171, p. 103336. ISSN: 0304-3878. DOI: 10.1016/j.jdeveco.2024.103336.

García, Mauricio, John E Candelo-Becerra, and Fredy E Hoyos (2019). "Quality and defect inspection of green coffee beans using a computer vision system". In: *applied sciences* 9.19, p. 4195.

Hansman, Christopher et al. (2020). "Vertical integration, supplier behavior, and quality upgrading among exporters". In: *Journal of Political Economy* 128.9, pp. 3570–3625.

Hoffmann, Vivian et al. (2023). "Upside risk, consumption value, and market returns to food safety". en. In: *American Journal of Agricultural Economics* 105.3, pp. 914–939. DOI: 10.1111/ajae.12349.

Huang, Guanru et al. (Sept. 2025). "Effects of Harvesting Periods and Cultivar on the Physicochemical and Sensory Properties of Two Coffee Bean Varieties". In: *Foods* 14.17, p. 3135. ISSN: 2304-8158. DOI: 10.3390/foods14173135. URL: http://dx.doi.org/10.3390/foods14173135.

Ivanic, Maros and Will Martin (Sept. 2018). "Sectoral Productivity Growth and Poverty Reduction: National and Global Impacts". In: *World Development* 109, pp. 429–439. ISSN: 0305-750X. DOI: 10.1016/j.worlddev.2017.07.004.

Jena, Pradyot Ranjan, Till Stellmacher, and Ulrike Grote (2017). "Can coffee certification schemes increase incomes of smallholder farmers? Evidence from Jinotega, Nicaragua". In: *Environment, Development and Sustainability* 19.1, pp. 45–66.

Kassie, Menale, Bekele Shiferaw, and Geoffrey Muricho (Oct. 2011). "Agricultural Technology, Crop Income, and Poverty Alleviation in Uganda". In: *World Development* 39.10, pp. 1784–1795. ISSN: 0305-750X. DOI: 10.1016/j.worlddev.2011.04.023.

Kijima, Yoko, Keijiro Otsuka, and Dick Sserunkuuma (Jan. 2011). "An Inquiry into Constraints on a Green Revolution in Sub-Saharan Africa: The Case of NERICA Rice in Uganda". In: *World Development* 39.1, pp. 77–86. ISSN: 0305-750X. DOI: 10.1016/j.worlddev.2010.06.010.

Leone, Fabrizio et al. (2025). "Market Structure, Vertical Integration and Farmers' Welfare in Costa Rica Coffee Industry". In.

Macchiavello, Rocco (2022). "Relational contracts and development". In: *Annual Review of Economics* 14.1, pp. 337–362.

Macchiavello, Rocco and Josepa Miquel-Florensa (Aug. 2019). *Buyer-Driven Upgrading in GVCS: The Sustainable Quality Program in Colombia*. en. SSRN Scholarly Paper. Rochester, NY.

Macchiavello, Rocco and Ameet Morjaria (Sept. 2015). "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports". en. In: *American Economic Review* 105.9, pp. 2911–2945. ISSN: 0002-8282. DOI: 10.1257/aer.20120141.

— (2022). *Acquisitions, management, and efficiency in Rwanda's coffee industry*. Tech. rep. National Bureau of Economic Research.

Manansala, Angelica J and Engr Charmaine C Paglinawan (2024). "Classification of Coffea Liberica Quality Using Convolution Neural Networks (Slim-CNN, YOLOv5, and VGG-16)". In: *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, pp. 1–6.

Minten, Bart et al. (2018). "Tracking the quality premium of certified coffee: evidence from Ethiopia". In: *World Development* 101, pp. 119–132.

Nindi, Tabitha Charles, Jonathan Bauchet, and Jacob Ricker-Gilbert (Dec. 2023). "Information and the trade-off between food safety and food security in rural markets: Experimental evidence from Malawi". In: *Journal of Economic Behavior & Organization* 216, pp. 170–186. ISSN: 0167-2681. DOI: 10.1016/j.jebo.2023.10.022.

Nugroho, Eko Dwi et al. (2025). “Development of YOLO-Based Mobile Application for Detection of Defect Types in Robusta Coffee Beans”. In: *Journal of Applied Informatics and Computing* 9.1, pp. 153–160.

Rao, Manaswini and Ashish Shenoy (Aug. 2023). “Got (clean) milk? Organization, incentives, and management in Indian dairy cooperatives”. In: *Journal of Economic Behavior & Organization* 212, pp. 708–722. ISSN: 0167-2681. DOI: 10.1016/j.jebo.2023.06.002.

Rodriguez-Clare, Andres (1996). “Multinationals, Linkages, and Economic Development”. In: *American Economic Review* 86.4, pp. 852–873.

Vrolijk, Kasper (2023). *Relational contracts in the Rwandan coffee chain*. Tech. rep. IDOS Policy Brief.

World Bank, The (2007). *World Development Report 2008: Agriculture for Development*. World Bank Group.

— (2019). *World development report 2020: Trading for development in the age of global value chains*. World Bank Group.