Constraints and Opportunities in Low- and Middle-Income Countries: Essays on Youth Fertility, Technology Adoption, and Credit Access

A dissertation submitted by

Leticia Cristina Donoso Peña

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics and Public Policy

Tufts University

May 2025

Advisers: Jenny Aker, Eoin McGuirk, Margaret McMillan

Abstract

This dissertation examines constrained settings in low- and middle-income countries and reveals opportunities for improving key economic outcomes in human capital accumulation, access to credit, and technology adoption. The first chapter studies the impact of teenage childbearing on education and labor force participation (LFP) in Ecuador—a country that, despite decades of policy efforts, consistently has some of the highest adolescent fertility rates in Latin America. Using a nationally representative survey on women's health, I exploit variation in menarche timing (the age of first menstrual period) as an instrument for the probability of teenage childbearing. I find no significant differences in educational attainment and LFP between teen mothers and other women in the sample, ruling out large negative effects. The results are not driven by differences across age cohorts or location; however, I find suggestive evidence that grandmother cohabitation can moderate these effects, allowing young mothers to continue their education and offering insights into potential pathways for human capital accumulation in a constrained setting.

In the second chapter (with Oscar Mitnik, Edgar Salgado, and Alejandro Tamola), I study how receiving a productive loan affects a business owner's future performance in the formal credit market, in the context of a credit supply expansion by El Salvador's Development Bank, Bandesal. We obtain individual credit bureau records for 80% of the loan recipients, both before and after receiving this publicly financed loan. Using a difference-in-differences strategy with multiple time periods and staggered treatment adoption, we compare cohorts of business owners who received their first Bandesal-funded loan at the beginning of the credit expansion to those who received it months later. We find that loan recipients are more likely to obtain new, high-quality loans, although their probability of default increases slightly. The effects appear to be driven mostly by first-time clients who had no credit history 24 months before receiving the loan, providing empirical evidence of how alleviating credit constraints can foster financial graduation for individual business owners.

In the third chapter (with Jenny C. Aker, Brian Dillon, and Anne Krahn), I evaluate whether willingness-to-pay (WTP) experiments—commonly used to elicit demand for a variety of products—generate persistent treatment effects over time. We address this question using a randomized controlled trial of a WTP experiment, combined with in-person and phone survey data over a four-year period. We find that a simple experiment leads to positive and persistent effects on both the adoption and usage of an improved storage technology, as well as the disadoption of traditional technologies. These results are primarily driven by households that experienced the product directly, rather than by informational effects or increased salience. Our findings suggest that failing to account for demand-elicitation experiments conducted at baseline may affect the external validity of the broader experiments in which

they are embedded. By measuring demand in a random subsample, researchers can test whether WTP elicitation affects downstream outcomes.

Acknowledgments

Many people have contributed to making this work possible.

First, I would like to thank my advisers, Professors Jenny Aker, Eoin McGuirk, and Margaret McMillan. Their technical expertise, generous availability, and unwavering support—delivered with patience and kindness—have been instrumental in both my training and the results of my work.

I also thank members of the Tufts faculty and NFEPP students for their comments and suggestions during seminar presentations and office hours, especially Professors Cynthia Kinnan, Kyle Emerick, Matthew Gudgeon, Cheryl Doss, Doug Gollin, Melissa McInerney, and Jeffrey Zabel, as well as my fellow peer and roommate Lorena Moreno.

I am deeply grateful as well for valuable feedback from colleagues outside Tufts, particularly Ivette Contreras, Oscar Mitnik, Edgar Salgado, and Beatrice Zimmerman.

Thank you all for the thoughtful discussions that have helped shape and bring about this work.

I could not have persevered without the support of family and friends.

First, I would like to thank my parents, Andrés and Leticia, who instilled in me the aspiration not to be perfect, but rather to become the best version of myself. Their work ethic, civic responsibility, and vocation for service are an inspiration to all their children.

To my grandmother Leticia, who is my fiercest fan, and my grandfather Julio, who celebrated even my smallest achievements while on earth, thank you for always supporting my academic and personal dreams. I am also certain that my grandparents Beatriz ("Betty") and Carlos ("Chalo") have been rooting for me from heaven, hoping I would become an economist just like Grandpa.

To my siblings, Juan Andrés and Daniela, who have offered timely advice, encouraging words, and unconditional love throughout this journey. To my nephew and nieces—Alberto, Luciana, Valentina, and Liliana—whose tender hugs have healed all my wounds and heartaches.

To my "DC" sisters, Ivette and Tanya, who have picked up my broken pieces and put me back together so many times I've lost count. Thank you for reminding me that I am enough.

To my NFEPP cohort—Anne, Sarah, and Manali—who have lent me strength and hope when it was in short supply. It has been a privilege to share this experience with you.

To all of my extended family and friends who have celebrated, suffered, and, most importantly, prayed leading up to this moment: may your kindness toward me be repaid tenfold.

Most importantly, I thank God for the work He continues to do in my life. Every little (and not so little) thing I do, I offer for His glory.

Contents

List of Figures

List of Tables

Chapter 1

The Impact of Teenage Childbearing on Educational Attainment and Labor Force Participation in Ecuador

by Leticia Donoso-Peña

1.1 Introduction

Adolescence is a pivotal stage in life, where choices and shocks -whether positive or negative- can have life-altering impacts. Teenage childbearing is one such shock that remains a pressing public policy issue due to its high prevalence in low- and middle-income countries; an estimated 46 births per 1,000 occur among adolescent girls in these regions, compared to 11 per 1,000 in high-income countries. Moreover, teenage childbearing is linked to poorer socioeconomic outcomes, including lower educational attainment, reduced income, and limited labor force participation. These associations have led to the development of numerous policy initiatives aimed at addressing its prevalence and long-term effects. (UNPD, 2021)

However, different theories yield different predictions regarding the effects of teenage childbearing. On one hand, teen motherhood can increase the opportunity cost of attending school or work, leading women to drop out of school or exit the labor force. (Ribar, 1994) Yet teenage childbearing can also serve as a positive turning point for troubled youth. (Fairfax, 2008; Brubaker and Wright, 2006) Moreover, since risk factors such as economic marginalization are associated with both teen motherhood and low educational attainment, effects could be null if adolescents face bleak economic prospects prior to childbearing. Alternatively, different baseline characteristics could result in heterogeneous impacts across groups. (Gorry, 2019; Diaz and Fiel, 2016)

Overcoming selection bias is therefore key to study the relationship between teenage childbearing and later-life outcomes. For this purpose, prior studies have used a variety of empirical methodologies, such as propensity score matching, sibling fixed effects and instrumental variables. (Diaz and Fiel, 2016; Heiland, Korenman, and Smith, 2019; Duncan et al., 2018; Zito, 2018; Kane et al., 2013; Yakusheva, 2011) Yet propensity score matching requires strong assumptions about no selection on unobservables, whereas sibling fixed effects require limiting the sample to sister pairs, thereby limiting generalizability. More recent work employs an instrumental variables approach, using miscarriage or abortion access as an instrument for the (non-likelihood) of childbirth. (Ashcraft and Lang, 2006; Fletcher and Wolfe, 2009; Gorry, 2019; Azevedo, Lopez-Calva, and Perova, 2012) However, access to abortion is limited in many low and middle-income countries.

This paper investigates the impact of teenage childbearing on long-term educational attainment and labor force participation ("LFP") in Ecuador, a country with one of the highest rates of adolescent fertility in Latin America. Using a nationally representative survey of women's reproductive health, I exploit the exogenous variation in a woman's fertility status - their reported age at first menstrual period, or menarche – as an instrument for teen motherhood, conditional on a number of socio-demographic and geographic characteristics, borrowing from Ribar (1994) and Klepinger, Lundberg, and Plotnick (1995).[1] Although certain environmental factors can play a role in menarche onset, medical literature suggests that it has a strong genetic component orthogonal to socioeconomic outcomes. (Dvornyk and Waqar-ul-Haq, 2012; Perry et al., 2014).

Previous literature has documented that early menarche can be associated with earlier sexual debut, hormonal changes, and child marriage (Field and Ambrus, 2008; Chari et al., 2017; Huang et al., 2019; Sabia and Rees, 2009), which poses a concern for potential violations of the exclusion restriction. However, child marriage is illegal in Ecuador, and given the low usage of contraception and restricted access to abortion, it is likely that the most relevant channel to affect educational attainment and LFP in this context is via early childbearing. Nevertheless, in addition to the instrumental variable strategy, I show a reduced-form model that can be interpreted as a "bundled treatment" for studying the effects of early menarche on education and labor outcomes.

Overall, my paper has three core findings. First, the first-stage relationship between age of menarche and adolescent childbearing is strong: women who experienced their first menstruation between the ages of 10-13 are 10.2 percentage points more likely to become a teen mother, a 31% increase over the mean rate of teen childbearing for women who had their menarche at later ages. Second, using menarche as an instrument for teen childbirth, I find no statistically significant differences in educational attainment and LFP between women who had children before the age of 19, and other women in my

[1] Cantet (2020) has used age of menarche as an instrument for teen pregnancy in the South African setting.

sample, rejecting large negative effects. These findings are supported by the reduced form estimates, which show no impact of early menarche across different outcomes, and are robust to a variety of specifications. The results stand in stark contrast with a naive ordinary least squares (OLS) comparison that shows teen mothers have, on average, three years less of education than non-teen moms, underscoring the importance of addressing selection into childbearing. Finally, I show these null effects are not driven by age cohorts or urban/rural localities, but I find suggestive evidence that women whose own mother lives in the same household have higher educational attainment, consistent with grandmother cohabitation playing a role in attenuating the potential negative effects of teenage childbearing observed in other contexts.

This paper makes three primary contributions. First, it presents new evidence of the relationship between teenage childbearing and educational attainment and labor force participation in an understudied context - most of the literature has focused on high income countries, primarily due to data availability. Second, it furthers the literature on the relationship between menarche, fertility, and later-life outcomes at the national level, and across different generations, arguing how menarche can be a relevant instrument for teenage childbearing in this context. Finally, while investigating mechanisms, it adds to the literature on how family structures, specifically grandmother cohabitation, can partially explain positive, yet not statistically significant effects, of teen childbearing on educational attainment.

The rest of the paper continues as follows. Section 1.2 discusses the context of Ecuador. Section 1.3 elaborates on the data sources and definitions. Section 1.4 delves into the empirical strategy and the required assumptions, section 1.5 discusses the results, and section 1.6 concludes.

1.2 Study Context

Ecuador is a middle income country with above average adolescent fertility rates: there were 63 births per 1,000 women ages 15-19 in 2021, compared to a rate of 16 in the United States, and 42.5 in the world. (UNPD, 2021) Despite decades of policy efforts attempting to reduce teen pregnancies, there were 37,027 live births to women ages 10-19 in 2023, roughly 15% of all births. In fact, over the last 15 years, there has been a 95% increase in births among girls in the 10 to 14 year age group, and a 9% increase in motherhood among girls aged between 15 and 18, as shown in Figure A1.

A Ministry of Health report suggested that the primary cause of adolescent pregnancy for girls aged 10-14 years old is sexual abuse and gender violence. According to Ecuadorian Law, gender violence encompasses physical, psychological, and sexual violence towards women, because of their gender. Although administratively under reported, a national survey conducted in 2019 found that 45% of adolescents aged 15 to 17 had experienced gender violence at least once in their lifetime. Within respondents 15 to 29 years old, 21% reported incidents at home, 18% at school, 19% at work, and 35% in social life. (INEC, 2019) (MSP, 2018) Underlying these staggering numbers is the socially accepted belief that men should have a dominant role in society, over women in particular. This view of masculinity, known as *machismo* in Latin America, permeates public and private lives, and is difficult to eradicate given cultural norms, and generations of oppression. These beliefs are not exclusive to men, for many women may believe that men are in their "right" to abuse them, given they probably saw their mothers or grandmothers being treated the same way.

In addition to sexual violence, an often cited cause for teenage childbearing in Ecuador are low usage of contraceptives among teenagers, either for lack of information, limited access, or overall reluctance of girls (or their partners) to use them. According to a national health survey conducted in 2018, around 88% of teens aged 15-18 had heard of at least one contraceptive method, but only 14% reported having used it at least once before. Although not all teenagers among the 15-18 age group would have been sexually active when the survey took place, these statistics are indicative that when a teenager reaches menarche and starts her sexual life, she has a higher probability of becoming pregnant.

Both gender violence and use of contraceptives can be aggravated by a teenager's low socioeconomic status, rendering them more vulnerable to a teen pregnancy. Poverty is therefore an important component of this issue. Since in Ecuadorian law, abortion is penalized except under specific circumstances,[2] most teenage pregnancies are not terminated, and result in births to young mothers.

Pregnancy in adolescence can have detrimental consequences for young mothers and their babies. According to the WHO, complications during childbirth are the second leading cause of mortality among adolescent girls aged 15-19 globally, and within low and middle-income countries, children born to mothers aged 20 years or younger have 50% higher mortality than children born to women between 20 and 29 years old, in addition to registering lower birth weight. (WHO, 2019, 2014a)

Besides these effects on health, teenage childbearing can lead to school dropout, resulting in lower education levels for young mothers and, in the long run, lower earnings. Alternatively, it can also foster a sense of purpose, motivating young mothers to pursue higher education in an effort to provide for their children. Anecdotal evidence from Ecuadorian news-reports suggest both forces can be at play. (El Universo, 2021)

A mitigating circumstance to these negative education effects could be the family structure in Ecuador. Similar to other Latin American countries, it is not uncommon for grandparents to live in the same house with the rest of the family. If, in addition, social norms dictate that women in the household take care of children, it is likely that young mothers who can rely on their mothers or another female relative for childcare support may be able to achieve higher education levels.

1.3 Data

1.3.1 Data Sources and Definitions

This paper uses data from the National Health and Nutrition Survey conducted in 2018 by the Ecuadorian National Statistics Office (INEC). The NHNS was conducted between November 2018-January 2019, as well as June-July 2019. A total of 43,311 households and 168,747 individuals were interviewed, collecting detailed information

[2] Article 150 of the *Código Orgánico Integral Penal* states the only exemptions are whether (1) the mother's life is in unavoidable danger, or (2) if the pregnancy is the consequence of rape to a woman with mental disabilities. (COIP, 2021)

on individuals' health and nutrition status. The survey had a special questionnaire for women ages 15-49 (considered of 'fertile age') which was administered to 48,700 women. The module collected detailed information on a woman's childbearing history, includes dates and ages of children, as well as age of menarche.

For this paper, I restrict my sample to women born between 1977 and 1997 (21-41 years of age), for a total of 24,987 women. Women aged 15-20 are excluded, as they are still completing their education.[3] Therefore, studying women in this age range allows me to capture teenage childbearing effects *after* a few years have passed since the birth of their child, when any remedial actions (such as finishing school later) have already taken place. On the other hand, excluding older women mitigates concerns on recall bias, specially regarding their reported age of menarche.

I define teenage childbearing as having had a child at 18 years or younger, the legal age in Ecuador. The age at first birth is constructed following the National Statistics Office guidelines, and is computed using the date of birth of both the mother and the first child. This reduces measurement error that could arise from recollection bias. The measurement does not take into account pregnancies that did not come to term, either due to miscarriage or abortion.[4]

1.3.2 Sample Characteristics

Selected sample characteristics are summarized in Table A1. Column 3 shows the raw mean for women who had a child by 18 years old or younger, and column 2 shows the raw mean for the controls. Column 4 shows the p-value for a regression of each variable on the binary indicator for teen motherhood, which includes cohort fixed effects and robust standard errors clustered at the parish level. Column 4 provides clear evidence of selection bias: teen mothers and the control group are statistically significantly different in most observable characteristics.

Focusing on the controls, on average, women in my sample are 30.5 years old,

[3]In Ecuador, as in most countries, adolescents are expected to graduate high school around 17-18 years old, and if they pursue university studies, most university careers take 4 to 5 years, such that by age 21-22 one would expect most of their educational attainment to be complete.

[4]The cut-off by 18 years old is slightly different from the one used by international organizations, who measure adolescent fertility rates from ages 15-19.

33% live in rural areas, and are distributed between the coastal (Costa, 34%) and the mountainous (Sierra, 42%) regions, the most populated in Ecuador, although 19% live in the Amazon region and 4% live in the Galapagos Islands. Around 81% of the sample identifies as "mestizo", or mixed race, whereas 11% is indigenous.

Overall, most of the women report being the spouse of the head of the household (54%), although 25% are daughters, and 13% are themselves the head of the household. Using the government's poverty definition, on average 23% of the women in the sample have unmet needs that qualify them as poor. In addition, 6% report being beneficiaries of the government cash transfer program, and 83% have an active cellphone.

On average, teen mothers are 28% more likely to live in rural areas compared to other women, and 13.7 percentage points more likely to be categorized as poor, that is 64% more than non teen mothers. They are also 2.4 percentage points more likely to be indigenous, which is 23% more than other women in the sample. In addition, they have 2.56 less years of education, and are 23 percentage points less likely to finish basic education and 30 percentage points less likely to complete high school, a reduction of 28% and 42% respectively, when compared to other women in the sample. The differences mentioned above are not only statistically significant, but also economically meaningful.

Panel D of Table A1 shows characteristics regarding women's reproductive health and practices. On average, women who were teen mothers in the sample have their first menstruation at age 12.64, slightly lower than non-teen mothers (12.96). On average, teen mothers are 7 percentage points less likely to have been aware of what menstruation was when it happened to them, and they are 5 percentage points less likely to have found out about it from their mother.

Although age of menarche varies slightly by the two groups, there is a 3.36 year difference for self-reported age for first sexual intercourse: teen mothers started having sexual relations at almost 16 years old on average, whereas other women did so at age 19. Interestingly, the reported age of contraceptive use is older than that of first sexual relation in both groups: teen mothers report being 19 years old when first started using contraceptives, while other women report being almost 22 years old, on average. In addition, teen mothers report earlier age of first marriage or non-marital relationship (18 vs 22 years old).

The average age at first birth in the sample in the entire sample is 20.76. For teen mothers, 83% of births occur between ages 16-18, while for other women most births occur between ages 19-26. On average, teen mothers have 3 children, while non-teen mothers have 2.

1.3.3 Data Limitations

Because the national survey is a cross-section conducted after women's fertility and education levels have been realized, there are data limitations that need to be accounted for.

First, there is little information *before* women's menarche, except for time-invariant characteristics, i.e. ethnicity and year of birth, included in my preferred specification. Information such as parent's education can only be observed for the subset of women whose parents live at home at the moment of the survey (roughly 20% of the sample). As a robustness check, I run reduced form estimates of my instrument on different education outcomes, with and without controlling for mother's education for this subsample, and find that the coefficients do not vary significantly.

Second, as with any survey asking retrospective questions, respondents could suffer from recollection bias. This is unlikely for the respondent's age at first birth, (enumerators were able to verify birth records for the children listed), but more so for the age of menarche, which for the oldest women in my sample would have occurred almost 30 years before. Although, as argued in Field and Ambrus (2008), menarche is a significant event in a woman's life, and therefore unlikely to be subject to recall bias, in my preferred specification I group women's responses into early (below or equal to the median) and late (over the median) menarche to account for this possibility.

Third, the responses to detailed questions regarding sexual experiences and reproductive health preferences could suffer from social desirability bias, as the respondent was interviewed at home, potentially in the presence of other family members. For instance, women could have modified their response regarding their age of sexual debut, to match what they expected was desirable by a family member or the enumerator. However, this is less likely to happen with age of menarche: as an adult, how early or late you experienced menarche does not have positive or negative implications, and there are no cultural norms attached to it in the study's context.

Finally, a concern is that teen mothers are underrepresented in the sample if this group has a higher mortality risk, compared to older mothers. However, this appears to be less relevant in this context: a national study of in-hospital maternal mortality in Ecuador examined public and private hospital records from 2015 to 2022, and found that although the country has a ratio of 18.94 maternal deaths per 100,000 live births, there were no statistically significant differences in the ratio by age cohorts. According to the study, mothers in the age range of 15-19 had an adjusted odds ratio of 1.1, compared to 0.98 and 0.92 of mothers ages 30-39 and 40-49, respectively. (Lapo-Talledo, 2024)

1.4 Empirical Strategy

1.4.1 Estimating Equation

To assess the impact of teenage childbearing on longer-term educational and labor market outcomes, I use an instrumental variables strategy that leverages age of menarche to predict the likelihood of teenage childbearing. The estimating equations are as follow:

$$Y_{i,c,m} = \beta_0 + \beta_1 TM_{i,c,m} + \beta_2 E_i + \psi_c \phi_m + \epsilon_{i,c,m} \quad \text{(IV, LATE)}$$

$$TM_{i,c,m} = \delta_0 + \delta_1 Z_{i,c,m} + \delta_2 E_i + \psi_c \phi_m + \mu_{i,c,m} \quad \text{(First-Stage)}$$

$$Z_{i,c,m} = 1(10 \leq age\ of\ menarche_{i,c,m} \leq 13) \quad \text{(Instrument)}$$

Where $Y_{i,c,m}$ is the outcome of individual i in age cohort c living in municipality m, $TM_{i,c,m}$ is the probability of having had a child at 18 years or younger (or the number of years as a teen mother), E_i are binary variables for different ethnicities, $\psi_c \phi_m$ are

4-year cohort by municipality fixed effects, and $\epsilon_{i,c,m}$, $\mu_{i,c,p}$ are error terms. Survey weights are included in all regressions.

The parameter of interest, β_1 captures the effect of teenage childbearing on the outcome of interest. If $\beta_1 < 0$, then having a child at 18 years or younger decreases the probability of completing high school, consistent with previous literature in the U.S., Brazil, Mexico, and South Africa (Branson and Byker, 2018; Ardington, Menendez, and Mutevedzi, 2015; Berthelon and Kruger, 2017; Arceo-Gomez and Campos-Vazquez, 2014; Narita and Montoya Diaz, 2016). Alternatively, $\beta_1 > 0$ would indicate an opposite effect, hinting that teen mothers could be motivated to complete their education to improve the outcomes of their children, such as in Azevedo, Lopez-Calva, and Perova (2012). It could also be the case that previous identification strategies were not able to fully account for selection bias, and once removed, there is no relationship between teen childbearing and a woman's educational attainment, that is, $\beta_1 = 0$.

In addition, to address concerns about violations to the exclusion restriction because of alternative channels outside of teenage childbearing, I estimate a reduced form model using the following equation:

$$Y_{i,c,m} = \gamma 0 + \gamma_1 Z_{i,c,m} + \gamma_2 E_i + \psi_c \phi_m + \epsilon_{i,c,m} \qquad \text{(Reduced Form, ITT)}$$

where the parameter of interest, γ_1, captures the effects of early menarche on outcomes, which can be interpreted as a "bundled treatment" of not only teenage childbearing, but child marriage, early sexual debut, among other channels that could be operating simultaneously.

1.4.2 Identifying Assumptions and Threats to Identification

For β_1 to be causally identified, the main assumptions are that (1) the instrument $Z_{i,c,m}$ is relevant (there is a first-stage), and (2) that potential outcomes are independent of $Z_{i,c,m}$, what is also known as the exclusion restriction. The estimation also relies on the conditional independence assumption as well as monotonicity, that is, that there are no defiers. I discuss these assumptions below.

First Stage

Identification of the IV model requires a strong correlation between age of menarche and the probability of teenage childbearing. As previously mentioned, Sabia and Rees (2009) and Huang et al. (2019) show a strong relationship between age of menarche and first sexual intercourse. Both the onset of menarche and sexual debut are necessary conditions for a pregnancy to occur. In contexts were knowledge of and access to contraceptives is relatively more limited, one expects a strong correlation between age of menarche and the age of the mother at the birth of her first child. This assumption is testable in the data, and as Section 1.5 will show, the instrument is relevant and the assumption holds.

Exclusion restriction

The exclusion restriction requires that (1) menarche impacts educational attainment and labor force participation only through teenage childbearing, and (2) that age of menarche is independent of potential outcomes, conditional on certain covariates (CIA).

First, unlike other contexts studied in the literature such as India or Bangladesh, in Ecuador, there are no societal norms that indicate girls should drop out of school after the onset of menarche. However, earlier menarche has been associated with earlier sexual debut and higher probability of child marriage, which could decrease educational attainment and LFP. (Field and Ambrus, 2008; Chari et al., 2017; Huang et al., 2019; Sabia and Rees, 2009) Although child marriage is illegal in Ecuador and therefore an unlikely channel, I rely on the reduced form model described in 1.4.1 to capture the intent-to-treat (ITT) estimates of an earlier menarche.

Second, to establish the independence of age of menarche from potential outcomes, it is necessary to understand the determinants of the onset of puberty and their connection to educational attainment. Medical literature indicates that menarche has a strong genetic component: an analysis of 400 genomes estimated the heritability of age at menarche to be between 50% and 70% (Kaprio et al., 1995; Campbell and Udry, 1995; Dvornyk and Waqar-ul-Haq, 2012). In an international study involving over 180,000 women and scientists from 166 institutions worldwide, Perry et al. (2014) identified 123 genetic variations associated with the timing of menarche. This study concluded that a wide and complex network of genetic factors is crucial in determining the timing of puberty. Furthermore, the discovery of "imprinted genes" suggests that the onset of puberty may be influenced unequally by the mother or father, depending on which parent the gene is inherited from. This indicates that within a family, one parent may have a more profound effect on the timing of their daughters' puberty than the other.

However, these findings do not exclude environmental factors from playing a role. One of the potential threats to identification is the relationship between nutrition in late childhood and menarche onset: overall, overweight and obesity is correlated with earlier ages of menarche, while chronic malnutrition is associated with later ages of menarche. Although medical literature has established a link between nutritional status and menarche onset, the relationship is complex, and the studies are not causal. To the best of my knowledge, only Barham, Macours, and Maluccio (2024) provides causal evidence on how nutrition can affect menarche timing. Exploiting the timing of a conditional cash transfer that improved the nutritional status of highly marginalized households in rural Nicaragua, they show that undernourished girls who received a positive nutrition shock had their menarche 2-3 months earlier on average.

In Ecuador, chronic malnutrition is a prevalent issue, especially in marginalized localities and among indigenous communities. A concern is that menarche timing is capturing extreme socioeconomic conditions in childhood. Although the evidence provided by Barham, Macours, and Maluccio (2024) suggests that even large nutrition shocks are unlikely to change menarche timing by years, I compare the distribution of adult height (which some consider a valid proxy for childhood nutrition) across menarche terciles and find no statistically significant differences among these groups. Nevertheless, I include ethnicity and cohort by municipality[5] fixed effects across all my specifications, as a proxy for baseline socioeconomic conditions. In addition, as a robustness check, I control for adult height and current poverty, and show that the reduced form estimates remain unchanged.

[5]Municipalities' size range from 2.7 million for Ecuador's capital city Quito, to small rural towns of 2,700 people.

Other environmental factors that the medical literature has documented include: geography, climate, and exposure to endocrine-disrupting chemicals. In addition, it has been widely documented that age of menarche is declining over time globally, although the reasons behind this phenomenon are not clear. I expect the cohort by municipality fixed effects to shut down the variation in age of menarche due to year of birth and geographical region.

1.5 Results

1.5.1 First-Stage

Figure 1.1 shows the kernel density plot and cumulative distribution of age at first birth in the sample, plotted for each of the age of menarche terciles (age 13 is the median age of menarche). Graphically, I observe that a later age of menarche is correlated with an older age at first birth, particularly above the median age of menarche. Table 1.1 summarizes the first-stage results, with panel A showing the effects of menarche on the probability of being a teen mother by 18 years old or younger. In column 1, I analyze how each additional year contributes to the probability of being a teen mother, using menarche at age 10 as reference. It is clear that after the median age of 13, the probability becomes statistically and economically significant. Column 2 shows that an additional year of age of menarche is associated with 3.6 percentage points decrease in the likelihood of becoming a mother at 18 years or younger. Because we may be concerned about recall bias in terms of the specific age of menarche, I create binary variables for ages 10-12 and 10-13, which are shown in columns 3 and 4, respectively. Overall, puberty onset by the age of 13 increases the probability of being a teen mother by 10.2 percentage points, and an F-statistic of 92.5. All specifications include ethnicity dummies, and cohort by municipality fixed effects.

In panel B, the dependent variable is the number of years as a teen mother, in an effort to study the effects along the intensive margin. I find that the results go in the same direction, with one more year of age of menarche decreasing the number of years as a teen mother by 0.105 years (roughly 1.26 months). Looking at column 1, the effect is more meaningful at age 13: while those who had their menarche at 13 decrease the number of years by 1.89 months, those who had their menarche at 14 decrease it by 3.80 months, up to 6 months when menarche occurred at age 16. Throughout the analysis, results are presented with the specification of column 4, but results are robust to the other specifications, included in the Appendix. Therefore, the evidence supports a strong correlation between age of menarche and teenage childbearing, and the first-stage requirement is satisfied in this context.

1.5.2 Main Results

The main results are summarized in Table 1.2. Panel A shows the naive ordinary least squares (OLS) estimates of regressing the likelihood of having a child by age 18, on different measures of education and labor force participation. Column 1 shows that being a teen mother is associated, on average, with a reduction of 25 percentage points (pp) in the likelihood of completing basic education (i.e. 10 years of schooling, similar to middle school in the U.S.), as well as a 33pp and 18pp reduction in the likelihood of completing high school or college, respectively. Total education attainment is therefore on average 3 years less than non-teen mothers, where 12.79 years of education is the mean for all other women who weren't teen mothers. Current labor force participation, however, appears to not be correlated with teenage childbearing, with a small negative non-significant coefficient of 0.03.

Panel C of Table 1.2 shows the results from the IV regressions. I find positive, yet not statistically significant, effects across all the educational outcomes, except for college completion. Having a child at 18 years or younger implies almost a 20pp increase in basic education completion, a 14pp increase in high school completion, and 1.42 more years of education, on average. The effects are large in magnitude, representing 11% to 25% of the control mean.

I start disentangling this puzzle by looking at the reduced form, in panel B, which shows the effects that early menarche has on education and labor force participation.

If one expects teenage childbearing to have a negative effect, the coefficients from the reduced form should be negative, indicating that an earlier menarche, through an increased likelihood in teenage childbearing, decreases educational attainment. However, panel B shows that this relationship is slightly positive and not statistically significant: the point estimates' range is relatively small (-0.009 to 0.145), approximate 1 to 3% of the control mean, and are not statistically significant.

There are two possible interpretations of these reduced form and 2SLS results. First, it could be the case that the model is correctly identified, and for the compliers, there is a positive yet not statistically significant effect of teenage childbearing on educational attainment and labor force participation. On the other hand, if the true effect is negative, but the instrument fails to capture all the bias, the reduced form could be displaying two opposing forces: the negative effect of teenage childbearing on education, while at the same time the positive effects of higher socioeconomic status before puberty (specifically, appropriate nutrition) that leads to an earlier menarche but is also associated with better educational attainment on average.

To check for the latter possibility, and taking into account my data limitations, I perform a series of robustness checks but find no evidence that this is the case. Table A2 summarizes these attempts for the reduced form estimates. Column 1 shows the estimates from my preferred specification for ease of comparison. First, following Field and Ambrus (2008), I control for current height (column 2), as a proxy for childhood nutrition. Then, in column 3, I control for current poverty as defined by the government's "unmet needs" measure.[6]. Finally, for the small subset of women who live with their mother, I include whether the mother completed primary school, as a proxy for socioeconomic status (column 4). None of these specifications change the direction of the estimates, and only basic education and years of education become marginally significant when including height (columns 2 and 3).

In addition, in Tables A3 and A4, I test the specification with slightly different samples for the reduced form and the 2SLS, respectively. In both tables, Panel B shows

[6]This could be considered a "bad" control, since it is not 'baseline' poverty and could have been affected by teenage childbearing. However, it could also be argued that current poverty is highly correlated with poverty at baseline.

the results for women born between 1977 and 1997, roughly ages 21-41, which is my preferred specification. Panel C includes women ages 21-41 *that completed primary school* (7 years of education) only, while panel D restricts the sample to *mothers* ages 21-41, that is, it excludes those women who never had children. Neither of these subsamples substantially change the reduced form results.

1.5.3 Mechanisms

What could explain these surprising results? There are several demographic, geographical, and social factors that could be moderating the effects for teenage mothers, which I discuss below.

First, a potential explanation is that the results are different between generations: younger cohorts of teen mothers may be subject to different gender and social norms regarding marriage, motherhood, and education. Globally, as gains to women's education are realized, women's years of education have increased, while marriage has gradually been delayed, and Ecuador is no exception, albeit at a slower rate. If this is the case, younger cohorts of teen mothers may exhibit higher levels of education, than their older counterparts.

Second, effects could be different between urban and rural areas: access to education, in particular for 8th grade onwards, has been traditionally lower in rural areas of Ecuador, as opposed to urban centers, where there is a larger school supply and shorter distances to school. Furthermore, geography may play a role in the onset of menarche: puberty in adolescent girls can be delayed in higher elevation areas. Since Ecuador's mountainous region has plenty of altitude variation, it could be that this influences the age of menarche, in turn affecting the probability of becoming a teen mother.

Thirdly, family's structure and social norms could play a mediating role in teen mothers' education. In Ecuador, contrary to many high income countries, it is common for young adults to live at their parent's home until they marry.[7]

Most women in my sample are the spouse of the household head (58%), the daugh-

[7]An exception is when attending college in a different city from home, but this is rare: national college enrollment is low (less than 30% of all young adults aged 18-24 years old were enrolled in college in 2024), and among enrolled students, unofficial estimates suggest around 16% students came from a different province https://shorturl.at/e4pYu

ter of the household head(20%), or the head themselves(14%). Although the data does not allow me to understand if these living decisions were made before or after the birth of their child, it is possible to see if being a teen mother is correlated with either arrangement. As shown in Table A5, teenage childbearing does not crowd-in grandmother cohabitation, nor affect the likelihood of being married. However, I find that education gains are larger for teen mothers where the grandmother cohabitates in the household, compared to those who are either the spouse of the household head, or the head themselves, either by running separate IV regressions (Figure 1.3) or interacting both the dependent variable and the instrument with a binary variable indicating grandmother cohabitation. (Table A7-A8)

1.6 Conclusion

Despite substantial policy efforts, teenage childbearing remains prevalent in many low- and middle-income countries, including Ecuador. While teenage motherhood is often linked to poorer educational and labor market outcomes, this paper finds that, after accounting for selection biases through an instrumental variables approach, there are no significant negative impacts on educational attainment and labor force participation for teenage mothers in Ecuador. These results challenge conventional assumptions and highlight the importance of understanding local contexts, such as family support structures, that may moderate potential outcomes. Future research should delve into the mechanisms, including the role of extended family cohabitation, and explore policy interventions that can support young mothers in achieving better socioeconomic outcomes.

Tables and Figures

Table 1.1: First-Stage

Panel A: Mother at 18 years or younger

	(1) Mother at 18 or younger	(2) Mother at 18 or younger	(3) Mother at 18 or younger	(4) Mother at 18 or younger
AM = 11	0.028			
	(0.032)			
AM = 12	-0.012			
	(0.030)			
AM = 13	-0.014			
	(0.030)			
AM = 14	-0.075**			
	(0.031)			
AM = 15	-0.136***			
	(0.031)			
AM = 16	-0.202***			
	(0.038)			
AM		-0.036***		
		(0.004)		
Menarche at ages 10-12			0.061***	
			(0.010)	
Menarche at ages 10-13				0.102***
				(0.011)
Observations	24987	24987	24987	24987
Mean Dep. Var. (AM=10)	0.412			
Mean Dep. Var.		0.332	0.332	0.332
F-stat		100.1	34.32	91.13
P-value		0.000	0.000	0.000

Panel B: Years as a teen mother

	(1) Number of years as teenmom	(2) Number of years as teenmom	(3) Number of years as teenmom	(4) Number of years as teenmom
AM = 11	0.043			
	(0.072)			
AM = 12	-0.124*			
	(0.065)			
AM = 13	-0.158**			
	(0.065)			
AM = 14	-0.317***			
	(0.065)			
AM = 15	-0.455***			
	(0.064)			
AM = 16	-0.500***			
	(0.068)			
AM		-0.105***		
		(0.006)		
Menarche at ages 10-12			0.198***	
			(0.019)	
Menarche at ages 10-13				0.278***
				(0.017)
Observations	24987	24987	24987	24987
Mean Dep. Var. (AM=10)	0.759			
Mean Dep. Var.		0.446	0.446	0.446
F-stat		270.6	107.1	276
P-value		0.000	0.000	0.000

Notes: This regression includes cohort fixed effects, and robust standard errors. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table 1.2: Main Specification: Extensive Margin

	(1) Completed Basic Education	(2) Completed High School	(3) Completed College	(4) Years of Education	(5) Currently works
Panel A: OLS					
Mother at 18yrs or younger	-0.251***	-0.328***	-0.177***	-2.993***	-0.029*
	(0.011)	(0.012)	(0.008)	(0.089)	(0.012)
Panel B: Reduced Form					
Menarche at ages 10-13	0.020	0.014	-0.003	0.145	-0.009
	(0.010)	(0.011)	(0.011)	(0.102)	(0.012)
Panel C: 2SLS					
Mother at 18yrs or younger	0.196	0.136	-0.032	1.418	-0.085
	(0.110)	(0.117)	(0.104)	(1.066)	(0.117)
2SLS Lower 95% C.I.	-0.020	-0.093	-0.236	-0.672	-0.315
2SLS Upper 95% C.I.	0.412	0.364	0.171	3.508	0.146
Kleibergen-Paap F-statistic	91.126	91.126	91.126	91.126	91.126
Kleibergen-Paap p-value	0.00	0.00	0.00	0.00	0.00
Observations	24987	24987	24987	24987	24987
Control Mean	0.78	0.70	0.22	12.79	0.56
Control SD	0.41	0.46	0.41	4.09	0.50

	Mother at 18 years or younger
Panel D: First-Stage	
Menarche at ages 10-13	0.102***
	(0.011)

Notes: Dependent variables for panels A, B, and C are the column titles. Dependent variable for Panel D is whether a woman was a mother by 18 years old. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table 1.3: Main Specification: Intensive Margin

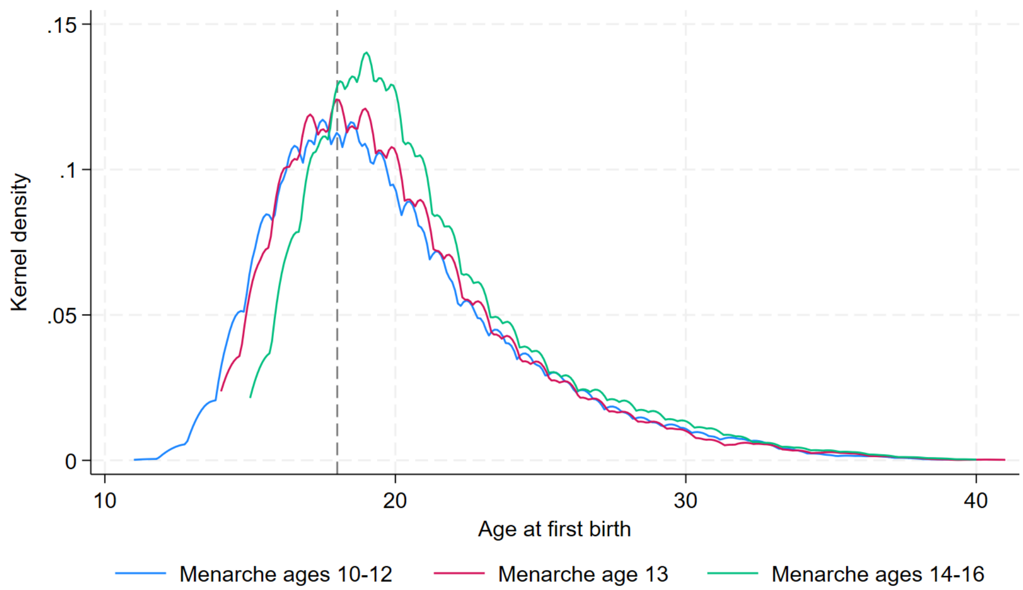
	(1) Completed Basic Education	(2) Completed High School	(3) Completed College	(4) Years of Education	(5) Currently works
Panel A: OLS					
Number of years as teenmom	-0.127***	-0.150***	-0.063***	-1.347***	-0.017**
	(0.005)	(0.005)	(0.003)	(0.043)	(0.005)
Panel B: Reduced Form					
Menarche at ages 10-13	0.020	0.014	-0.003	0.145	-0.009
	(0.010)	(0.011)	(0.011)	(0.102)	(0.012)
Panel C: 2SLS					
Number of years as teenmom	0.072	0.050	-0.012	0.522	-0.031
	(0.039)	(0.042)	(0.038)	(0.378)	(0.043)
Observations	24987	24987	24987	24987	24987
Control Mean	0.756	0.665	0.210	12.476	0.555
Control SD	0.430	0.472	0.407	4.135	0.497
Lower 95% C.I.	-0.004	-0.032	-0.087	-0.219	-0.116
Upper 95% C.I.	0.148	0.132	0.063	1.263	0.053
Kleibergen-Paap F-statistic	275.965	275.965	275.965	275.965	275.965
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000

	Mother at 18 years or younger
Panel D: First-Stage	
Menarche at ages 10-13	0.278***
	(0.017)

Notes: Dependent variables for panels A, B, and C are the column titles. Depedent variable for Panel D is whether a woman was a mother by 18 years old. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Figure 1.1: Distribution of age at first birth by age of menarche tercile

Panel A: Kernel Density



Panel B: Cumulative Distribution Function

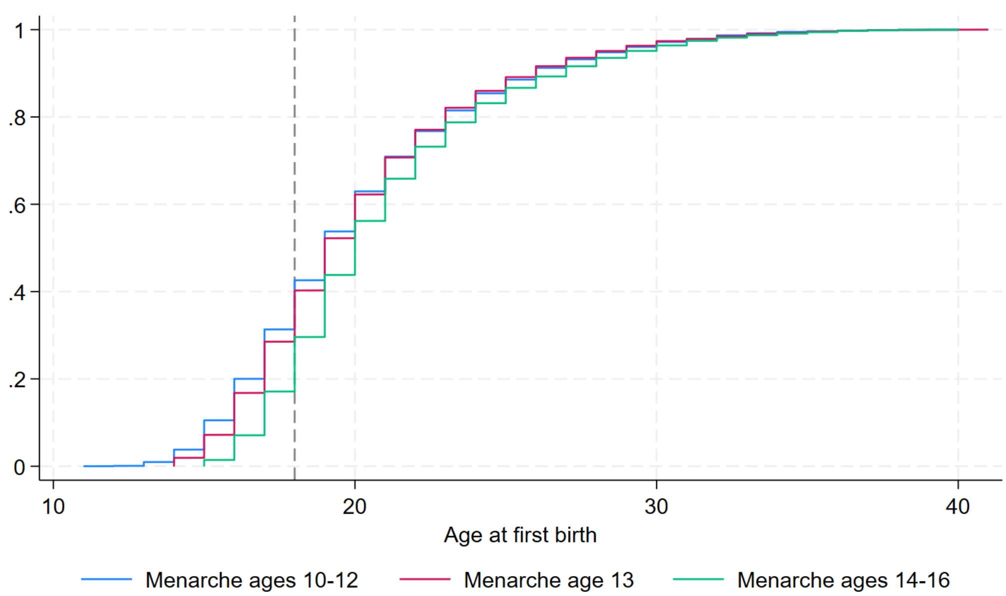
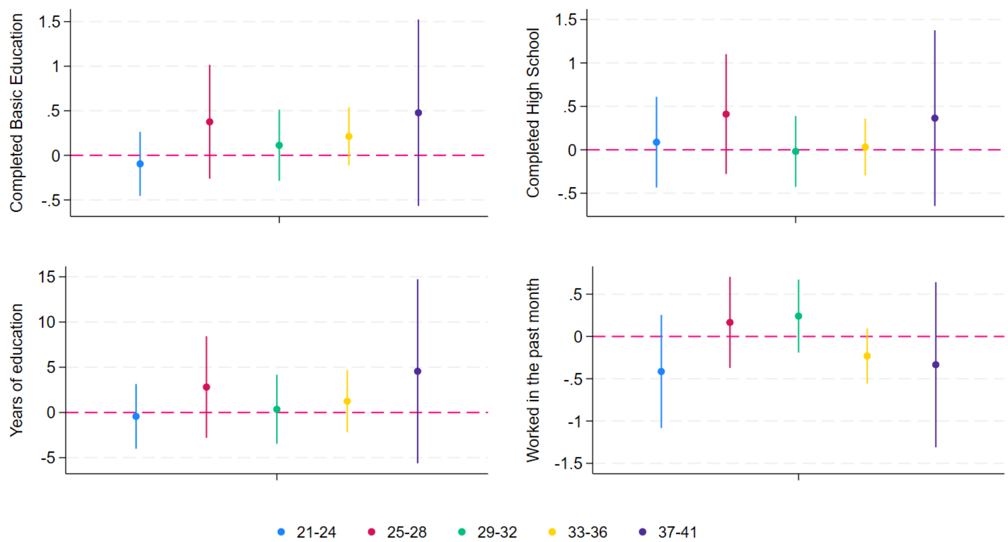


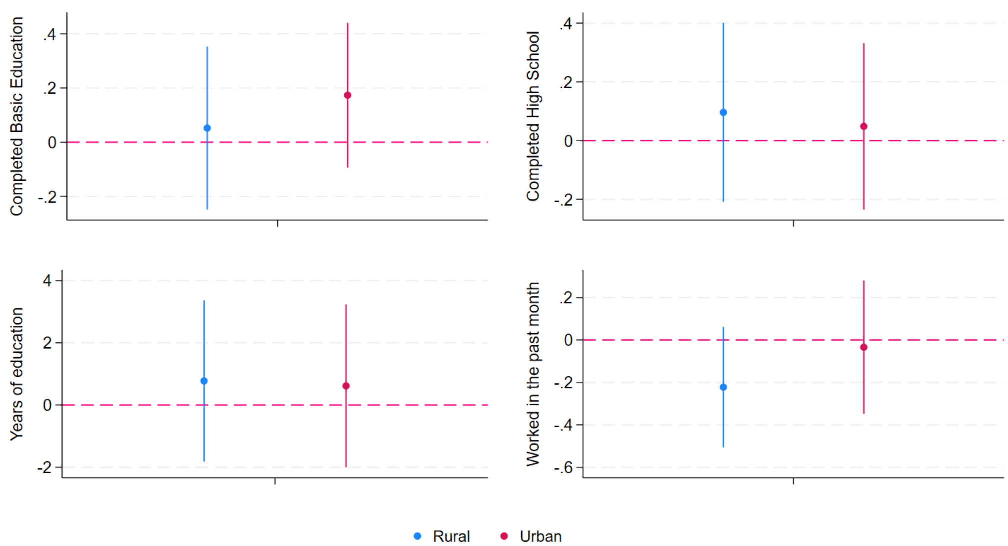
Figure 1.2: Mechanisms: 2SLS Estimates for Different Samples

Panel A: By Age Cohort



Instrumented: Mother at 18yrs or younger. Includes ethnicity and municipality fixed effects.
Robust standard errors. F-stat: 16.30 13.96, 24.77, 52.80 5.41

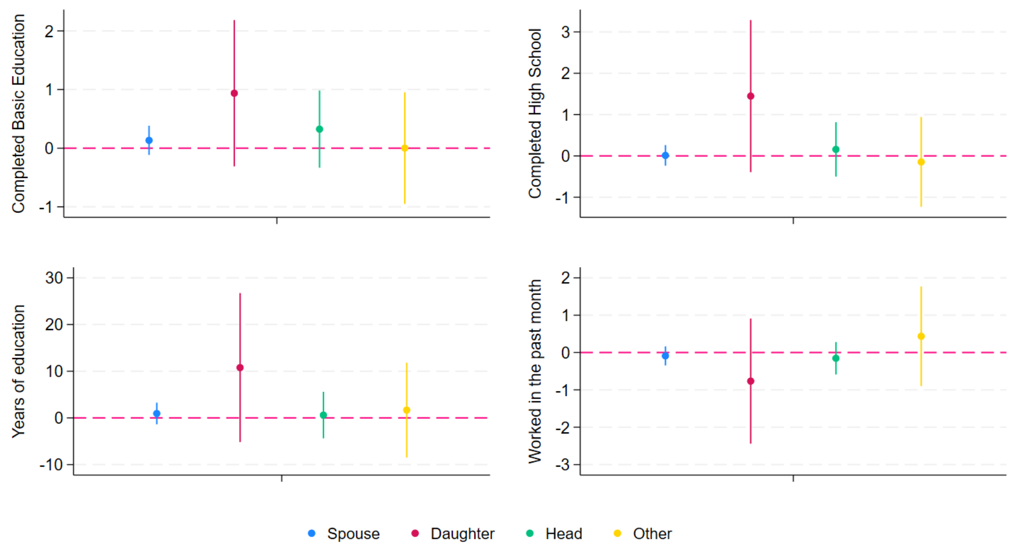
Panel B: By Location



Instrumented: Mother at 18yrs or younger. Includes ethnicity and cohort by municipality fixed effects.
Robust standard errors, C.Intervals at 95%. F-stat: 47.17, 53.09

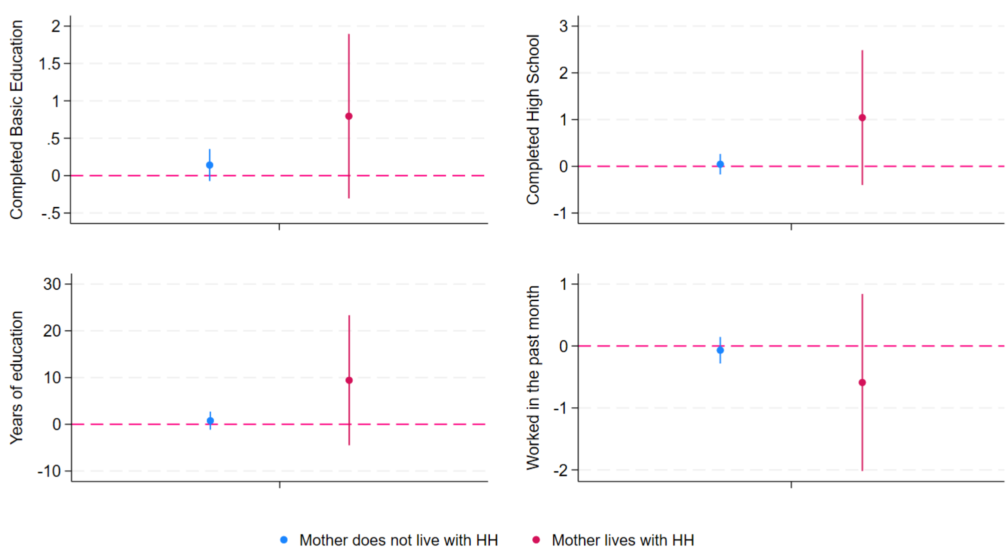
Figure 1.3: Mechanisms: 2SLS Estimates for Different Samples

Panel A: By Relationship to the Household Head



Instrumented: Mother at 18yrs or younger. Includes ethnicity and cohort by municipality fixed effects.
Robust standard errors, C.Intervals at 95%. F-stat: 70.97 5.03, 9.95, 3.59

Panel B: Whether respondent's mother lives in the household



Instrumented: Mother at 18yrs or younger. Includes ethnicity and cohort by municipality fixed effects.
Robust standard errors, C.Intervals at 95%. F-stat: 89.16, 6.30

Chapter 2

Does Increasing Productive Credit Availability Improve Financial Inclusion? Causal Evidence from El Salvador

by Leticia Donoso-Peña, Oscar Mitnik, Edgar Salgado, and Alejandro Tamola

2.1 Introduction

A policy priority is to provide productive loans to small and medium business owners, since they are key in the economic development of emerging markets. In 2018, the International Finance Corporation (IFC) alone facilitated 406.7 billion USD in loans reaching more than 10 million businesses (International Finance Corporation, 2020). The vast literature on credit access for financially constrained business owners (including, but not limited to, microfinance) underscores the popularity of loan programs as a policy tool, even as their overall effects remain modest and nuanced, benefiting some but not all recipients. (Banerjee, Karlan, and Zinman, 2015; Banerjee et al., 2015, 2019; Meager, 2019).

Yet are these programs "filling a hole or digging a hole" for future loans? (Brailovskaya, Dupas, and Robinson, 2024) In many cases, business owners are initially excluded from formal credit markets, either because of insufficient credit supply, or because of information asymmetries, which are difficult for lenders to untangle. (Karlan and Zinman, 2009) On the one hand, business owners could be "good creditors", and receiving a loan can act as a credential mechanism, that enables them to receive other formal loans. (Agarwal, 2018; Azevedo et al., 2019; Frisancho, 2012) On the other hand, loan recipients could quickly become over-indebted, fall into predatory creditors, and fail to repay their debts. (Azevedo et al., 2019; Brailovskaya, Dupas, and Robinson, 2024)

In the context of a credit supply expansion in El Salvador, this paper studies how receiving a productive loan affects a business owner's future performance in the formal credit market. We employ a difference-in-differences strategy with staggered treatment timing to compare cohorts of business owners who received their first public-funded loan in the beginning of the credit supply expansion to cohorts who received it later on.

Between 2015 and 2018, El Salvador's national development bank Bandesal increased credit availability in the economy by 34%, disbursing in total USD 104 million in productive loans to 18,716 registered firms and individuals. Bandesal operates as a second-tier lender, acting through regulated financial institutions, who screen and select customers to sign loans for working capital and investment purposes (otherwise defined as productive loans). The largest share of their portfolio is concentrated in credit cooperatives, although private banks also play a role.

We leverage loan-level information from Bandesal's productive credit portfolio, and combine it with confidential individual credit bureau records that report aggregate monthly data on credit usage and quality across different debt sources (banks, non-banking financial institutions, credit cards, auto, telecoms, etc.). Importantly, these aggregates exclude the Bandesal-funded loan. We are able to find records for 80% of Bandesal's productive credit portfolio.

Exploiting variation in treatment timing, we use a difference in differences strategy with multiple time periods to estimate the effect of the first public-funded loan on future formal credit market performance. The identifying assumption is that, in the absence of the public-funded productive loan, and conditional on baseline characteristics (such as prior credit history), the performance in formal credit markets for all cohorts would have been parallel.

Overall, we find that the probability of obtaining additional banking productive loans increases in the first 1.5 years after the intervention, but the effects fade out in the later periods, representing an average increase of 39% over the mean of the never treated cohorts. The result is accompanied by a slight increase in the probability of default (i.e., having debt past due in the last 30 days), but the overall past-due amount is relatively small. Importantly, the increase in banking productive loans is not accompanied by an increase in consumption or other (credit card, auto, telecommunications) loans, which we interpret as business owners being able to secure "better quality" loans in the formal credit market. The effects appear to be driven by individuals who did not have active loans 24 months prior to obtaining the public-funded loan. Although we find these results optimistic, we cannot observe if business owners resort to informal credit markets to pay-off their formal credit market debts, and therefore our findings must be interpreted with caution.

Our paper has four important contributions. First, we show that in credit-constrained settings, expanding credit supply and relying on local financial institutions to allocate these loans can be effective in increasing access to more credits. Recent papers have shown how novel screening methods can give access to previously uncredited individuals (Azevedo et al., 2019; Liu, Lu, and Xiong, 2022; Chioda et al., 2024). In our setting, letting financial institutions take the lead yields comparable results. Second, while some studies are limited to studying the main lender's portfolio, we can document how accessing a publicly financed-loan can impact participation in the formal credit market *net of the initial loan*, effectively showing how debt sources change after treatment. Third, we provide evidence that publicly-financed loans can act as a credential mechanism to obtain other loans, especially since we find stronger effects for business owners without any credit history in the past 24 months, similar to Agarwal (2018); Azevedo et al. (2019); Frisancho (2012). Finally, we provide causal evidence showing how development banks can alleviate small and medium enterprises credit constraints, and how the effects persist in the short and medium term, adding to the literature on the role of financial intermediaries in development (Levine, Loayza, and Beck, 2000).

The remainder of the paper is organized as follows. Section 2.2 describes the context of the credit supply expansion, Section 2.3 details our data sources and describes our sample, Section 2.4 explains the empirical strategy, and Section 3.3 presents the key results. Section 2.6 concludes.

2.2 Study Context

The government of El Salvador signed an agreement with the Inter-American Development Bank to inject $104 millions into the Salvadorian development bank, *Bandesal*. Bandesal operates as a second-tier lender acting through regulated financial institutions. These institutions screen and select customers to sign loans for both working capital and investment for acquiring equipment, vehicles, among other productive assets. 41% of the portfolio amount is concentrated in 7 financial institutions (86% of all credits)

The funds aimed to fund micro and small and enterprises. Figure 2.1 shows the share of the $104 million in Bandesal's portfolio in the period 2015 to 2018. Bandesal's financing alone peaked in 2013, and from 2015 the capital injection led to a surge in financing for small and micro enterprises, which otherwise would have been underserved by Bandesal alone given the financial context of the country during those years.

2.3 Data

2.3.1 Data Sources and Definitions

We construct our database using administrative data from El Salvador's Development Bank Bandesal, confidential individual records obtained from a credit bureau, and complement it with additional publicly-available administrative data.

Bandesal data. We obtained loan-level information for all productive loans (classified as either working capital or investment loans) financed by Bandesal from 2015 to 2018. As a second-tier lender, Bandesal collects data from the financial institutions it works with to monitor its portfolio performance. Data includes loan and some individual characteristics, including official identification number (NIT or DUI), business size, municipality, sector of the economy in which the person or business operates (primary, secondary, tertiary), gender (if applicable), among others. Loan information is limited to amount borrowed, date, interest rate, term to maturity, and in some cases, the type of guarantee required. It is important to note that these characteristics were recorded *when the loan was obtained*. In addition, the administrative database does not record the individuals that applied for a loan with the financial institution but were rejected.

Credit bureau data. Using individuals' national identification number, the credit bureau found records for 80% of Bandesal's portfolio. We obtained monthly aggregated data for individuals' credit usage and quality. The credit bureau gathers data from different sources of credits, reporting information on credits at banking institutions (distinguishing between productive, consumption, and real-estate loans), non-banking institutions (i.e. credit cooperatives), and other debt types, such as commercial loans, credit cards, vehicles, and telecommunications. Importantly, the credit bureau identified Bandesal's financed-loans and provided aggregated outcomes *net* of these loans. Given legislation changes in El Salvador that modified credit bureau reporting practices, the sample is restricted to information from January 2015 to December 2018.

To simplify our analysis, we consolidate our outcomes of interest into four debt sources: (1) banking-productive loans, (2) non-banking financial institution loans, (3) banking-consumption loans, and (4) other debt, which includes commercial loans, credit cards, vehicles, and telecommunications. Given their small amount and insignificance in this context, we ignore banking real-estate loans.

We construct five main outcomes to measure credit usage and quality *net* of the Bandesal-funded loan:

- The probability of having debt [$\Pr(Debt > 0)$] is a binary variable equal to 1 if the individual has an active reference (debt) in that particular time period.

- Probability of debt past-due greater than 30 days [$\Pr(DebtPastDue > 30days)$] is also a binary variable that equals 1 if the individual has delayed repayment for 30 or more days.

- Outstanding amount of debt in real USD, is the amount owed by the individual in that time period. When an individual does not have any active debt, this amount is zero. The values are winsorized at the 1% and 99% percentile, and are reported in 2015 USD.

- Outstanding amount of debt past-due in real USD, is the amount owed by the individual for 30 days or more. If an individual has not delayed repayment, this amount is zero. The values are winsorized at the 1% and 99% percentile, and are reported in 2015 USD.

- Share of Banking Productive Loans is obtained dividing the outstanding amount of banking productive debt over the total amount of outstanding debt each quarter. This percentage serves as a proxy to measure shifts towards "better quality" debt sources. Banking productive loans are generally known for better loan conditions in terms of maturity and interest rate (as opposed to credit cards, for example).

We complement the data set with publicly available municipality-level administrative data, such as municipality poverty levels from the 2005 census, as well as monthly homicide rates.

2.3.2 Sample Characteristics

We aggregate the data to build an individual by quarter panel spanning 2015Q1 to 2018Q4.[1] Our sample is restricted to individuals who (1) were successfully matched with individual credit bureau data using their national identification number (80% of Bandesal's portfolio), (2) received a loan from Bandesal between 2015Q3 to 2018Q4[2], and (3) did not receive additional loans from Bandesal in subsequent quarters.

Table 2.1 shows descriptive statistics for individuals in our sample, by cohort, *when they received their loan*. Our final sample includes 8,446 individuals, and 113,719 observations. Focusing on the last column, 57% of business owners are female, and 88% are classified as micro-enterprises, that is, having from 0 to 5 employees. 83% of our sample belongs to the tertiary sector, involved in retail and service activities, such as transportation. In addition, around 51% of beneficiaries are located in urban areas, and only 18% of the sample is located in the capital city's metropolitan area.

Using the credit bureau data, we construct a measure of credit history reviewing the individual's financial activity 6 months before obtaining the Bandesal-funded loan. Panel C of Table 2.1 shows the lagged values for different debt sources one quarter before the Bandesal-funded loan was received ($t-1$). Overall, we find that 16% of our sample had an active banking productive credit, 15% a non-banking productive credit, 24% a banking-consumption credit, and 37% had other types of debt. Default rates are low, ranging from 0% in banking productive loans, to 6% in other debt sources.

In addition, we use the last 24 months of credit history before receiving the Bandesal-funded loan and construct four different "profiles" of customers depending on their credit history. We find that 24% of our sample has no credit history, that is, no active references nor outstanding past due amounts, 41% have good credit history, having an

[1]When collapsing the monthly data, we take the mean value of outstanding debt, and the maximum measure for all binary variables (probability of additional loans, probability of default) for each individual-quarter.

[2]For justification, see Section 2.4.1

active reference but no outstanding past due amounts, 14% have 'curing' credit history, as they have an active reference, an outstanding past-due amount in the past 24 months, but no longer have any outstanding past-due amounts in the last 6 months, and 21% have a bad credit history, having an active reference, with outstanding past-due amounts in the last 6 months.

2.4 Empirical Strategy

We are interested in studying how receiving a publicly financed productive loan impacts a business owner's performance in the formal credit market, using a difference in differences (DiD) design with multiple time periods and staggered treatment adoption. Recent developments in the DiD literature (summarized by Roth et al. (2023)) warn that standard two-way fixed effects (TWFE) models are not adequate when treatment effects are allowed to be heterogeneous, which could be in our setting: it may well be the case that some business owners benefit from the public-financed loan ("fill the hole"), while for others it leads to worse formal financial market outcomes ("dig a hole").

Therefore, we adopt Callaway and Sant'Anna (2021)'s proposed estimator for average treatment effects. In this section, we follow the forward-engineering philosophy laid out in Baker et al. (2025), and proceed in four steps. First, we set-up a potential outcomes framework with staggered treatment timing; second, we establish our target parameters (aggregate average treatment effects), third, we discuss our identifying assumptions, and fourth, we detail our choices when using the user-written command that accompanies Callaway and Sant'Anna (2021).

2.4.1 Potential Outcomes Framework

First, we borrow the potential outcomes framework from Baker et al. (2025) for staggered treatment adoption:

$$Y_{i,t} = \sum_{g \in \mathcal{G}} Y_{i,t}(g)\mathbf{1}\{G_i = g\} \tag{2.1}$$

where $Y_{i,t}$ is the outcome of interest (in our study, the probability of obtaining another productive loan) for individual i in quarter $t = 1, 2, ..., T$. We index potential outcomes by the quarter the treatment begins, $g : Y_{i,t}(g)$, such that G_i is the time the treatment begins for each individual, and $\mathcal{G}$ represents the set of all treatment times (or cohorts).

In our context, all units are eventually treated, because they receive a Bandesal-funded loan at a given time. As in Baker et al. (2025), we assume that a "never treated" group exists by dropping the last two cohorts that are treated (those who received the loan in 2018Q3 and 2018Q4) and T therefore denotes the subset of data that is used for our analysis ($T = 14$). We also drop the first two cohorts ($G_i = 1$ and G_i=2), given that $G_i = 1$ does not have any pre-treatment data, and $G_i = 2$ only has data for one quarter before treatment (effectively, cohorts 2015Q1 and 2015Q2). In addition, Bandesal received the first tranche of the IDB financing in September 2015, so loan recipients from 2015Q3 onwards are the beneficiaries of the credit supply expansion.

In practice, this means every cohort is observed at least two periods before and two periods after treatment, and the timeframe for our analysis is from 2015Q1 to 2018Q2.

2.4.2 Target Parameters

Second, we define our target parameters of interest as the group-average treatment effects as laid out in Callaway and Sant'Anna (2021) and Baker et al. (2025):

$$ATT(g,t) = E_\omega[Y_{i,t} - Y_{i,t=g-1}|G_i = g] - E_\omega[Y_{i,t} - Y_{i,t=g-1}|G_i > max\{g,t\}] \qquad (2.2)$$

Where each $ATT(g, t)$ is the average treatment effect of receiving a Bandesal-funded loan in quarter g relative to never (or not yet) receiving it by time period t, among units that belong to cohort g. In this set-up, t is the post and $g - 1$ is the pre period, while $G_i = g$ is the treated and $G_i > t$ is the comparison group. Our goal is to estimate the overall average treatment effect across all business owners, as well as understand how these effects vary over event-time $e = t - g$. Therefore, we choose two ways of aggregating the ATTs from equation (2.2):

$$ATT_{simple} = \sum_{g,t} w_{\omega,g,t} ATT(g,t) \tag{2.3}$$

$$ATT_{event}(e) = \sum_{g<\infty} w^{event}_{\omega,g,e} ATT(g, g+e) \tag{2.4}$$

where $w_{\omega,g,t}$ are ω-weighted group and time-specific (non-negative) weights that sum up to one, while each weight $w^{event}_{\omega,g,e}$ gives the share of a group $G = g$ among treated units that have been exposed to treatment for exactly e periods. We also report a simple average of all post-treatment event times, $ATT_{event}(e)$, when $e \geq 0$.

2.4.3 Identifying Assumptions

For causal identification of the aggregated parameters described in Section 2.4.2, we require the no anticipation and parallel trends assumptions.

First, defining $Y_{i,t}(\infty)$ to denote never-treated potential outcomes, the no anticipation assumption states that for all units i that are eventually treated and all pre-treatment periods t, $Y_{i,t}(g) = Y_{i,t}(\infty)$, that is, that there is no causal effect before the treatment takes place. Because Bandesal is a second-tier lender, final beneficiaries receive their loans from intermediary financial institutions and never directly from Bandesal. Therefore, it is unlikely that individual business owners modified their behavior in anticipation of the increase in credit supply.

Second, the key identifying assumption is that the average outcome among treated and control individuals, in absence of the loan, would have been parallel. Because our sample includes the entire productive credit portfolio of Bandesal, it is possible that other baseline characteristics can determine changes in untreated potential outcomes, particularly prior credit history. Therefore, we condition our parallel trends assumption on a vector of characteristics X_i, including: gender, location (urban vs. rural, as well as inside the capital city metropolitan area), an indicator for micro enterprise, for operating in the service sector, and the financing objective of the loan (investment vs. working capital). Importantly, we include lagged variables that encompass credit history 6 months before obtaining the Bandesal-funded loan: probability of having outstanding debts, debt amounts in USD, the probability of having debt-past-due in the last 30 days, and its corresponding outstanding amount, for the four types of debt sources: banking-productive, non-banking, consumption, and other types of loans. Table 2.1 shows the mean of these variables across cohorts.

We formally describe our assumption as in Baker et al. (2025). For every eventually-treated group g, not-yet-treated group g' and time periods t, such that $t \geq g$ and $g' > t$, and every covariate value X_i:

$$E_\omega[Y_{i,t}(\infty) - Y_{i,t-1}(\infty)|G_i = g, X_i] = E_\omega[Y_{i,t}(\infty) - Y_{i,t-1}(\infty)|G_i = g', X_i] \qquad (2.5)$$

The conditional parallel trends assumption also relies on a strong overlap condition, which requires that the conditional probability of being in the treatment group g, given observed characteristics X_i that influence potential outcome growth in the absence of treatment, remains bounded away from both zero and one across the sample. Figure 2.2 shows the overlap of the distribution of the probability of being treated for cohorts 2015Q3-2018Q2, compared to the two groups in our sample who we treat as never treated: 2018Q3-Q4. Although this is not the exact overlap in our estimation (given that we use a not-yet-treated comparison group as well), it helps visualize how the overlap condition works.

2.4.4 Estimation

We rely on the CSDID Stata package to estimate the group-time average treatment effects. We use the the default doubly-robust (dripw) estimator based on stabilized inverse probability weighting and ordinary least squares. (Sant'Anna and Zhao, 2020) This option is generally preferred because offers robustness and improved efficiency compared to relying solely on the propensity score. Following Roth (2024), we specify the option that requires the estimation of long-gaps, for comparability to conventional event-study plots.

2.5 Results

2.5.1 Main Results

The validity of our identification strategy requires similar trends between treatment and control before the intervention. Figures 2.3 and 2.4 report the quarterly effects before and after receiving the Bandesal-funded loan for the unconditional and conditional model, respectively. Each figure is divided into four panels that represent our four main outcomes. Panel A shows the effects for the probability of having additional banking productive loans, Panel B shows the outstanding amount of debt in real USD in that given time period, Panel C shows the results for the probability of having debt past due for more than 30 days, and Panel D the outstanding amount of debt past due in real USD.

Figure 2.3 shows the *unconditional* model, where we observe a violation to the parallel trends assumption, specially in Panel A. Figure 2.4 shows the model *conditional on the vector of baseline characteristics* discussed in Section 2.4. As our preferred specification, the discussion of all the results is based on this model.

Visually, in Figure 2.4 we verify the presence of parallel trends before treatment for all outcomes. In three out of four panels, the p-value of the pre-trend test is greater than 0.4, and we therefore conclude that there are no pre-trends for outcomes shown in Panels A, B, and D. For Panel C, the probability of having debt past due, we reject the null that all values are equal to zero with a p-value of 0.00. However, the amount of debt-past-due exhibits parallel trends (p-value of 0.96), and magnitudes are relatively small. Therefore, we proceed with caution when interpreting these effects causally. In addition, the event study plots show wide confidence intervals the farther the time periods are from treatment, because only one cohort is observed in the earliest (or latest) period.

The aggregate average treatment effects are summarized in table 2.2, with our five main outcomes represented in each column. Overall, we find that receiving a Bandesal-funded loan increases business owners' likelihood of receiving additional loans from banking-productive sources by 8.2 percentage points, which represents a 39% increase compared to the control mean of never treated cohorts (2018Q3-Q4). In Figure 2.4 we observe the dynamics over time of this effect: the probability of additional banking productive loans increases in the first five periods after treatment (in our setting, 1.5 years), but fades out after that. Although the estimates in the latest periods show zero impacts, it is important to note that these belong to the first cohorts, which can be observed at 10 or more periods after treatment. Fewer observations explain the loss of precision and wider confidence intervals.

To understand the intensive margin of this effect, we turn to column 2 of Table 2.2. We see that receiving a Bandesal-funded loan increases the outstanding amount of debt by USD $1,366.94, an 88% increase over the control mean. Considering the median amount of the Bandesal-funded loan in our sample was USD $3,000, we find that although the intervention allows beneficiaries to increase the amount they borrow from the formal credit market, it does not surpass the amount received from publicly-funded sources.

Importantly, this increase in banking productive loans does not come at the expense of increasing other types of less-favorable debt. We observe this by studying changes to the share of banking productive loans out of the total debt portfolio of business owners. As shown in column 3 of table 2.2, receiving a Bandesal-funded loan increases the share of banking productive loans by 6.8 percentage points, a 40% increase when compared to the never treated groups. Furthermore, figure 2.5 shows that Bandesal-funded loans do not increase the probability of obtaining consumption or other loans (credit card, auto, commercial, etc.), respectively. We interpret this as a shift towards better-quality debt, as banking productive loans tend to have lower interests and longer maturities than consumption and commercial loans.

Finally, the increase in the probability of new productive loans is accompanied by a slight increase in the probability of having debt-past due for more than 30 days, as shown in column 4 of Table 2.2. Although receiving a Bandesal-funded loan increases the probability of default by 2.2 percentage points, the average outstanding amount past-due is USD $7.60, as shown on column 5. [3]

In general, our main results indicate higher participation in the formal credit market for clients served by a Bandesal-funded loan. This increase in participation materializes itself in the first quarter after receiving the Bandesal-funded loan, it is sustained for approximately 1.5-2 years, and is not accompanied by large default rates. Although we find these results optimistic, our study does not allow us to observe if business owners resort to informal credit markets to pay-off their formal credit market debts, and therefore our findings must be interpreted with caution.

2.5.2 Heterogeneity by Credit History

How heterogeneous is the increase in credit access? In this section, we explore whether past credit history implies a differential response to the greater credit availability brought by Bandesal.

We define a time horizon of 24 months to classify customers with certain types of credit history. In the first group, termed *No History* (24% of the sample) we group individuals with no registered financial activity in the 24 months preceding treatment. The second category, *Good History* (41%), groups individuals with financial activity in the 24 months before treatment and no reported unpaid debt. The third category, *Curing History* (14%), groups customers that at any point in the 24 months before treatment delayed repayment but managed to get back on track in the remaining period before treatment. The final category, *Bad History* (21%), groups customers with reported unpaid debt by the time they were treated.

Table 2.3 shows the results of this exercise. Column 1 indicates that the probability of obtaining new banking productive loans is significantly higher for individuals with no credit history in the past 24 months: it is an increase of 17.9 percentage points, where the never treated cohorts in that category have no active references by construction. The effects are larger in magnitude for this group than for those with good credit history (3.6 pp), curing credit history (9.2 pp), and bad credit history (9.6 pp). On the intensive

[3] As a reference, International Monetary Fund estimates show that the average ratio of defaulting loans (payments of interest and principal past due by 90 days or more) to total gross loans (total value of loan portfolio) was 2% between 2015 and 2018 for El Salvador. (World Bank, 2025)

margin, it appears that the group with no prior credit history benefits the most of the intervention: it increases the outstanding loan amount by USD \$1,496 (column 4), and the share of banking productive loans increases by 16.4 percentage points (column 3). Reassuringly, the higher increase in the probability of obtaining new debt is not linked with higher default rates for this group, as shown in column 2. Figure 2.6 plots the aggregated dynamic-event average treatment effects for each group, along with their corresponding 95% confidence intervals for ease of comparison.

Taken together, these results indicate that the increase in the probability of accessing new productive loans is relatively more important for business owners that enter the formal credit market because of the Bandesal-funded loan. However, average effects remain economically significant for beneficiaries with any type of credit history.

2.6 Conclusion

This paper examines whether providing publicly funded productive loans to small and medium business owners fills a crucial gap—serving as a credential for accessing higher-quality formal credit—or whether it inadvertently "digs a hole" by leading borrowers into over-indebtedness and deteriorating credit profiles. In essence, the study investigates whether credit expansions can effectively integrate previously underserved borrowers into the formal financial system while avoiding adverse consequences.

To tackle this question, we implement a difference-in-differences strategy with multiple time periods and staggered treatment timing. Leveraging detailed administrative records alongside individual credit bureau data, the analysis compares early borrowers with those treated later in the credit expansion. The design harnesses variation in treatment timing, conditions on baseline characteristics (including 6-month credit history), and uses improved doubly robust estimator based on inverse probability of tilting and weighted least squares. (Sant'Anna and Zhao, 2020)

Overall, we find that recipients of Bandesal-funded loans are more likely to secure subsequent high-quality banking loans, with an increase of up to 8 percentage points. This effect is most pronounced among borrowers with no credit history in the

24 months preceding treatment, suggesting the loan functions effectively as a financial credential—supporting the "filling a hole" hypothesis. However, despite these optimistic results, caution is warranted in interpretation. The analysis is limited to formal credit market data, and it does not capture potential shifts to the informal credit sector. As a result, further research is needed to understand the broader financial behaviors of borrowers, including any movement towards informal borrowing channels.

Tables and Figures

Table 2.1: Baseline Characteristics By Cohort (Mean)

	2015Q3	2015Q4	2016Q1	2016Q2	2016Q3	2016Q4	2017Q1	2017Q2	2017Q3	2017Q4	2018Q1	2018Q2	2018Q3 (NT)	2018Q4 (NT)	Total
A. Characteristics at the time of receiving the Bandesal-funded loan															
Female	0.58	0.54	0.62	0.74	0.75	0.67	0.61	0.59	0.48	0.50	0.39	0.36	0.47	0.53	**0.57**
San Salvador	0.20	0.15	0.16	0.12	0.15	0.11	0.12	0.16	0.19	0.21	0.23	0.32	0.22	0.26	**0.18**
Urban	0.59	0.53	0.52	0.44	0.43	0.38	0.47	0.52	0.54	0.55	0.59	0.63	0.54	0.56	**0.51**
Micro-enterprise	0.83	0.83	0.86	0.91	0.92	0.93	0.86	0.86	0.87	0.88	0.93	0.88	0.85	0.71	**0.88**
Service Sector	0.85	0.88	0.84	0.75	0.87	0.86	0.77	0.83	0.85	0.82	0.82	0.85	0.81	0.76	**0.83**
Investment Loan*	0.70	0.81	0.67	0.30	0.30	0.33	0.57	0.47	0.37	0.29	0.17	0.19	0.21	0.11	**0.37**
B. Municipality-by-quarter Homicide Rate (per 100,000)															
0-25	0.11	0.11	0.09	0.06	0.09	0.07	0.07	0.05	0.04	0.05	0.05	0.05	0.05	0.05	**0.07**
25-50	0.19	0.16	0.15	0.16	0.14	0.17	0.12	0.13	0.14	0.13	0.11	0.16	0.18	0.17	**0.15**
50-75	0.25	0.22	0.25	0.27	0.31	0.34	0.32	0.28	0.33	0.31	0.33	0.33	0.24	0.32	**0.29**
75-100	0.19	0.24	0.27	0.23	0.23	0.20	0.27	0.29	0.27	0.30	0.28	0.22	0.24	0.19	**0.24**
100-150	0.15	0.13	0.13	0.15	0.12	0.10	0.10	0.11	0.11	0.12	0.13	0.13	0.13	0.11	**0.12**
More than 150	0.12	0.14	0.10	0.13	0.12	0.11	0.12	0.14	0.10	0.10	0.10	0.10	0.17	0.15	**0.12**
C. Lagged (t-1) Credit History															
Total Banking Productive Debt															
Pr(Debt¿0)$_{t-1}$	0.14	0.17	0.15	0.11	0.09	0.11	0.21	0.18	0.19	0.18	0.14	0.14	0.29	0.31	0.16
Pr(Debt Past Due¿30 days)$_{t-1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00
Outst. amount in USD$_{t-1}$	2124.08	2044.31	3146.29	1765.16	1761.43	961.13	2858.10	1462.24	2650.61	1171.76	1084.31	1517.18	5893.34	3242.06	2340.64
Outst. amount past-due in USD$_{t-1}$	1.12	7.95	3.45	0.96	6.71	0.56	0.70	5.83	2.04	0.95	7.91	10.93	8.25	0.00	4.57
Total Non-Banking Productive Debt															
Pr(Debt¿0)$_{t-1}$	0.16	0.14	0.16	0.15	0.13	0.15	0.14	0.15	0.16	0.13	0.14	0.16	0.17	0.17	0.15
Pr(Debt Past Due¿30 days)$_{t-1}$	0.02	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.01	0.01	0.01	0.01
Outst. amount in USD$_{t-1}$	829.25	723.26	899.89	1037.02	773.78	838.34	844.88	849.95	1000.23	579.30	428.30	985.68	1056.03	1205.70	860.55
Outst. amount past-due in USD$_{t-1}$	14.16	0.61	2.58	3.17	5.87	1.38	1.41	1.72	32.05	1.72	7.28	8.47	9.85	5.16	6.54
Total Banking Consumption Debt															
Pr(Debt¿0)$_{t-1}$	0.26	0.24	0.25	0.21	0.21	0.23	0.23	0.26	0.23	0.28	0.24	0.24	0.23	0.25	0.24
Pr(Debt Past Due¿30 days)$_{t-1}$	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01
Outst. amount in USD$_{t-1}$	2225.23	1689.34	1891.50	1456.50	1680.90	1682.05	1676.01	2010.59	2004.28	1418.02	1203.92	1135.48	2214.61	1891.02	1711.61
Outst. amount past-due in USD$_{t-1}$	24.78	1.30	3.50	1.07	35.46	10.90	6.87	3.79	4.03	13.53	10.65	7.24	16.08	12.33	11.48
Other Debt (credit cards, auto, etc.)															
Pr(Debt¿0)$_{t-1}$	0.35	0.36	0.38	0.35	0.35	0.36	0.42	0.40	0.39	0.34	0.37	0.38	0.34	0.44	0.37
Pr(Debt Past Due¿30 days)$_{t-1}$	0.04	0.06	0.05	0.06	0.07	0.07	0.06	0.07	0.06	0.07	0.08	0.09	0.07	0.05	0.06
Outst. amount in USD$_{t-1}$	1446.51	1229.63	1353.95	967.08	935.21	885.66	1814.46	1425.04	1508.01	858.28	1150.24	1279.03	1333.17	1903.08	1227.32
Outst. amount past-due in USD$_{t-1}$	38.76	27.20	25.67	22.99	49.72	27.77	30.53	18.46	25.82	31.75	46.31	30.69	43.10	26.75	33.18
Observations	6,846	7,059	8,316	13,390	11,382	9,229	5,432	5,992	5,655	5,558	9,814	8,246	13,230	3,570	**113,719**

Notes: This table shows the mean by cohort for selected sample characteristics. Each column represents one treatment cohort from 2015q3 to 2018q2. We also include information about the last two cohorts, are never treated (NT) since the analysis runs until 2018q2. The last column reports the overall mean for the entire sample. Panel A shows information *at the time of receiving the Bandesal-funded loan*, Panel B reports the average homicide rate per 100,000 inhabitants, at the municipality-quarter level. Panel C shows the lagged (t-1) values for formal credit market outcomes for four different sources of debt. *Investment loan is a binary variable that refers to the type of Bandesal-funded loan the individual received: for investment as opposed to a working capital loan.

Table 2.2: Aggregate Average Treatment Estimates, Conditional

	(1) Pr(Debt> 0)	(2) Outstanding Amount of Debt in Real USD	(3) Share of Banking Productive Loans	(4) Pr(Debt-Past-Due > 30 days)	(5) Outstanding Past-Due Amount of Debt in Real USD
ATT_{simple}	0.101***	1342.012***	0.084***	0.014***	4.715***
	(0.008)	(140.069)	(0.007)	(0.001)	(0.452)
Pre: $ATT_{event}, e < 0$	0.002	167.707	0.009	-0.001	0.051
	(0.026)	(798.287)	(0.031)	(0.001)	(1.044)
Post: $ATT_{event}, e \geq 0$	0.082***	1366.941***	0.068***	0.022***	7.599***
	(0.013)	(237.642)	(0.011)	(0.002)	(0.792)
Observations	113,719	113,719	113,719	113,719	113,719
Control Mean (NT)	0.21	1,560.43	0.17	0.00	0.80
Control SD (NT)	0.41	7,776.89	0.35	0.05	13.34
P-value (Pre-trend Test)	0.62	0.67	0.42	0.00	0.96

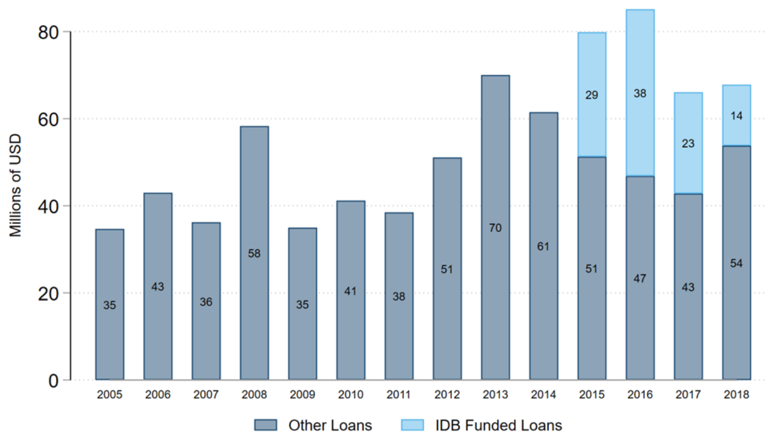
Notes: This table shows the aggregate ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. We show the 'simple' ATT for all groups across all periods, as well as the dynamic ATT, that is, an average of all negative (pre) and non-negative (post) event times. Column 1-5 show the ATT for different outcomes: the probability of obtaining additional banking-productive loans, the outstanding amount of debt in real USD, the share of banking productive loans out of the total outstanding loans in a given period, the probability of having debt past due in the last 30 days, and the outstanding amount of debt past-due in real USD, respectively. To ease interpretation, we report the mean and standard deviation for the control group that was never treated (the two last cohorts in our data). Standard errors in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table 2.3: Heterogeneity by Credit History: Aggregate Average Treatment Estimates

	(1)	(2)	(3)	(4)
	Pr(Debt> 0)	Debt-Past-Due > 30 days	Share of Banking Productive Loans	Outstanding Amount of Debt in USD
A. No prior credit history:				
Pre: $ATT_{event}, e < 0$	-0.004	-0.001	-0.003	-5.808
	(0.003)	(0.001)	(0.002)	(4.569)
Post: $ATT_{event}, e \geq 0$	0.179***	0.016***	0.164***	1495.995***
	(0.011)	(0.003)	(0.010)	(227.627)
B. Good credit history:				
Pre: $ATT_{event}, e < 0$	-0.003	-0.000	0.005	-75.766
	(0.010)	(0.001)	(0.010)	(129.065)
Post: $ATT_{event}, e \geq 0$	0.036**	0.023***	0.011	1061.336*
	(0.014)	(0.003)	(0.013)	(435.822)
C. Curing credit history:				
Pre: $ATT_{event}, e < 0$	-0.053	0.006	0.037	879.511
	(0.050)	(0.004)	(0.034)	(764.312)
Post: $ATT_{event}, e \geq 0$	0.092*	0.026***	0.073*	-382.919
	(0.044)	(0.006)	(0.036)	(1329.687)
D. Bad credit history:				
Pre: $ATT_{event}, e < 0$	0.021	-0.002	0.012	225.667
	(0.022)	(0.004)	(0.014)	(367.931)
Post: $ATT_{event}, e \geq 0$	0.096*	0.030***	0.102*	1353.527
	(0.049)	(0.007)	(0.046)	(745.428)

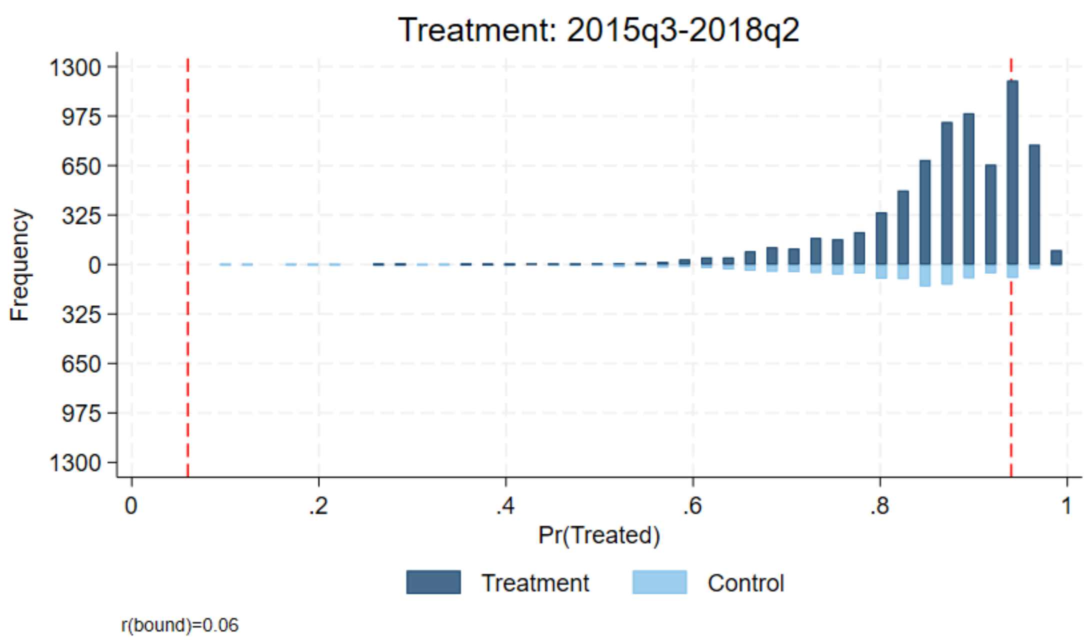
Notes: This table shows the aggregate ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. We show the dynamic ATT, that is, an average of all negative (pre) and non-negative (post) event times. Column 1-4 show the ATT for different outcomes: the probability of obtaining additional banking-productive loans, the probability of having debt past due in the last 30 days, the share of banking productive loans out of the total outstanding loans in a given period, and the outstanding amount of debt in USD, respectively. Each panel reports the estimates for different samples: individuals with no prior credit history (Panel A), good credit history (Panel B), curing credit history (Panel C), and bad credit history (Panel D). Standard errors in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Figure 2.1: Bandesal's Productive Loan Portfolio



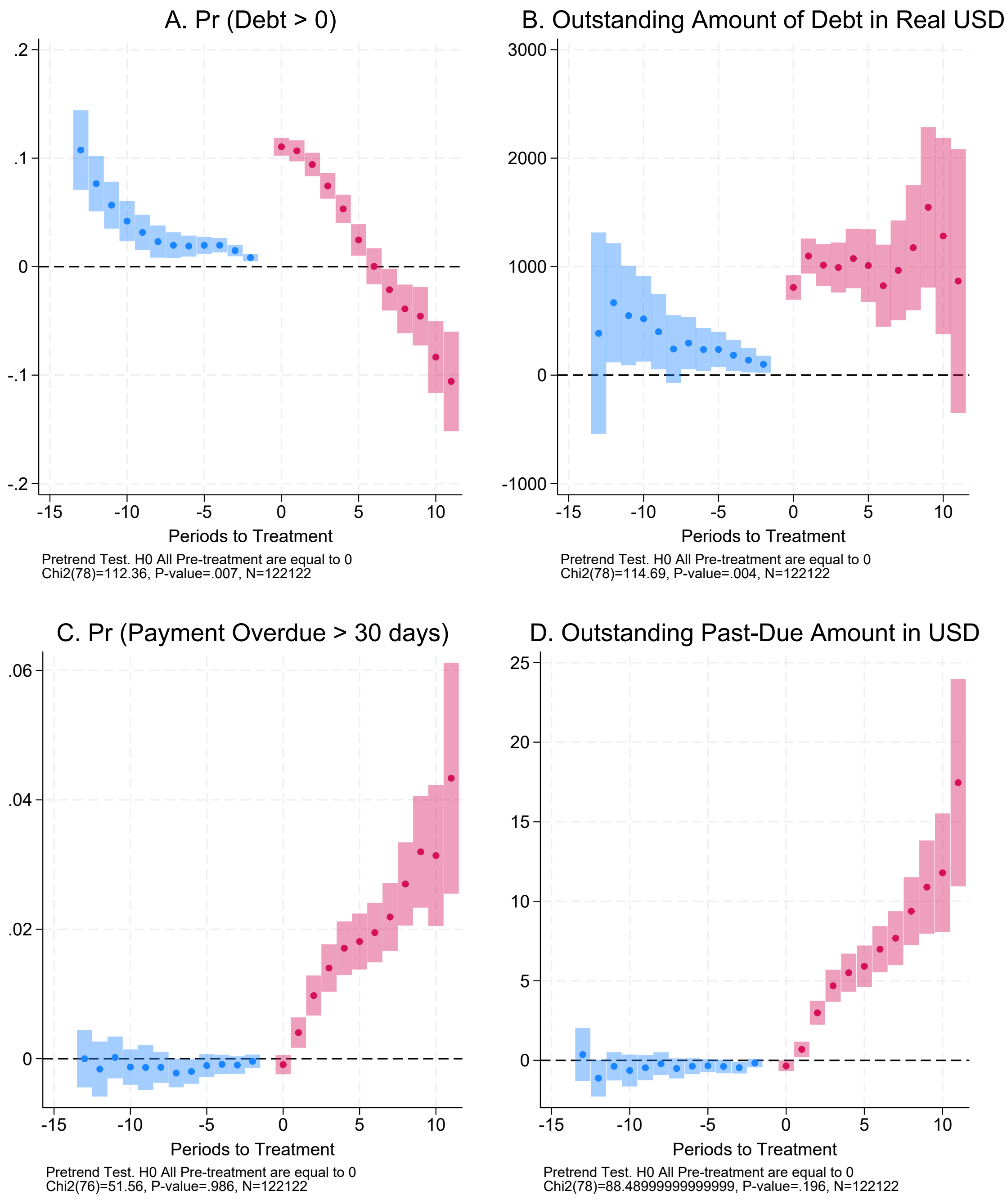
Notes: This figure shows the productive loan portfolio of El Salvador's National Development Bank, Bandesal, from 2005 to 2018. Each bar represents the total amount of loans given out in a year, measured in millions of USD. The light blue represents the total amount of loans that were financed by the Inter-American Development Bank, while the dark blue represents loans that were financed by other lenders.

Figure 2.2: Probability of Being Treated - Overlap



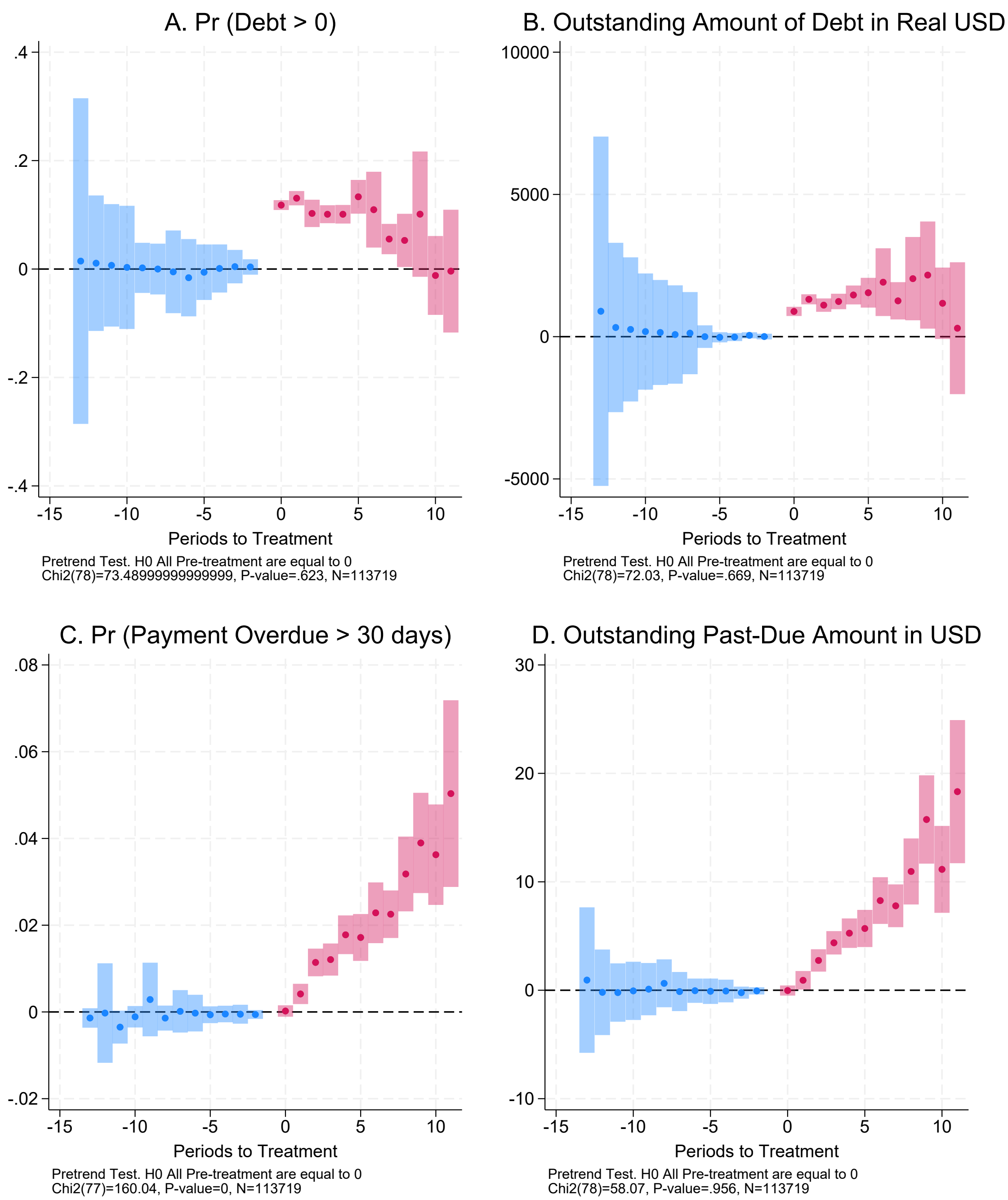
Notes: This figure shows the distribution of the probability of being treated, defining treatment group as cohorts 2015Q3-2018Q2, and control group as cohorts 2018Q3-2018Q4. The probability of being treated is estimated using a logit estimator regressing the treatment indicator on baseline characteristics at the time the Bandesal-loan was received, including 6 months of prior credit history, gender, location, indicator for micro-enterprise, tertiary sector and financing loan objective, as well as municipality homicide rates. The dotted red lines indicate the optimal bound selected using Crump et al. (2009).

Figure 2.3: Event Study: Banking Productive Loans, Unconditional



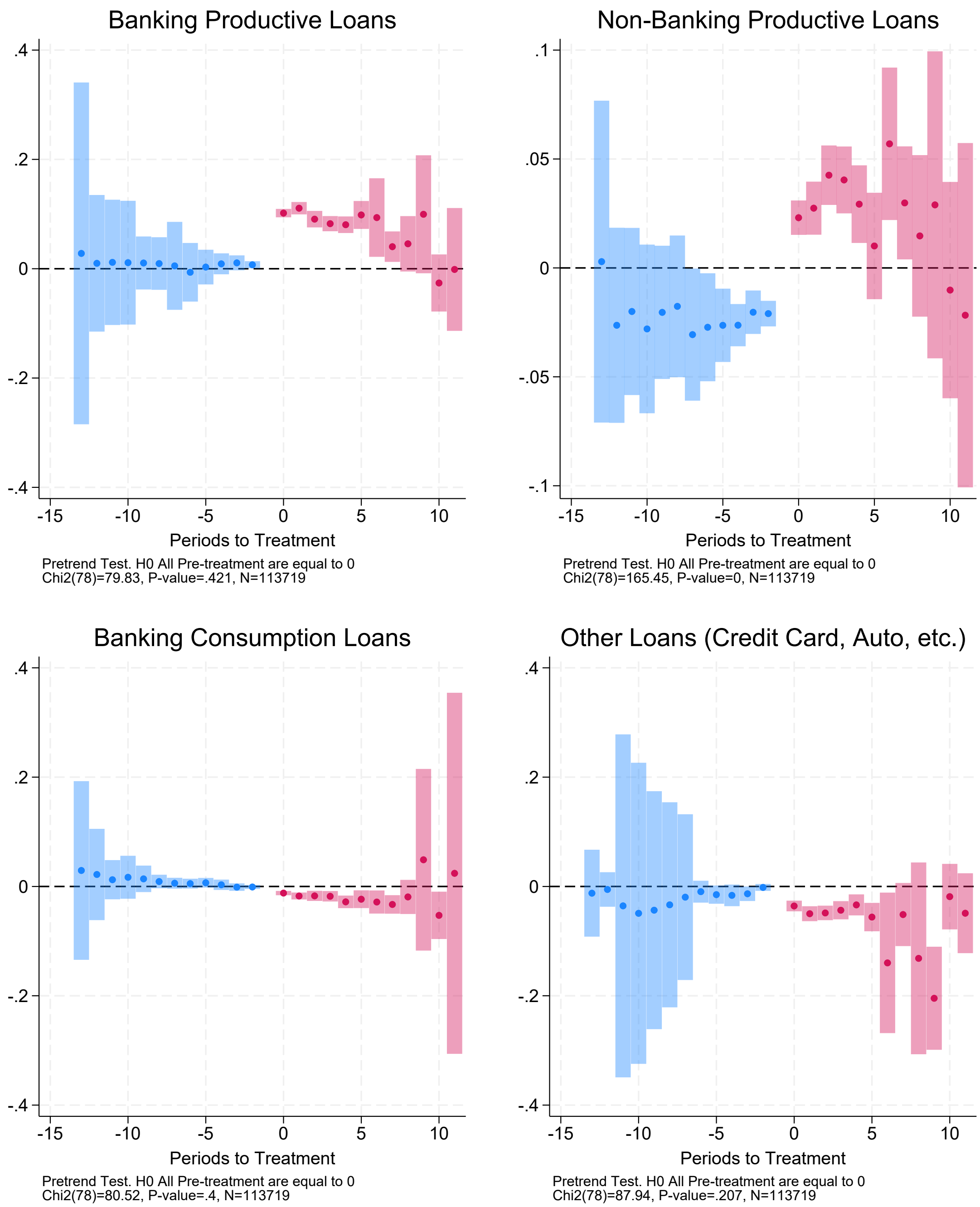
Notes: This figure shows the unconditional event study estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure 2.4: Event Study: Banking Productive Loans, Conditional



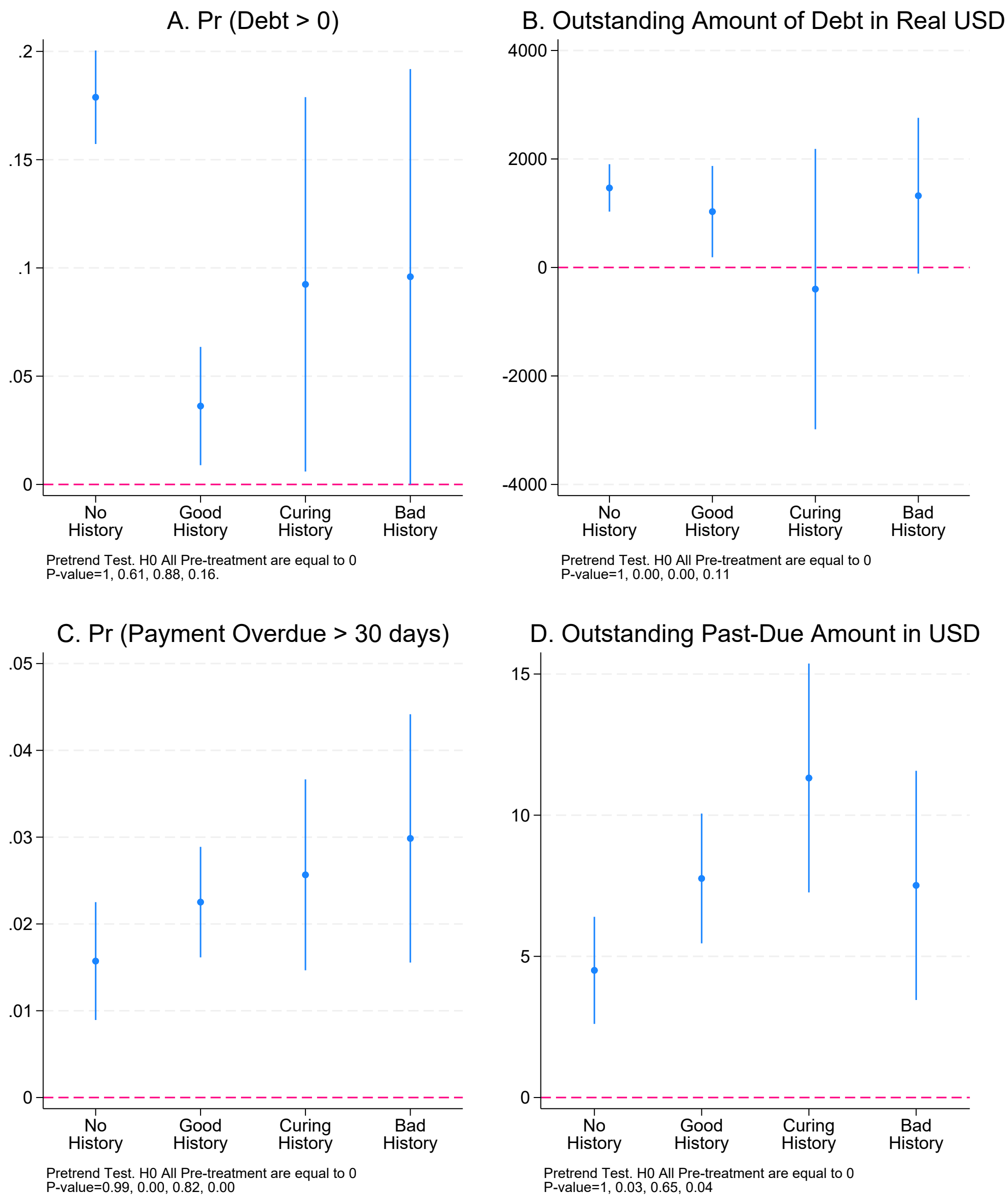
Notes: This figure shows the conditional event study estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Controls include: gender, location, an indicator for micro enterprise, for operating in the service sector, and the financing objective of the loan (investment vs. working capital), as well as lagged variables for past 6 months of credit history. Each panel is an outcome variable within the category of banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively.The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure 2.5: Event Study: Share of Different Debt Sources



Notes: This figure shows the event study estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is a different debt source: (A) banking productive, (B) non-banking productive (such as credit cooperatives), (C) banking consumption, and (D) other debt (including credit cards, auto, commercial, and telecommunications). The outcome variable (share of debt source) is constructed dividing the outstanding amount of USD in a given debt source, over the total outstanding amount of USD for each individual.The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure 2.6: Heterogeneity by Credit History



Notes: This figure shows the aggregated treatment effect for all post-periods ($ATT_{event}, e \geq 0$) using the doubly-robust estimation method from Callaway and Sant'Anna (2021) with not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. Each panel includes four separate regressions for four different samples: (1) no credit history, (2) good credit history, (3) curing credit history, and (4) bad credit history in the past 24 months (for more information on these definitions, see Section 2.3. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Chapter 3

It's all Fun and Games? The Persistent Treatment Effects of Willingness-to-Pay Experiments

by Jenny C. Aker, Brian Dillon, Leticia Donoso-Peña, and Anne Krahn*

*We are grateful for comments from seminar participants at Oxford University, Tilburg University, Yale University, the University of Urbana-Champagne Illinois, CSAE and NEUDC. This research is funded by J-PAL and CEGA's Agricultural Technology Adoption Initiative (ATAI) with generous support from the Bill and Melinda Gates Foundation and UK Aid.

3.1 Introduction

Willingness-to-pay (WTP) experiments have been widely used to assess demand for a variety of goods and services in behavioral, development and environmental economics, ranging from insecticide-treated bednets to human capital goods and market information services. Most studies use such experiments in one of four ways: to elicit WTP on its own, without additional interventions or analyses (Aker, Prina, and Welch, 2020; Channa et al., 2019; Grimm et al., 2020; Burchardi et al., 2021; Dizon-Ross and Jayachandran, 2023; Burchardi et al., 2024); as a means of allocating access to a particular good or service at *baseline*, with follow-up measures of downstream outcomes (**?**Shukla, Pullabhotla, and Baylis, 2022; Meriggi, Bulte, and Mobarak, 2021; Berry, Fischer, and Guiteras, 2020; Janzen et al., 2021; Berry and Mukherjee, 2019; Lybbert et al., 2018; Hoffmann, 2018); to estimate demand at baseline, *before* implementing policy interventions (**?**Hidrobo et al., 2022; Butera et al., 2022; Andor et al., 2023); or to assess demand *after* a given policy or intervention has taken place (Hoffmann, Barrett, and Just, 2009; Ashraf, Berry, and Shapiro, 2010; Dupas, 2014; Ben Yishay et al., 2017; Bensch and Peters, 2017).

Beyond estimating demand curves, such experiments can address important barriers to technology adoption, either by providing *information* on the technology, improving *access* to it, or increasing its *salience*. As a result, the act of participating in a WTP experiment can induce changes in behavior, thereby affecting downstream outcomes and potentially confounding external validity. Yet few, if any, studies estimate the stand-alone treatment effects of such experiments.[2]

This paper aims to fill this gap by estimating the dynamic impacts of a demand elicitation experiment. We randomly assign villages to participate in a WTP experiment for an improved storage technology, and then collect data on a number of outcomes over 3.5 years.[3] Overall, we find that the WTP experiment significantly modified

[2] A subset of studies compare different types of WTP experiments, or compare the WTP experiment with a subsidized price. For example, Shukla, Pullabhotla, and Baylis (2022) compare the impacts of being able to purchase a technology via a WTP experiment with fixed and free prices. However, there is no pure control, so they are unable to estimate the impact of the WTP experiment on its own.

[3] The specific technology is the Purdue Improved Cowpea Storage (PICS) bag, a hermetically-sealed bag that can kill pests without the use of dangerous pesticides, namely, rat poison.

households' behavior in the short-, medium- and long-term: 92% of households owned the technology nine months after the game, dropping to 67% more than three years later, primarily because the technology fully depreciated. As a result, farmers in WTP villages disadopted other storage technologies, including dangerous pesticides, and suffered fewer storage losses. These results persisted 3.5 years after the initial experiment, beyond the traditional "shelf-life" of the technology. We do not find impacts on other downstream outcomes, nor does the experiment crowd in additional demand for the new technology.

What are the channels through which the effects of WTP experiment persist? The experiment could have affected behavior via *information* about the product, *experience* with it (Dupas and Miguel, 2017), increased *salience* at a key moment or greater *enthusiasm* due to the nature of the game. To partially disentangle the first two mechanisms, we exploit the random draw price as an instrument for winning the technology. We find that the results are significantly stronger for winners, suggesting that our findings are partially driven by experience with the product, rather than information. We not find evidence that the experiment made storage decisions more salient. Finally, there is some suggestive evidence that the game generated enthusiasm in the product amongst winners, but had mixed effects on non-winners.

These results suggest that baseline WTP could potentially affect the treatment effects of follow-on experiments. We are able to test for this directly, as our WTP experiment was embedded in a larger study. We show that the treatment effects of the larger study are only statistically significant in areas where the demand elicitation experiment was implemented.

Our paper's main contribution is to quantify the treatment effects of commonly-used WTP experiments on adoption outcomes. The standard approach in the literature is fourfold. A first subset of studies use WTP experiments to elicit willingness to pay for a specific good at baseline, and not conduct additional analyses (Aker, Prina, and Welch, 2020; Channa et al., 2019; Grimm et al., 2020; Burchardi et al., 2021; Dizon-Ross and Jayachandran, 2023; Burchardi et al., 2024). A second set of studies elicits WTP after providing information, credit or subsidies for the product, thus assessing the impact of specific interventions on WTP (Hoffmann, Barrett, and Just, 2009; Ashraf, Berry, and Shapiro, 2010; Dupas, 2014; Ben Yishay et al., 2017; Bensch and Peters, 2017). A third subset of studies uses demand elicitation as a mechanism to induce random exposure to the technology, and measures the impact of this variation on downstream outcomes (e.g., (Shukla, Pullabhotla, and Baylis, 2022; Meriggi, Bulte, and Mobarak, 2021; Berry, Fischer, and Guiteras, 2020; Janzen et al., 2021; Berry and Mukherjee, 2019; Lybbert et al., 2018; Hoffmann, 2018). A final set of studies elicits WTP at baseline, prior to the additional interventions (Hidrobo et al., 2022; Butera et al., 2022; Andor et al., 2023). Our study is unique in that it randomly varies access to the WTP experiment as compared with a pure control.[4] Similar to the literature on the impact of being surveyed on respondents' behavior (Zwane et al., 2011; Treurniet, 2023), we show that demand elicitation experiments can have substantial impacts on technology adoption and downstream outcomes.

Our study also fits into the broader literature on the dynamic and persistent impacts of short-run interventions. Such studies find that one-time policies can have persistent effects on adoption, in part due to learning and spillovers (Dupas, 2014; Bensch and Peters, 2017; Carter, Laajaj, and Yang, 2021; Deutschmann, 2024; Balew, Bulte, and Kassie, 2024). By using in-person and phone surveys over a four-year period, we are able to document the dynamic and long-run treatment effects of a one-time experiment. Yet we do not find evidence of learning, perhaps because baseline knowledge of the product was high and beliefs were accurate.

Finally, our findings suggest that demand elicitation experiments can potentially alter the external validity of the studies in which they are embedded, especially if conducted prior to an intervention. This problem has an easy fix: By measuring demand in a random subsample at baseline, researchers could test whether WTP elicitation affects adoption outcomes.

The remainder of the paper is organized as follows. Section 3.2 describes the ex-

[4] Shukla, Pullabhotla, and Baylis (2022) assign villages to one of three treatments: a WTP auction, a fixed price or a free treatment, and compare their impacts on WTP later. There is no pure control.

perimental design, data and empirical strategy, whereas Section 3.3 presents the key results. Section 3.4 concludes.

3.2 Experimental Design and Data

3.2.1 The WTP Experiment

Our primary intervention is a WTP experiment embedded in a household survey. The experiment was a two-stage, incentive compatible variant of the Becker-Degroot-Marschak (BDM) (Becker, DeGroot, and Marschak, 1964) mechanism designed to elicit respondents' WTP for an improved storage technology. The technology is a hermetically-sealed bag that can store commodities such as cowpeas and maize without pesticides, and lasts *at least* three agricultural seasons before needing to be replaced (Aker, Dillon, and Welch, 2023; Omotilewa et al., 2018).[5]

After presenting the respondent with the technology and explaining its attributes, the respondent was asked whether he or she would be willing to purchase the good at a series of prices in increasing order from zero to above market price.[6] The respondent was informed that, after the price sequence, one price would be randomly drawn, and he or she would have the opportunity to purchase the good at the drawn price if his or her maximum WTP was at or above that price. The enumerator explained the process in detail and confirmed the respondent's maximum WTP prior to the random draw. If the respondent won, the sale took place after a short "cooling off" period, approximately one hour. Unlike other studies that have had a substantial share of "decliners" – i.e., those participants who refused to pay the drawn price if they won – all of the respondents paid in our context.[7] Respondents were not compensated for their participation in the game nor the survey.[8]

The WTP experiment was implemented over a four-week period in November and

[5]Other commonly-used storage technologies are nylon bags with pesticides, which last one year, and 20-kg plastic jugs, which can last up to five years.

[6]At the time of the survey, the average market price for the technology was 1000 CFA (2 USD), but was available on fewer than 10% of markets. We included 5000 CFA in the price list in order to set the intercept.

[7]As noted in Maffioli, McKenzie, and Ubfal (2023), rates of "decliners" in WTP experiments can vary widely, from 0 to 54%. While our study had no decliners, 1.5% of participants refused to play the game.

[8]Unlike other studies using the BDM mechanism, we did not not play a practice round using another product. However, we had done substantial piloting the BDM experiment during a prior study.

December 2020, immediately after the harvest and during the baseline of a broader study designed to address barriers to the adoption of the hermetically-sealed bag. The WTP experiment was also timed to ensure that the technology was salient for households as they were making storage decisions.[9] Data collection activities are provided in Appendix Table C1.

How might the WTP experiment affect demand for a new technology? There are a number of ways. First, as the game provides *information* about the technology and how to use it, this could address an important barrier to adoption, especially for those who had never heard about the technology. Second, since the technology is an experience good, those who win the game are able to learn by using it. Third, the timing of the experiment (immediately after the harvest) may make storage more *salient*, thereby bringing storage expenditures top of mind, at least for the first year. And finally, similar to other behavioral experiments, playing the game might generate *excitement* about the technology, thereby encouraging sustained usage (Janzen et al., 2021). While all of these factors would potentially *increase* demand for the technology, the WTP experiment could also dampen demand if households purchased at subsidized prices, thereby reducing the "sunk cost effect" (Ashraf, Berry, and Shapiro, 2010; Cohen and Dupas, 2010), or if non-winners were loss averse and disappointed by not winning. In addition, since the technology is semi-durable, if the experiment did not crowd in additional demand, it could reduce market demand and hence supply on local markets.

3.2.2 Experimental Design

In November 2019, we identified 220 villages in the Maradi and Zinder regions of Niger. Villages were stratified by region and market size before being randomly assigned to either the treatment (*WTP*) or control (*no WTP*).[10]

[9]The WTP elicitation was conducted in the context of a larger experiment, which provided information to farmers and traders about the improved storage technology and its relative costs with alternative technologies. These interventions took place approximately one year after the baseline WTP experiment, and the experiments were cross-randomized with the WTP experiment.

[10]The randomization resulted in 107 villages assigned to treatment and 113 to control. The slight imbalance was due to the modifications to the village sample during the baseline. A total of 226 villages were originally identified during the census and randomly assigned to treatment and control. Six villages were dropped from the sample, as they were located in Nigeria. Given the stratification, the randomization ensured balance between WTP and non-WTP villages' access to markets, and hence to traders who had sold hermetically sealed bags at baseline.

Within each village, we conducted a random walk and randomly chose 12 households per village, stratified by gender, interviewing either the primary male or primary female within the household. As over 99% of households planted cowpea and hence had the potential to store, there were no inclusion criteria in the sample. This resulted in a sample of 2,639 households.

3.2.3 Data

Household Surveys The data in this paper are comprised of two in-person household surveys and one phone survey. The first in-person survey took place in December 2020. The survey asked questions about agricultural production, storage, knowledge of and experience with the storage technology. Enumerators conducted the WTP experiment at this time.[11]

A second in-person survey was conducted in September 2021, nine months after the baseline survey and WTP experiment. These surveys took place immediately prior to the harvest (and the broader experiment), and asked a limited number of questions about households' agricultural storage and their use of the improved storage technology.

A third in-person survey was conducted in December 2022, with a phone survey in March 2024, approximately two and 3.5 years after the initial intervention, respectively. The final in-person survey asked a number of questions about production, storage and marketing. The phone survey primarily asked about households' production and storage behaviors. The type and timing of each survey, along with the number of observations, is provided in Table C1.

Baseline Balance and Attrition Table C2 presents the baseline characteristics for our sample. Focusing on the control group, the average respondent age was 41 years and 68% of the households owned a mobile phone. Almost all households harvested cowpea in the prior agricultural season, producing 160 kg on average. 72% stored cowpea, storing an average of 94 kg. The primary means of storage were normal bags (27%) and

[11] The baseline survey was originally scheduled for March 2020 but was suspended due to the COVID-19 pandemic. The December 2020 survey thus followed strict health protocols. While schools were closed in Niger between March and June 2020, no other significant lockdowns or border closures occurred that may have affected agricultural production, storage or marketing.

plastic jugs (40%), with only 7% storing cowpea in hermetically-sealed bags in the prior season. Conditional on storage, respondents spent 977 CFA (1.95 USD) per 100 kg of cowpea stored in the past agricultural season, regardless of the technology used. 67% of respondents had heard of the improved storage technology, and 24% had previously used it. Households also had remarkably accurate beliefs about the depreciation rates of the traditional and "new" technologies. The treatment and control groups are similar along observable dimensions: Of the 30 variables tested, only 1 coefficient was statistically significant at the 10% level.

Of the original 2,639 households, 89% were tracked across all survey rounds (Table C3). The only differential attrition occurs during the first follow-up survey, whereby households in WTP villages were 2 percentage points more likely to attrit than those in the control.[12] As this survey round is not the primary focus of this paper, we do not report the corrections for differential attrition for that survey round.

3.2.4 Empirical Strategy

We estimate the impact of being assigned to the WTP experiment using the following specification:

$$Y_{iv} = \alpha_1 + \beta_1 WTP_v + \delta_1 Y_{i0} + \theta_s + \epsilon_{iv} \tag{3.1}$$

where Y_{iv} is the outcome of interest (such as adoption, usage and storage behavior) for individual i in village v. WTP is village-level assignment to the WTP experiment, θ_s is stratification (region and market size) fixed effects and Y_{i0} is the baseline value of the outcome variable.[13] We cluster our standard errors at the village level. For surveys that took place after the broader experiment (the second and third survey rounds), we control for the additional treatment in a "short" model, and report the results from the fully interacted "long" model in the Appendix.[14]

[12] The relatively lower response rate in the 2022 survey was due to a data collection error, whereby data were not collected from two villages.

[13] While some outcomes have low autocorrelation (e.g., storage losses and duration), others have high autocorrelation (e.g., storage practices). Thus, we use an ANCOVA specification for the results in this paper, but also conduct robustness checks using a first-differenced specification. We also control for gender.

[14] Muralidharan, Romero, and Wüthrich (2023) show that t-tests using fully saturated models provide valid inferences, whereas t-tests using "short" models can yield higher power if the interaction terms are zero. Although none of our

To disentangle whether the impacts are primarily due to information about or experience with the technology, we use the variation from the randomly-drawn price during the WTP game as an instrument for the likelihood of winning the technology, using the following specification:

$$Y_{iv} = \alpha_2 + \beta_2 \widehat{Won\ Game}_{iv} + \delta_2 Y_{i0} + \theta_s + \epsilon_{iv} \tag{2}$$

$$Won\ Game_{iv} = a + bDrawn\ Price_{iv} + \delta_3 Y_{i0} + \theta_s + \nu_{iv} \tag{3}$$

where *Won Game* is a binary variable equal to 1 if individual i won the game, 0 otherwise; and *Drawn Price* is the randomly drawn price. The coefficient on β_2 is thus the treatment effect for the compliers, those who obtained information on the technology and had access to it via the game.[15]

3.3 Results

3.3.1 Willingness to Pay for the Technology

Figure 3.1 shows the inverse demand curves derived from the WTP experiment. Overall, mean WTP for the entire sample is 563 CFA (USD 1), about 56% of the average retail sales price, and 42.5% of those who played the game won the technology.[16] Demand is relatively inelastic at lower prices (Figure C1): take-up is universal when the technology is free and drops to 92% when the price goes to 250 CFA (USD .25). It then falls substantially after this point, dropping to 72% when the price goes to 400 CFA (USD .80), and 38% when the price crosses the 600 CFA (USD 1.20) threshold. There are no significant differences in WTP by either region (Panel B) or gender (Panel C), despite the fact that cowpea production is higher in one region, and women are traditionally

interaction terms are statistically significant, the medium- and long-term results presented in this paper should be interpreted as composite treatment effects.

[15] As we are controlling for the baseline outcome variable in each specification, the estimation of our first stage changes slightly for each regression. As a result, the F-statistic from our first stage is not constant across all regressions. As a robustness check, we also estimate these regressions controlling for baseline WTP, and find similar effects.

[16] Only 10% of markets located near villages in our sample sold the technology at baseline.

more credit-constrained in this context.[17]

Nine months after the initial experiment, 92% of those who played and won the game still owned the technology, with 67% reporting storing in it during the prior agricultural season (Table C5, Panel A).[18] Households did not use the technology due to the timing of the game (e.g., households who played soon after the harvest were more likely to store) or the quantity produced. More than three years later, ownership amongst those who played the game fell to 63% (Panel B), primarily due to deterioration of the bag, yet 81% of those who had the bag still used it. As the average duration of the technology is three years, this suggests that households continued to use it beyond its "shelf-life."[19]

3.3.2 Impacts of the WTP Experiment Over Time

Table 3.1 shows the impact of the WTP experiment on households' storage choices, storage and health outcomes. In the first year after the intervention, the WTP experiment reduced the likelihood of purchasing an additional hermetically-sealed bag, but did not affect storage losses or the duration of storage, perhaps because only 2/3 of households had used it during the prior season (Panel A). Yet two years after the experiment, there was a significant impact upon households' storage behavior (Panel B): Households in WTP villages were 13% points more likely to store in the improved technology and 14% points less likely to use traditional storage technologies (Columns 1 and 2). As a result, they were 10% points less likely to use pesticides during storage and less likely to suffer storage losses, although the latter is not statistically significant at conventional levels. These effects are large in magnitude, ranging from 25-34% of the control mean. Yet there were no impacts on storage duration, health outcomes associated with pesticide consumption or demand for an additional product.[20]

Many of these results persisted 3.5 years after the experiment (Panel C): Households

[17] Appendix Table C4 shows the correlates of respondents' WTP. The only statistically significant correlates of WTP are mobile phone ownership (as a proxy for wealth) and the amount of cowpea produced in the year of the survey.

[18] Among the 8% who no longer owned the technology, the primary reason was that it had been destroyed.

[19] While usage is slightly correlated with WTP in the short-term, this does not persist for owning the technology in the longer-term (Figure C2). Visual inspection of the bags in a subset of villages suggested that they were still in "good" condition, namely, that there were no holes or tears.

[20] There were no heterogeneous effects on storage losses, duration or additional purchases by key characteristics, such as gender, mobile phone ownership or agricultural production (TableC6).

in WTP villages continued to use the new technology and shifted away from traditional technologies and pesticide use. The magnitudes of these effects are smaller, in part due to a significant increase in adoption in the control group over time, from 7% at baseline to over 46% 3.5 years later.[21] Unlike other studies on technology adoption (Omotilewa, Ricker-Gilbert, and Ainembabazi (2019)), the experiment did not crowd in new purchases of the technology after the third year (Figure C4). [22]

3.3.3 Information, Experience, Salience or "Gaming"?

What explains the persistent effects of this game? As outlined above, the experiment could have encouraged sustained adoption via a number of pathways. We attempt to address each of these in turn.[23]

Information. While the WTP provided farmers with information about the technology, information does not seem to be a primary barrier to adoption. At baseline, over 67% of households had heard about the technology, and households had fairly accurate beliefs about its effectiveness. This was confirmed at follow-up: Living in a WTP village did not improve participants' knowledge about the technical aspects of the technology as compared to those in control villages (Table C8, Panel A).[24]

Experience. While all of those who played the game received information about it, only a subset won and hence were able to experience it. We measure this effect by estimating equation (2), the results of which are reported in Table 3.2. Overall, the results are consistent with those in Table 3.1, but stronger in magnitude. In the short-term (Panel A), households who played and won the game were significantly less likely to purchase an additional technology and were less likely to suffer from storage losses. Two years later, households continued to use the new technology (Panel B), thus

[21] We do not find spillovers across villages that could potentially explain the increase in adoption in control villages. While there was an overall increase in the availability of the bags on the market, this was not correlated with WTP treatment status.

[22] The results in Table 3.1 (Panels B and C) are all conditional on the other treatment and hence are composite effects. We therefore report the results of the fully interacted model in Table C7. With one exception, none of the interaction terms are statistically significant, and the coefficients on the WTP variable are large and statistically significant in the medium-term, and large but imprecise in the longer-term.

[23] Janzen et al. (2021) note three other mechanisms through which a behavioral game could influence behavior, namely, the length of the game, interactions with enumerators (who administer the game and survey questions) and the payout. Our game lasted an average of 30 minutes, and different enumerators were used for each survey round. Our results are also robust to the inclusion of enumerator fixed effects for a given round.

[24] Experiencing the product also did not improve households' knowledge (Panel B).

shifting away from using traditional technologies and pesticides. These results persisted for most outcomes after three years (Panel C). Winners were also significantly less likely to purchase an additional technology by the third year (Panel C, Column 6), in part because they kept the technology.[25] These results are largely robust to estimating the effects on the sample that excludes the other treatments (Table C9). Taken together, this suggests that experience with the technology was a key driver of the persistent effects of the experiment.

Salience. The timing of the experiment after the harvest may have brought storage costs to farmers' attention at an opportune time and made them more *salient*. If farmers were inattentive to future benefits of the technology, then the game could have brought expenditures "top-of-mind" at the time of storage. In theory, this could have increased farmers' WTP and future usage, if credit constraints did not bind (Berkouwer and Dean, 2022).

We assess the importance of salience by estimating the impact of the *timing* of the experiment on adoption and usage. As the baseline survey was conducted over one month, there was natural variation as to whether the WTP game was played immediately after the harvest in a given village – the time when storage decisions were made – or one month later.[26] We thus first assess the correlation between experiment timing and WTP (Table C4), and find that there is not a strong correlation. We then estimate heterogeneous effects of the WTP experiment by the date of baseline (Table C6). Overall, while the timing of the game is correlated with short-term usage - households who played the game right after the harvest were 18% points more likely to store in the bag in the following year - this seems to be mechanical, as households had access to the bag before purchasing other storage technologies. Survey timing was not correlated with WTP, the likelihood of winning or other storage outcomes. These findings suggest that the game did not make storage expenditures more salient.

Behavioral Biases. Given the body of literature on the use of behavioral games to

[25] Survey evidence suggests that those who won the technology had a positive experience with it: 80% of respondents thought the bag was high-quality and effective in preventing storage losses. Within WTP villages, there were no effects on storage duration or pesticide use by the bid price (Figure C3).

[26] In theory, the timing of the game could have affected both salience and liquidity. Since the experiment was conducted immediately after the harvest in all villages, liquidity constraints should have been less of a concern.

spur adoption, the WTP experiment could have affected adoption in two ways. First, the nature of game itself could have generated enthusiasm about the technology, thus encouraging sustained adoption and usage amongst winners. For non-winners, however, the effects are ambiguous. If the game also generated enthusiasm about the product, then it could have spurred adoption amongst non-winners. Yet if participants were loss-averse, the game could have had negative effects on adoption (Armah and Schwab, 2019).

To test for this, we would ideally want a treatment arm that provided information and access to the technology *without* the game (such as demonstrations and subsidized distribution) or a placebo behavioral experiment (Janzen et al., 2021). In the absence of this setup, we test for the impact of the game on behavioral factors in two ways. First, to assess the impact upon "winners", we estimate equation (2) using outcome data on households' experience with the game – i.e., if they remembered the game and their drawn price - to measure whether this differed by winning status. Second, to compare non-winners in WTP villages with non-winners in control villages, we construct a bootstrapped sample of non-winners whose bid distribution matches that of the entire WTP sample, and compare that bootstrapped sample with the pure control. We also conduct a similar exercise by constructing a bid-specific propensity score for the probability of winning, and estimate a regression using this propensity score as a weight.[27]

Less than one year after the experiment, "winning" households had more accurate recollections about the game: They were 5 percentage points more likely to accurately recall whether they had won and 63 percentage points more likely to recall the draw price (not shown). These results were slightly larger for those for whom the consumer surplus was higher (in other words, the bid price was substantially higher than the drawn price). This suggests that the game may have generated enthusiasm amongst those who won.

Appendix Table C10 shows the impacts of the game on non-winners for a subset

[27] In this regression, observations for non-winners in WTP villages are weighted by the inverse of 1 minus the propensity score, whereas observations in non-WTP villages receive a weight of 1. More information on these methodologies are available upon request.

of outcomes. While the effects are mixed, they are primarily negative, especially for purchasing the bag. Nine months and two years after the study, the likelihood of purchasing a PICS bag was lower amongst non-winners, on the order of magnitude of 10-11 percentage points, with statistically significant effects. Thus, while there may have been a positive behavioral effect of the game for winners, this effect may have been *negative* for non-winners, as least in the medium-term.

3.3.4 Impacts of Baseline Demand Elicitation on Broader Treatment Effects

Given that WTP experiments are embedded in larger studies, a key question is whether and how such experiments can affect the treatment effects of complementary interventions. As our WTP elicitation was conducted in the context of a baseline survey of a larger experiment, we are able to test for this directly.[28] We thus use the same specification as that in equation (1), but replacing the "WTP" variable with the treatment assignment variable from the broader experiment, splitting the sample.

Table 3.3 reports the results from this regression. There are economically and statistically significant effects after the second year in WTP areas, whereby the information treatment increased households' likelihood of using and purchasing the new storage technology, as well as reduced storage losses (Panel A). These results did not persist until the third year (Panel C). Yet in the non-WTP areas, there are no statistically significant impacts of the treatment (Panels B and D). Methodologically, our findings suggest that WTP elicitation experiments can affect the parameter estimates of the larger study in which they are embedded.[29] Consequently, it may be preferable in some contexts to elicit demand in a random subsample, thereby testing whether the WTP experiment induces changes in later behavior.

[28] As outlined above, the broader study was designed to address the impact of supply- and demand-side constraints on the adoption of new storage technologies, in particular by providing information to both traders and farmers. The treatment effects from this broader experiment are available in the fully interacted model in Table C7.

[29] We estimate the results on a split sample to demonstrate how results might differ between RCTs conducted with and without baseline WTP elicitation. In practice, there may not be variation in the demand experiment at baseline.

3.4 Conclusion

We revisit the use of WTP experiments in economics, using a different approach than in past studies: Varying access to the game at baseline and documenting the persistent impacts of the experiment over time. Overall, we find that the impacts of such experiments can be large and persistent: Despite a significant increase in technology adoption of the control over time, we find use of the technology 3.5 years after the experiment. These effects were stronger for those who experienced the good, namely, game-winners.

Despite these persistent effects on adoption, there were few effects on other downstream outcomes, such as storage losses, duration or illnesses. This could be due the fact that the magnitude of storage losses was relatively low (as alternative storage technologies are highly effective in reducing losses), or due to the idiosyncratic nature of agricultural production and storage. In addition, the experiment did not crowd in demand for a new technology after three years.

We provide suggestive evidence that experience with the product seems to drive persistent adoption, similar to the distribution of other experience goods (Bensch and Peters (2017). A key question is whether we would see similar results if we simply provided information and access to technology in an interactive manner, but without the game. While we are unable to fully test whether the game itself had an impact, we provide suggestive evidence that it was a memorable experience for winners.

Similarly, we cannot test whether we would observe similar results for different technologies, such as those that need to be purchased on an annual basis, or those that may last for a longer duration. Unsurprisingly, this would be highly context-dependent: Our technology is one that was relatively well-known, and, while available on local markets, was difficult to obtain on local markets, at least at the outset. This differs from goods that are completely new, or would not be readily available.

These findings suggest that researchers should be aware of the potentially large impacts of baseline WTP elicitation on adoption measures and other outcomes, and design experiments in such a way that allow these effects to be captured.

Tables and Figures

Table 3.1: Intent to Treat (ITT) Effects of the WTP Experiment

Panel A: 9 months after WTP game

	(1) Purchased PICS bag last year	(2) HH suffered storage losses	(3) HH stored until hot season
WTP Assignment	-0.15***	-0.02	0.02
	(0.02)	(0.02)	(0.02)
Observations	2,393	2,164	2,164
Control Mean	0.270	0.200	0.720
Control SD	0.440	0.400	0.450

Panel B: 2 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
WTP Assignment	0.13***	-0.14***	-0.09***	-0.02	0.02	-0.01	0.01
	(0.03)	(0.03)	(0.02)	(0.01)	(0.02)	(0.03)	(0.02)
Observations	2,249	2,249	2,249	2,249	2,249	2,249	2,249
Control Mean	0.380	0.530	0.300	0.0800	0.760	0.370	0.0600
Control SD	0.490	0.500	0.460	0.270	0.430	0.480	0.240

Panel C: 3.5 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH store until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
WTP Assignment	0.09**	-0.11***	-0.05**	-0.02*	-0.05	-0.03	-0.01
	(0.04)	(0.03)	(0.02)	(0.01)	(0.05)	(0.04)	(0.01)
Observations	2,354	2,354	2,354	2,354	2,354	2,354	2,354
Control Mean	0.460	0.460	0.250	0.0700	0.560	0.440	0.0300
Control SD	0.500	0.500	0.430	0.250	0.500	0.500	0.170

Notes: All panels show the results of estimating equation (1). We control for gender and stratification fixed effects, as well as the baseline value of the outcome variable. For Panels B and C, we control for the additional treatment implemented after the WTP experiment. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table 3.2: Local Average Treatment Effects (LATE)

Panel A: 9 months after WTP game

	(1) Purchased PICS bag last year	(2) HH suffered storage losses	(3) HH stored until hot season
Won WTP game	-0.11***	-0.17***	0.02
	(0.03)	(0.03)	(0.05)
Observations	1,135	1,010	1,010
Control Mean	0.16	0.20	0.69
Control SD	0.36	0.43	0.46
F-stat on instrument	1684	1423	1413

Panel B: 2 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
Won WTP game	0.29***	-0.27***	-0.15***	-0.03	0.04	-0.04	-0.03
	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)
Observations	1,068	1,068	1,068	1,068	1,068	1,068	1,068
Control Mean	0.37	0.51	0.29	0.09	0.75	0.34	0.10
Control SD	0.48	0.50	0.45	0.28	0.43	0.47	0.31
F-stat on instrument	1438	1419	1425	1417	1423	1422	1421

Panel C: 3.5 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
Won WTP game	0.14***	-0.08*	-0.08**	0.00	0.01	-0.10**	-0.02
	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)	(0.01)
Observations	1,128	1,128	1,128	1,128	1,128	1,128	1,128
Control Mean	0.44	0.39	0.20	0.03	0.48	0.38	0.02
Control SD	0.50	0.49	0.40	0.17	0.50	0.49	0.15
F-stat on instrument	1746	1752	1772	1757	1767	1777	1720

Notes: All panels show the results of estimating equation (2). We control for gender, the individual's maximum WTP at baseline, stratification fixed effects and the baseline value of the outcome variable. For Panels B and C, we control for the other treatment implemented after the WTP experiment. Robust standard errors clustered at the village level in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table 3.3: Treatment Effects of Broader Experiment by WTP Status

Panel A: Two years after WTP game, in WTP Areas

VARIABLES	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag
Any treated group (=1)	0.12**	-0.05	-0.04	-0.05**	0.04	0.09**
	(0.05)	(0.04)	(0.03)	(0.02)	(0.03)	(0.04)
Observations	1,081	1,081	1,081	1,081	1,081	1,081
Control Mean	0.390	0.390	0.390	0.390	0.390	0.390
Control SD	0.490	0.490	0.490	0.490	0.490	0.490

Panel B: Two years after WTP game, in non-WTP Areas

VARIABLES	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag
Any treated group (=1)	0.02	-0.00	0.01	0.02	0.00	0.02
	(0.05)	(0.04)	(0.05)	(0.02)	(0.03)	(0.05)
Observations	1,168	1,168	1,168	1,168	1,168	1,168
Control Mean	0.390	0.390	0.390	0.390	0.390	0.390
Control SD	0.490	0.490	0.490	0.490	0.490	0.490

Panel C: 3.5 years after WTP game, in WTP Areas

VARIABLES	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag
Any treated group (=1)	0.06	-0.08	-0.05	-0.02	0.01	-0.02
	(0.06)	(0.05)	(0.04)	(0.02)	(0.08)	(0.06)
Observations	1,146	1,146	1,146	1,146	1,146	1,146
Control Mean	0.470	0.470	0.470	0.470	0.470	0.470
Control SD	0.500	0.500	0.500	0.500	0.500	0.500

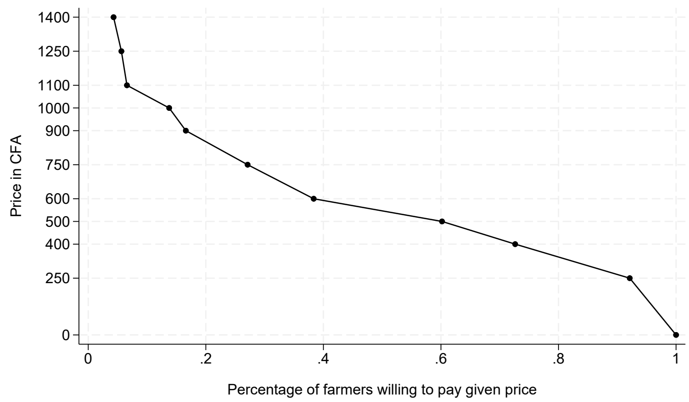
Panel D: 3.5 years after WTP game, in non-WTP Areas

VARIABLES	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag
Any treated group (=1)	-0.06	0.01	-0.04	-0.03	-0.01	-0.06
	(0.07)	(0.05)	(0.06)	(0.02)	(0.08)	(0.07)
Observations	1,208	1,208	1,208	1,208	1,208	1,208
Control Mean	0.470	0.470	0.470	0.470	0.470	0.470
Control SD	0.500	0.500	0.500	0.500	0.500	0.500

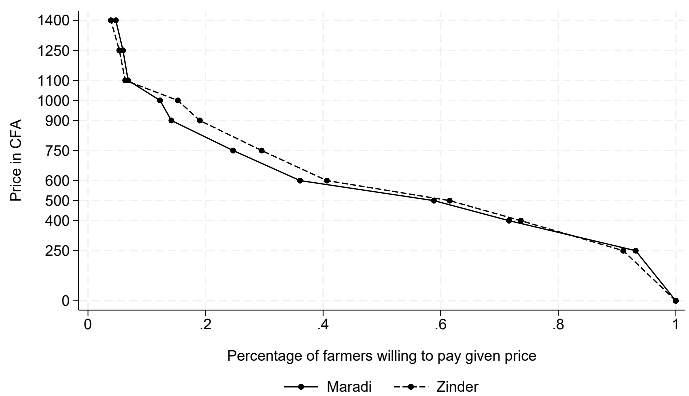
Notes: All panels show the results of regressing the outcome variable on a binary variable for other treatments. We control for gender, stratification fixed effects and the baseline value of the outcome variable. Robust standard errors clustered at the village level in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Figure 3.1: Inverse Demand Curves

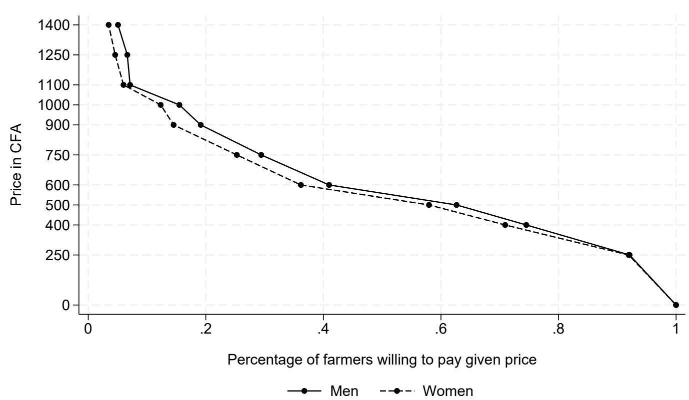
Panel A: Overall Demand



Panel B: By Region



Panel C: By Gender



Notes: Panel A displays the inverse demand curve in the entire sample, where an individual's WTP is reported on the vertical axis and percentage of individuals reporting a given WTP is reported on the horizontal axis. Panel B displays the inverse demand curves by region. Panel C displays the inverse demand curves by gender.

Appendices

Appendix A

Table A1: Summary Statistics

	(1) Observations	(2) Control	(3) Teen Mothers	(4) P-value
Panel A: Demographics				
Age in years	24,987	30.50	29.76	0.00
Height	24,443	154.62	153.72	0.00
Ethnicity:				
Mixed race	24,987	0.81	0.76	0.00
Indigenous	24,987	0.11	0.14	0.00
Rural	24,987	0.33	0.42	0.00
Household has unmet needs (poverty)	24,987	0.23	0.38	0.00
Has cellphone	24,987	0.83	0.73	0.00
Cash Transfer Beneficiary (BDH)	24,987	0.06	0.16	0.00
Region:				
Costa	24,987	0.34	0.38	0.04
Sierra	24,987	0.42	0.34	0.00
Amazon	24,987	0.19	0.26	0.00
Galapagos Island	24,987	0.04	0.03	0.19
Panel B: Family Structure				
Respondent is spouse of HH head	24,987	0.54	0.67	0.00
Respondent is daughter of HH head	24,987	0.25	0.10	0.00
Respondent is head of household	24,987	0.13	0.17	0.00
Number of children	24,987	1.63	2.91	0.00
Respondent's mother lives in the household	24,987	0.27	0.12	0.00
Panel C: Educational Attainment				
Total years of education	24,987	12.79	9.92	0.00
Completed basic education (y=10)	24,987	0.78	0.53	0.00
Completed high school (y=13)	24,987	0.70	0.38	0.00
Panel D: Sexual and Reproductive Health				
Age of Menarche	24,987	12.96	12.64	0.00
Aware of what menstruation was when it happened	24,987	0.73	0.65	0.00
Mother explained about menstruation	24,987	0.72	0.67	0.00
Age of first sexual relation	22,539	18.92	15.66	0.00
Age when first used contraceptive	22,050	21.59	18.97	0.00
Age of first marriage or non-marital relationship	20,019	22.06	17.84	0.00
Age at first birth	21,756	22.64	16.65	0.00

Notes: Column 1 shows the number of observations. Column 2 and 3 show the mean of the control and teen mothers in the sample, respectively. Column 4 shows the p-value of the coefficient from a regression of the dependent variable on an indicator variable for teen mother. This regression includes cohort fixed effects, and robust standard errors. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table A2: Robustness checks for Reduced Form, Main Specification

	(1)	(2)	(3)	(4)
Completed Basic Education				
Menarche at ages 10-13	0.020	0.023*	0.023*	0.029
	(0.010)	(0.010)	(0.010)	(0.015)
Respondent's height		0.011***	0.010***	0.004***
		(0.001)	(0.001)	(0.001)
Poverty (unmet needs)			-0.222***	-0.134***
			(0.011)	(0.023)
Mother completed primary school				0.151***
				(0.022)
Completed High School				
Menarche at ages 10-13	0.014	0.016	0.016	0.038*
	(0.011)	(0.011)	(0.011)	(0.019)
Respondent's height		0.013***	0.011***	0.004**
		(0.001)	(0.001)	(0.002)
Poverty (unmet needs)			-0.264***	-0.195***
			(0.012)	(0.026)
Mother completed primary school				0.174***
				(0.026)
Years of Education				
Menarche at ages 10-13	0.145	0.196	0.196*	0.379*
	(0.102)	(0.101)	(0.099)	(0.182)
Respondent's height		0.137***	0.122***	0.065***
		(0.007)	(0.007)	(0.013)
Poverty (unmet needs)			-2.454***	-2.050***
			(0.097)	(0.214)
Mother completed primary school				1.947***
				(0.237)
Observations	24987	24439	24439	5029

Table A3: Reduced form for different samples

	(1) Completed Basic Education	(2) Completed High School	(3) Completed College	(4) Years of Education	(5) Currently works
Panel A: All women 21-41 (preferred specification)					
Menarche at ages 10-13	0.020	0.014	-0.003	0.145	-0.009
	(0.010)	(0.011)	(0.011)	(0.102)	(0.012)
Panel B: Women 21-41, that completed primary school					
Menarche at ages 10-13	0.019	0.012	-0.005	0.136	-0.005
	(0.010)	(0.011)	(0.008)	(0.072)	(0.009)
Panel C: Reduced form: Mothers ages 21-41					
Menarche at ages 10-13	0.017	0.007	-0.007	0.072	-0.002
	(0.013)	(0.012)	(0.008)	(0.093)	(0.012)
Observations	24987	24987	24987	24987	24987
Control Mean	0.756	0.665	0.210	12.476	0.555
Control SD	0.430	0.472	0.407	4.135	0.497
Kleibergen-Paap F-statistic	92.503	92.503	92.503	92.503	92.503
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000
Observations	22928	22928	22928	22928	22928
Control Mean	0.798	0.703	0.222	12.958	0.556
Control SD	0.401	0.457	0.416	3.673	0.497
Kleibergen-Paap F-statistic	71.134	71.134	71.134	71.134	71.134
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000
Observations	21478	21478	21478	21478	21478
Control Mean	0.756	0.666	0.210	12.487	0.552
Control SD	0.429	0.472	0.408	4.125	0.497
Kleibergen-Paap F-statistic	69.091	69.091	69.091	69.091	69.091
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000

height

Notes: Each specification includes ethnicity and cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table A4: 2SLS for different samples

	(1) Completed Basic Education	(2) Completed High School	(3) Completed College	(4) Years of Education	(5) Currently works
Panel A: 2SLS: All women 21-41 (preferred specification)					
Mother at 18yrs or younger	0.196	0.136	-0.032	1.418	-0.085
	(0.109)	(0.100)	(0.081)	(0.896)	(0.083)
Panel B: 2SLS Women 21-41, that completed primary school					
Mother at 18yrs or younger	0.188	0.112	-0.048	1.314	-0.051
	(0.106)	(0.107)	(0.082)	(0.758)	(0.084)
Panel C: 2SLS Mothers ages 21-41					
Mother at 18yrs or younger	0.131	0.054	-0.050	0.548	-0.015
	(0.097)	(0.091)	(0.058)	(0.715)	(0.093)
Observations	24987	24987	24987	24987	24987
Control Mean	0.756	0.665	0.210	12.476	0.555
Control SD	0.430	0.472	0.407	4.135	0.497
Kleibergen-Paap F-statistic	92.503	92.503	92.503	92.503	92.503
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000
Observations	22928	22928	22928	22928	22928
Control Mean	0.798	0.703	0.222	12.958	0.556
Control SD	0.401	0.457	0.416	3.673	0.497
Kleibergen-Paap F-statistic	71.134	71.134	71.134	71.134	71.134
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000
Observations	21478	21478	21478	21478	21478
Control Mean	0.756	0.666	0.210	12.487	0.552
Control SD	0.429	0.472	0.408	4.125	0.497
Kleibergen-Paap F-statistic	69.091	69.091	69.091	69.091	69.091
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000

Notes: Each specification includes ethnicity and cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table A5: Mediating Factors: IV for Family Structure

	Spouse of HH Head	Daughter of HH Head	Head of Household	Other relationship to HH Head	Mother lives at Home
Mother at, 18yrs, or younger	0.214	-0.068	-0.132	-0.013	0.027
	(0.121)	(0.106)	(0.098)	(0.062)	(0.110)
2SLS Lower 95% C.I.	-0.024	-0.275	-0.324	-0.134	-0.188
2SLS Upper 95% C.I.	0.451	0.138	0.059	0.109	0.242
Kleibergen-Paap F-statistic	91.126	91.126	91.126	91.126	91.126
Kleibergen-Paap p-value	0.00	0.00	0.00	0.00	0.00
Observations	24987	24987	24987	24987	24987
Control Mean	0.64	0.16	0.13	0.06	0.18
Control SD	0.48	0.37	0.34	0.25	0.38

Notes: Dependent variables are the column titles. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table A6: Mediating Factors: ITT for Family Structure

	Spouse of HH Head	Daughter of HH Head	Head of Household	Other relationship to HH Head	Mother lives at Home
Menarche at ages 10-13	0.022	-0.007	-0.014	-0.001	0.003
	(0.012)	(0.011)	(0.010)	(0.006)	(0.011)
2SLS Lower 95% C.I.	-0.002	-0.028	-0.032	-0.014	-0.019
2SLS Upper 95% C.I.	0.046	0.014	0.005	0.011	0.025
Observations	24987	24987	24987	24987	24987
Control Mean	0.58	0.20	0.15	0.07	0.21
Control SD	0.49	0.40	0.36	0.26	0.40

Notes: Dependent variables are the column titles. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table A7: Moderating factor: IV for Mother lives in the household

	Completed Basic Education	Completed High School	Completed College	Years of Education	Currently works
Mother at, 18yrs, or younger	0.136	0.028	-0.014	0.624	-0.065
	(0.107)	(0.109)	(0.089)	(0.966)	(0.109)
Teen mom * Mom at home	0.533	0.974	-0.207	7.066	-0.197
	(0.481)	(0.642)	(0.616)	(5.866)	(0.629)
Mother lives in the household	0.095	0.067	0.155	1.128	0.071
	(0.076)	(0.095)	(0.091)	(0.870)	(0.094)
Observations	24987	24987	24987	24987	24987
Control Mean	0.782	0.700	0.219	12.794	0.557
Control SD	0.413	0.458	0.414	4.089	0.497
Kleibergen-Paap F-statistic	4.171	4.171	4.171	4.171	4.171
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000	0.000

Notes: Dependent variables are the column titles. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Table A8: Moderating factor: ITT for Mother lives in the household

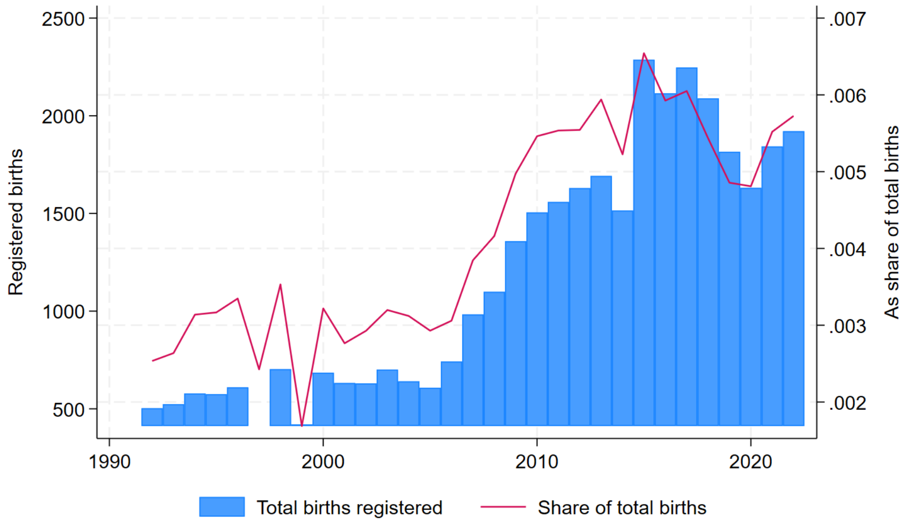
	Completed Basic Education	Completed High School	Completed College	Years of Education	Currently works
Menarche at ages 10-13	0.017	0.004	-0.002	0.085	-0.008
	(0.013)	(0.013)	(0.011)	(0.116)	(0.014)
Menarche 10-13 * Mom at home	0.009	0.035	-0.007	0.219	-0.002
	(0.019)	(0.022)	(0.027)	(0.208)	(0.028)
Mother lives in the household	0.132***	0.168***	0.134***	1.801***	0.060**
	(0.017)	(0.019)	(0.023)	(0.182)	(0.023)
Observations	24987	24987	24987	24987	24987
Control Mean	0.688	0.579	0.159	11.728	0.561
Control SD	0.463	0.494	0.366	4.239	0.496

Notes: Dependent variables are the column titles. Each specification includes ethnicity and 4-year cohort by municipality fixed effects, as well as survey weights. Robust standard errors in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Figure A1: Number of Live Births in Ecuador (1992-2022)

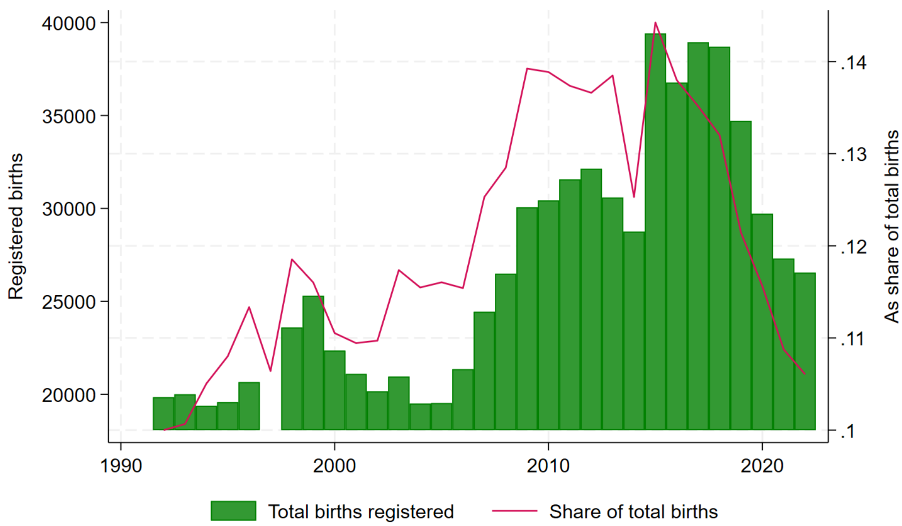
Panel A: Children Born to Mothers Age 10-14 (1992-2022)

5



Panel B: Children Born to Mothers Age 15-18 (1992-2022)

5



Source: Administrative records published by Ecuador's "Instituto Nacional de Estadísticas y Censos (INEC)" .

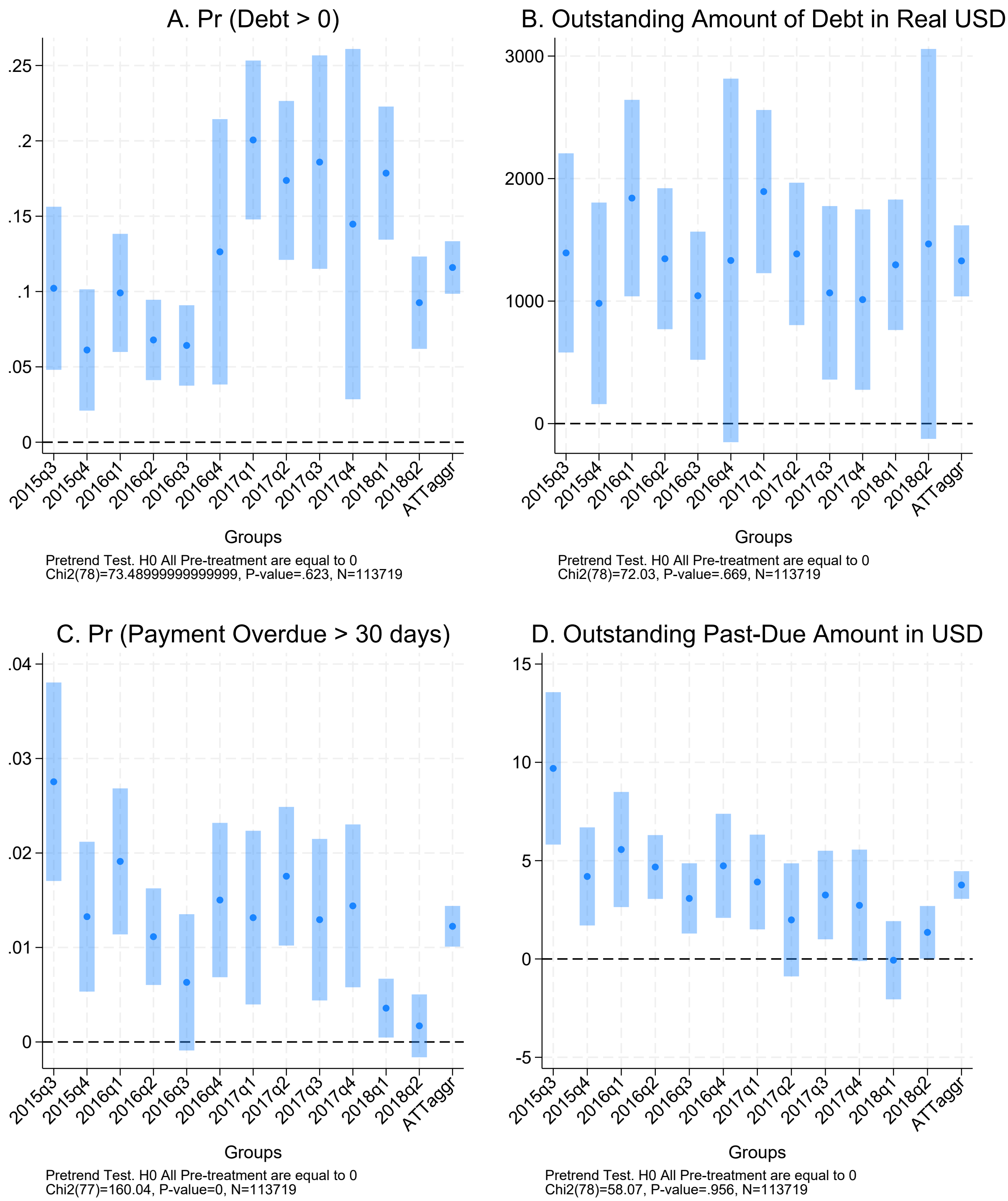
Appendix B

Table B1: Aggregate Average Treatment Estimates, Unconditional

	(1)	(2)	(3)	(4)	(5)
	Pr(Debt> 0)	Outstanding Amount of Debt in Real USD	Share of Banking Productive Loans	Pr(Debt-Past-Due > 30 days)	Outstanding Past-Due Amount of Debt in Real USD
ATT_{simple}	0.051***	1009.241***	0.047***	0.014***	4.556***
	(0.006)	(127.612)	(0.005)	(0.001)	(0.433)
Pre: $ATT_{event}, e < 0$	0.037***	329.295*	0.035***	-0.001	-0.391
	(0.006)	(136.524)	(0.005)	(0.001)	(0.245)
Post: $ATT_{event}, e \geq 0$	0.014	1054.958***	0.020**	0.020***	6.969***
	(0.008)	(190.983)	(0.007)	(0.002)	(0.729)
Observations	122,122	122,122	122,122	122,122	122,122
Control Mean (NT)	0.21	2,331.47	0.17	0.00	0.88
Control SD (NT)	0.41	11,018.33	0.35	0.05	13.66
P-value (Pre-trend Test)	0.007	0.004	0.000	0.986	0.196

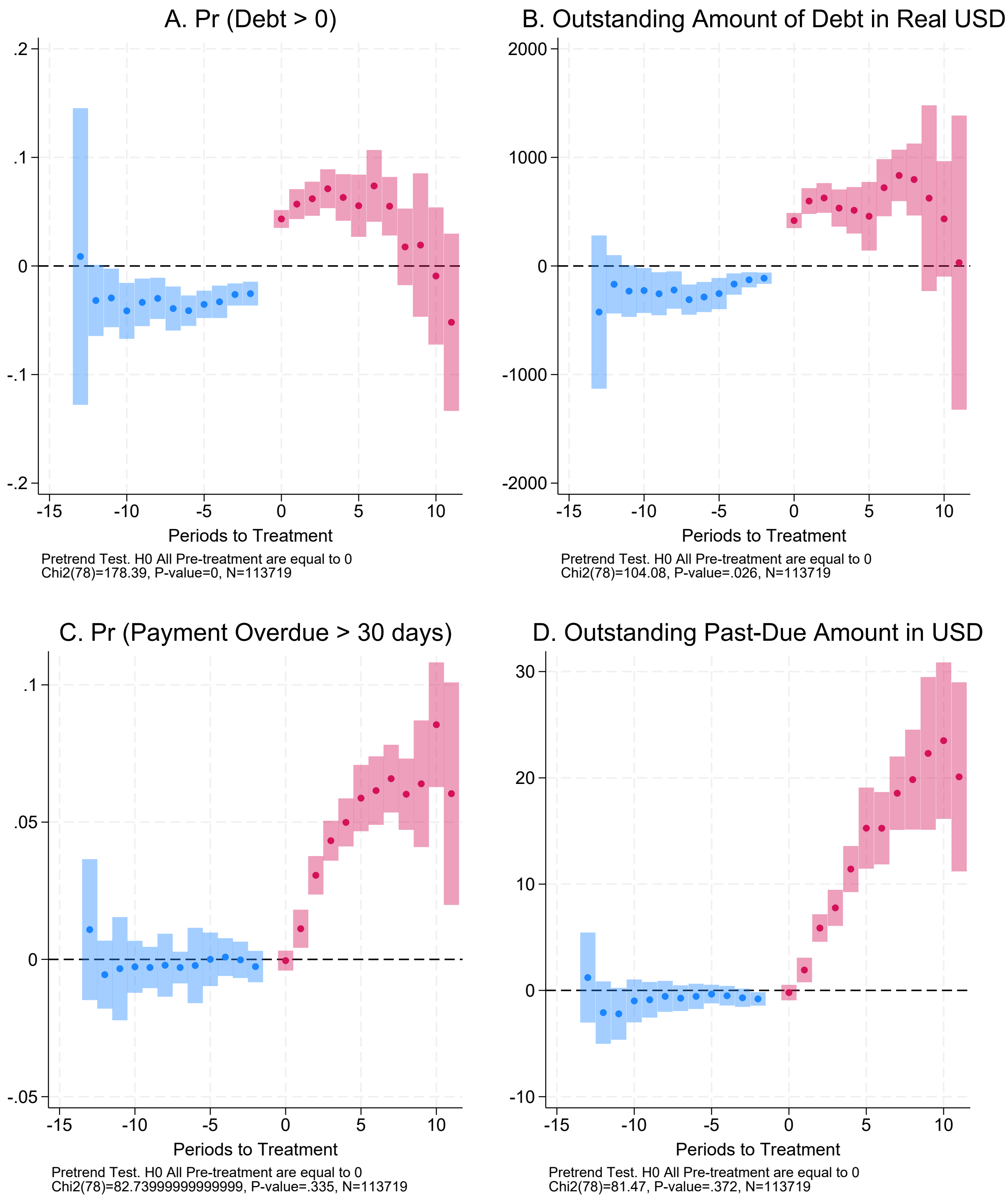
Notes: This table shows the unconditional aggregate ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. We show the 'simple' ATT for all groups across all periods, as well as the dynamic ATT, that is, an average of all negative (pre) and non-negative (post) event times. Column 1-5 show the ATT for different outcomes: the probability of obtaining additional banking-productive loans, the outstanding amount of debt in real USD, the share of banking productive loans out of the total outstanding loans in a given period, the probability of having debt past due in the last 30 days, and the outstanding amount of debt past-due in real USD, respectively. To ease interpretation, we report the mean and standard deviation for the control group that was never treated (the two last cohorts in our data). Standard errors in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Figure B1: Group Aggregate ATTs: Banking Productive Loans



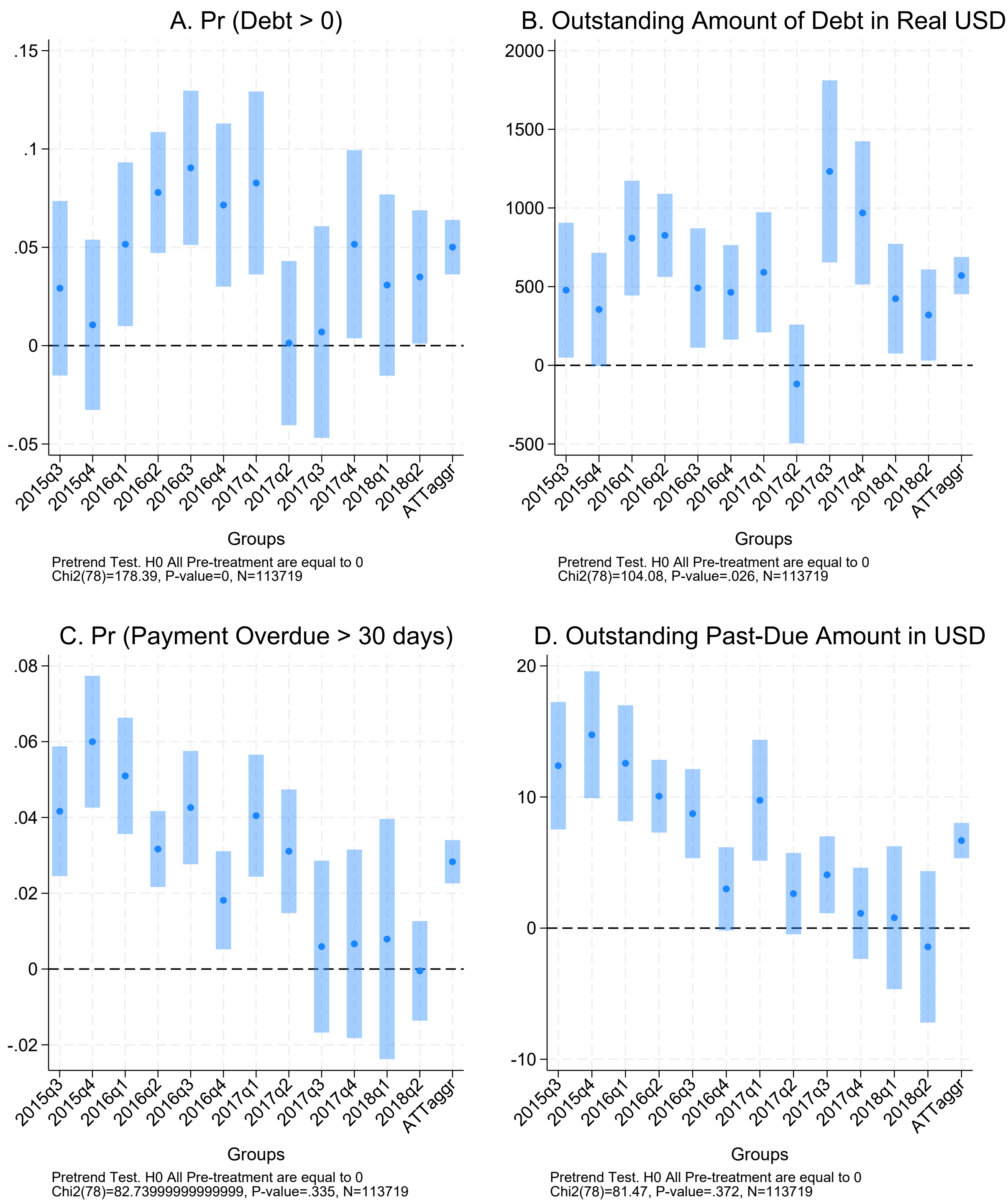
Notes: This figure shows the group ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B2: Dynamic (Event) Aggregate ATTs: Non-Banking Productive Loans



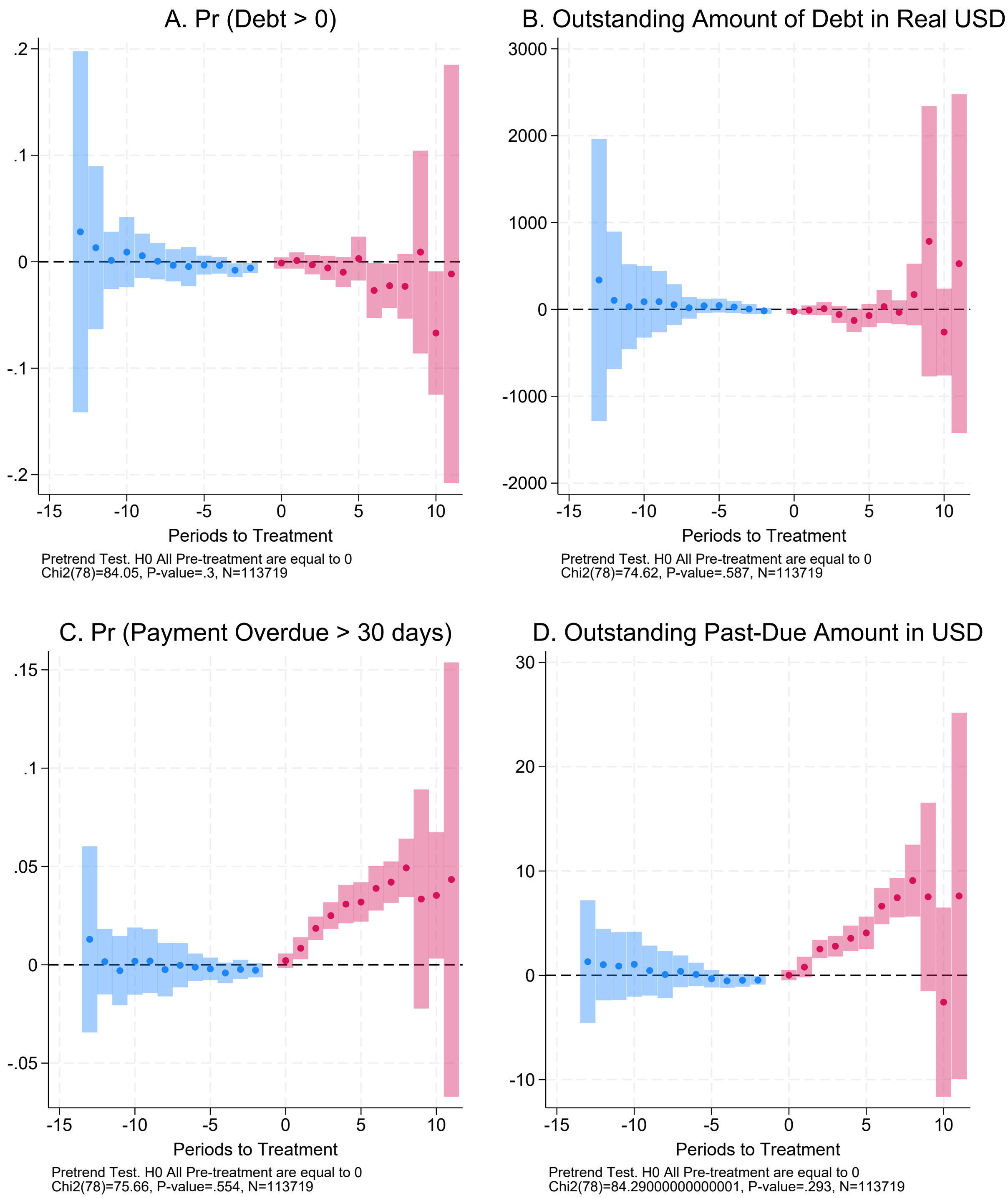
Notes: This figure shows the dynamic-event ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of non-banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B3: Group Aggregate ATTs: Non-Banking Productive Loans



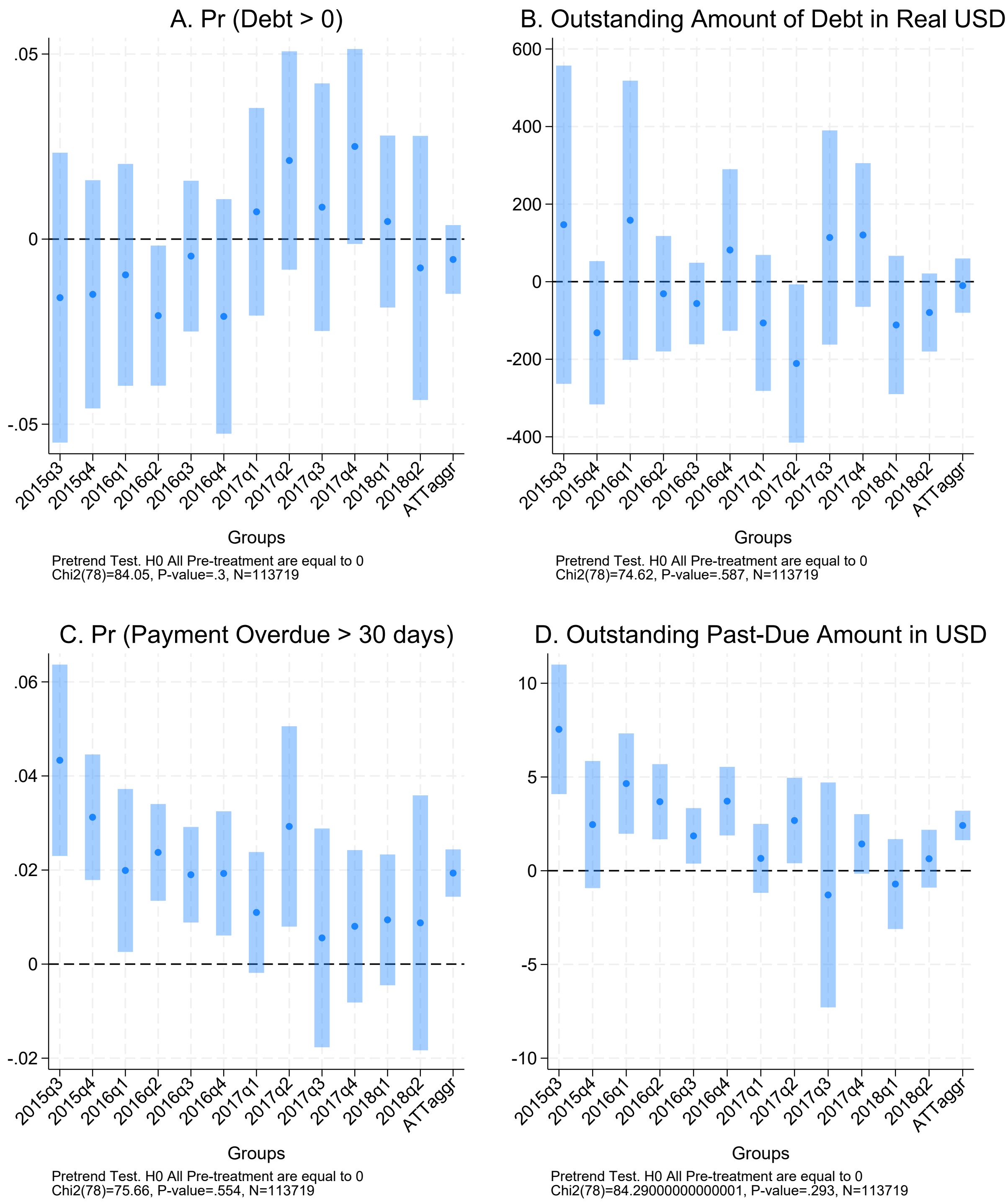
Notes: This figure shows the group ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of non-banking productive loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B4: Dynamic (Event) Aggregate ATTs: Consumption Loans



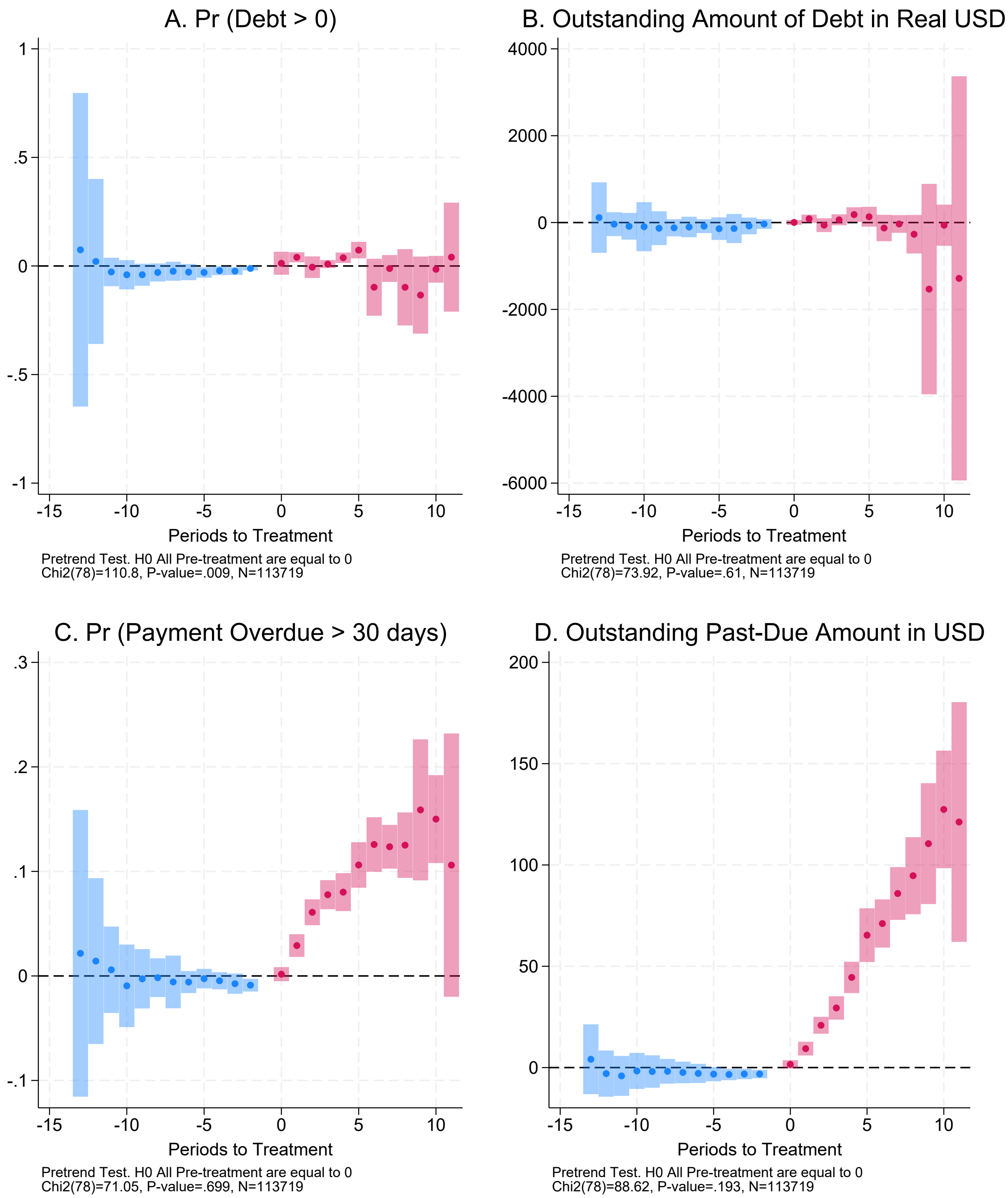
Notes: This figure shows the dynamic-event ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of consumption loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B5: Group Aggregate ATTs: Consumption Loans



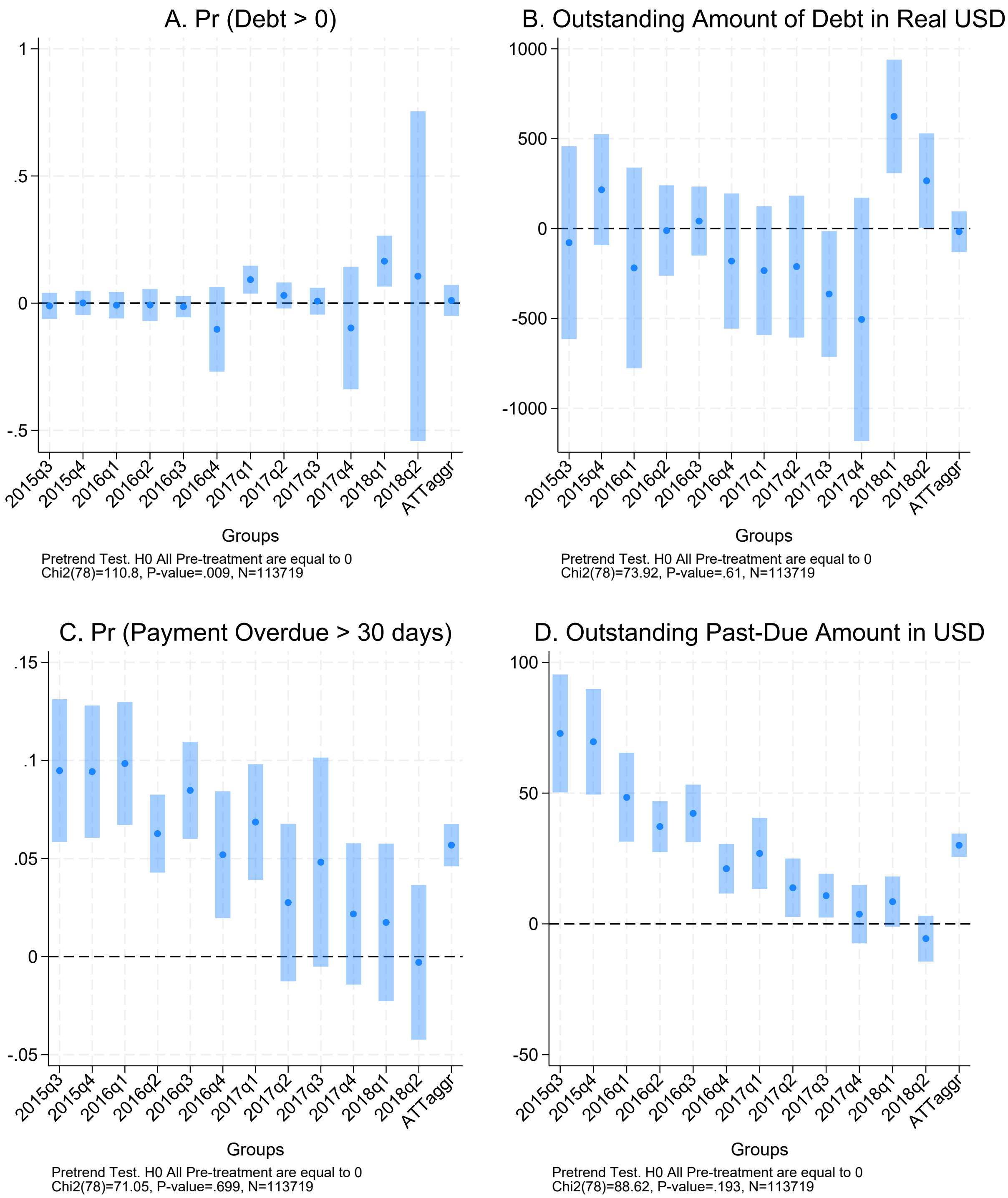
Notes: This figure shows the group ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of consumption loans: (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B6: Dynamic (Event) Aggregate ATTs: Other Loans



Notes: This figure shows the dynamic-event ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of other loans (credit cards, auto, telecommunications): (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Figure B7: Group Aggregate ATTs: Other Loans



Notes: This figure shows the group ATT estimates with staggered treatment timing using the doubly-robust estimation method from Callaway and Sant'Anna (2021) using not-yet-treated units as the comparison group. Each panel is an outcome variable within the category of other loans (credit cards, auto, telecommunications): (A) the probability of obtaining additional loans, (B) the probability of having debt past due in the last 30 days, (C) the share of loans out of the total outstanding loans in a given period, and (D) the outstanding amount of debt in USD, respectively. The point estimate is reported by the circles, and 95% confidence intervals are reported in the vertical lines.

Appendix C

Table C1: Data Collection

Survey Round	Dates	Observations
Baseline Survey	December 2020	2,639
Midline Survey	September 2021	2,326
Endline Survey	December 2022	2,249
Phone Survey	March 2024	2,354

Notes: Each number is the total sample size of households found by survey round.

Table C2: Baseline Balance

	(1) Control	(2) WTP	(3) N
Age	41.24	-0.64	2639
	(15.19)	(0.59)	
Female	0.50	0.00	2639
	(0.50)	(0.00)	
Household owns a cell phone	0.68	0.01	2639
	(0.47)	(0.03)	
Quantity in kg of cowpea harvested in 2020/2021	161.02	-7.08	2639
	(222.11)	(16.23)	
Household sold cowpea during the 2019/2020 harvest	0.64	-0.01	2639
	(0.48)	(0.03)	
Total number of markets where cowpea was sold during last harvest	0.60	0.02	2639
	(0.65)	(0.04)	
Stored cowpea in 2020/2021	0.72	-0.01	2639
	(0.45)	(0.03)	
Stored cowpea in any bag	0.34	0.00	2639
	(0.47)	(0.02)	
Stored in normal bags	0.27	0.00	2639
	(0.44)	(0.02)	
Stored in PICS bags	0.07	0.00	2639
	(0.25)	(0.01)	
Stored in bidon	0.40	0.01	2639
	(0.49)	(0.03)	
Number of PICS bags bought	0.15	-0.02	2639
	(0.86)	(0.04)	
Price per unit of PICS bags (CFA)	1020.25	-47.94	152
	(187.34)	(32.48)	
Total expenses (CFA) on cowpea storage per 100kg (CFA)	944.65	51.08	1874
	(3679.20)	(142.89)	
Respondent has heard about PICS bags	0.67	-0.01	2639
	(0.47)	(0.03)	
Respondent has used PICS bags at some point	0.24	-0.03	2639
	(0.43)	(0.03)	
9 month subj. depreciation rate, trad bags + pesticides	0.33	-0.02	2639
	(0.38)	(0.03)	
9 month subj. depreciation rate, PICS	0.03	0.01*	2639
	(0.11)	(0.01)	

Notes: Column 1 presents the mean of the dependent variable for villages not assigned to the willingness-to-pay (WTP) game (standard deviation in parentheses), Column 2 reports the coefficient from a regression of the dependent variable on an indicator variable for WTP (standard error in parentheses), controlling for strata fixed effects. Robust standard errors clustered at the village level presented in parentheses. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table C3: Attrition

	9 Months after WTP Game	2 Years after WTP Game	3.5 Years after WTP Game
	(1)	(2)	(3)
	September	December	March
WTP Assignment	0.02*	0.02	0.00
	(0.01)	(0.02)	(0.02)
Female(=1)	-0.05***	-0.03**	-0.01
	(0.01)	(0.01)	(0.01)
Zinder(=1)	-0.01	0.04*	-0.03
	(0.01)	(0.02)	(0.02)
Market size	-0.00	-0.01	0.01
	(0.01)	(0.02)	(0.02)
Other treatment assignment		-0.03	-0.03
		(0.03)	(0.02)
Dependent Variable Control Mean	0.08	0.14	0.11
R-Squared	0.01	0.01	0.00
Observations	2639	2639	2639

Notes: Columns 1-3 show the coefficients from regressing a variable for attrition on the WTP assignment indicator for each survey round, as well as stratification (region and market size) and gender controls. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table C4: Correlates of Willingness to Pay

	(1) Maximum WTP
Female	-15.02 (38.02)
Age	-1.44 (0.98)
Owns mobile phone	104.10*** (36.67)
Number of letters respondent can read	-2.18 (6.15)
Self-assessment of math skills	11.89 (27.31)
Stored cowpea in 2020	86.14 (75.44)
Quantity in KG of cowpea harvested in 2020/2021	0.16** (0.08)
Quantity in KG of cowpea harvested in 2019/2020	-0.00 (0.10)
Stored cowpea in any bag	-16.51 (76.46)
Total expenses on cowpea storage	0.02 (0.02)
Respondent has used PICS bags at some point	54.43 (36.20)
9 month subj. depreciation rate, traditional bags	-8.52 (38.59)
9 month subj. depreciation rate, trad bags + pesticides	179.54 (152.64)
9 month subj. depreciation rate, PICS	271.55 (424.35)
WTP experiment took place soon after the harvest	-26.64 (35.73)
Zinder region	-24.63 (41.15)
Mean Maximum WTP	562.44
Observations	1068
R-Squared	0.06

Notes: Maximum WTP values are in CFA. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table C5: Summary Statistics

	Mean	Std. Dev.	N
Panel A: 2021			
Respondent won the technology	0.42	0.49	1,122
Respondent still owns it	0.92	0.27	468
Technology destroyed	0.58	0.5	38
Respondent stored in it	0.67	0.47	430
Panel B: 2024			
Respondent still owns it	0.63	0.48	510
Technology destroyed	0.88	0.32	168
Respondent stored in it	0.81	0.39	320

Notes: This table displays summary statistics collected during the surveys nine months and 3.5 years after the WTP experiment.

Table C6: Heterogeneous Effects

Outcome is:	Household suffered losses					Household stored until hot season				
Het var is:	Female	Cell phone	Store in traditional technologies	Produce ¿ 100kg	Early WTP	Female	Cell phone	Store in traditional technologies	Produce ¿ 100kg	Early WTP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 9 months after WTP game										
WTP treatment	-0.06**	-0.02	-0.02	-0.02	-0.02	0.06*	-0.00	-0.02	0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
Het variable	-0.05**	0.00	0.01	-0.03	0.01	0.04	0.03	0.07**	0.09***	0.05
	(0.02)	(0.02)	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)
WTP × Het var	0.06*	-0.02	0.00	-0.01	-0.02	-0.07*	0.05	0.07	0.03	0.03
	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Panel B: 2 years after WTP game										
WTP treatment	-0.01	-0.02	-0.01	-0.03	-0.02	0.01	0.01	0.06*	-0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Het variable	0.01	0.01	0.04**	-0.01	-0.09***	-0.07***	0.03	0.10***	0.08***	-0.08
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.05)
WTP × Het var	-0.02	0.00	-0.02	0.01	-0.01	0.02	0.01	-0.06*	0.04	0.04
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Panel C: 3.5 years after WTP game										
WTP treatment	-0.02*	-0.01	-0.02	-0.00	-0.01	-0.06	-0.09	-0.04	-0.10*	-0.09
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.05)	(0.07)	(0.06)	(0.05)	(0.06)
Het variable	-0.01	0.01	-0.01	0.01	0.04*	-0.05**	0.03	0.05	-0.01	-0.23***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.08)
WTP × Het var	0.01	-0.01	0.01	-0.02	-0.02	0.01	0.06	-0.02	0.10*	0.09
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.06)	(0.05)	(0.06)	(0.11)

Notes: All panels show the results of regressing the outcome of interest on their village's WTP assignment interacted with the variable specified in the column title. We add stratification fixed effects. For Panels B and C, we control for the other treatment implemented after the WTP experiment. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table C7: Fully Interacted Model

Panel A: 2 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH store until hot season	(6) Purchased PICS bag
WTP Assignment	0.10*	-0.16**	-0.13***	0.04	0.03	-0.08
	(0.06)	(0.07)	(0.05)	(0.04)	(0.06)	(0.06)
Any treated group (=1)	0.03	-0.03	-0.02	0.01	0.02	0.01
	(0.04)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)
WTP*Treatment Status	0.02	0.01	0.01	-0.02*	-0.00	0.03
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Observations	2,249	2,249	2,249	2,249	2,249	2,249
Control Mean	0.379	-0.0437	0.0671	0.00875	0.0933	0.321
Control SD	0.486	0.667	0.528	0.386	0.554	0.492

Panel B: 3.5 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH store until hot season	(6) Purchased PICS bag
WTP Assignment	0.12	-0.13	-0.12**	-0.00	0.04	-0.05
	(0.09)	(0.09)	(0.06)	(0.03)	(0.12)	(0.09)
Any treated group (=1)	-0.00	-0.04	-0.07**	-0.01	0.03	-0.05
	(0.06)	(0.06)	(0.04)	(0.02)	(0.08)	(0.06)
WTP*Treatment Status	-0.01	0.01	0.03	-0.01	-0.04	0.00
	(0.03)	(0.03)	(0.02)	(0.01)	(0.04)	(0.03)
Observations	2,354	2,354	2,354	2,354	2,354	2,354
Control Mean	0.461	-0.126	0.0229	-0.0115	-0.106	0.395
Control SD	0.499	0.720	0.530	0.379	0.705	0.545

Notes: All panels show the results of regressing the outcome of interest on their WTP assignment, the other treatment and the interaction between the two. We control for gender, stratification fixed effects and the baseline values of the outcome variable. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Table C8: Effects on Learning, 2 years after WTP game

Panel A: Intent to Treat (ITT)

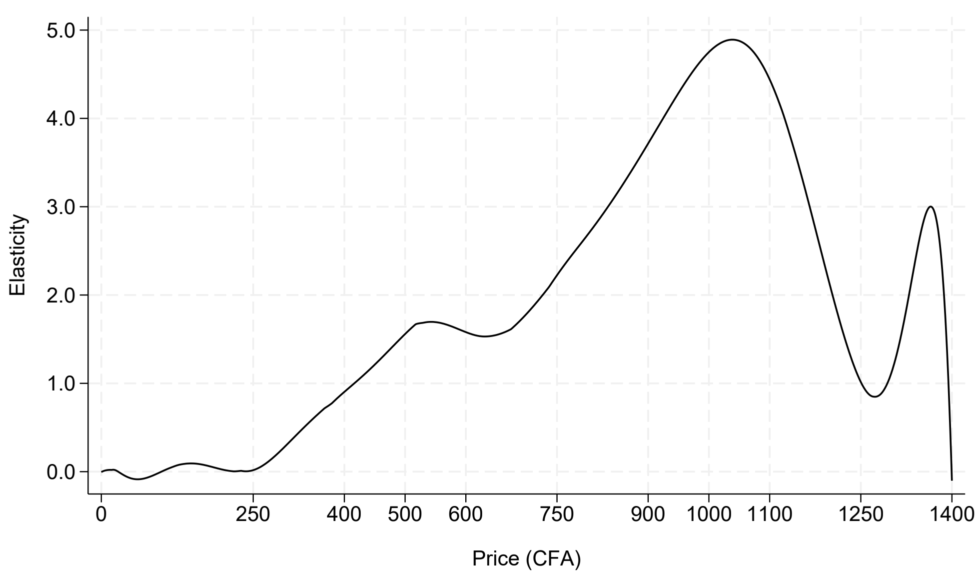
	(1) Knowledge score about PICS bag (out of 4)
WTP Assignment	0.01
	(0.04)
Observations	2,249
Control Mean	3.309
Control SD	0.756

Panel B: Local Average Treatment Effect (LATE)

	(1) Knowledge score about PICS bag (out of 4)
Won WTP game	0.02
	(0.07)
Observations	1,068
Control Mean	3.230
Control SD	0.830
F-stat on instrument	1420

Notes: Panel A shows the results of estimating equation (1), while Panel B shows the results of estimating equation 92). We control for gender and stratification fixed effects. As we do not have a baseline measure of knowledge, we also control for the baseline measure of beliefs about the technology. We cluster our standard errors at the village level. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

Figure C1: Elasticity of Demand



Notes: Demand elasticities are calculated by local polynomial regression using an Epanechnikov kernel, following Berry, Fischer, and Guiteras (2020).

Table C9: LATE Effects in Non-Treatment Villages

Panel A: 9 months after WTP game

	(1) Purchased PICS bag last year	(2) HH suffered storage losses	(3) HH stored until hot season
Won WTP game	-0.12**	-0.14**	0.04
	(0.05)	(0.06)	(0.08)
Observations	329	329	329
Control Mean	0.160	0.230	0.690
Control SD	0.370	0.420	0.460
F-stat on instrument	381.1	384.8	393

Panel B: 2 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
Won WTP game	0.39***	-0.28***	-0.16**	-0.01	0.02	0.04	-0.02
	(0.10)	(0.09)	(0.06)	(0.05)	(0.08)	(0.09)	(0.05)
Observations	291	291	291	291	291	291	291
Control Mean	0.270	0.540	0.320	0.130	0.720	0.250	0.0800
Control SD	0.440	0.500	0.470	0.330	0.450	0.440	0.270
F-stat on instrument	263.3	251	265.3	256.5	260.5	262.7	266

Panel C: 3.5 years after WTP game

	(1) Stored in PICS bags	(2) Stored in traditional technologies	(3) Used pesticides for cowpea storage	(4) HH suffered storage losses	(5) HH stored until hot season	(6) Purchased PICS bag	(7) Household experienced health symptoms
Won WTP game	0.22***	-0.24***	-0.05	-0.02	-0.03	0.01	-0.03
	(0.08)	(0.08)	(0.06)	(0.03)	(0.09)	(0.07)	(0.03)
Observations	323	323	323	323	323	323	323
Control Mean	0.390	0.460	0.230	0.0500	0.500	0.350	0.0100
Control SD	0.490	0.500	0.420	0.220	0.500	0.480	0.120
F-stat on instrument	320.3	313.8	339.2	336.4	329.4	321.3	334.3

Notes: All panels show the results of estimating equation (2). We control for gender, the individual's maximum WTP at baseline, stratification fixed effects and the baseline value of the outcome variable. For Panels B and C, we control for the other treatment implemented after the WTP experiment. Robust standard errors clustered at the village level in parenthesis. ***, **, * denote statistical significance at the 1, 5, 10 percent levels, respectively.

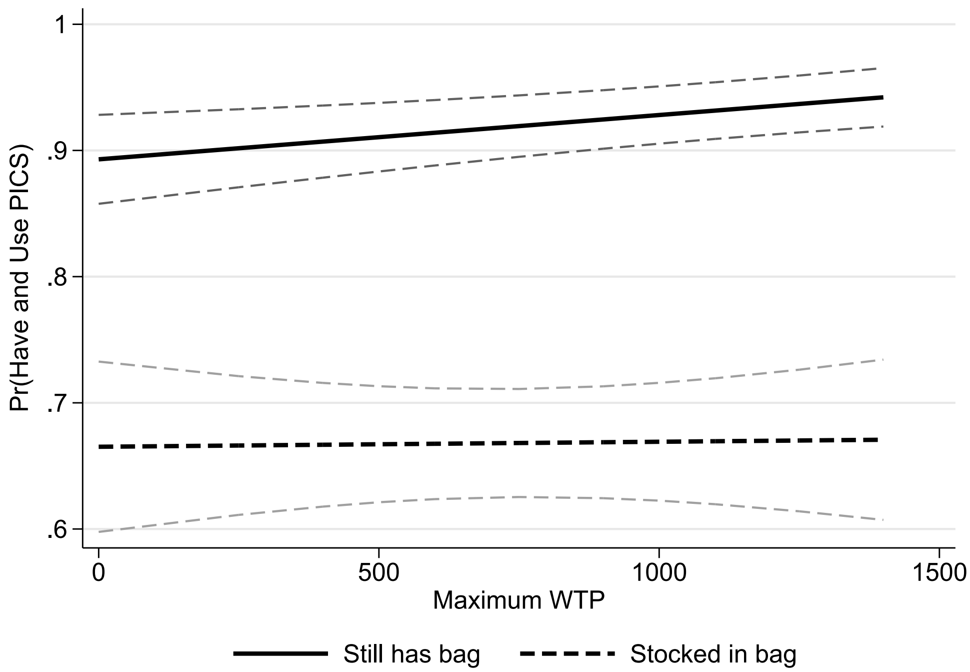
Table C10: Effects of Playing the Game

	(1) Purchased PICS bag last year	(2) HH suffered storage losses	(3) HH stored until hot season
Panel A: 9 Months after the Experiment			
Bootstrapping	-0.11***	-0.01*	-0.03
Propensity Score Matching	-0.10***	0.05	-0.02
Panel B: Two Years after the Experiment			
Bootstrapping	-0.03	0.00	-0.02
Propensity Score Matching	-0.11*	0.02	-0.04
Panel B: 3.5 Years after the Experiment			
Bootstrapping	-0.02	-0.01	-0.06
Propensity Score Matching	-0.07	-0.06	-0.02

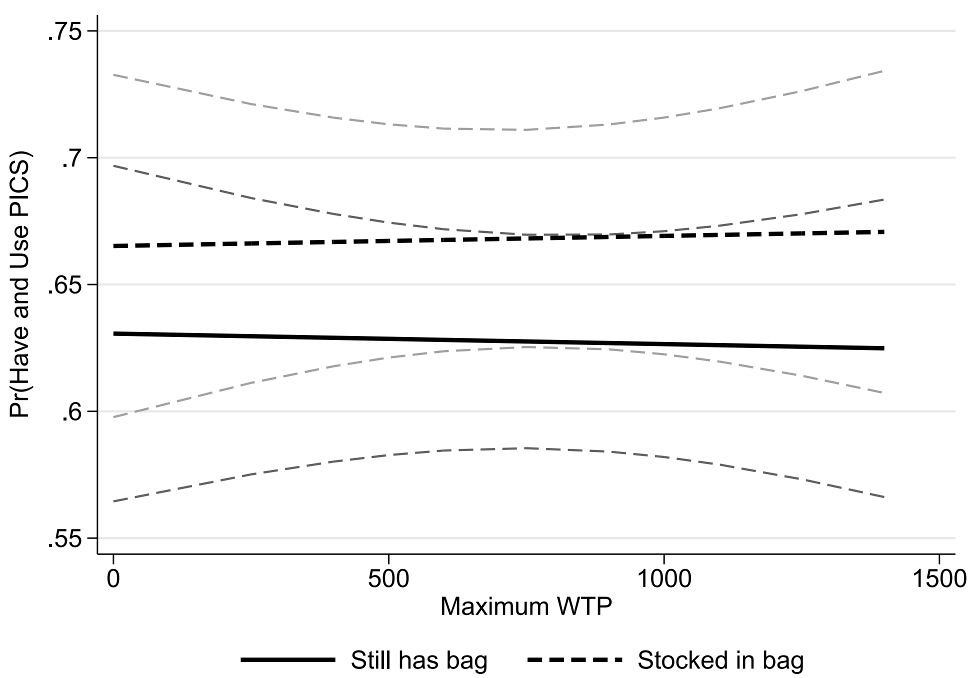
Notes: This shows the results of a regression of inverse probability using the propensity score on non-winners in WTP villages and non-winners in non-WTP villages. Robust standard errors clustered at the village level in parenthesis. *, **, and *** denote statistical significance at the 0.1, 0.05, and 0.001 level, respectively.

Figure C2: Ownership and Usage by Max. Willingness to Pay Price

Panel A: 9 months after WTP game



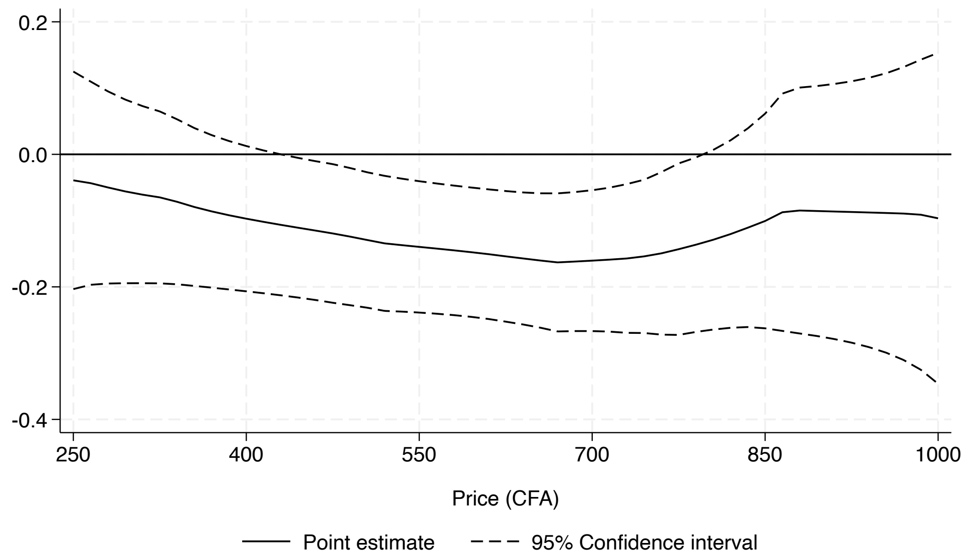
Panel B: 3.5 years after WTP game



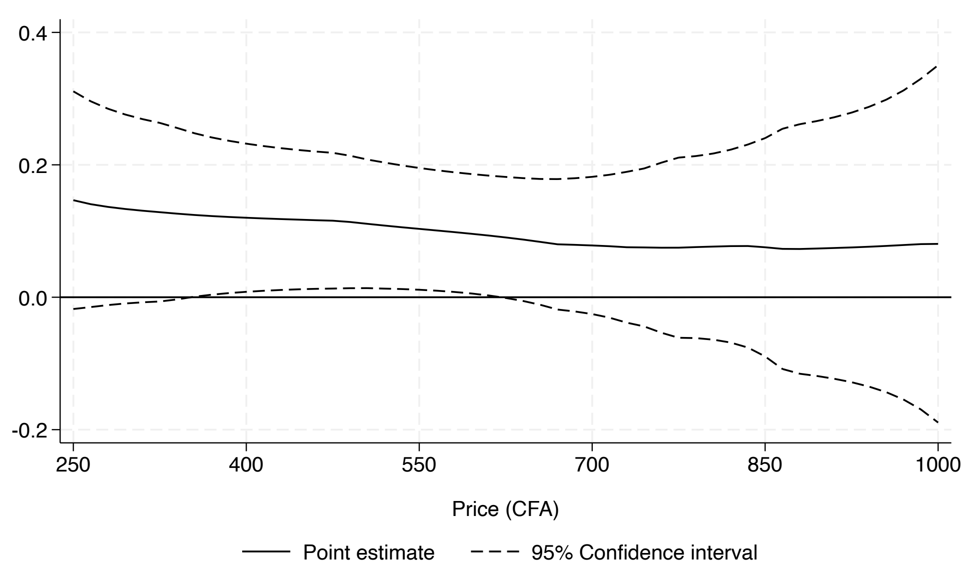
Notes: Panels A and B display how ownership (solid black line) and usage (dashed black line) varies by revealed maximum willingness to pay price. Regressions include stratification (region and market size) fixed effects.

Figure C3: Kernel IV Estimates of Treatment Effects in the Long-Term

Panel A: Storing in PICS Bag

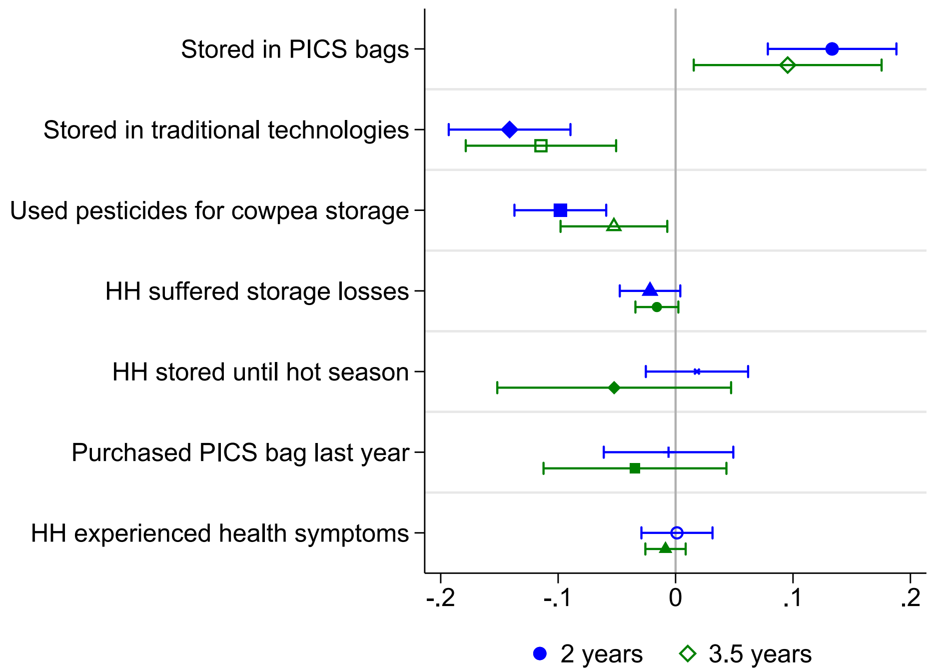


Panel B: Using Pesticides for Storage



Notes: Panel A displays the estimated reduction in reported household diarrhea as a function of willingness-to-pay (WTP). Panel B displays the estimated season in which cowpea was sold as a function of WTP. In both graphs, these are two-stage least squares estimates at a WTP of 250 CFA to 1000 CFA in increments of 15 CFA. Observations are weighted by their distance from the evaluation WTP using Silverman's rule of thumb for the bandwidth and an Epanechnikov kernel. Standard errors are clustered at the village level.

Figure C4: Effects over Time



Notes: Panel A displays the coefficients on the "WTP" variable for all regressions, controlling for strata fixed effects and clustering the s.e. at the village level. The second panel shows the coefficients from the same regressions, but only for the last two survey rounds.

Bibliography

Agarwal, Sumit. 2018. *Financial Inclusion under the Microscope.* Washington, D. C: International Monetary Fund.

Aker, Jenny C, Brian Dillon, and C Jamilah Welch. 2023. "Demand, supply and long-term adoption: Evidence from a storage technology in West Africa." *Journal of Development Economics* :103129.

Aker, Jenny C., Silvia Prina, and C. Jamilah Welch. 2020. "Migration, Money Transfers, and Mobile Money: Evidence from Niger." *AEA Papers and Proceedings* 110:589–93.

Andor, Mark, Lorenz Goette, Michale K. Price, Anna Schulze Tilling, and Lukas Tomberg. 2023. "Differences in How and Why Social Comparison and Real-Time Feedback Impact Resource Use: Evidence from a Field Experiment." *NBER Working Paper 31845* .

Arceo-Gomez, Eva O. and Raymundo M. Campos-Vazquez. 2014. "Teenage Pregnancy in Mexico: Evolution and Consequences." *Latin American Journal of Economics* 51 (1):109–146. URL http://www.laje-ce.org/previous-issues-results-en?docid=5928.

Ardington, Cally, Alicia Menendez, and Tinofa Mutevedzi. 2015. "Early Childbearing, Human Capital Attainment, and Mortality Risk: Evidence from a Longitudinal Demographic Surveillance Area in Rural KwaZulu-Natal, South Africa." *Economic Development and Cultural Change* 63 (2):281–317. URL https://www.jstor.org/stable/10.1086/678983.

Armah, Ralph and Benjamen Schwab. 2019. "Does Loss Aversion affect Improved Storage Technology Adoption? Evidence from a Field Experiment in Ghana." *Unpublished Working Paper* .

Ashcraft, Adam and Kevin Lang. 2006. "The Consequences of Teenage Childbearing." URL https://www.nber.org/papers/w12485.

Ashraf, Nava, James Berry, and Jesse M Shapiro. 2010. "Can higher prices stimulate product use? Evidence from a field experiment in Zambia." *American Economic Review* 100 (5):2383–2413.

Azevedo, Joao Pedro, Luis F. Lopez-Calva, and Elizaveta Perova. 2012. *Is the Baby to Blame? An Inquiry into the Consequences of Early Childbearing*. Policy Research Working Papers. The World Bank. URL http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-6074.

Azevedo, Viviane, Jeanne Lafortune, Liliana Olarte, and José Tessada. 2019. "Does Formal Credit Lead to More Financial Inclusion or Distress? Results Using a Strict Scoring Rule among Marginal Clients in Paraguay." URL https://idbinvest.org/en/publications/report-does-formal-credit-lead-more-financial-inclusion-or-distress.

Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2025. "Difference-in-Differences Designs: A Practitioner's Guide." URL http://arxiv.org/abs/2503.13323. ArXiv:2503.13323.

Balew, Solomon, Erwin Bulte, and Menale Kassie. 2024. "Short-run subsidies and long-run willingness to pay: Learning and anchoring in an agricultural experiment in Ethiopia." *American Journal of Agricultural Economics* :ajae.12498.

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan. 2019. "Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?" Tech. Rep. w26346, National Bureau of Economic Research, Cambridge, MA. URL http://www.nber.org/papers/w26346.pdf.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics* 7 (1):22–53. URL https://pubs.aeaweb.org/doi/10.1257/app.20130533.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. 2015. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Jour-*

nal: Applied Economics 7 (1):1–21. URL https://pubs.aeaweb.org/doi/10.1257/app.20140287.

Barham, Tania, Karen Macours, and John A Maluccio. 2024. "Experimental Evidence from a Conditional Cash Transfer Program: Schooling, Learning, Fertility, and Labor Market Outcomes after 10 Years." *Journal of the European Economic Association* :jvae005URL https://academic.oup.com/jeea/advance-article/doi/10.1093/jeea/jvae005/7585333.

Becker, G. M., M. H. DeGroot, and J. Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9:226–232.

Ben Yishay, Ariel, Andrew Fraker, Raymond Guiteras, Giordano Palloni, Neil Buddy Shah, Stuart Shirrell, and Paul Wang. 2017. "Microcredit and willingness to pay for environmental quality: Evidence from a randomized-controlled trial of finance for sanitation in rural Cambodia." *Journal of Environmental Economics and Management* 86:121–140.

Bensch, Gunther and Jörg Peters. 2017. "One-off subsidies and long-run adoption – Experimental evidence on improved cooking stoves in Senegal." (256217).

Berkouwer, Susanna B. and Joshua T. Dean. 2022. "Credit, Attention, and Externalities in the Adoption of Energy Efficient Technologies by Low-Income Households." *American Economic Review* 112 (10):3291–3330.

Berry, James, Greg Fischer, and Raymond Guiteras. 2020. "Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana." *Journal of Political Economy* 128 (4):1436–1473.

Berry, James and Priya Mukherjee. 2019. "Pricing Private Education in Urban India: Demand, Use and Impact." URL https://www.povertyactionlab.org/sites/default/files/research-paper/Pricing-Private-Education_Feb2019.pdf. Working Paper.

Berthelon, Matias and Diana I. Kruger. 2017. "Does adolescent motherhood affect education and labor market outcomes of mothers? A study on young adult women in Chile during 1990–2013." *International Journal of Public Health* 62 (2):293–303. URL http://link.springer.com/10.1007/s00038-016-0926-5.

Brailovskaya, Valentina, Pascaline Dupas, and Jonathan Robinson. 2024. "Is Digital Credit Filling a Hole or Digging a Hole? Evidence from Malawi." *The Economic Journal* 134 (658):457–484. URL https://academic.oup.com/ej/article/134/658/457/7296122.

Branson, Nicola and Tanya Byker. 2018. "Causes and consequences of teen childbearing: Evidence from a reproductive health intervention in South Africa." *Journal of Health Economics* 57:221–235.

Brubaker, Sarah Jane and Christie Wright. 2006. "Identity Transformation and Family Caregiving: Narratives of African American Teen Mothers." *Journal of Marriage and Family* 68 (5):1214–1228. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1741-3737.2006.00324.x.

Burchardi, Konrad, Jonathan de Quidt, Selim Gulesci, Benedetta Lerva, and Stefano Tripodi. 2021. "Testing willingness to pay elicitation mechanisms in the field: Evidence from Uganda." *Journal of Development Economics* 152.

Burchardi, Konrad, Jonathan de Quidt, Selim Gulesci, and Munshi Sulaiman. 2024. "Borrowing Constraints and Demand for Remedial Education: Evidence from Tanzania." *The Economic Journal* 134:2621–2637.

Butera, Luigi, Robert Metcalfe, William Morrison, and Dmitry Tubinsky. 2022. "Measuring the Welfare Effects of Shame and Pride." *American Economic Review* 112:122–68.

Callaway, Brantly and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225 (2):200–230. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407620303948.

Campbell, Benjamin C. and J. Richard Udry. 1995. "Stress and age at menarche of mothers and daughters." *Journal of Biosocial Science* 27 (2):127–134. URL https://www.cambridge.org/core/journals/journal-of-biosocial-science/article/abs/stress-and-age-at-menarche-of-mothers-and-daughters/E2352C0F118DA028CEE9A8B7BEA81E3F.

Cantet, Natalia. 2020. "The Effects of Teenage Pregnancy on Schooling and Labor

Force Participation: Evidence from Urban South Africa." *Unpublished manuscript.* URL https://papers.nataliacantet.com/Teenage%20pregnancy%20-%20Cantet.pd.

Carter, Michael, Rachid Laajaj, and Dean Yang. 2021. "Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique†." *American Economic Journal: Applied Economics* 13 (2):206–229.

Channa, Hira, Amy Z Chen, Patricia Pina, Jacob Ricker-Gilbert, and Daniel Stein. 2019. "What drives smallholder farmers' willingness to pay for a new farm technology? Evidence from an experimental auction in Kenya." *Food Policy* 85:64–71.

Chari, A. V., Rachel Heath, Annemie Maertens, and Freeha Fatima. 2017. "The causal effect of maternal age at marriage on child wellbeing: Evidence from India." *Journal of Development Economics* 127:42–55. URL https://www.sciencedirect.com/science/article/pii/S0304387817300123.

Chioda, Laura, Paul Gertler, Sean Higgins, and Paolina Medina. 2024. "FinTech Lending to Borrowers with No Credit History." Tech. Rep. w33208, National Bureau of Economic Research, Cambridge, MA. URL http://www.nber.org/papers/w33208.pdf.

Cohen, Jessica and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *The Quarterly Journal of Economics* 125 (1):1–45.

Crump, R. K., V. J. Hotz, G. W. Imbens, and O. A. Mitnik. 2009. "Dealing with limited overlap in estimation of average treatment effects." *Biometrika* 96 (1):187–199. URL https://academic.oup.com/biomet/article-lookup/doi/10.1093/biomet/asn055.

Deutschmann, Joshua. 2024. "Recognizing a good deal: short-term subsidies and the dynamics of public service use." *Working Paper* .

Diaz, Christina J. and Jeremy E. Fiel. 2016. "The Effect(s) of Teen Pregnancy: Reconciling Theory, Methods, and Findings." *Demography* 53 (1):85–116. URL https://www.jstor.org/stable/24756858.

Dizon-Ross, Rebecca and Seema Jayachandran. 2023. "Detecting Mother- Father Differences in Spending on Children: A New Approach Using Willingness- to-Pay Elicitation†." *American Economic Review: Insights* 5:445449.

Duncan, Greg J., Kenneth T. H. Lee, Maria Rosales-Rueda, and Ariel Kalil. 2018. "Maternal Age and Child Development." *Demography* 55 (6):2229–2255. URL https://www.jstor.org/stable/45048125.

Dupas, Pascaline. 2014. "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment." *Econometrica* 82 (1):197–228.

Dupas, Pascaline and Edward Miguel. 2017. "Impacts and determinants of health levels in low-income countries." In *Handbook of economic field experiments*, vol. 2. Elsevier, 3–93.

Dvornyk, Volodymyr and Waqar-ul-Haq. 2012. "Genetics of age at menarche: a systematic review." *Human Reproduction Update* 18 (2):198–210. URL http://academic.oup.com/humupd/article/18/2/198/620595/Genetics-of-age-at-menarche-a-systematic-review.

Fairfax, Colita Nichols. 2008. "Edin, K., & Kefalas, M. (2005). *Promises I Can Keep: Why Poor Women Put Motherhood before Marriage* . Berkeley: University of California Press. (293 pp., $24.95-hb, ISBN-0-520-24113-4)." 18 (3):392–396. URL http://www.tandfonline.com/doi/abs/10.1080/10911350802427480.

Field, Erica and Attila Ambrus. 2008. "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh." *Journal of Political Economy* 116 (5):881–930. URL https://www.journals.uchicago.edu/doi/10.1086/593333.

Fletcher, Jason M. and Barbara L. Wolfe. 2009. "Education and Labor Market Consequences of Teenage Childbearing: Evidence Using the Timing of Pregnancy Outcomes and Community Fixed Effects." *The Journal of Human Resources* 44 (2):303–325. URL https://www.jstor.org/stable/20648899.

Frisancho, Veronica. 2012. "Signaling Creditworthiness in Peruvian Microfinance Markets: The Role of Information Sharing." *The B.E. Journal of Economic Analysis & Policy* 12 (1). URL https://www.degruyter.com/document/doi/10.1515/1935-1682.3178/html.

Gorry, Devon. 2019. "Heterogeneous Consequences of Teenage Childbearing." *Demography* 56 (6):2147–2168. URL https://read.dukeupress.edu/demography/article/56/6/2147/168069/Heterogeneous-Consequences-of-Teenage-Childbearing.

Grimm, Michael, Luciane Lenz, Jörg Peters, and Maximiliane Sievert. 2020. "Demand for Off-Grid Solar Electricity: Experimental Evidence from Rwanda." *Journal of the Association of Environmental and Resource Economists* 7:417–454.

Heiland, Frank, Sanders Korenman, and Rachel A. Smith. 2019. "Estimating the educational consequences of teenage childbearing: Identification, heterogeneous effects and the value of biological relationship information." *Economics & Human Biology* 33:15–28. URL https://linkinghub.elsevier.com/retrieve/pii/S1570677X18301412.

Hidrobo, Melissa, Giordano Palloni, Daniel O. Gilligan, Jenny C. Aker, and Natasha Ledlie. 2022. "Paying for Digital Information: Assessing Farmers' Willingness to Pay for a Digital Agriculture and Nutrition Service in Ghana." *Economic Development and Cultural Change* 70 (4):1367–1402.

Hoffmann, Bridget. 2018. "Do Non-Monetary Prices Target the Poor? Evidence from a Field Experiment in India." *Journal of Development Economics* 133:15–32.

Hoffmann, Vivian, Christopher B. Barrett, and David R. Just. 2009. "Do Free Goods Stick to Poor Households? Experimental Evidence on Insecticide Treated Bednets." *World Development* 37:607–617.

Huang, Jian, Wim Groot, John G. Sessions, and Yinyen Tseng. 2019. "Age of Menarche, Adolescent Sexual Intercourse and Schooling Attainment of Women." *Oxford Bulletin of Economics and Statistics* 81 (4):717–743. URL https://onlinelibrary.wiley.com/doi/10.1111/obes.12284.

Janzen, Sarah, Nicholas Magnan, Conner Mullally, Soye Shin, I. Bailey Palmer, Judith Oduol, and Karl Hughes. 2021. "Can Experiential Games and Improved Risk Coverage Raise Demand for Index Insurance? Evidence from Kenya." *American Journal of Agricultural Economics* 103 (1):338–361.

Kane, Jennifer B., S. Philip Morgan, Kathleen Mullan Harris, and David K. Guilkey. 2013. "The Educational Consequences of Teen Childbearing." *Demography*

50 (6):2129–2150. URL https://www.jstor.org/stable/42919973.

Kaprio, J., A. Rimpelä, T. Winter, R. J. Viken, M. Rimpelä, and R. J. Rose. 1995. “Common genetic influences on BMI and age at menarche.” *Human Biology* 67 (5):739–753.

Karlan, Dean and Jonathan Zinman. 2009. “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment.” *Econometrica* 77 (6):1993–2008. URL https://www.jstor.org/stable/25621388.

Klepinger, Daniel H., Shelly Lundberg, and Robert D. Plotnick. 1995. “Adolescent Fertility and the Educational Attainment of Young Women.” *Family Planning Perspectives* 27 (1):23–28. URL https://www.jstor.org/stable/2135973.

Lapo-Talledo, German Josuet. 2024. “Nationwide study of in-hospital maternal mortality in Ecuador, 2015-2022.” 48:e5.

Levine, Ross, Norman Loayza, and Thorsten Beck. 2000. “Financial intermediation and growth: Causality and causes.” *Journal of Monetary Economics* 46 (1):31–77. URL https://www.sciencedirect.com/science/article/pii/S0304393200000179.

Liu, Lei, Guangli Lu, and Wei Xiong. 2022. “The Big Tech Lending Model.” Tech. Rep. w30160, National Bureau of Economic Research, Cambridge, MA. URL http://www.nber.org/papers/w30160.pdf.

Lybbert, Travis J, Nicholas Magnan, David J Spielman, Anil K Bhargava, and Kajal Gulati. 2018. “Targeting technology to increase smallholder profits and conserve resources: experimental provision of laser land-leveling services to Indian farmers.” *Economic Development and Cultural Change* 66 (2):265–306.

Maffioli, Alessandro, David McKenzie, and Diego Ubfal. 2023. “Estimating the Demand for Business Training: Evidence from Jamaica.” *Economic Development and Cultural Change* 72.

Meager, Rachael. 2019. “Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments.” *American Economic Journal: Applied Economics* 11 (1):57–91. URL http://www.aeaweb.org/articles?id=10.1257/app.20170299.

Meriggi, Niccolò F., Erwin Bulte, and Ahmed Mushfiq Mobarak. 2021. "Subsidies for technology adoption: Experimental evidence from rural Cameroon." *Journal of Development Economics* 153.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich. 2023. "Factorial designs, model selection, and (incorrect) inference in randomized experiments." *Review of Economics and Statistics* :1–44.

Narita, Renata and Maria Dolores Montoya Montoya Diaz. 2016. "Teenage motherhood, education, and labor market outcomes of the mother: Evidence from Brazilian data." *EconomiA* 17 (2):238–252. URL https://linkinghub.elsevier.com/retrieve/pii/S1517758016300790.

Omotilewa, Oluwatoba J., Jacob Ricker-Gilbert, and John Herbert Ainembabazi. 2019. "Subsidies for Agricultural Technology Adoption: Evidence from a Randomized Experiment with Improved Grain Storage Bags in Uganda." *American Journal of Agricultural Economics* 101 (3):753–772.

Omotilewa, Oluwatoba J., Jacob Ricker-Gilbert, John Herbert Ainembabazi, and Gerald E. Shively. 2018. "Does improved storage technology promote modern input use and food security? Evidence from a randomized trial in Uganda." *Journal of Development Economics* 135:176–198.

Perry, John R. B., Sheila Ulivi, Erin K. Wagner, Najaf Amin, Laura J. Bierut, Enda M. Byrne, Jouke-Jan Hottenga, Daniel L. Koller, Massimo Mangino, Tune H. Pers, Laura M. Yerges-Armstrong, Jing Hua Zhao, Irene L. Andrulis, Hoda Anton-Culver, Femke Atsma, Stefania Bandinelli, Matthias W. Beckmann, Javier Benitez, Carl Blomqvist, Stig E. Bojesen, Manjeet K. Bolla, Bernardo Bonanni, Hiltrud Brauch, Hermann Brenner, Julie E. Buring, Jenny Chang-Claude, Stephen Chanock, Jinhui Chen, Georgia Chenevix-Trench, J. Margriet Collée, Fergus J. Couch, David Couper, Andrea D. Coviello, Angela Cox, Kamila Czene, Adamo Pio D'adamo, George Davey Smith, Immaculata De Vivo, Ellen W. Demerath, Joe Dennis, Peter Devilee, Aida K. Dieffenbach, Alison M. Dunning, Gudny Eiriksdottir, Johan G. Eriksson, Peter A. Fasching, Luigi Ferrucci, Dieter Flesch-Janys, Henrik Flyger, Tatiana Foroud, Lude Franke, Melissa E. Garcia, Montserrat García-Closas, Frank Geller, Eco E. J. de Geus, Graham G. Giles, Daniel F. Gudbjartsson, Vilmundur Gudnason, Pascal Guénel, Suiqun Guo, Per Hall, Ute Hamann, Robin Haring, Catharina A. Hartman, Andrew C. Heath, Albert Hofman, Maartje J. Hooning, John L. Hopper, Frank B. Hu, David J. Hunter, David Karasik, Douglas P. Kiel, Julia A. Knight, Veli-Matti Kosma, Zoltan Kutalik, Sandra Lai, Diether Lambrechts, Annika Lindblom, Reedik Mägi, Patrik K. Magnusson, Arto Mannermaa, Nicholas G. Martin, Gisli Masson, Patrick F. McArdle, Wendy L. McArdle, Mads Melbye, Kyriaki Michailidou, Evelin Mihailov, Lili Milani, Roger L. Milne, Heli Nevanlinna, Patrick Neven, Ellen A. Nohr, Albertine J. Oldehinkel, Ben A. Oostra, Aarno Palotie, Munro Peacock, Nancy L. Pedersen, Paolo Peterlongo, Julian Peto, Paul D. P. Pharoah, Dirkje S. Postma, Anneli Pouta, Katri Pylkäs, Paolo Radice, Susan Ring, Fernando Rivadeneira, Antonietta Robino, Lynda M. Rose, Anja Rudolph, Veikko Salomaa, Serena Sanna, David Schlessinger, Marjanka K. Schmidt, Mellissa C. Southey, Ulla Sovio, Meir J. Stampfer, Doris Stöckl, Anna M. Storniolo, Nicholas J. Timpson, Jonathan Tyrer, Jenny A. Visser, Peter Vollenweider, Henry Völzke, Gerard Waeber, Melanie Waldenberger, Henri Wallaschofski, Qin Wang, Gonneke Willemsen, Robert Winqvist, Bruce H. R. Wolffenbuttel, Margaret J. Wright, Dorret I. Boomsma, Michael J. Econs, Kay-Tee Khaw, Ruth J. F. Loos, Mark I. McCarthy, Grant W. Montgomery, John P. Rice, Elizabeth A. Streeten, Unnur Thorsteinsdottir, Cornelia M. van Duijn, Behrooz Z. Alizadeh, Sven Bergmann, Eric Boerwinkle, Heather A. Boyd, Laura Crisponi, Paolo Gasparini, Christian Gieger, Tamara B. Harris, Erik Ingelsson, Marjo-Riitta Järvelin, Peter Kraft, Debbie Lawlor, Andres Metspalu, Craig E. Pennell, Paul M. Ridker, Harold Snieder, Thorkild I. A. Sørensen, Tim D. Spector, David P. Strachan, André G. Uitterlinden, Nicholas J. Wareham, Elisabeth Widen, Marek Zygmunt, Anna Murray, Douglas F. Easton, Kari Stefansson, Joanne M. Murabito, and Ken K. Ong. 2014. "Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche." *Nature* 514 (7520):92–97. URL https://www.nature.com/articles/nature13545.

Ribar, David C. 1994. "Teenage Fertility and High School Completion." *The Re-*

view of Economics and Statistics 76 (3):413–424. URL https://www.jstor.org/stable/2109967.

Roth, Jonathan. 2024. "Interpreting Event-Studies from Recent Difference-in-Differences Methods." URL https://arxiv.org/abs/2401.12309.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics* 235 (2):2218–2244. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407623001318.

Sabia, Joseph J. and Daniel Rees. 2009. "The Effect of Sexual Abstinence on Females' Educational Attainment." *Demography* 46 (4):695–715. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831358/.

Sant'Anna, Pedro H.C. and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of Econometrics* 219 (1):101–122. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407620301901.

Shukla, Pallavi, Hemant K. Pullabhotla, and Kathy Baylis. 2022. "Trouble with Zero: The Limits of Subsidizing Technology Adoption." *Journal of Development Economics* 158.

Treurniet, Mark. 2023. "The Impact of Being Surveyed on Agricultural Technology Adoption." *Economic Development and Cultural Change* 71 (3):941–960.

UNPD, World Population Prospects. 2021. "Data Bank, Gender Statistics." URL https://genderdata.worldbank.org/en/indicator/sp-ado-tfrt.

World Bank. 2025. "Bank Non-Performing Loans to Gross Loans for El Salvador." https://fred.stlouisfed.org/series/DDSI02SVA156NWDB. Retrieved from FRED, Federal Reserve Bank of St. Louis.

Yakusheva, Olga. 2011. "In High School and Pregnant: The Importance of Educational and Fertility Expectations for Subsequent Outcomes." *Economic Inquiry* 49 (3):810–837. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1465-7295.2010.00313.x.

Zito, Rena Cornell. 2018. "Children as Saviors? A Propensity Score Analysis of the Impact of Teenage Motherhood on Personal Transformation." *Youth & Society* 50 (8):1100–1122. URL http://journals.sagepub.com/doi/10.1177/0044118X16653872.

Zwane, Alex, Jonathon Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, William Devoto, Bruno Crepon, and Abhijit Banerjee. 2011. "Being Surveyed can Change later Behavior and Related Parameter Estimates." *PNAS* 108 (5):395–424.