ABSTRACT

Title of dissertation:	THE ROLE OF PRICE INFORMATION IN AGRICULTURAL MARKETS: EVIDENCE FROM RURAL PERU
	Eduardo Nakasone, Doctor of Philosophy, 2014
Dissertation directed by:	Professor Kenneth Leonard Department of Agricultural and Resource Economics

Agriculture is a source of livelihood for 86% of households in rural areas, many of whom rely on their crops for income. However, many farmers in isolated areas do not have access to reliable market price information that can inform them about the most profitable opportunities on where to sell their products.

This dissertation presents new evidence on the role of price information in farmers' marketing outcomes. I use data from a field experiment in the central highlands of Peru. A group of farmers in randomly selected villages was provided with mobile phones, through which they received detailed price information for seventeen relevant crops in six regional markets. I find that those provided with the information received 13-14% higher prices for their products. This effect was larger for perishable crops and for more risk-averse households. Information also made farmers more likely to participate in commercial activities and sell their crops (rather than allocating them for self-consumption).

These results were not driven by other mobile phone benefits as the phones distributed to the

farmers were restricted to only receive the price SMS during the period of the intervention. They are not driven by production decisions either because the intervention took place after planting decisions had already been made. Finally, I also investigate the possibility of information spillovers by examining marketing outcomes of households who did not receive the information but lived in villages where others did. I do not find any significant effects among households in this group.

The Role of Price Information in Agricultural Markets: Evidence from Rural Peru

by

Eduardo Nakasone

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Kenneth Leonard, Chair/Advisor
Professor Raymond Guiteras
Professor John Hoddinott
Professor Vivian Hoffmann
Professor Kenneth McConnell
Professor Sergio Urzua (Dean's Representative)

Acknowledgments

I would like to thank my advisor, Ken Leonard, for his guidance and feedback throughout my graduate studies. I am also grateful to the members of my advisory committee: Raymond Guiteras, Vivian Hoffmann, John Hoddinott, Ted McConnell, and Sergio Urzua for their insightful comments. I owe a great deal of gratitude to Maximo Torero and all of my colleagues at IFPRI for their feedback and support. In particular, I would like to thank Alan de Brauw, Miguel Almanzar, and Manuel Hernandez for their multiple advice on my research. I am also grateful to several people, who, in addition to providing practical advice and useful feedback in many cases, gave me the emotional support without which I would not have been able to complete this program. I would like to thank Veronica Frisancho, Rodrigo Salcedo, Pablo Suarez, James Doyle, Silvia Luperdi, Jorge de la Roca, Matthew McClellan, Luis Camacho, Ledda Macera, Romina Ordoñez, Katya Vasilaky, Kanishka Kacker, and John Roberts for helping me through this difficult task. Most importantly, I would like to thank my family and, in particular, my parents for their unwavering support and encouragement throughout my life.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Imperfect information has adverse consequences on market performance and welfare. Such imperfections seem to be especially prevalent in developing countries, where communication technologies and infrastructure are often deficient. Within developing countries, information imperfections may be particularly acute in agricultural markets and primarily affect small farmers[1]. In particular, these small farmers — usually living in remote areas and without access to adequate infrastructure — may be less informed about market conditions. As they sell their products to middlemen, they face one considerable disadvantage (among others): better informed traders can exploit information asymmetries to pay lower farm gate prices. Therefore, enhanced market price information should increase farmers'

[1]For example, Mitra & Sarkar (2003) investigate the potato markets in West Bengal. They find that, while farmers earn very small profits for (that even become negative when imputing for factors of production, such as family labor), traders have exhibit substantial mark-ups for their commercial activity. They argue that these differences in profit margins are likely to be a consequence of traders' informational advantage. In another example, Fafchamps & Hill (2008) analyze coffee markets in Uganda. The authors find that while increases in international coffee prices readily translate into higher export and wholesale prices, but imply much smaller increments in the prices paid to farmers. They posit that traders take advantage of farmers' ignorance about price movements.

sales prices. However, the evidence has been somewhat mixed: while some authors find a positive impact, others do not find any[2].

In this paper, I provide new experimental evidence about the role of agricultural information on marketing outcomes. I conducted a field experiment in the central highlands of Peru by randomly allocating price information among agricultural households in 58 villages. Twenty six villages were assigned to the treatment group, while the others remained as controls. Within villages in the treatment group, I randomly provided cell phones to 111 households. I collected detailed price information for seventeen different crops by quality in six different relevant markets. Those who received cell phones were sent price information through Short Message Service (SMS) for the four months immediately after the rainy season in the highlands. This is the period in which farmers have already harvested their crops and make most of their sales decisions. Therefore, the intervention allows me to capture the effects of price information on marketing strategies, isolated from any production decisions. To make information more digestible — rather than providing a massive number of SMS — farmers only received information for the crops they harvested.

The intervention also ensured that the farmers benefited only from enhanced market price information. In general, mobile phones provide users with a wide array of commercial benefits, besides access to price information (e.g. they facilitate coordination, direct bargaining of sales conditions with clients; arrangements with input providers; collaboration with other producers, etc.). To measure the impact of price information isolated from

[2]Svensson & Yanagizawa (2009), Goyal (2010), Courtois & Subervie (2013) and Nyarko, Hildebrandt, Romagnoli & Soldani (2013) find positive impacts; Fafchamps & Minten (2012), Aker & Fafchamps (2013), Camacho & Conover (2011) and Mitra, Mookherjee, Torero & Visaria (2013) do not find any effect. This evidence is discussed in more detail in Section 2.

these parallel benefits of mobile phones, the devices provided to farmers were only able to receive SMS and calls from a phone number managed by the project. Participants were able to keep the devices as unrestricted pre-paid phones with no further obligation after this period.

Within this setting, I test four hypotheses. First, I test whether market price information leads to higher sales prices. For this purpose, I compare the prices of the beneficiaries who directly received the price information through their cell phones with those of households in the control villages. Second, I test if access to market price information increases farmers' market participation. In contrast to previous papers that focus on farmers who were already participating in commercial activities, I intentionally sampled both farmers who sell their production and who only grow their crops for self-consumption. This allows me to estimate the effect of information on the probability of engaging in any commercial activity. Third, I analyze potential heterogeneous effects by crop types and households' risk aversion, to assess the conditions under which information may be more effective. Fourth, I investigate if there are any spillover effects of information by analyzing the marketing outcomes of households who did not receive the treatment but lived in villages where others were treated. Farmers in this group might have been exposed indirectly to the price information, even when they did not receive it directly.

This paper presents four main contributions to the literature relating price information and farmers' agricultural market performance. First, I present the first experimental estimate of the effect of information on the extensive margin of market participation. Second, I am able to isolate the short-run effect of information on farmers' marketing strategy by phasing the timeline of the intervention. Third, as opposed to some previous work that has focused on Information and Communication Technologies in general, the nature of the intervention allows me to disentangle the sole effect of market price information stripped from any other benefits. Fourth, in contrast to previous papers which restrict attention to households that had previous access to a certain technology (e.g. previous cell phone ownership, radio, etc.), this intervention encompassed the provision of such technology. This allows me to explore to what extent the focus on certain subpopulations may have led previous results, since households with previous access to technology tend to be wealthier and more educated.

My results suggest that price information has a large and sizable impact: farmers who received the information directly experienced 13%-14% increases in their sales prices. This result is robust to different specifications and variations in the sample. The effect is mostly driven by increases in prices for relatively more perishable products (for which information is more valuable), and for more risk-averse households (which might have been more affected by price market uncertainty). I find no differential effects by previous ownership of a cell phone. This suggests that those less familiar with this technology can also benefit from a price dissemination policy.

Among households who received the price information, there was an increase of about 12% in the probability of engaging in a commercial transaction for their crops. Thus, the information had a large effect on the extensive margins of sales. On the intensive margin (traded volumes, conditional on any sales), I find large but not statistically significant impacts on the information group. Due to the timing of the intervention, these results are not driven by changes in households' crop choices or output levels. I do not find any evidence to support the presence of spillover effects: there are no apparent price benefits to farmers who did not receive the information directly but were in villages where someone else did. This result is consistent across multiple specifications for social interaction (e.g. geographic distance, crop restrictions, etc.).

As with any limited scale study, results should be interpreted in context. First, when comparing these results to previous findings in the literature, the context should be considered. In this paper, the considerable increases in the probability of commercial transactions reveal that this area still had ample room for further participation in agricultural markets (where price information may potentially have larger impacts). Further, the effect of information may depend on the crops under analysis. My results suggest that information is more important on perishable crops.

The remainder of this dissertation is organized in five chapters. Chapter 2 discusses some of the related literature on the impact of market price information in rural areas of developing countries. Chapter 3 describes the RCT in the central highlands of Peru. Chapter 4 presents a simple theoretical model to frame some of the impact of information on marketing outcomes. Chapter 5 presents the empirical strategy and the results. Finally, Chapter 6 concludes.

Chapter 2

The Effects of Price Information on Market Performance: A Review of the Literature

This chapter presents a brief discussion of the recent literature that analyzes the impact of market price information on market performance in developing countries. A first group of papers have analyzed the availability of mobile phone service to improve the functioning of rural markets in developing countries[1]. Mobile phones can facilitate timely access to market prices and unexploited opportunities to sell or buy goods[2]. In this spirit, Jensen

[1]For a recent review of the impact of mobile phones on agricultural development, see Nakasone, Torero & Minten (forthcoming).

[2]Other papers have also discussed the impact of public phone coverage. For example, Chong, Galdo & Torero (2009) exploit large expansions of the public phone coverage in rural Peru. They combine the roll-out of public phones with household surveys, and find that this policy led to increases in per-capita agricultural income of 17-21%. Beuermann (2011) analyzes the same policy and finds similar results: access to a public phone led to a 19.5% increase in agricultural profitability. In contrast to these findings, Futch & McIntosh (2009) find no impact on farm gate prices of a program that installed village phones in Rwanda. The

(2007) analyzes the introduction of mobile phone service among fishermen in Kerala. He finds that this led to compliance with the law of one price across different markets, fewer wasted fish and a reduction in prices. On the demand side, this price reduction increased consumers' surplus. On the supply side, price reductions were dwarfed by growth in sale volumes (from reduced wastage), so fishermen's surplus increased as well. Aker (2010) studies the rollout of mobile service coverage in Niger. Using data from national markets and traders, she finds that mobile service reduced price dispersion between millet markets and increased middlemen's profits. Later studies show that these benefits might not have translated into improvements for farmers, though. In a complementary study, Aker & Fafchamps (2013) find that mobile phones did not lead to increases in cowpea prices for producers in the same context. However, the authors do find evidence of reduced intra-annual price variability. Muto & Yamano (2009) use a household panel dataset to identify the impact of cell phone coverage on farmers' participation in maize and banana markets in Uganda. They find that mobile coverage has a positive impact on the sales of bananas but no effect for maize. They argue that these results might be driven by the higher perishability of the former crop compared to the latter. Molony (2008) argues that the existence of credit relationships between farmers and traders may prevent the former from improving their marketing strategies when provided with mobile phones.

While these studies analyze the introduction of a technology that potentially allow households to find out prevailing market prices, a second group of papers investigate the

authors envisaged an experimental evaluation of this program. However, because the actual phone assignment deviated considerably from the original design, they caution about the interpretation of their results.

effect of the direct provision of such information. Svensson & Yanagizawa (2009) study the impact of the Market Information System (MIS) in Uganda, which disseminated agricultural prices through radio stations. The authors compare households with and without radios in districts that were and were not covered by MIS, and find that access to information increased farm-gate prices for maize by 10%-15%. Goyal (2010) investigates the impact of internet kiosks installed by a large processor in Madhya Pradesh, India; which provided soybean price information. She finds that this led to an increase of 1-3% in the prices received by farmers. It also increased the farmers' land allocated to soybeans by 19%, suggesting a substitution away from other crops.

Fafchamps & Minten (2012) conduct a field experiment, where a group of farmers was provided with one-year free subscriptions to an SMS-based agricultural information service (Reuters Market Light) in Maharashtra, India. The subscriptions were randomly allocated among farmers who already had cell phones in this region. The service included price information in different markets, and also encompassed weather forecasts and crop advisory. They find that such service did not lead to increases in agricultural prices for those who received it. Second, Mitra et al. (2013) study the impact of price information on potato farmers in West Bengal, India. They test the efficacy of two alternative strategies for market price dissemination: a private one where a group of randomly selected farmers received SMSs with this information and a public one where prices where posted in public notice boards in some villages. The authors find that neither of these strategies improved farmers' market performance.

A couple of recent papers analyze the impact of ESOKO, a program that provides farmers with market prices in Ghana. To account for selection into the program, Courtois &

Subervie (2013) use propensity score matching techniques; and find that those who received information from ESOKO increased their sales prices by 13% for maize and 10% for groundnuts. Nyarko et al. (2013) also analyze the impact of this program through a randomized controlled trial. Their results suggest that, after one year of exposure, farmers with information experienced a 7%-11% increase in the prices they receive for yam. However, they do not find any significant impact for other crops, such as maize and cassava.

This research is also related to the literature of information spillover effects, which has been present for a while. Though applied to a consumer problem, in the early sixties, Stigler had already noted that: "Information is pooled when two buyers compare prices: if each buyer canvasses *s* sellers, by combining they effectively canvas 2*s* sellers, duplications aside[...] in fact, pooling can be looked upon as a cheaper form of search" Stigler (1961, p. 219). Previous literature has highlighted role of neighbors and social networks in agricultural technology adoption and learning (e.g. Bandiera & Rasul 2006, Besley & Case 1994, Foster & Rosenzweig 1995, Munshi 2004, Vasilaky & Leonard 2013, Magnan, Spielman, Lybbert & Gulati 2013).

One of the difficulties of such analysis is the reflection problem proposed by Manski (1993): when someone behaves in a similar way to their peers, it might be because of the influence that peers have on him, or because he shares similar characteristics with his peers (and, thus, behaves similarly). To circumvent this problem, some studies exploit variation from random assignment to assess changes among peers of the subjects that receive a treatment. For example, Oster & Thornton (2012) analyze the impact of providing school girls with menstrual cups on their peers' subsequent adoption of this technology. Bobonis & Finan (2009) use experimental evidence from the allocation of a large education-related conditional cash transfer program (i.e. PROGRESA / Oportunidades) to assess the impact of eligible children's schooling decisions on the enrollment decisions of other children that lived in the same communities but were ineligible for the program. Giné & Mansuri (2011) investigate the impact of a voter awareness campaign in Pakistan that was implemented in randomly selected clusters. They find significant increases of female turnout among treated households. Importantly, they also find similar increases in turnout among women that were not part of this campaign but lived in communities where others were.

In this line, I exploit the fact that the treatment was randomized at the village level in the first stage. In particular, I investigate the marketing outcomes of households who did not receive any price SMS, but lived in villages where others did. The idea is that those in this group might have been exposed indirectly to the price information, even when they did not receive such information directly.

Chapter 3

An Intervention to Assess the Role of Price Information on Agricultural Markets in Rural Peru

The main problem with disentangling the causal effect of the impact of agricultural information on marketing decisions is the endogenous nature of this relationship. In a non-experimental setting, assume that one finds that access to information leads to better sales outcomes. This relationship could be driven by any number of factors and not necessarily by the information itself. For example, the ones seeking information may be precisely those who find more profitable to do so, may have better entrepreneurial skills, or may be more market-oriented. In this sense, this relationship would be merely correlational and not causal.

To tackle this obstacle, I conducted a field experiment. The experiment randomly allocated cell phones to some farmers in the central highlands of Peru. Through these cell phones, I provided price information in nearby markets for the main crops in this region. Farmers received this information for four months, throughout the period during which they sell most of their agricultural production. Hence, the intervention provides me with exogenous variation in access to information among similar households. The objective is to investigate whether this information leads to better marketing outcomes.

The intervention took place in the five provinces of the Mantaro Valley in the Central Highlands of Peru (Figure 3.1). The Mantaro valley is one of the most productive agricultural areas in Peru, and is usually considered the country's "food pantry". The five provinces in the intervention encompass 158 thousand hectares of agricultural land, and host about 75 thousand farmers. While there are some with large extensions of land, farmers in this area are predominantly small: 53.2% of the farmers grow their crops in less than 0.5 hectares, and 90% of them have less than 3 hectares (National Statistics Institute 2013). Most households diversify their production, and grow several crops across multiple plots. The most important crops are potatoes, *olluco* (a popular Andean tuber), barley, wheat, peas, lima beans, and animal fodder (e.g. forage grass, alfalfa, ryegrass, etc.). The area is also characterized by a high rotation of crops between years as part of their soil and pest management strategies[1].

While large farmers grow crops for strictly commercial purposes, small landholders have to decide whether to allocate their crops to household self-consumption or to sell them

[1] For example, one of the most common rotations is to grow potatoes one year, another Andean tuber during the second year, lima beans or barley during the third year, let the land fallow for a year, and start this rotation again.

(or a combination of both). Farmers that decide to sell their production have several options. The first one is to sell them to agricultural middlemen (locally known as *acopiadores*). These are itinerant traders that visit villages (usually with a truck), purchase crops from several farmers, and resell them to wholesale buyers in local markets or in the capital city. While there is no precise information about the middlemen in this area, there is a widespread belief that they face limited competition and earn considerable profits[2]. The farmers' second option is to bypass the middlemen, and sell their harvests directly in local markets (but incur transportation and transaction costs). There are several markets in the Mantaro Valley. Two of them are permanent markets that operate daily (Huancayo, Juaja, and Tarma), while some others are weekly *ferias*. The size of these markets or ferias is also quite variable, and some of them are very specialized (for example, some focus exclusively on cattle or particular agricultural products). However, there are multiple buyers and sellers (both middlemen and farmers) operating in all of them, fostering a relatively competitive environment.

Farmers who decide to sell grade their harvest according to their quality: prices are higher for first quality (usually larger, with better appearance, without any insect damage, etc.) than for second, third, or fourth qualities. There are implicit agreements between

[2]For example, a Peruvian Minister of Agriculture argued in an interview that "individual sales of small quantities of agricultural production to middlemen implies giving up most of the value added of the crop" (Eguren 2008). The government has promoted different policies to reduce the bargaining power of the middlemen in the Mantaro Valley. For example, the Ministry of Agriculture launched the "Potato Train" (*Tren Papa*) in 2008, a program through which the railway company would sporadically provide free transportation of potato harvests from Jauja (one of the Mantaro Valley's largest cities) to Lima's main wholesale market. While the program was unsuccessful, it was explicitly implemented to reduce middlemen's profit margins in this area and to increase farmers' incomes. Similarly, Escobal, Ponce & Hernandez-Asensio (2010) describe how local municipalities in Jauja have organized annual festivals (*raymis*) for different crops. These festivals gather local producers of a particular crop in a single space and aim to attract larger wholesale buyers, who can purchase farmers' production directly and bypass the middlemen.

buyers and sellers of about these qualities, and both parties are able to readily identify them. Households who sell their production are required to grade and sort their harvest. Some households also grade the share of their production for self-consumption, but usually they do not.

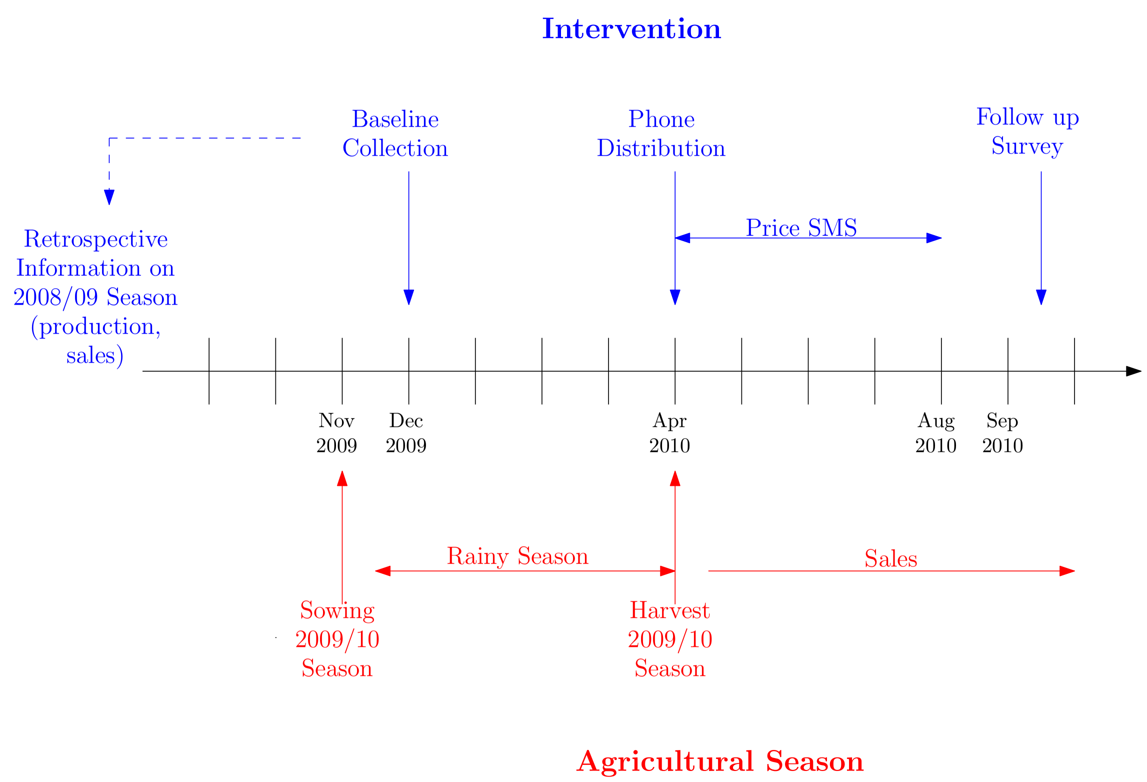
An important characteristic of this area is the agricultural year (see Figure 3.2). Farmers in the highlands of Peru usually sow their crops around mid-November, at the start of the rainy season. The rainy season typically extends until March or April. The growing periods across products, but harvest is generally between late March and May. For farmers without irrigation, this is their only cropping cycle in the year and an important source of income. Those with irrigation can start an additional cropping cycle in May or June. However, even those with irrigation take advantage of the rainy season, which yields their largest production in the year.

I selected 58 villages in the Mantaro Valley that met the following criteria in the 2007 Peruvian Census: (a) were in the highlands, (b) were in a rural area, (c) had at least 60 households, (d) were connected to the electricity grid, (e) had at most 35% of cell phone coverage[3]. Data from a random sample of households in each of these villages was collected in December 2009, when the rainy season had already started and farmers had already sown their crops for the 2009/2010 agricultural cycle. I collected information about socio-economic characteristics (household composition, education, income, expenditures, etc.), agricultural land, social networks (participation in organizations) and location (GPS

[3]While the rates in the 2007 Census were substantially lower, I found that cell phone penetration had already reached about 50% during the intervention.

Figure 3.1: Location of the Intervention

Figure 3.2: Timeline of the Intervention



location of dwelling and main agricultural plot). Importantly, I gathered retrospective data about their previous (2008/2009) agricultural cycle: production, sales volume, prices, and marketing decisions. The questionnaire also asked them which products they had already planted for the 2009/2010 season.

The baseline survey included 790 households in the 58 villages where the intervention took place. Rather than randomly allocating the cell phones among the full roster of households, the villages were assigned either to a treatment or a control groups in a first stage (Figure 3.3). This initial assignment of treatment by cluster has two advantages. First, it minimized the risk of contamination of the control group: if treatment and control households were in the same village, this would increase the possibility of beneficiaries passing price information along to control households. Second, this provides a framework to investigate the existence of spillover effects in the treatment villages.

The budget of the project allowed for the purchase of 112 mobile phones (out of which one was connected to a computer and used to deliver the SMS messages to farmers). I randomly sorted the 58 villages in my sampling framework, and allocated one cell phone per each four households (rounded up). The 111 mobile phones available for distribution determined a treatment group of 26 villages and a control group of 32 villages.

There were 410 households in the treatment villages, from which 111 were randomly selected to receive a cell phone. The devices were basic inexpensive phones[4] and were handed out even to households who already had one. The devices were distributed in early

[4]The devices cost about $25 each. Considering an average monthly household expenditure of $165 at baseline, they represented about 15% of their monthly expenditure or 1.3% of their annual expenditure.

Figure 3.3: Location of Markets, Treatment and Control Villages

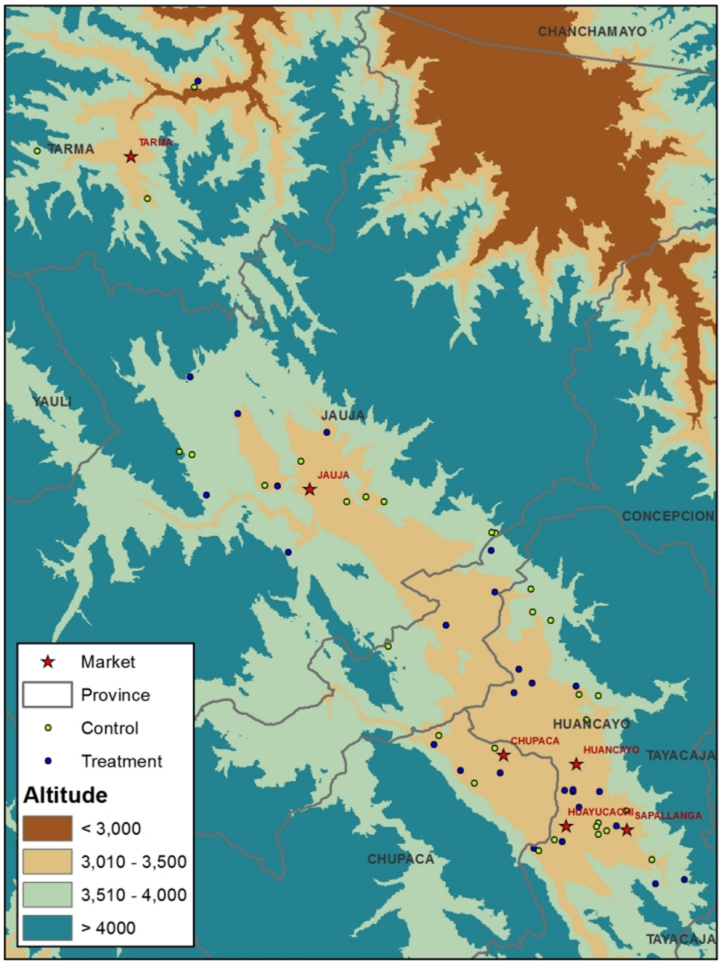


Table 3.1: Calendar of Price Distribution by Permanent Markets and Ferias

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Permanent							
Huancayo	X		X		X		
Tarma	X		X		X		
Jauja			X				X
Ferias							
Chupaca						X	
Huayucachi	X						
Zapallanga				X			

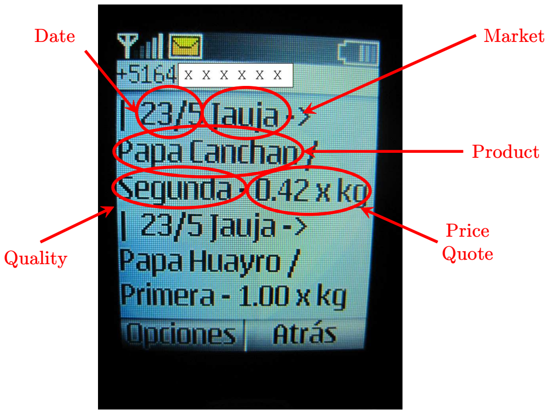
April, during the early harvest. For four months (mid-April to mid-August), a team of undergraduate students collected price information of 17 different products (by quality): peas, lima beans, barley, four types of corn, two types of *olluco*, and eight types of potato. The information was gathered in three permanent markets (Huancayo, Jauja and Tarma) and three weekly *ferias* (Chupaca, Huayucachi and Zapallanga). The calendar of price distribution is presented in Table 3.1. Once the information was collected, it was compared with the list of products that households planted for the 2009/2010 season according to the baseline information. During the same morning, only the information of the relevant products for each participant was sent through SMS to the number of the cell phone the intervention provided. An example of a text message with price information is presented in Figure 3.4a. The text message included the date, market, product, quality, and price quote.

I tried to ensure that participants understood the information they were being sent. Along with the devices, the participants were provided with two manuals. The first one explained how to use the cell phone[5]. The second one had explanations on the price infor-

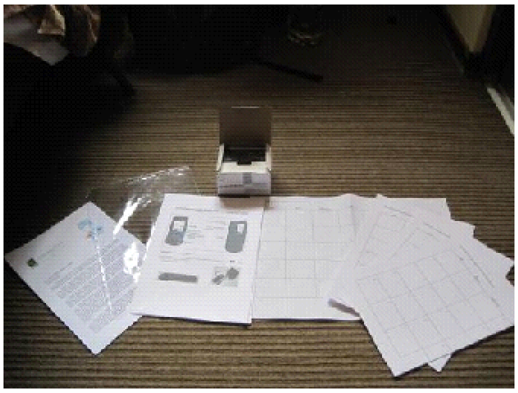
[5]Beneficiaries were expected to be able to use a cell phone, either by themselves or to have someone else

Figure 3.4: Cell Phone and Price Distribution

(a) Example of Price SMS



(b) Distribution kit: Cell Phones, Manuals and Charts



mation that would be sent out. It included a calendar with the weekdays in which information for each market would be distributed and detailed instructions on how to read the text messages with the prices. They also received a chart to help them keep track of the prices they received (Figure 3.4b). The team went through the manuals with each participant and answered any questions or doubts they had.

The participants were informed of an important service restriction: during the first few months (until late August), their mobiles would only receive calls and text messages from a number authorized by the project. Through this restriction, I can rule out any other potential uses of the mobile phone that could drive the results (i.e. communication with input providers, collusion with other producers, coordination with traders, etc.). In this way, the treatment does not encompass the full advantages of a mobile phone, but only

in the household help them. However, just in case, they were also provided with a manual - with pictures and detailed instructions - of their basic functions (how to charge them, how to know if there are any new text messages, how to open them, etc.).

being able to receive price information in different markets. Participants were also required to answer periodic calls to check if there were any problems with the devices, whether the price SMS were being delivered appropriately, and whether they had any problems reading the information. All in all, besides being able to receive periodic check-up calls, these devices did not have any capabilities beyond those of a pager during the intervention period. However, after August, full capabilities of the cell phones would be restored and they would operate as regular pre-paid phones. Participants were told they would be able to keep the devices without any further obligation. These phones were distributed to all selected households, even to those who already owned one. No one who was offered a cell phone declined to participate in the project.

In September 2010, a follow-up survey was conducted. The questionnaire included information about production, sales volumes and prices in the 2009/2010 agricultural season. This provides me with a panel of households, where I can compare the outcomes of the 2008/2009 (before the intervention) and 2009/2010 agricultural season (after the intervention) among those who received the intervention *via-à-vis* those who did not.

Chapter 4

A Model of Price Bargaining under Asymmetric Information

This chapter presents a simple theoretical model to investigate the role that information deficiencies can play in agricultural marketing decisions. A farmer harvests $\overline{Q}$ units of his crop, and has to negotiate with a trader the quantity and payment involved in a sales transaction. At the time of the negotiation, they both face uncertain market prices for an agricultural product. For simplicity, assume there are two possible states of nature: the market price is either high (p_H) with probability λ or low (p_L) with probability $(1-\lambda)$, with $0 < \lambda < 1$, $p_H > p_L$. Before the market price is unveiled both parties establish a contract that determines the sales quantity (s_i) and total payment (Y_i) for each state of nature, where $i = \{H, L\}$. Assume that the farmer offers a contract to the trader, establishing combinations Y_H, s_H (if the market price is high) and Y_L, s_L (if the market price is low). The trader can either accept or reject it. If he rejects the contract, there is no sale. If he does accept it, the parties verify if the state was H or L, and the corresponding combination is enforced.

The farmer is paid Y_i, supplies the trader with s_i units of his crop, and keeps $\overline{Q} - s_i$ units for self-consumption. The trader sells the s_i units at market price p_i. Suppose that the farmer's utility function is $\frac{Y_i^{1-\beta}}{(1-\beta)} + a(\overline{Q} - s_i)$: he exhibits constant relative risk aversion over his monetary income (where β is the parameter of relative risk-aversion, $0 < \beta < 1$) and a constant marginal utility from self-consumption. The trader sells the s_i units he bought from the farmer at market price p_i. Assume the trader is risk neutral. His profits are given by $p_i s_i - Y_i$.

Note that the contract is based on the ability of the farmer and trader to observe the actual realization of p_i. I examine the solutions in two information settings. In the first one, both the trader and the farmer observe p_i, providing a benchmark case with symmetric information. These results are compared to a second model with asymmetric information, where farmers are uninformed but traders do know the market price. In this second case, the farmers have to rely on the traders to reveal the realized value of p_i.

4.1 Symmetric Information

First, I consider the case where both the farmer and the trader observe the market prices after establishing the contract. The farmer's objective is to maximize his utility, subject to the trader's individual rationality constraints (which should bind in each revealed state).

$$\underset{s_H,Y_H,s_L,Y_L}{MAX}\ \lambda\left[\frac{Y_H^{1-\beta}}{(1-\beta)}+a(\overline{Q}-s_H)\right]+(1-\lambda)\left[\frac{Y_L^{1-\beta}}{(1-\beta)}+a(\overline{Q}-s_L)\right] \tag{4.1a}$$

$$s.t.\ p_H s_H - Y_H \geq 0 \tag{4.1b}$$

$$p_L s_L - Y_L \geq 0 \tag{4.1c}$$

In this case, it is straightforward to see that the farmer will push the trader to his reservation utility in both market price scenarios, so constraints (4.1b) and (4.1c) bind with equality. The farmer's optimal contract is given by:

$$s_H^{SI} = p_H^{\frac{1-\beta}{\beta}}\frac{1}{a^{1/\beta}};\quad Y_H^{SI} = p_H^{1/\beta}\frac{1}{a^{1/\beta}} \qquad \text{if the price is high} \tag{4.2a}$$

$$s_L^{SI} = p_L^{\frac{1-\beta}{\beta}}\frac{1}{a^{1/\beta}};\quad Y_L^{SI} = p_L^{1/\beta}\frac{1}{a^{1/\beta}} \qquad \text{if the price is low} \tag{4.2b}$$

In this case, the implicit farm-gate prices in the contract ($r_i^{SI} = \frac{Y_i^{SI}}{s_i^{SI}} = p_i$ for $i = L,H$) are precisely those prevailing in the market.

4.2 Asymmetric Information

Now suppose that there is asymmetric information. The contract is established before the market price is known by the agents. However, once the market price is unveiled, only the trader can observe it and the farmer has to rely on what the trader reports to him. The farmer knows that the trader has an incentive to cheat: if the realized price turns out to be high, he will lie and tell him it was low. Thus, the farmer's objective is to establish a contract that encourages the trader to reveal the state of nature truthfully.

The farmer maximizes his expected utility

$$\underset{Y_H, Y_L, s_H, s_L}{MAX} \quad \lambda \left[\frac{Y_H^{1-\beta}}{(1-\beta)} + a(\overline{Q} - s_H) \right] + (1-\lambda) \left[\frac{Y_L^{1-\beta}}{(1-\beta)} + a(\overline{Q} - s_L) \right] \tag{4.3a}$$

subject to Individual Rationality (IR) constraints that ensure that the trader is provided with his reservation utility under both states of nature (so he would be willing to accept the contract):

$$p_H s_H - Y_H \geq 0 \tag{4.3b}$$

$$p_L s_L - Y_L \geq 0 \tag{4.3c}$$

and the following Incentive Compatibility (IC) constraints that incentivize the trader to reveal the true market prices:

$$p_H s_H - Y_H \geq p_H s_L - Y_L \tag{4.3d}$$

$$p_L s_L - Y_L \geq p_L s_H - Y_H \tag{4.3e}$$

IC constraint (4.3d) states that if p_H is the prevailing market price, the trader is better off revealing the true outcome (and enforcing combination s_H, Y_H) rather than cheating (and enforcing combination s_L, Y_L). Constraint (4.3e) works analogously for low market prices. In an optimum, (4.3c) and (4.3d) should bind with equality, while (4.3b) and (4.3e) are slack conditions of the problem[1].

[1]Note that, constraints (4.3d) and (4.3c) imply that: $p_H s_H - Y_H \geq p_H s_L - Y_L \geq p_L s_L - Y_L \geq 0$, so (4.3b)

The optimal contract establishes the following quantity (q), payment (Y) and implicit prices ($r = \frac{Y}{q}$) if the trader declares low market prices:

$$
\begin{aligned}
s_L^{AI} &= p_L^{\frac{1-\beta}{\beta}} \frac{1}{a^{1/\beta}} \left[\frac{(1-\lambda)p_H}{P_H - \lambda p_L} \right]^{1/\beta} \leq p_L^{\frac{1-\beta}{\beta}} \frac{1}{a^{1/\beta}} = s_L^{SI} && \text{(4.4a)} \\
Y_L^{AI} &= p_L^{1/\beta} \frac{1}{a^{1/\beta}} \left[\frac{(1-\lambda)p_H}{P_H - \lambda p_L} \right]^{1/\beta} \leq p_L^{\frac{1}{\beta}} \frac{1}{a^{1/\beta}} = Y_L^{SI} && \text{(4.4b)} \\
r_L^{AI} &= p_L = r_L^{SI} && \text{(4.4c)}
\end{aligned}
$$

Analogously, the optimal contract establishes the following outcomes if the trader reveals a high price:

$$
\begin{aligned}
s_H^{AI} &= p_H^{\frac{1-\beta}{\beta}} \frac{1}{a^{1/\beta}} \left[1 + \frac{(p_H - p_L)}{p_H} \left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L} \right)^{1/\beta} \right] \geq p_H^{\frac{1-\beta}{\beta}} \frac{1}{a^{1/\beta}} = s_H^{SI} && \text{(4.4d)} \\
Y_H^{AI} &= p_H^{\frac{1}{\beta}} \frac{1}{a^{1/\beta}} = Y_H^{SI} && \text{(4.4e)} \\
r_H^{AI} &= \frac{p_H}{1 + \frac{(p_H - p_L)}{p_H} \left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L} \right)^{1/\beta}} \leq p_H = r_H^{SI} && \text{(4.4f)}
\end{aligned}
$$

Under the optimal contract with asymmetric information, the farmer uses the implicit farm-gate prices and quantities as instruments to find out the true state of nature. If prices are high, the trader gets a price premium: while $Y_H^{AI} = Y_H^{SI}$, the farmer increases the sales quantity leading to a lower farm-gate price (i.e. $r_H^{AI} \leq r_L^{AI}$). These informational rents induce the trader to reveal that the market price is high. In contrast, when market prices are low, the trader cannot exploit any informational rents: the farm gate price remains p_L.

is redundant. To solve the problem, I initially solve the problem ignoring (4.3e). It can be shown later that an optimal solution complies with this constraint.

However, the farmer reduces the quantity he sells under asymmetric information. If the trader wants to lie and claim that prices are low (when they are actually high), the farmer limits his supply to reduce the trader's profits, reducing his incentives to cheat.

In these lines, the comparison of both models predicts that farm-gate prices for farmers would be higher under symmetric than asymmetric information. Solving for the expected farm-gate price $E[r^{SI} - r^{AI}]$ yields:

$$\begin{aligned} E[r^{SI} - r^{AI}] &= \lambda p_H + (1-\lambda)p_L - \left[\lambda r_H^{AI} + (1-\lambda) r_L^{AI}\right] \\ &= \lambda \left[p_H - \frac{p_H}{1 + \frac{(p_H - p_L)}{p_L}\left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L}\right)^{1/\beta}} \right] \geq 0 \end{aligned} \tag{4.5}$$

Also note that the difference between r^{SI} and r^{AI} depends on the degree of risk aversion of the farmer. Note that $\frac{(1-\lambda)p_L}{p_H - \lambda p_L} < 1$, so $E[r^{SI} - r^{AI}]$ shrinks as β decreases. For example, if the farmer is risk neutral ($\beta \to 0$), then $E[r^{SI} - r^{AI}] \to 0$. Intuitively, this implies that more-risk averse farmers are willing to offer larger premia to insure from price uncertainty and for the traders to reveal the actual realization of market prices. In terms of the model, this is equivalent to differentiating (4.5) with respect to β.

$$\begin{aligned} \frac{dE[r^{SI} - r^{AI}]}{d\beta} &= -\frac{1}{\beta^2} \lambda p_H \frac{(p_H - p_L)^2}{p_L^2} \frac{1}{\left[1 + \frac{(p_H - p_L)}{p_L}\left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L}\right)^{1/\beta}\right]^2} \\ & \quad \left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L}\right)^{1/\beta} Ln\left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L}\right) \geq 0 \end{aligned} \tag{4.6}$$

In terms of sales volumes, the predictions of the model are more ambiguous. In particular, under asymmetric information, the farmer would offer larger sales volumes if the trader reports high prices, but smaller volumes otherwise. The impact on the expected sales quantity $E[s^{SI} - s^{AI}]$ will depend on the difference between p_H and p_L, and the beliefs about the price distribution (λ).

$$\lambda s_H^{SI} + (1-\lambda)s_L^{SI} - \left[\lambda s_H^{AI} + (1-\lambda)s_L^{AI}\right] =$$
$$\frac{1}{a^{1/\beta}} \left\{ \underbrace{-\lambda p_H^{\frac{1-\beta}{\beta}} \left[\frac{(p_H - p_{L)}}{p_H} \left(\frac{(1-\lambda)p_L}{p_H - \lambda p_L} \right)^{1/\beta} \right]}_{<0} + \underbrace{(1-\lambda)p_L^{\frac{1-\beta}{\beta}} \left[1 - \left(\frac{(1-\lambda)p_H}{P_H - \lambda p_L} \right)^{1/\beta} \right]}_{>0} \right\} \tag{4.7}$$

So far, I have assumed that there are no restrictions on the quantities that farmers can allocate to self-consumption. This makes sense when products can be temporarily stored or can be transformed into by-products with larger shelf-life[2]. However, product perishability of other crops (e.g. produce) imposes natural limits on how much can be destined to self-consumption. In this line, consider the following additional restrictions to the farmers' optimizations in (4.1a) and (4.3a):

$$\bar{Q} - s_H \leq c \tag{4.8a}$$

$$\bar{Q} - s_L \leq c \tag{4.8b}$$

where c is the maximum amount that a farmer can self-consume of his crop. These restric-

[2]For example, potatoes can be transformed into a flour (know as *chuño* or *tunta*) by exposing them to freezing temperatures overnight and to intense sunlight during the day. Olluco and corn can also be dried for later consumption.

tions entail three possibilities. The first one is that neither (4.8a) nor (4.8b) is binding. In this case, the farmer's choice would be given by (4.2a)-(4.2b) if there is symmetric information and by (4.4a)-(4.4f) if there is asymmetric information.

The second possibility is that (4.8b) is binding, while (4.8a) is not (which is more likely if the difference between p_H and p_L is relatively large)[3]. Denote $\{s_L^{AI'}, Y_L^{AI'}, r_L^{AI'}\}$, $\{s_H^{AI'}, Y_H^{AI'}, r_H^{AI'}\}$ as the optimal sales volumes, payments and per-unit prices in this situation. In this case, the contract involves:

$$s_L^{AI'} = \bar{Q} - c \tag{4.9a}$$

$$Y_L^{AI'} = p_L(\bar{Q} - c) \tag{4.9b}$$

$$r_L^{AI'} = p_L \tag{4.9c}$$

$$q_H^{AI'} = p_H^{\frac{1-\beta}{\beta}} \frac{1}{a^{1/\beta}} + (\bar{Q} - c)\frac{(p_H - p_L)}{p_H} \tag{4.9d}$$

$$Y_H^{AI'} = p_H^{\frac{1}{\beta}} \frac{1}{a^{1/\beta}} \tag{4.9e}$$

$$r_H^{AI'} = \frac{p_H}{1 + (a/p_H)^{1/\beta}(\bar{Q} - c)(p_H - p_L)} \tag{4.9f}$$

Note that without restrictions on self-consumption, the farmer was able to use two instruments to have the trader truthfully reveal the market prices: on the one hand, he would threat to cut back his supply if the trader revealed low prices and, on the other, he would reward the trader with lower per-unit farm-gate prices for his purchase if he disclosed high market prices. However, when $\bar{Q} - s_L \leq c$ is binding, he cannot cut back his sales volumes:

[3]Note that restriction (4.8a) cannot bind if (4.8b) does not.

even when he would like to supply less than $s_L^{AI'} = \bar{Q} - c$, he cannot keep large volumes for self-consumption before the crop goes bad. Thus, he can only provide the trader with an information price premium for him to reveal if market prices are high ($r_H^{AI'} < r_H^{AI}$).

The third possibility is that both (4.8b) and (4.8a) are binding. In this case, the farmers supply is completely restricted by his self-consumption restrictions, such that $s_L^{AI''} = s_H^{AI''} = (\bar{Q} - c)$, $Y_L^{AI''} = Y_H^{AI''} = p_L(\bar{Q} - c)$, and $r_H^{AI''} = r_L^{AI''} = p_L$. Thus, he cannot alter his supply at either price level and is unable to provide any incentives to the trader to reveal the actual market prices.

4.3 Discussion of Model Predictions

Overall, these results highlight the differences in the farmer's marketing outcomes under both information settings. Under symmetric information, the optimal quantities are traded and the farmer sells his production for the actual market prices. However, when there is asymmetric information, the trader has an incentive to cheat by telling the farmer that market prices are low when they are actually high. There are two (costly) mechanisms for the farmer to elicit this information: he offers the trader with informational rents (through lower farm-gate prices) when market prices are high, and he can alter the quantities he supplies depending on the trader's revelation of the state of nature. While the model leads to ambiguous predictions regarding sales volumes, it shows that farm-gate prices should increase when the farmer is supplied with market information. These increases should be larger for more risk-averse farmers, who are willing to offer larger informational premia in order to cope with the price uncertainty. In addition, perishability limits the farmer's ability to restrict the quantity he offers to the trader. Thus, the impact of market price information should be larger for perishable products but sales volumes should not increase as much as with perishable products. The following chapter provides an empirical analysis of these predictions.

Chapter 5

Price Information and Farmers' Marketing Outcomes: An Empirical Application for Peru

This chapter analyzes the impact of the intervention described in 3. In particular, I analyze whether the direct provision of price information through SMS increases farmers' sales prices, market participation, and sales volumes. For this purpose, I compare the changes in marketing outcomes between the 2008/2009 and 2009/2010 agricultural seasons of those who received the SMS directly to those in control villages (where no one received the SMS). I also investigate heterogeneous treatment effects by product perishability, the degree of risk aversion, and previous cell phone ownership. Additionally, I test for the presence of information spillover effects within villages by comparing the outcomes of households in control villages with those of farmers those who did not receive the SMS but lived in a village where someone else did.

Throughout the analysis, consider the definitions of the following variables:

- Info takes a value of 1 if the household is in a treated village and received the price SMS. It takes a value of zero otherwise.
- Spill takes a value of 1 if the household is in a treated village but did not receive the price SMS (i.e. excludes *Info=1*) . It takes a value of zero otherwise.

The remaining households (i.e. those with Info=0 and Spill=0) are those in control villages.

5.1 Baseline Comparisons

In this section, I show that the randomization process delivered three similar groups: those who directly received price information, those who lived in treated villages but did not receive information directly, and those in control villages. I compare the baseline characteristics of those who received information and those in the spillover group with respect to the control group, with the following Ordinary Least Squares (OLS) Regression:

$$Y_{i0} = \alpha_0 + \alpha_1 \text{Info}_i + \alpha_2 \text{Spill}_i + \mu_i \tag{5.1}$$

where Y_{i0} is a characteristic of the ith household before the intervention and μ_i is a zero-mean household-specific error term. The coefficients α_1 and α_2 provide estimates of the differences in Y_{i0} of the Info and Spill groups relative to the control group. Sample means of the information, spillover and control groups — as well as estimates for Equation (5.1) — are presented in Table 5.1. The sample is relatively well balanced in terms of characteristics of the household head (age; gender; years and level of education), land, household expenditure, and cell phone ownership (prior to the intervention).

I also analyze the crop distribution in the information, spillover and control groups. Table 5.2 compares the proportion of households that cultivated seventeen important crops during the 2008/2009 (baseline) agricultural season. The first three columns present the proportion of households that grew each crop in each of the three groups. The last two columns report the differences of the information and spillover groups, relative to the controls. The approach to estimate the differences among groups in the proportion of households growing is similar to that in Equation (5.1). However, because this variable is binary, I estimate marginal effects from a probit model rather than OLS. The sample is not balanced across all seventeen crops: arguably it would be difficult to achieve the same composition across such a large number of crops[1]. However, the differences in the share of households growing each product are small and only significant for two products (one variety of olluco and one of potato). In any case, all the subsequent analysis will include crop controls.

Next, I present the baseline differences in production, sales volumes and prices among the three groups of interest. Because the sample is not stratified by crop, I cannot draw any inferences from a specific agricultural product. As a matter of fact, if I were to restrict my sample to households who produced the most popular crop in the region (Yungay potato), the sample size would drop by more than half. Thus, I use the full sample to estimate average differences. Because estimations with the full sample would entail comparing the

[1] Morgan & Rubin (2012) shows that, given k variables over which we check balance, the chance of finding a statistical difference between two groups at significance α is $1-(1-\alpha)^k$. Therefore, if $\alpha = 0.1$ and we test 34 hypotheses, the probability that we reject at least one of them is 0.97.

Table 5.1: Household Characteristics in Baseline

	Info	Spill	Control	Diff[1]	
	(I)	(S)	(C)	(I)-(C)	(S)-(C)
HH Head Characteristics					
Age	50.53	51.45	49.91	0.62	1.54
	(12.81)	(15.62)	(14.54)	(1.69)	(1.77)
Head is male	0.86	0.80	0.84	0.02	-0.05
	(0.34)	(0.40)	(0.37)	(0.04)	(0.03)
Years of education	7.45	6.89	7.51	-0.06	-0.62
	(3.92)	(4.14)	(4.03)	(0.50)	(0.44)
Primary [2]	0.45	0.46	0.45	0.00	0.01
	(0.50)	(0.50)	(0.50)	(0.05)	(0.05)
Secondary [2]	0.42	0.37	0.40	0.02	-0.03
	(0.50)	(0.48)	(0.49)	(0.06)	(0.05)
Technical [2]	0.06	0.06	0.06	0.00	0.00
	(0.24)	(0.23)	(0.23)	(0.03)	(0.02)
College [2]	0.04	0.03	0.05	-0.01	-0.02
	(0.19)	(0.18)	(0.22)	(0.02)	(0.02)
Any member has Cell Phone [2]	0.46	0.50	0.51	-0.05	-0.01
	(0.50)	(0.50)	(0.50)	(0.07)	(0.06)
Log PC HH Exp	4.69	4.61	4.70	-0.01	-0.09
	(0.48)	(0.49)	(0.45)	(0.08)	(0.06)
Log Land	8.37	8.17	8.32	0.05	-0.15
	(1.36)	(1.50)	(1.50)	(0.36)	(0.36)
Has land with irrigation [2 3]	0.28	0.29	0.26	0.02	0.03
	(0.45)	(0.45)	(0.44)	(0.11)	(0.10)
N	111	299	380		

[1] For the first three columns, the means and standard deviations of each variable in the information, spillover and control groups are reported. In the last two columns, the differences were calculated using the following regression: $Y_i = \alpha_1 Info_i + \alpha_2 Spill_i + \mu_i$. Regression standard errors are reported in parentheses.
[2] In the case of discrete variables the linear regression was replaced for a probit model.
[3] The variable takes a value of one if the household has at least one plot with irrigation.

Significance levels of the differences between the treatment and spillover groups (with respect to the control group) denoted by: *** 1%, ** 5%, * 10% .

Table 5.2: Crop Composition in Baseline

	Treat	Spill	Control	Difference[1]	
	(T)	(S)	(C)	(T)-(C)	(S)-(C)
Peas	0.16	0.12	0.16	0.00	-0.04
	(0.37)	(0.33)	(0.24)	(0.08)	(0.06)
Barley (common)	0.30	0.24	0.23	0.06	0.00
	(0.46)	(0.43)	(0.42)	(0.08)	(0.08)
Lima Beans	0.12	0.09	0.17	-0.06	-0.08
	(0.32)	(0.29)	(0.38)	(0.08)	(0.08)
Corn - White	0.42	0.37	0.26	0.17	0.12
	(0.50)	(0.48)	(0.44)	(0.12)	(0.13)
Corn - Cusqueado	0.03	0.03	0.03	0.00	0.00
	(0.16)	(0.17)	(0.17)	(0.02)	(0.02)
Corn - Cusqueno	0.05	0.02	0.01	0.03	0.01
	(0.21)	(0.15)	(0.10)	(0.02)	(0.01)
Corn - San Jeronimo	0.04	0.04	0.02	0.01	0.01
	(0.19)	(0.19)	(0.15)	(0.02)	(0.02)
Olluco - Yellow	0.07	0.06	0.09	-0.02	-0.04
	(0.26)	(0.23)	(0.29)	(0.04)	(0.04)
Olluco - Dotted	0.03	0.01	0.04	-0.02	-0.03
	(0.16)	(0.12)	(0.21)	(0.02)	(0.02)
Potato - Yellow	0.03	0.02	0.01	0.01	0.01
	(0.16)	(0.15)	(0.11)	(0.02)	(0.02)
Potato - Andean	0.02	0.05	0.03	-0.01	0.02
	(0.13)	(0.21)	(0.17)	(0.03)	(0.04)
Potato - Canchan	0.07	0.03	0.07	0.01	-0.03
	(0.26)	(0.18)	(0.25)	(0.03)	(0.02)*
Potato - Huayro	0.02	0.03	0.02	0.00	0.01
	(0.13)	(0.18)	(0.14)	(0.02)	(0.02)
Potato - Perricholi	0.25	0.21	0.35	-0.10	-0.14
	(0.44)	(0.41)	(0.48)	(0.15)	(0.14)
Potato - Peruanita	0.05	0.04	0.01	0.04	0.03
	(0.23)	(0.20)	(0.11)	(0.03)	(0.03)
Potato - Unica	0.01	0.00	0.04	-0.03	-0.03
	(0.09)	(0.06)	(0.19)	(0.02)	(0.02)*
Potato - Yungay	0.41	0.47	0.45	-0.04	0.02
	(0.49)	(0.50)	(0.50)	(0.09)	(0.10)
N	111	299	380		

[1] For the first three columns, the proportion of households that grew each crop is reported (standard deviation in parentheses). In the last two columns, the differences were calculated using a probit model: $Prob[Crop_{ic} = 1] = \Phi(\alpha_1 Info_i + \alpha_2 Spill_i)$ for each crop c. Regression standard errors are reported in parentheses.

Significance levels of the differences between the treatment and spillover groups (with respect to the control group) denoted by: *** 1%, ** 5%, * 10% .

prices of households that grow low-value potatoes with higher-value produce, I include crop controls. Thus, I estimate the following regression:

$$Y_{ic0} = \alpha_0 + \alpha_1 \text{Info}_i + \alpha_2 \text{Spill}_i + \delta_c D_c + \varepsilon_i + \mu_{ic} \tag{5.2}$$

where Y_{ic0} is the marketing outcome (production, probability of sales, volume of sales and price) of the ith household in the baseline (2008/2009 season) for crop c and D_c is an indicator variable for each crop. The equation allows for correlation of error terms within the same household (across crops) through ε_i, which is distributed i.i.d. with zero-mean and variance σ_ε^2. It also includes a zero-mean error term μ_{ic} that varies by household and crop. The results for Equation (5.2) are presented in Table (5.3). For each outcome, the first column reports estimates using crop controls. The second one includes the same crop controls with additional quality controls[2]. Overall, they show that households did not exhibit significant differences among treatment statuses before the intervention.

5.2 The Effect of Information on Agricultural Prices

I calculate the impact of the treatment on agricultural prices through a Difference-in-Differences (DID) model, including crop (and quality) controls and random effects at the

[2]Note that about 16% of the observations drop out of the production regression when we control for quality. This is because farmers do not necessarily sort all their harvest. Production that is sold is necessarily graded by quality. However, households who do not sell (i.e. allocate their harvests to self-consumption, seed, by-products, etc.) do not necessarily do so.

Table 5.3: Agricultural Production and Sales Comparison in Baseline

	Log Production		Prob Sales[1]		Log Sales Volume		Log Price	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Info	0.10	0.10	0.02	0.02	0.07	0.05	-0.02	-0.01
	(0.12)	(0.13)	(0.04)	(0.04)	(0.15)	(0.16)	(0.04)	(0.04)
Spill	-0.06	0.04	-0.01	0.02	0.02	-0.00	0.01	0.00
	(0.09)	(0.10)	(0.03)	(0.03)	(0.11)	(0.12)	(0.03)	(0.03)
Constant	4.72	5.37	0.47	0.81	5.56	5.88	-0.15	-0.01
	(0.13)***	(0.14)***	(0.05)***	(0.06)***	(0.19)***	(0.18)***	(0.07)**	(0.06)
Product Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quality Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2426	2059	2426	2059	1048	1046	1048	1046
Households	790	700	790	700	500	500	500	500

[1] This variable takes a value of one if the household has sold at least some of his harvest of a certain product. Marginal effects calculated from a Probit Model.

Significance levels denoted by: *** 1%, ** 5%, * 10% .

household level. Namely, I estimate the following regression:

$$log(P_{ict}) = \alpha\text{Info}_i + \theta\text{Spill}_i + \gamma t + \beta_1\text{Info}_i t + \beta_2\text{Spill}_i t + \delta_c\text{D}_c + \varepsilon_i + \mu_{ict} \tag{5.3}$$

where P_{ict} is the net price at which household i sells crop c in period t. If the household sells this crop to a middleman, P_{ict} is the gross sales price. If he sells this crop directly in the market, I subtract the transportation cost from the gross price. The variable t takes a value of zero for for the 2008/2009 season (before the treatment) and a value of one for the 2009/2010 season (after the treatment). Info_i and Spill_i are the (time-invariant) treatment statuses for each household. D_c are indicator variables for crop c. Additionally, the error term has two components. The first one ε_i accounts for the fact that the errors within the same household are not independent from one another (ε_i is i.i.d with mean zero and variance σ_ε^2). The second one (μ_{ict}) is a zero-mean idiosyncratic error that varies across households, crops and time. For consistency, this specification requires ε_i to be uncorrelated with other explanatory variables. This is a plausible assumption in this setting because of the random assignment of the treatment[3]. Additionally, standard errors are clustered at the village level to allow for any covariate shocks.

[3]Fixed-effects (FE) estimates exploit the variability of sales prices within the household across both periods of time. The nature of the data does not seem suitable for this type of regression for two reasons. First, households vary the crops they harvest between periods. Changes in crops are not systematically correlated with the treatment (by design of the field experiment and as shown in Table 5.9). However, a FE regression would exploit variation in prices of potentially very different crops within individuals (and not averages). Second, households do not necessarily sell in both periods of the data. A FE regression would only take into account households that have sold their production in both periods. Because this group of households is likely to have more marketing experience, a FE specification would overestimate the impact of the treatment (see Table 5.5). Because the treatment is randomly assigned, I prefer a Random-Effects specification. However, households' decision whether to sell or not their harvests still imposes a selection problem. I will discuss this problem later, and calculate sample selection bounds under alternative assumptions.

Table 5.4: DID Estimation for Prices

	(1)	(2)
Info	0.00	-0.02
	(0.076)	(0.064)
Spill	0.04	0.03
	(0.069)	(0.062)
t	0.13**	0.15***
	(0.056)	(0.058)
Info x t	0.13*	0.14*
	(0.076)	(0.085)
Spill x t	-0.01	-0.02
	(0.064)	(0.069)
Constant	-0.10**	0.02
	(0.050)	(0.047)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	2,125	2,111
Number of households	601	600

Regressions include household random effects. Standard errors are clustered at the village level.

Significance levels denoted by: *** 1%, ** 5%, * 10%.

In this framework, β_1 captures the average effect of information over all crops for which I observe any sales. β_2 provides an analogous estimate for those who may have benefited from information spillovers, relative to the control group. The results of this estimation are reported in Table (5.4). They suggest that there were sizable impacts for those who benefited directly from the information: prices at which they were able to sell their production increased by 13% (with crop controls) to 14% (with crop and quality controls). The results show little evidence of spillover effects at treated villages: the estimates for β_2 are small and statistically insignificant.

Estimations in Table (5.4) include all observations for which there are price data. However, I only observe households' prices for crops that were sold. For example, if a farmer harvested a crop for sole subsistence (self-consumption) purposes, I cannot observe the outcome variable. Due to random assignment, the treatment is not correlated with the households' decision whether to sell or not in the baseline (as shown in Table 5.3). However, as I will show later, the information did make households more likely to sell in the follow-up survey. This induces a sample selection bias in my results.

The most direct approach to account for sample selection would be to estimate a system of equations, where one of them determines the decision to sell and the other would measure the impact of information on prices. However, I would need to determine at least one variable that influences households' decision to sell but does not impact directly its sales price (only through its sales decision). Because sales decisions are arguably based on the potential price that a household might be able to charge, I do not have any variables that would credibly meet this exclusion restriction. Thus, rather than estimating effects corrected for this selection bias, I try to sign the bias and to construct some bounds for the treatment effects.

In general, the sample selection issue is created by the subsample of households in the treatment group who decide to sell their harvests, but would not have chosen to do so in the absence of the treatment. Intuitively, these households probably have less marketing experience than the group who would have sold regardless of the intervention. Thus, I will argue that, if anything, these households are less likely to have fully benefited from the treatment and sell their production at lower prices than the "always sellers". Ideally, I would like to observe the sales history of the households in a time series to discern new from regular sellers, but I do not have this information and will rely on some proxies.

I implement alternative approaches. The simplest one is to estimate equation (5.3) with

Table 5.5: Price Regression for
Households with Sales in Both Periods[1]

	(1)	(2)
Info	-0.05	-0.04
	(0.069)	(0.055)
Spill	0.05	0.04
	(0.063)	(0.057)
t	0.10*	0.13**
	(0.051)	(0.057)
Info x t	0.19***	0.19**
	(0.071)	(0.087)
Spill x t	-0.00	-0.01
	(0.064)	(0.077)
Constant	-0.08*	0.04
	(0.049)	(0.045)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	1,579	1,567
Households	311	311

[1] Includes households who sold in both periods, regardless of the product and quality.
All regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

a sample of households that sold (at least one product) in both periods. If households with commercial activity in the baseline (before the intervention) are believed to be more likely to sell even in the absence of the treatment, this might be indicative of the impact of "always sellers". The results are reported in the first two columns of Table 5.5. Information seem to have had a larger impact on those with sales in both periods (around 19%). The larger coefficient for the treatment variable suggests that the effect on this group might be larger than the one for new sellers. The estimates for the spillover groups remain small and indistinguishable from zero.

In a second estimation, I use the sales information in the follow-up survey and construct two subsamples: one with households who sold (at least one product) in both periods and another one with those who only reported any sales in the follow-up. Here, I take the estimates of the first group as a proxy to the impact on the "always sellers" and the second one as the effect on new sellers (those more likely to sell only because they were encouraged by the treatment). I estimate the following regression for all households in the follow-up and, then, for both of the subsamples described:

$$log(P_{ic,t=1}) = \beta_1 \mathrm{Info}_i + \beta_2 \mathrm{Spill}_i + \delta_c D_c + \varepsilon_i + \mu_{ic} \tag{5.4}$$

These results are presented in Table (5.6). The first two columns just confirm that the results using all the observations in the follow-up survey are almost identical to those estimated with the panel dataset. The effect is 13% for households in the information group and are close to zero for those in the spillover group. This is not surprising because prices were not significantly different among groups in the baseline. The subsequent columns show the differences in the price effect for the group who sold in both periods and the one who only sold after the treatment. The effect for those who sold in both periods is 14%-15%, but the one for new sellers is considerably lower (2%-4%) and statistically insignificant. If the treatment is differentially attracting more of the "new sellers", this would suggest that, if anything, my previous estimates are downwardly biased and underestimate the overall effect of the information.

I also estimate bounds for the impact of the treatment effect. Following Horowitz & Manski (2000), I assume different potential values for the prices of the crops that were not sold. First, I simulate what would have happened if non-sellers would not have benefited

Table 5.6: Cross Sectional (Follow-up) Regression for Prices

	All[1]		Both Periods[2]		Follow-up Only[3]	
	(1)	(2)	(3)	(4)	(5)	(6)
Info	0.13***	0.13***	0.14***	0.15**	0.04	0.02
	(0.04)	(0.05)	(0.05)	(0.06)	(0.11)	(0.08)
Spill	0.03	0.01	0.05	0.04	-0.06	-0.07
	(0.04)	(0.03)	(0.04)	(0.04)	(0.10)	(0.08)
Constant	0.19***	0.35***	0.19***	0.32***	0.23**	0.42***
	(0.04)	(0.05)	(0.04)	(0.06)	(0.09)	(0.10)
Product Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Quality Dummies	No	Yes	No	Yes	No	Yes
Observations	1,075	1,063	876	866	199	197
Number of households	411	406	311	307	100	99

[1] Includes all households who sold at least part of the production in the follow-up surveyWhile c, regardless of their sales in the baseline.
[2] Includes households who sold at least part of the production of any of their crops both in the baseline and the follow-up survey.
[3] Includes households who did not sell any of the production of their crops in the baseline, but did in the follow-up survey.

Significance levels denoted by: *** 1%, ** 5%, * 10%. Regressions include household random effects. Standard errors are clustered at the village level.

at all from the information intervention. To assess this possibility, I calculate the average sales price for each crop in the control group in the follow-up survey, and impute this value to all unsold crops in both the information and spillover groups. With this imputed data, I estimate Equation (5.4). Effectively, this procedure averages the positive effect of the intervention among those who sold their crops with a zero-effect on those who did not sell their harvests. Column (1) of Table (5.7) shows the result of this exercise: even under this pessimistic scenario, the information group would have experienced a 5% increase in their sales prices. Even more pessimistically, I also simulate what would have happened if those who did not sell their crops experienced losses of -2%, -4%, and -6% of their crop prices with respect to the control group. The results are shown in Columns (2)-(4) of (5.7): even with losses of 2 and 4%, the treatment group would have increased their sales prices by 3 and 4% respectively. Only when we assume a loss of 6% among those without sales, the coefficient is no longer statistically significant (though positive).

The advantage of Horowitz & Manski's (2000) approach is that it does not require any exclusion restriction (i.e. a variable that determines the probability of sales but does not affect directly the sales price) and relies on very few distributional assumptions. However, it is possible to narrow the bounds of the effect with some additional assumptions. In particular, I follow Angrist, Bettinger, Bloom, King & Kremer (2002) and Angrist, Bettinger & Kremer (2006), who propose non-parametric bounds with sample selection[4]. Denote

[4]In particular, Angrist et al. (2006) investigate the impact of a lottery that provided high school students with vouchers to attend private schools in Colombia. The authors analyze the impact of this lottery on students' high school exit-examination scores. They find that the vouchers made students more likely to graduate high school. Because only students who graduated took the exit-examination, this creates a sample selection problem. In this line, their selection problem is quite similar to this one.

Table 5.7: Bounds under Alternative Price Assumptions for Non-Sellers (Cross Section, Follow-up)

Non-seller gains [1]	(1)	(2)	(3)	(4)
	0%	-2%	-4%	-6%
Info	0.05**	0.04**	0.03*	0.02
	(0.017)	(0.017)	(0.017)	(0.018)
Spill	0.01	0.00	-0.00	-0.01
	(0.014)	(0.014)	(0.013)	(0.013)
Constant	0.23***	0.24***	0.25***	0.25***
	(0.015)	(0.015)	(0.015)	(0.016)
Product Dummies	Yes	Yes	Yes	Yes
Households	755	755	755	755
Observations	2,784	2,784	2,784	2,784

[1] The dependent variable of this regression is: (a) the reported price if the crop was sold, or (b) (1-X)% of the mean of the sales price in the control group for unsold crops.

Significance levels denoted by: *** 1%, ** 5%, * 10%. Regressions include household random effects. Standard errors are clustered at the village level.

p_{ic} as the the latent price at which household i would have sold crop c in the follow-up. The household's decision to sell is represented by the indicator variable $s_{ic} = \{0,1\}$. Then, the observed prices are determined by $P_{ic} = s_{ic}p_{ic}$. In this case, there are two treatment groups: those who receive the information directly (Info$_i$) and those in the spillover group (Spill$_i$). The potential outcomes for a household that receives one of these treatments ($T =$ Info$_i$, Spill$_i$) are p^T_{ic}, s^T_{ic}, while the potential outcomes for those in the control group are p^C_{ic}, s^C_{ic} . Also define $q_T(\theta)$ as the θ-quantile of the distribution of p^T_{ic} and $q_C(\theta)$ as the θ-quantile of the distribution of p^C_{ic}.

The estimation of the bounds for self-selection requires three assumptions. The first one is that the treatment is never harmful (i.e. $p^T_{ic} \geq p^C_{ic}$). In this line, the price information should not reduce the price at which households are able to sell their crops. The second one is monotonicity of selection status, which assumes that the treatment can only affect the selection in one direction. Partly, because households decide whether to sell or not based on the price they would get, I will assume that the treatment affects the sales decision positively (i.e. $s_{ic}^T \geq s_{ic}^C$). These two assumptions are somewhat mild in this context. The third one is probably the most restrictive one: I will assume that the treatment is a quantile preserving transformation. This assumption implies that for any $\theta > \theta_0$ — where $q_C(\theta_0) = 0$ — the following condition holds: $P\left[p_{ic}^T \geq q_T(\theta) \,|\, p_{ic}^C \geq q_C(\theta)\right] = 1$. What is assumed here is that "when the potential outcome in the comparison state is above a certain quantile in its own distribution, then the potential outcome in the treatment state is also above that quantile in its own distribution" (Duflo, Glennerster & Kremer 2007). In other words, a household in the control group would have fared as well in the treatment group quantile-wise.

Under these assumptions, Angrist et al. (2006) derive the following inequality:

$$\underbrace{E\left[P_{ic}^T | P_{ic}^T > q_T(\theta)\right] - E\left[P_{ic}^C | P_{ic}^C > q_C(\theta)\right]}_{\text{Upper bound}} \geq \underbrace{E\left[p_{ic}^T - p_{ic}^C | p_{ic}^C > q_c(\theta), s_i^C = 1\right]}_{\text{Treatment Effect (cond on } \theta)} \geq \underbrace{E\left[P_{ic}^T | P_{ic}^T > q_C(\theta)\right] - E\left[P_{ic}^C | P_{ic}^C > q_C(\theta)\right]}_{\text{Lower bound}} \tag{5.5}$$

Conditional on being above quantile θ, the first term of this expression provides the upper bound of the treatment. Intuitively, this component balances the proportion of crops sold in the treatment and control groups. For example, if $\theta = 0.7$, it compares prices in the top 30% of the treatment group with the top 30% of the control group. On the other hand, the last term provides a lower bound by comparing all prices in the treatment and control groups that are above a certain threshold determined by $q_C(\theta)$, i.e. quantile θ of the price distribution in the control group. The term in the middle is the treatment effect (conditional on being above the θ quantile of the control group), which lays somewhere between these two bounds.

To construct these bounds, I will make an additional adjustment. Because my sample is comprised of 17 crops by quality, the sales of low-value products (e.g. potatoes, ollucos) are likely to be in the bottom of the price distribution even if they are sold at a relatively good price. Similarly, even if a farmer gets a relatively low price for a higher-value crop (e.g. produce), this sale might be in the uppermost part of the distribution only because prices for this product tend to be higher. Thus, any quantile of the distribution would trim a disproportionate number of low-value crops. To standardize these differences, I calculate the residuals of a regression of prices on crop and quality indicator variables. The adjusted prices are then: $\tilde{P}_{ic} = log(P_{ic}) - \hat{\gamma}_c\text{crop}_c - \hat{\lambda}_c\text{quality}_c$, where coefficients $\hat{\gamma}_c$ and $\hat{\lambda}_c$ are estimated through an OLS regression. In this sense, $\tilde{P}_{ic}$ can be interpreted as deviations from the mean for each quality-crop[5].

I will apply this estimation framework to the follow-up survey and compare: (a) the Information and the Control groups, (b) the Spillover and the Control groups. In general, 62% of households' crops were not sold in the follow-up survey. However, there are important differences among groups. The proportion of unsold crops was 65% in the control

[5]Note that in the original framework discussed above, $P_{ic} = s_{ic}p_{ic} \geq 0$, where $s_{ic} = \{0,1\}$. In contrast, now the observed (adjusted) price is not bounded by zero (i.e. because it measures deviations from the mean, it can take positive or negative values). Nevertheless, it is relatively straightforward to accommodate for this fact. Consider the new latent variable $\tilde{p}_{ic}$. Then, I construct a variable $\tilde{P}_{ic}$ which takes the value $\tilde{p}_{ic}$ if $s_{ic} = 1$, and any other number considerably smaller than $\text{Min}\{\tilde{P}_{ic}\}$ if $s_{ic} = 0$.

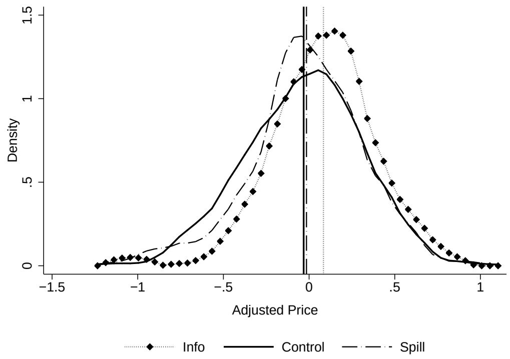
group (i.e. $\theta_0 = 65$) and was slightly smaller in the spillover group (62%). In contrast, it was considerably smaller in the information group, where only 54% of households' crops were not sold. This highlights the extent of the selection problem.

I illustrate the calculation of the upper and lower bounds of the treatment effects at $\theta = 65$. For the lower bound, I would need to trim the distributions of the information and spillover groups by dropping all prices below $q_c(65)$. In this case, $q_c(65)$ is the minimum price across all groups, so no trimming takes place. The lower bound at $\theta = 65$ is represented by the actual price distributions (Figure 5.1a). The estimation of the upper bound of the effect at $\theta = 65$ requires balancing of the proportion of products sold in each group (i.e. comparing the top 35% of the price distributions for each group). At $\theta = 65$, no observations are trimmed from the control group in the upper bound. For the spillover group, this calculation requires dropping the observations between the 62nd and 65th percentiles. This is graphically depicted in Figure (5.1b), where the price distribution of the spillover group shifts slightly to the right. The trimming is much more significant for the information group: we need to drop observations between the 54th and 65th percentile of the distribution. Graphically, there is a pronounced shift to the right in this price distribution.

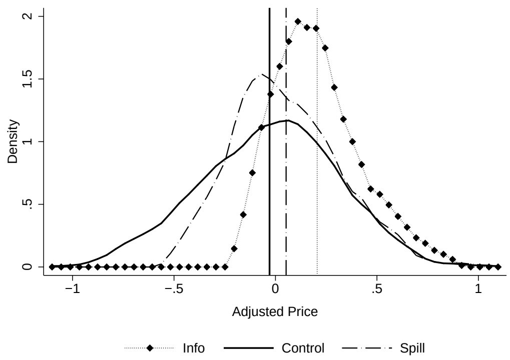
I follow the same procedure for different values of θ. Table (5.8) presents the estimation of the lower and upper bounds θ=65, 70, 75, 85, 90. These estimates do not allow to precisely pin down the effect, but do provide an idea of their maximum and minimum possible values. In general, the upper and lower bounds for those who received the information directly differ by 10 to 14 points at each level of θ. The upper bound remains large (between 21 and 8 points) and statistically significant, and all the lower bounds are positive except when $\theta = 90$. This is consistent with the effect being smaller when the informa-

Figure 5.1: Price Distributions in Endline (Observed and Quantile-Adjusted)

(a) Distribution of Prices in Endline (Observed)



(b) Distribution of Prices in Endline (Adjusted at $\theta = 65$)



Adjusted prices are calculated, as the residuals of a regression of prices on crop and quality dummies, i.e. $\tilde{P}_{ic} = P_{ic} - \hat{\gamma} Crop_c - \hat{\lambda} Quality_c$. Figure 5.1a is the distribution of $\tilde{p}_{ic}$ in the endline and a grahic representation of the results in Table 5.4. Figure 5.1b is the distribution in the endline, with the following quantile restrictions: $\tilde{p}^C_{ic} > q_C(65)$, $\tilde{p}^T_{ic} > q_C(65)$ and $\tilde{p}^S_{ic} > q_C(65)$; where $\tilde{p}^C_{ic}$, $\tilde{p}^S_{ic}$ and $\tilde{p}^T_{ic}$ are the adjusted prices in the control, spillover and treatment groups and $q_C(65)$ is the 65-th percentile of the adjusted price distribution of the control group. Figure 5.1b is the distribution of prices with the following restrictions: $\tilde{p}^C_{ic} > q_C(65)$, $\tilde{p}^T_{ic} > q_T(65)$ and $\tilde{p}^S_{ic} > q_S(65)$.

tion group is compared to those in the control group that would have sold at a high price anyway. For the spillover group, the bounds are much tighter: the difference between the upper and lower bounds is between 1 and 6 points; but all these estimates are considerably smaller and not statistically significant for the most part. Figure (5.2) graphically present the bounds for the information and spillover groups for values of θ between 0.65 and 0.9.

All in all, the evidence suggests that there are positive and significant effects on prices for those who received the information. Among the sample of those who sold their crops, the effect is between 13% and 14%. While sample selection poses a problem for my estimations, if anything, this selection is likely to have a negative impact on these estimates. The upper bound of this effect can even reach 20%.

5.3 Effects on Production and Sales

Next, I estimate the impact of the treatment on households' crop choice, agricultural production and sales volumes. Note that — because of its timing — the intervention should not have altered households' crop choice or production volumes. As discussed in the model, however, the impact on sales volumes is ambiguous.

To determine the impact on the farmers' crop choice, I estimate the following regression with all observations in the sample:

$$C_{ict} = D_c \left[\lambda_c \text{Info}_i + \theta_c \text{Spill}_i + \delta_c t + \beta_c \text{Info}_i t + \gamma_c \text{Spill}_i t\right] + \varepsilon_i + \mu_{ict} \tag{5.6}$$

where C_{ict} is an indicator variable that takes the value of 1 if household i produces crop c in

Table 5.8: Upper and Lower Bounds of Treatment Effect on Prices for Alternative Values of θ[1]

	$\theta = 65$		$\theta = 70$		$\theta = 75$	
	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound
Info	0.21***	0.09**	0.18***	0.05	0.16***	0.02
	(0.030)	(0.040)	(0.029)	(0.033)	(0.023)	(0.029)
Spill	0.07**	0.01	0.05**	0.00	0.03	-0.02
	(0.026)	(0.028)	(0.022)	(0.023)	(0.023)	(0.022)
Constant	-0.03	-0.03	0.05***	0.05***	0.12***	0.12***
	(0.020)	(0.020)	(0.019)	(0.018)	(0.018)	(0.018)
Observations	976	1,063	836	948	699	822
Households	395	406	366	387	329	365

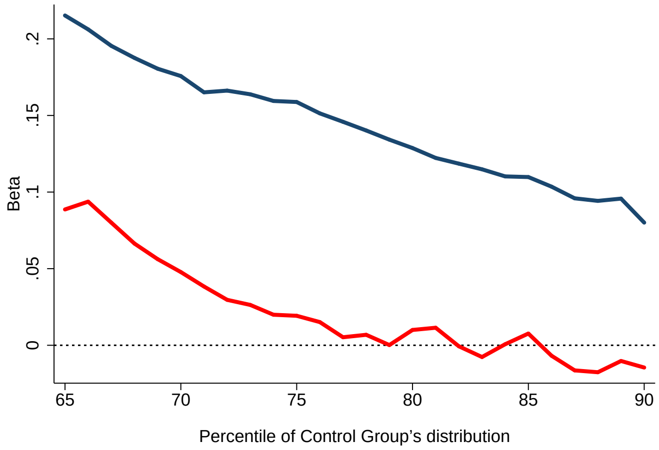
	$\theta = 80$		$\theta = 85$		$\theta = 90$	
	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound
Info	0.13***	0.01	0.11***	0.01	0.08**	-0.01
	(0.025)	(0.029)	(0.033)	(0.026)	(0.037)	(0.031)
Spill	0.02	0.00	0.02	0.00	0.01	-0.00
	(0.027)	(0.026)	(0.025)	(0.025)	(0.027)	(0.027)
Constant	0.20***	0.20***	0.27***	0.27***	0.37***	0.37***
	(0.020)	(0.020)	(0.020)	(0.021)	(0.024)	(0.024)
Observations	559	638	414	492	279	322
Households	281	303	227	252	162	180

[1] The dependent variable of these regressions is $\tilde{P}_{ic} = P_{ic} - \hat{\gamma} Crop_c - \hat{\lambda} Quality_c$, where $\hat{\gamma}$ and $\hat{\lambda}$ are estimated through OLS regressions. The upper and lower bounds are estimated through $\tilde{P}_{ic} = \beta_1 Info_i + \beta_2 Spill_i + \varepsilon_i + \mu_{ic}$, restricting the sample for each level of θ.

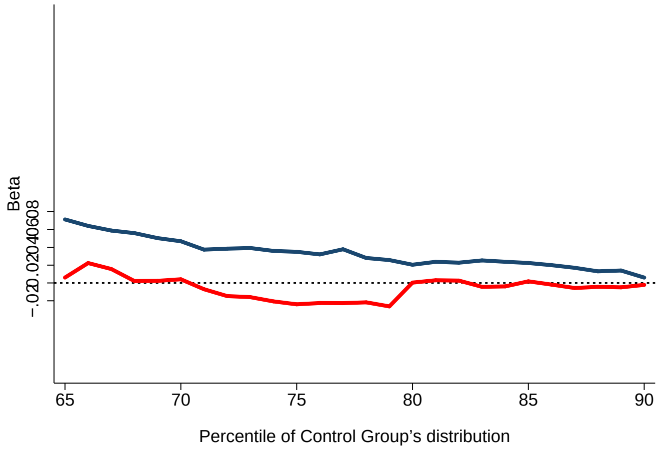
Significance levels denoted by: *** 1%, ** 5%, * 10%. Standard errors are clustered at the village level.

Figure 5.2: Lower and Upper Bounds of Treatment Effects for Prices[1]

(a) Information Group



(b) Spillover Group



[1]These graphs provide a graphic representation of the coefficients in Table 5.8 for all integer values of θ between 0.65 and 0.9.

period t, and 0 otherwise; and D_c is a set of indicator variables for each crop. The DID coefficients for each crop (i.e. β_c and γ_c) are presented in the first two columns of Table (5.9). Albeit significant in two products for the information and two for the spillover groups, overall it seems that the intervention did not alter households' overall crop choice: they are statistically insignificant for the all the other crops (fifteen and sixteen in the information and spillover groups, respectively). I also present the results from a random-effects probit following the same specifications, which also suggest little impact of the intervention on households' overall crop choice.

To determine the impact on households' output, I estimate a Differences-in-differences regression with household random effects, following the same strategy outlined in Equation (5.3). The results are presented in Table (5.10). The first column shows that the DID estimate is close to zero (-1%) and statistically insignificant. Because of the timing of the intervention, I did not expect the price information to induce any change in the composition or amount of harvests. Taken together, these estimates are reassuring that the intervention did not systematically alter households' production. This lends some credibility to the interpretation of the results in this paper: the effect comes from changes in marketing rather than productive outcomes.

Next, I estimate the impact of price information on sales decisions, by analyzing three variables: the extensive margin (i.e. whether they decide to sell or not a particular product), the allocation decision (i.e. where to sell), and the extensive margin (i.e. the sales volumes, conditional on selling). For the extensive margin, I estimate a difference-in-differences

Table 5.9: DID Coefficients for Crop Composition

	RE Linear Reg[1]		RE Probit[2]	
	Info	Spill	Info	Spill
Peas	0.02	0.00	0.02	0.00
	(0.041)	(0.027)	(0.050)	(0.066)
Barley	-0.04	0.02	-0.04	0.02
	(0.056)	(0.043)	(0.061)	(0.053)
Lima Beans	0.00	-0.04	0.00	-0.04
	(0.041)	(0.037)	(0.042)	(0.037)
Corn - White	-0.06	-0.03	-0.06	-0.03
	(0.062)	(0.054)	(0.058)	(0.051)
Corn - Cusqueado	0.12***	0.03**	0.12	0.03
	(0.046)	(0.016)	(0.078)	(0.093)
Corn - Cusqueno	-0.01	-0.01	-0.01	-0.01
	(0.029)	(0.014)	(0.083)	(0.047)
Corn - San Jeronimo	0.00	-0.01	0.00	-0.01
	(0.019)	(0.015)	(0.074)	(0.063)
Olluco - Yellow	-0.04	-0.05***	-0.04	-0.05
	(0.026)	(0.019)	(0.113)	(0.116)
Olluco - Dotted	-0.01	0.00	-0.01	0.00
	(0.030)	(0.025)	(0.092)	(0.025)
Potato - Yellow	0.03	0.00	0.03	0.00
	(0.028)	(0.010)	(0.139)	(0.032)
Potato - Andean	0.03**	0.03	0.03	0.03
	(0.017)	(0.022)	(0.043)	(0.051)
Potato - Canchan	-0.01	-0.02	-0.01	-0.02
	(0.028)	(0.020)	(0.110)	(0.122)
Potato - Huayro	0.01	-0.02	0.01	-0.02
	(0.025)	(0.017)	(0.055)	(0.074)
Potato - Perricholi	0.08	-0.02	0.08	-0.02
	(0.063)	(0.027)	(0.066)	(0.031)
Potato - Peruanita	0.02	-0.00	0.02	-0.00
	(0.020)	(0.010)	(0.017)	(0.040)
Potato - Unica	-0.06	-0.06	-0.06	-0.06
	(0.052)	(0.051)	(0.041)	(0.059)
Potato - Yungay	0.02	-0.02	0.02	-0.02
	(0.051)	(0.053)	(0.053)	(0.047)

[1] The DID coefficients are β_c and γ_c estimated through the regression: $C_{ict} = D_c\left[\lambda_c \text{Info}_i + \theta_c \text{Spill}_i + \delta_c t + \beta_c \text{Info}_i t + \gamma_c \text{Spill}_i t\right] + \varepsilon_i + \mu_{ict}$, where C_{ict}=1 if household i planted crop c in period t (and C_{ict}=0, otherwise) and D_c are indicator variables for each crop.

[2] Marginal effects from a random effects probit, following the same specification.

All regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

Table 5.10: DID Estimation for Production

	(1)	(2) [1]
Info	0.06	0.12
	(0.242)	(0.244)
Spill	-0.10	-0.04
	(0.217)	(0.217)
t	-0.51***	-0.39***
	(0.147)	(0.128)
Info x t	0.01	-0.01
	(0.193)	(0.183)
Spill x t	-0.01	0.01
	(0.176)	(0.168)
Constant	5.45***	5.82***
	(0.191)	(0.249)
Observations	5,236	4,212
Households	789	755
Product Dummies	Yes	Yes
Quality Dummies	No	No

[1] Note that the regression with quality controls has around one thousand fewer observations. This is because most households do not grade their crops when they will not sell them. Thus, the relevant estimates for the production are those in Column 1. However, I present this additional regression for completeness.

All regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

regression for the probability of incurring any sales of a certain product:

$$s_{ict} = \lambda \text{Info}_i + \theta \text{Spill}_i + \delta t + \beta \text{Info}_i t + \gamma \text{Spill}_i t + \alpha_c \text{Crop}_{ict} + \varepsilon_i + \mu_{it} \tag{5.7}$$

where s_{ict} takes a value of one if household i decides to sell crop c in period t, and zero otherwise. The first two columns of Table (5.11) present the results of this regression, and suggests that information increased the probability of selling a product by 12% among those in the information group. Columns 3 and 4 analyze the same outcome using a random-effects probit, and show similar impacts. These shifts are the rationale behind the sample selection adjustments in the previous section: better price information encourages households who would have not sold before to sell (at least part of) their harvests; and this impact is quite large.

I also disaggregate these sales decisions to investigate where households sell. I construct a decision variable for each product with three categories: (a) no sales, (b) sales to middlemen, and (c) direct sales in markets. I estimate a multinomial probit on these three categories to assess households' sales choices. The results are shown in Table (5.12). They show that the previous increases in participation in commercial activities are mostly driven by direct sales in markets. However, on average, there is no change in the proportion of households selling to middlemen.

This suggests that most of the increase in participation in commercial activities are explained by households who did not sell before and sell directly in markets after the intervention. However, one would also have expected that farmers with better information would have shifted their sales away from middlemen to sell directly in markets, but the

Table 5.11: Difference-in-Differences for Probability of Sales

	Probability of Sales			
	Linear Regression		RE Probit[1]	
	(1)	(2)	(3)	(4)
Info	0.01	0.01	0.02	0.00
	(0.057)	(0.065)	(0.058)	(0.058)
Spill	-0.03	-0.02	-0.03	-0.02
	(0.043)	(0.052)	(0.050)	(0.047)
t	-0.12***	-0.01	-0.14***	-0.01
	(0.033)	(0.039)	(0.023)	(0.030)
Info x t	0.12**	0.10	0.14***	0.12**
	(0.057)	(0.064)	(0.052)	(0.057)
Spill x t	0.08	0.07	0.09*	0.08
	(0.058)	(0.065)	(0.050)	(0.050)
Constant	0.48***	0.90***		
	(0.068)	(0.110)		
Observations	5,236	4,212	5,236	4,212
Households	789	755	789	755
Product Dummies	Yes	Yes	Yes	Yes
Quality Dummies	No	Yes	No	Yes

[1] Marginal effects.

Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

Table 5.12: Impact on Alternative Sales Possibilities[1]

	(1) No Sale	(2) Sale to Middleman	(3) Sale in Market
Info x t	-0.11**	-0.01	0.12**
	(0.056)	(0.053)	(0.048)
Spill x t	-0.08	0.01	0.06
	(0.058)	(0.040)	(0.047)
Observations		5,236	
Households		789	

[1] The estimates are calculated through a Random-Effects Multinomial Probit for three categories (no sales, sales to middlemen, and direct sales in markets), with the following variables: $Info_i$, $Spill_i$, $Info_i t$, $Spill_i t$ and crop controls. This table presents marginal effects of $Info_i t$ and $Spill_i t$.

Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

data does not support this idea. Anecdotal evidence suggest that farmers increased their bargaining power with middlemen and were able to achieve higher farm-gate prices. These higher prices encouraged them not to sell their crops directly in markets. For example, during some informal interviews, farmers expressed that they preferred to receive price information through text messages rather than phone calls. They argued that, when needed, they could show text messages to traders, and "prove" that they are aware of the prevailing market prices.

In the endline, I also collected information about the time that the average time farmers spent bargaining with traders during their transactions. I regress the bargaining time (in minutes) on the treatment dummies and crop-quality controls. This sample is restricted only to households with any sales to middlemen during the post-intervention period: $\text{Time}_{ic} = \beta_1 \text{Info}_i + \beta_2 \text{Spill}_i + \gamma D_c + \varepsilon_i + \mu_{ic}$. Results in Table (5.13) show that the effect of information on bargaining time is positive. The coefficient is not statistically significant, but is quite large in magnitude: it increases bargaining time by more than one third compared to the control group.

On the extensive margin, I estimate a regression similar to Equation (5.3), where the dependent variable is the log of sales volume. This estimation only includes those crops for which at least some output was sold and is therefore the impact within those who decide to sell their crops. Table (5.14) presents the results of this estimation. While the point estimates of the DID are quite large (around 19%), they are not statistically significant at conventional levels. It is certainly possible that — in a similar mechanism as the one described for price behavior — the treatment is attracting "marginal" farmers whose sales volumes are smaller because they are less experienced in marketing their products. I

Table 5.13: Bargaining time with Middlemen (in minutes)[1]

	(1)	(2)
Info	3.30	3.69
	(4.104)	(4.013)
Spill	-0.13	-0.23
	(1.169)	(1.041)
Constant	8.98***	8.60***
	(1.790)	(1.960)
Observations	592	592
Product Dummies	Yes	Yes
Quality Dummies	No	Yes

[1] The sample is restricted to crops sold to middlemen at endline (for which we have information about bargaining time).

Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

estimate bounds for the effect on sales volumes following the same approach as the previous section. The distributions of sales volumes in the baseline and follow-up surveys are included in Figure (5.3). It shows that the distribution of sales volumes in the endline becomes more variable and considerably widens in the endline. The estimated bounds are presented in Figure (5.4): they are quite wide, ranging from 17% to 60% for the information and from 17% to 37% for the spillover group. This seems to be consistent with a very noisy distribution that does not allow to accurately detect the effect even with potentially large impacts.

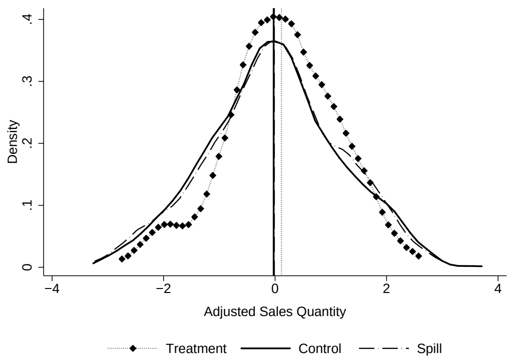
Table 5.14: Difference-in-Differences
Estimations for Sales Volume

	Log(Sales Volume)	
	(1)	(2)
Info	0.05	0.01
	(0.209)	(0.222)
Spill	-0.00	-0.01
	(0.212)	(0.221)
t	-0.43***	-0.39***
	(0.143)	(0.125)
Info x t	0.19	0.19
	(0.178)	(0.182)
Spill x t	0.19	0.19
	(0.168)	(0.160)
Constant	5.68***	5.87***
	(0.269)	(0.308)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	2,122	2,108
Number of households	600	599

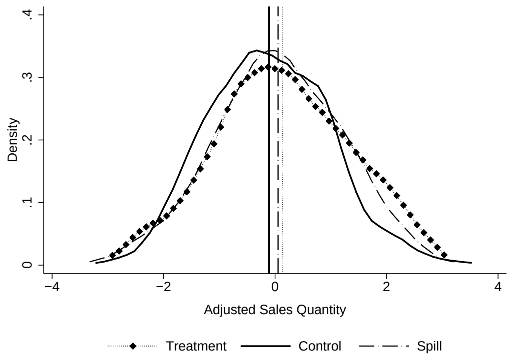
Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

Figure 5.3: Distributions of Sales Volumes in Baseline and Endline

(a) Distribution of Adjusted Sales Volumes in Baseline



(b) Distribution of Adjusted Sales Volumes in Endline

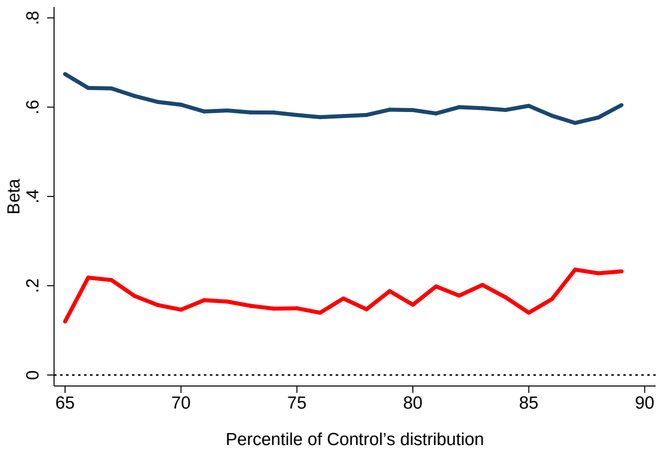


To standardize quantity differences, adjusted quantities are calculated, as the residuals of a regression of prices on crop and quality dummies, i.e. $\tilde{Q}_{ic} = Q_{ic} - \hat{\gamma}Crop_c - \hat{\lambda}Quality_c$. Figure 5.3a is the distribution of $\tilde{q}_{ic}$ in the baseline and is a graphic representation of the estimates in Column 6 of Table 5.3. Figure 5.3b is the distribution of $\tilde{p}_{ic}$ in the endline and a grahic representation of the results in Column 4 of Table 5.11.

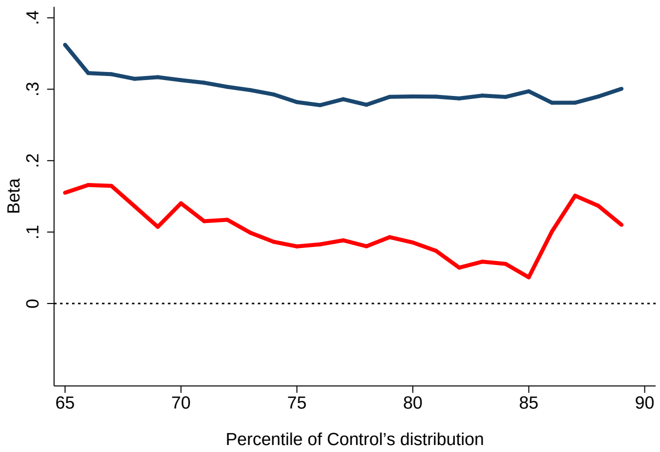
Figure 5.4: Lower and Upper Bounds of Treatment Effects for Sales Volumes[1]

(90% Confidence Intervals in Parentheses)

(a) Information Group



(b) Spillover Group



[1]These graphs provide a graphic representation of the coefficients and confidence intervals of the bounds for all integer values of θ between 0.65 and 0.9.

5.4 Heterogeneous Treatment Effects

5.4.1 Differences by Perishability of Crop

The model in Chapter 4 predicts that improvements of market price information should have different effects for relatively perishable and non-perishable products. The idea is that there is a limit on the amount of the latter that farmers can self-consume before they spoil. This imposes a limit to which farmers can use supply restrictions to obtain better prices from traders. Thus the model predicts that price increases should be larger for perishable products but sold quantities should not increase as much.

This is consistent with previous work that finds that the impact of price information should be more valuable for farmers who sell more perishable crops (e.g. Muto & Yamano 2009, Aker & Fafchamps 2013)[6]. While there might be a set of other factors in play (e.g. market structure, context-specific features, etc.), it is possible that differences between perishable and non-perishable crops might also explain why Mitra et al. (2013) do not find any impact of their price transmission intervention with farmers growing potato (a relatively less perishable crop) in India.

To test for this possibility, I examine the degree of perishability within the seventeen crops in the sample. All in all, there are two that are clearly more perishable than others: lima beans and green peas (which spoil much more quickly than maize, barley, potatoes or *olluco*). To capture differences in the effect for these groups, I use the following variation

[6]For example, Muto & Yamano (2009, p. 1887) argue that: "although the increased flow of information can potentially benefit marketing of all kinds of crops, we expect that it has a larger impact on perishable products than cereals because the prices of perishable products depend heavily on freshness at the time of exchange."

of equation (5.3):

$$\begin{aligned}\log(\mathrm{Y}_{ict}) = &\alpha_1 \mathrm{Info}_i + \alpha_2 \mathrm{Spill}_i + \gamma_1 \mathrm{t} + \beta_1 \mathrm{Info}_i \mathrm{t} + \beta_2 \mathrm{Spill}_i \mathrm{t} + \\ &\lambda \mathrm{Perish}_c + \theta_1 \mathrm{Perish}_c \mathrm{Info}_i + \theta_2 \mathrm{Perish}_c \mathrm{Spill}_i + \\ &\gamma_2 \mathrm{Perish}_c \mathrm{t} + \delta_1 \mathrm{Perish}_c \mathrm{Info}_i \mathrm{t} + \delta_2 \mathrm{Perish}_c \mathrm{Spill}_i \mathrm{t} + \varepsilon_i + \mu_{ict}\end{aligned} \tag{5.8}$$

where Y_{ict} is either prices or sales volumes, $\mathrm{Perish}_c = 1$ for lima beans and green peas and $\mathrm{Perish}_c = 0$ for all other crops. In the case of those who received price information directly, the DID estimators for (relatively) non-perishable and perishable crops are β_1 and $(\beta_1 + \delta_1)$, respectively. Analogously, the DID estimators for the spillover group are β_2 and $(\beta_2 + \delta_2)$.

The results of the regression for prices is reported in the first two columns of Table (5.15). Arguably due to sample sizes, β_1 and δ_1 are not statistically significant, but the large size of δ_1 seem to suggest that the impact on perishable products is larger than the one on non-perishable ones, and the sum of $\beta_1 + \delta_1$ (i.e. the total effect on perishable products) is significant at conventional levels. Columns 3 and 4 present the coefficients of the regression for sales volumes. In contrast to the results for prices, these estimates suggest that, if anything, the increases in sales volumes are smaller for non-perishable products.

5.4.2 Differences by (Previous) Cell Phone Ownership

As explained previously, one of the differences with previous papers exploiting RCTs in this area is that I do not restrict the treatment to those who already had a cell phone. In this spirit, this intervention did not exclude anyone who could have participated, but would have been unable to do so because they did not own a mobile phone. In fact, the devices were distributed regardless of previous ownership, and actually about half of my sample

Table 5.15: Effects by Product Perishability[1]

	Log(Price)		Log(Sales Vol.)	
	(1)	(2)	(3)	(4)
Info x t	0.10	0.09	0.19	0.20
	(0.092)	(0.106)	(0.214)	(0.209)
Spill x t	-0.06	-0.07	0.20	0.20
	(0.083)	(0.099)	(0.204)	(0.193)
Info x t x Perish [2]	0.13	0.16	-0.02	-0.17
	(0.135)	(0.158)	(0.545)	(0.569)
Spill x t x Perish [2]	0.17	0.13	-0.11	-0.16
	(0.240)	(0.211)	(0.519)	(0.514)
Product Dummies	Yes	Yes	Yes	Yes
Quality Dummies	No	Yes	No	Yes
Observations	2,125	2,111	2,122	2,108
Households	601	600	600	599

[1] Estimation results for the following regression: $Y_{ict} = \alpha_1 \text{Info}_i + \alpha_2 \text{Spill}_i + \gamma_1 t + \beta_1 \text{Info}_i t + \beta_2 \text{Spill}_i t + \lambda_1 \text{Perish}_c + \theta_1 \text{Perish}_c \text{Info}_i + \theta_2 \text{Perish}_c \text{Spill}_i + \gamma_2 \text{Perish}_c t + \delta_1 \text{Info}_i \text{Perish}_c t + \delta_2 \text{Spill}_i \text{Perish}_c t + \varepsilon_i + \mu_{ict}$, where Y_{ict} is either Log(Price) or Log(Sales Volume). The Table reports estimates for β_1, β_2, δ_1 and δ_2.

[2] Perishable Products: lima beans and green peas. All other crops (i.e. all types of maize, barley, olluco and potatoes) are considered less perishable.

Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

already had one prior to the intervention. I estimate the following variation of Equation ((5.3) to estimate the impact among those who already had a cell phone and those who did not:

$$\begin{aligned} Log(P_{ict}) \quad = \quad & \alpha_o \text{Info}_i + \alpha_1 \text{Spill}_i + \gamma_0 t + \beta_1 \text{Info}_i t + \beta_2 \text{Spill}_i t + \lambda_0 \text{Mob}_i + \lambda_1 \text{Info}_i \text{Mob}_i + \\ & \lambda_2 \text{Spill}_i \text{Mob}_i + \gamma_1 \text{Mob}_i t + \theta_1 \text{Info}_i \text{Mob}_i t + \theta_2 \text{Spill}_i \text{Mob}_i t + \varepsilon_i + \mu_{ict} \end{aligned} \tag{5.9}$$

where Mob_i=1 if there was a mobile phone in the household before the intervention, and Mob_i=0 otherwise. The coefficients θ_1 and θ_2 measure whether there were any differential effects between both groups. The estimate on the former sample is roughly the one I would have obtained had my treatment been randomized only among those with mobile service. Indeed, because of variations in the intervention and the information provided, it is not strictly comparable to those in Fafchamps & Minten's (2012) study. However, they do provide an idea of what would have happened had my intervention (with the variations in the treatment) been restricted like theirs.

These results are presented in Table 5.16. The coefficients are similar in both groups, providing evidence that the selection among those who previously owned a mobile phone is not driving my results. Thus, I posit that the differences in my results and the ones in Fafchamps & Minten (2012) are more likely to arise from differences in the particular contexts of the interventions, the relevance of information provided, or how the information was displayed. However, I cannot distinguish among these competing hypotheses.

Table 5.16: Effect by (Previous) Cell Phone Ownership

	(1)[1]	(2)[2]
Info x t (β_1)	0.14*	0.15
	(0.085)	(0.096)
Info x t x Mobile (θ_1)		-0.01
		(0.055)
Spill x t (β_2)	-0.02	-0.05
	(0.069)	(0.074)
Spill x t x Mobile (θ_2)		0.10
		(0.084)
$\beta_1 + \theta_1$		0.14
		(0.082)*
$\beta_2 + \theta_2$		0.05
		(0.080)
Observations	2,111	2,111
Households	600	600
Product Dummies	Yes	Yes
Quality Dummies	Yes	Yes

[1] Regression: $Log(P_{ict}) = \alpha_o\text{Info}_i + \alpha_1\text{Spill}_i + \gamma_0 t + \beta_1\text{Info}_i t + \beta_2\text{Spill}_i t + \lambda_0\text{Mob}_i + \lambda_1\text{Info}_i\text{Mob}_i + \lambda_2\text{Spill}_i\text{Mob}_i + \gamma_1\text{Mob}_i t + \theta_1\text{Info}_i\text{Mob}_i t + \theta_2\text{Spill}_i\text{Mob}_i t + \varepsilon_i + \mu_{ict}$, where $\text{Mob}_i = 1$ if household i owned a mobile phone before the intervention, and $\text{Mob}_i = 0$ otherwise.

[2] Corresponds to the results shown in Table 5.4

Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

5.4.3 Effects by Risk Aversion

The baseline questionnaire included a set of questions to elicit risk aversion, where the head of household had to select an option among alternative lotteries. I use a simplified version of a Binswanger's (1980) task to elicit risk preferences[7]. Each respondent is presented with five alternative lotteries. Each of them provides a low and a high payoff to be realized with an equal probability (through a coin toss). While I intended to play these games with cash, unanticipated problems during the fieldwork prevented me from doing so, and the games were only played hypothetically.

My measure of risk aversion is based on the following game, where the respondent is asked to choose one of these lotteries (in Peruvian *Soles*), based on the potential outcome of a coin toss[8]:

1. Get S/. 0.50 without playing
2. Get S/. 0.80 if it comes up head or S/. 0.40 if it comes up tails
3. Get S/. 1.10 if it comes up head or S/. 0.30 if it comes up tails
4. Get S/. 1.40 if it comes up head or S/. 0.20 if it comes up tails
5. Get S/. 1.70 if it comes up head or S/. 0.10 if it comes up tails

More risk averse households should choose the first option (i.e. not to play and receive S/. 0.50 with certainty), less risk averse ones should choose lotteries between two and four, and the least risk averse would choose the fifth option. For simplicity, I construct an

[7]To determine households' risk aversion, I use a set of questions developed by Castillo, Petrie & Torero (2008) and Castillo, Petrie & Torero (2010). Also, the 50/50 gambles involved in each lottery keep the task as simple as possible. Eckel & Grossman (2008) use similar questions to elicit risk aversion. They compare the results of this simplified tasks with more complex ones, such as Holt & Laury's (2002) methods. Their results suggest that simplified tasks might be more suitable for subjects with lower numerical skills.

[8]Note that the expected value and risk (measured as the standard deviation) of the payoffs of each lottery increase linearly from options 1 to 5: there are progressive increases of 0.1 in the $\bar{x}$ and of 0.2 in the σ.

Table 5.17: Risk Distribution in Baseline

	Info	Spill
Risk Averse ($RA_i = 1$)	-0.05	-0.01
(Options 1-3)[1]	(0.043)	(0.036)
Observations	749	

[1] Marginal effects from a probit, where the dependent variable takes a value of 1 if household chose options 1, 2 or 3 (more risk averse); and 0 if it chose options 4 or 5 (less risk averse).

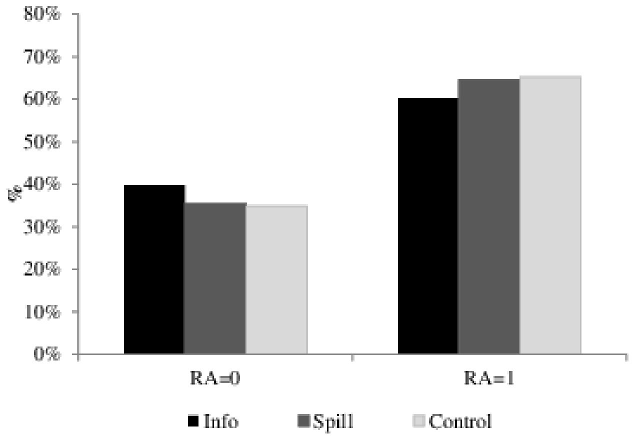
Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

indicator variable for risk aversion: $RA_i = 1$ for households who choose lotteries 1, 2 or 3; and $RA_i = 0$ for those who choose options 4 or 5[9]. Because this information was gathered in the baseline — and, thus, before intervention — there is no reason to believe that RA_i would be correlated with the group assignment. Figure (5.5) shows the distribution of RA_i by treatment assignment. Graphically, no large differences in the proportion of households in each group of risk aversion are apparent. I confirm this through a probit regression for RA_i in Table (5.17), which shows that there are no significant differences in risk aversion between treatment groups.

Even when my games were only played hypothetically, there are reasons to believe that they can be used to capture heterogeneous treatment effects for households with different degrees of risk aversion. First, Castillo, Cotla, Petrie & Torero (2013) use similar modified Binswanger's in a large representative sample of around 13 thousand households in the

[9] Albeit arbitrary, changes in the cut-off point for variable RA_i do not alter the results that follow. Though not shown, the same qualitative results are found, for example, when $RA_i = 1$ for lotteries 1 or 2 and $RA_i = 0$ for lotteries 3, 4 or 5.

Figure 5.5: Distribution of Risk Answers by Group [a/.]



[a/.] This is the distribution of answers to risk questions by group. The risk questions were based on a hypothetical game. Households are assigned to $\mathrm{RA}_i = 1$ if they choose options 1, 2 or 3 from the following lotteries based on a coin toss (and to $\mathrm{RA}_i = 0$, otherwise):

1. S/. 0.50 without playing
2. S/. 0.80 if it comes up heads and S/. 0.40 if it comes up heads
3. S/. 1.10 if it comes up heads and S/. 0.30 if it comes up heads
4. S/. 1.40 if it comes up heads and S/. 0.20 if it comes up heads
5. S/. 1.70 if it comes up heads and S/. 0.10 if it comes up heads

Peruvian highlands (including the Mantaro valley). One subsample of individuals was randomly assigned to play their games hypothetically, while other was offered cash payments. They find no statistically significant differences between both groups. Second, even if the hypothetical nature of the lotteries would have induced any misreporting, this misreporting would have to be different between the treatment and control groups. This is unlikely because these preferences were elicited before the treatment was assigned.

To estimate the heterogeneous treatment effects by risk aversion, I initially estimate the following regression:

$$\begin{aligned} log(\mathrm{P}_{ict}) \quad = \quad & \alpha_o \mathrm{Info}_i + \alpha_1 \mathrm{Spill}_i + \gamma_0 \mathrm{t} + \beta_1 \mathrm{Info}_i \mathrm{t} + \beta_2 \mathrm{Spill}_i \mathrm{t} + \lambda_0 \mathrm{Info}_i \mathrm{RA}_i + \lambda_1 \mathrm{Spill}_i \mathrm{RA}_i \\ & + \tau \mathrm{RA}_i + \gamma_1 \mathrm{RA}_i \mathrm{t} + \theta_1 \mathrm{Info}_i \mathrm{RA}_i \mathrm{t} + \theta_2 \mathrm{Spill}_i \mathrm{RA}_i \mathrm{t} + \delta_c D_c + \varepsilon_i + \mu_{ict} \qquad (5.10) \end{aligned}$$

where $\mathrm{RA}_i = 1$ if household chose options 1-3 for the risk elicitation questions, and $\mathrm{RA}_i = 0$ otherwise. I report these results in Table (5.18). Consistent with my previous results, the impact on households in the spillover group remain are low and statistically insignificant in all cases. However, this estimation would imply that less risk-averse household would have accrued most of the benefits of the intervention ($\beta_1 = 0.21$) within the information group. In contrast, more risk-averse ones would have experienced little or no gains ($\beta_1 + \theta_1 = 0.07$). This result would be somewhat counterintuitive and contrary to the the predictions of the model in Chapter (4). However, more careful examination indicates that this is not necessarily the case.

I estimate a DID regression for households' sales decisions. The estimation framework is similar to equation (5.10), but the dependent variable is an indicator for sales ($\mathrm{S}_{ict} =$

Table 5.18: Heterogeneous Effects by Risk Aversion [1]

	Log(P_{ict})	
	(1)	(2)
Info x t (β_1)	0.21***	0.21***
	(0.072)	(0.067)
Info x RA x t (θ_1)	-0.14	-0.11
	(0.106)	(0.095)
Spill x t (β_2)	0.01	0.03
	(0.073)	(0.065)
Spill x RA x t (θ_2)	-0.02	-0.08
	(0.106)	(0.084)
$\beta_1 + \theta_1$	0.07	0.10
	(0.106)	(0.116)
$\beta_2 + \theta_2$	-0.02	-0.05
	(0.088)	(0.092)
Observations	2,063	2,049
Households	575	574
Product Dummies	Yes	Yes
Quality Dummies	No	No

and would

[1] Estimation of the following equation: $Log(P_{ict}) = \alpha_o \text{Info}_i + \alpha_1 \text{Spill}_i + \gamma_0 t + \beta_1 \text{Info}_i t + \beta_2 \text{Spill}_i t + \lambda_0 \text{RA}_i + \lambda_1 \text{Info}_i \text{RA}_i + \lambda_2 \text{Spill}_i \text{RA}_i + \gamma_1 \text{RA}_i t + \theta_1 \text{Info}_i \text{RA}_i t + \theta_2 \text{Spill}_i \text{RA}_i t + \sum_c \delta_c D_c + \varepsilon_i + \mu_{ict}$.

Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

1 if household i sells product c in period t and $S_{ict} = 0$ otherwise) rather than prices[10]. Table (5.19) shows these estimates. The results suggest that the impact of information on farmers' participation in commercial activities (described Section 5.3) was mostly driven by risk-averse households: while the effect among less risk-averse households was 4% (not significant), more risk-averse ones increased their probability of sales by 16%. Intuitively, if risk-averse households were the ones for which market price uncertainty was a more significant barrier to market participation, this would be the group that would react the most to the information treatment.

These results have important implications for the analysis of price effects by risk aversion. The sample selection problem (described in Section 5.2) is more severe for risk-averse households. Among less-risk averse households in the endline, the proportions of unsold crops were 53% and 59% in the information and control groups, respectively. In contrast, these proportions were 55% and 68% for more risk-averse farmers. This implies that the treatment makes me much more likely to observe a price for risk-averse rather than non risk-averse households (6% vs. 13%).

To address this problem, I construct bounds similar to those in Equation (5.5). I take the deviations from the mean of each crop and quality for each $log(P_{ic})$ in the endline to standardize prices. I estimate the upper and lower bounds with the standardized variable $\tilde{P}_{ic}$ considering six groups: households in information, spillover and control groups, categorized by RA. For each percentile $\theta \in [68, 90]$, I construct samples for the upper and lower

[10]The estimates for a random-effects probit on the sales decision is not reported, but yielded results similar to those of the linear specification.

Table 5.19: Effect of Information on Sales Decision, by Risk Aversion

	(1)	(2)
Info x t (β_1)	0.04	0.04
	(0.08)	(0.09)
Info x t x RA (θ_1)	0.12	0.08
	(0.08)	(0.09)
Spill x t (β_2)	0.10	0.10
	(0.11)	(0.12)
Spill x t x RA (θ_2)	-0.05	-0.07
	(0.09)	(0.11)
$\beta_1 + \theta_1$	0.16**	0.12*
	(0.067)	(0.068)
$\beta_2 + \theta_2$	0.06	0.03
	(0.046)	(0.500)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	5,031	4,058
Households	748	716

[1] Results from the following regression: $S_{itc} = \alpha_o \text{Info}_i + \alpha_1 \text{Spill}_i + \gamma_0 t + \beta_1 \text{Info}_i t + \beta_2 \text{Spill}_i t + \lambda_0 \text{RA}_i + \lambda_1 \text{Info}_i \text{RA}_i + \lambda_2 \text{Spill}_i \text{RA}_i + \gamma_1 \text{RA}_i t + \theta_1 \text{Info}_i \text{RA}_i t + \theta_2 \text{Spill}_i \text{RA}_i t + \sum_c \delta_c D_c + \varepsilon_i + \mu_{ict}$, where $S_{itc} = 1$ if household i sells (at least some) of its harvest of crop c in period t, and $S_{itc} = 0$ otherwise.

Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

bounds, and estimate the following regression:

$$\tilde{P}_{ic} = \alpha_0 \text{Info}_i + \alpha_1 \text{Spill}_i + \gamma_0 \text{RA}_i + \gamma_1 \text{RA}_i + \delta_0 \text{Info}_i \text{RA}_i + \delta_1 \text{Spill}_i \text{RA}_i + \varepsilon_i + \mu_{ic} \quad (5.11)$$

Coefficient α_0 compares the outcomes in the information and control groups among those with $RA_i = 0$, while $\alpha_0 + \delta_0$ contrast the outcomes among those with $RA_i = 1$. The analogous estimators for the spillover group are α_1 and $\alpha_1 + \delta_1$. I present estimates of the upper and lower bounds for selected values of θ in Table (5.20) and plot these coefficients in Figure (5.6). There are larger differences between the upper and lower bounds for more risk-averse households, suggesting that sample selection would have underestimated the previous impact on this group. Within the information group, the upper bound for more-risk averse households is consistently higher for all values of θ. For values $\theta > 75$, both the upper and lower bounds of the effect are larger for more risk-averse households (unambigually showing larger effects within this group). Within the spillover households, both the upper land lower bounds remain relatively small, suggesting limited impacts for both risk-averse and non risk-averse households in this group.

5.4.4 Additional Regressions for Spillover Effects

The results in the previous sections do not support the presence of strong spillover effects. One possibility is that villages are somewhat broad areas for information exchange: if the nearest neighbor with information is still too far away, there might be no possibility for communication. In this spirit, I provide some estimates that restrict the spillover effect through geographic distances. I collected the GPS position of each household in the base-

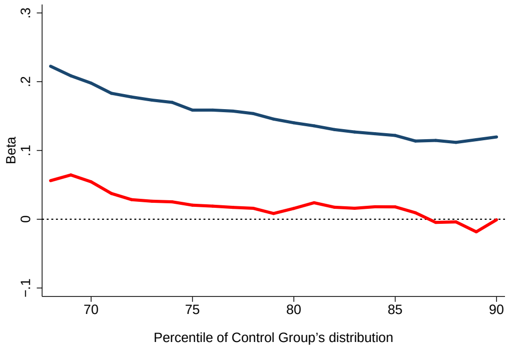
Table 5.20: Upper and Lower Bounds of Treatment Effect on Prices by Risk Aversion for Alternative Values of θ

	$\theta = 68$		$\theta = 70$	
	Upper Bound	Lower Bound	Upper Bound	Lower Bound
Info (β_1)	0.16***	0.15***	0.16***	0.11***
	(0.03)	(0.04)	(0.03)	(0.03)
Info x RA (θ_1)	0.06	-0.09	0.04	-0.06
	(0.04)	(0.06)	(0.04)	(0.04)
Spill (β_2)	0.05	0.04	0.04	0.03
	(0.03)	(0.05)	(0.03)	(0.04)
Spill x RA (θ_2)	0.04	-0.04	0.03	-0.02
	(0.05)	(0.07)	(0.04)	(0.04)
$\beta_1 + \theta_1$	0.22***	0.06	0.20***	0.05
	(0.04)	(0.06)	(0.03)	(0.05)
$\beta_2 + \theta_2$	0.09***	0.00	0.07***	0.00
	(0.03)	(0.04)	(0.03)	(0.03)
Observations	874	1,037	814	986
Households	366	392	355	382

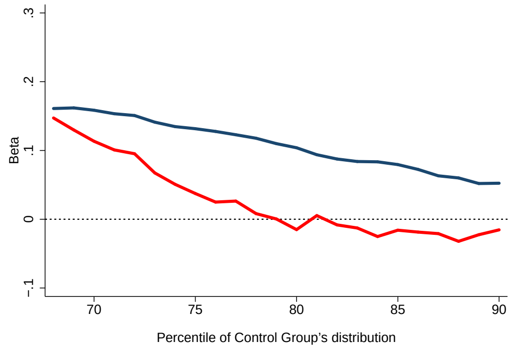
	$\theta = 80$		$\theta = 90$	
	Upper Bound	Lower Bound	Upper Bound	Lower Bound
Info (β_1)	0.10***	-0.02	0.05*	-0.02
	(0.03)	(0.04)	(0.03)	(0.03)
Info x RA (θ_1)	0.04	0.03	0.07**	0.01
	(0.04)	(0.03)	(0.03)	(0.04)
Spill (β_2)	0.03	-0.02	0.03	0.03
	(0.03)	(0.03)	(0.03)	(0.03)
Spill x RA (θ_2)	-0.01	-0.00	-0.03	-0.04
	(0.04)	(0.03)	(0.04)	(0.04)
$\beta_1 + \theta_1$	0.14***	0.02	0.12***	0.00
	(0.03)	(0.03)	(0.04)	(0.04)
$\beta_2 + \theta_2$	0.02	-0.03	0.00	-0.02
	(0.03)	(0.03)	(0.03)	(0.03)
Observations	543	665	274	339
Households	272	307	159	185

Figure 5.6: Upper and Lower Bounds of the Information Effect, by Risk Aversion

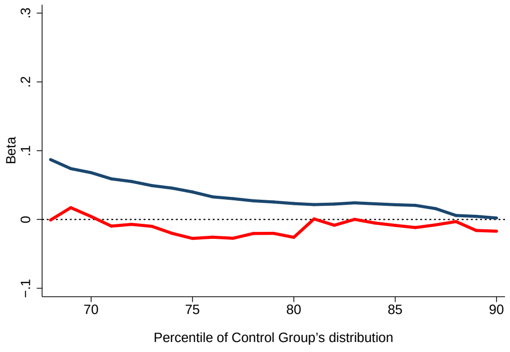
(a) Info, RA=1



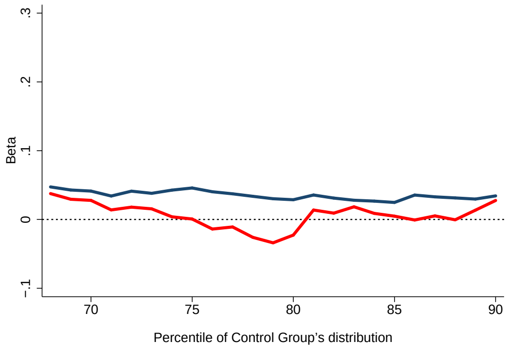
(b) Info, RA=0



(c) Spillover, RA=1



(d) Spillover, RA=0



The dependent variable of these graphs is $\tilde{P}_{ic}$: the deviation from the mean of each crop-quality for each observation P_{ic} in the endline. Each point in the graph is obtained from the following regression: $\tilde{P}_{ic} = \alpha_0 \text{Info}_i + \alpha_1 \text{Spill}_i + \gamma_0 \text{RA}_i + \gamma_1 \text{RA}_i + \delta_0 \text{Info}_i \text{RA}_i + \delta_1 \text{Spill}_i \text{RA}_i + \varepsilon_i + \mu_{ic}$, constructing different samples for all integer values of $\theta \in [68, 90]$ (see Section 5.2).

Figure 5.6b plots the values of α_0 for different levels of θ. Figure 5.6a plots the values of $\alpha_0 + \delta_0$. Analogously, Figures 5.6d and 5.6c plot the coefficients of α_1 and $\alpha_1 + \delta_1$, respectively.

line that allow me to control for this. For each of the households that lived in a treated village but did not receive the market price information, I estimate the distance to its closest neighbor who directly received the price information. I construct quartiles with the distances to the nearest source of information. Denote D_q as dummy variables for each of these quartiles, where $q = 1$ is the group with closest neighbors that directly received the information and $q = 4$ is the most distant. I calculate the following regression:

$$Log(\mathrm{P}_{ict}) = \alpha \mathrm{Info}_i + \sum_{q=1}^{4} \theta_q (D_q \mathrm{Spill}_i) + \delta t + \beta \mathrm{Info}_i \mathrm{t} + \sum_{q=1}^{4} \gamma_q (D_q \mathrm{Spill}_i \mathrm{t}) + \varepsilon_i + \mu_{ict} \quad (5.12)$$

The estimates for γ_q are shown in Table (5.21). If the geographic distance were to play an important role in price transmission, then we would expect $\gamma_1 > \gamma_2 > \gamma_3 > \gamma_4$. However, all the coefficients are still small and not statistically different from zero., and do not suggest this pattern.

Another possibility is that lack of spillover effects is driven by crop differences between the group that directly received the information and the one that could potentially benefit from them indirectly. For example, a farmer in treated village might be getting price information for a certain crop. Because there are 17 different relevant crops in the sample, his neighbors (who are not receiving the information) might be harvesting a different product. To account for this, I construct a variable Match_{ict} for households in the spillover group. $\mathrm{Match}_{ict} = 1$ if any other household in the farmer i's village is directly receiving price information for crop c in period t, and $\mathrm{Match}_{ict} = 0$ otherwise. I estimate the following equation:

Table 5.21: Effects by Distance to Nearest Neighbor with Information[1]

	Log (Price)	
	(1)	(2)
Q_1 x t	0.01	-0.01
	(0.074)	(0.081)
Q_2 x t	-0.05	-0.02
	(0.076)	(0.075)
Q_3 x t	-0.01	-0.08
	(0.093)	(0.099)
Q_4x t	0.04	0.03
	(0.104)	(0.092)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	2,125	2,111
Households	601	600

[1] The quartiles to the nearest neighbor with information are created with the distance of each household in the spillover group (i.e. in a treated village but did not receieve the price SMS) to the closest household that directly received the price information.

[2] The results in this table correspond to the following regression: Log(P_{ict})= $\alpha \text{Info}_i + \Sigma_{q=1}^{4} \theta_q (D_q \text{Spill}_i)$ $+\delta t + \beta \text{Info}_i t + \Sigma_{q=1}^{4} \gamma_q (D_q \text{Spill}_i t) + \Sigma_c \lambda_c \text{Crop}_c + \varepsilon_i + \mu_{ict}$. Columns 1 and 2 report the esimtates for γ_q, without and with crop quality controls.

Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

$$\begin{aligned} Log(P_{ict}) &= \alpha \text{Info}_i + \delta_0 \text{Spill}_i + \delta_1 \text{Spill}_i \text{Match}_{ict} + \gamma \text{t} + \\ & \delta \text{Info}_i \text{t} + \theta_0 \text{Spill}_i \text{t} + \theta_1 \text{Spill}_i \text{Match}_{ict} \text{t} + \varepsilon_i + \mu_{ict} \end{aligned} \tag{5.13}$$

If the previous results — where there was no evidence of spillover effects for the transmission of prices — were driven by product differences, we would expect $\theta_0 = 0$ and $\theta_1 > 0$. However, the results in Table 5.22 show that both coefficients are small and not statistically significant. This additional piece of evidence seems to confirm the absence of spillover effects and the idea that farmers do not share the market information they receive privately with others.

Table 5.22: Effects by Crop Match to Households with Direct Information[1]

	Log(Price)	
	(1)	(2)
Treatment	0.00	-0.02
	(0.076)	(0.065)
t	0.13**	0.15***
	(0.056)	(0.058)
Treatment x t	0.13*	0.14*
	(0.076)	(0.085)
Spill	0.01	-0.03
	(0.094)	(0.075)
Spill x t	0.04	0.04
	(0.101)	(0.090)
Spill x Match	0.03	0.08
	(0.074)	(0.048)
Spill x t x Match	-0.06	-0.07
	(0.083)	(0.061)
Constant	-0.10**	0.02
	(0.050)	(0.046)
Product Dummies	Yes	Yes
Quality Dummies	No	Yes
Observations	2,125	2,111
Households	601	600

[1] Match=1 if the household in the spillover group is in a village where another household has directly received market price information for crop c in year t. It takes a value of zero otherwise.

Regressions include household random effects. Standard errors are clustered at the village level. Significance levels denoted by: *** 1%, ** 5%, * 10% .

Chapter 6

Conclusion

The objective of this paper is to analyze the effect of agricultural price information on marketing outcomes. For this purpose, I set up a RCT where I give access to market price information to farmers in the central highlands of Peru. I find that households with access to information are able to get better prices for their crops: their sales prices increase by 13%-14% relative to those of their counterparts. These results are affected by a sample selection mechanism. However, it seems likely that households who marginally decide to sell (and would have not sold in the absence of the information) would have achieved lower prices than those who would have sold anyway. In this spirit, I construct bounds for the effect of information which suggest that the impact could potentially have been larger.

On the extensive margin, I find a positive and significant impact of information (about 14%) on the probability that households with information engage in commercial transactions for their crops. On the intensive margin, my estimates are positive, but not statistically significant.

Direct beneficiaries might have shared the information they received with their neighbors and lead to indirect gains by others. To test for this possibility, I examine the marketing outcomes of households who did not receive the information but lived in villages where others did. All in all, I do not find any significant impact on marketing outcomes among households in this group.

While in line with the evidence presented by Nyarko et al. (2013) and Courtois & Subervie (2013), my results contrast other studies who analyze programs that directly provide farmers with price information and find a limited effect of information on farmers' sales prices (i.e. Fafchamps & Minten 2012, Mitra et al. 2013). While there are many possibilities to explain such divergence, there are three features in these papers that might help explain why my results are quite different. First, Mitra et al. (2013) focus on potato markets, which is a relatively less perishable crop. As shown in the model (and later validated in my results), the potential benefits for farmers are smaller for these crops. Second, agricultural market conditions in their region of study in India are probably very different from those in the ones in the central highlands of Peru. In explaining their small effects, the authors argue that farmers in this market have little opportunities to sell their crops directly in markets. These markets are usually characterized by long-term relationships between large buyers and sellers[1]. Thus, their results might not necessarily apply to other more fluid agricultural markets. Third, Fafchamps & Minten (2012) focus on farmers that are much more market-oriented [2]. Therefore, they might have targeted a population that was

[1]In their analysis, Mitra et al. (2013) conclude that "mandis in West Bengal still feature decentralized trades between large buyers and local traders engaged in bilateral long term personalized relationships. This creates entry barriers for farmers or other newcomers who intend to sell in these markets".

[2]The authors argue that "unlike Aker (2008) or Muto (2009) who focus on poorly developed agricultural markets, we focus on a part of India where small scale commercial farming has been on the rise, with a growing emphasis on horticulture for urban domestic consumption".

relatively more informed about market prices even in the absence of an intervention.

All in all, this suggests that price information can have very different impacts depending on the crops harvested, the market conditions and the population under analysis. From a policy perspective, this suggests that, while there seem to be opportunities to improve farmers' welfare by enhancing market information, the environment in which price dissemination programs are implemented should be carefully considered. Probably more research is required to fully understand the impact of price information.

Bibliography

Aker, J. (2010), 'Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger', *American Economic Journal - Applied Economics* **2**(3), 46–59.

Aker, J. C. & Fafchamps, M. (2013), 'How does mobile phone coverage affect farm-gate prices? Evidence from West Africa', *World Bank Economic Review* . Forthcoming.

Angrist, J., Bettinger, E., Bloom, E., King, E. & Kremer, M. (2002), 'Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment', *American Economic Review* **92**(5), 1535–1558.

Angrist, J., Bettinger, E. & Kremer, M. (2006), 'Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia', *American Economic Review* **96**(3), 847–862.

Bandiera, O. & Rasul, I. (2006), 'Social Networks and Technology Adoption in Northern Mozambique', *The Economic Journal* **116**(514), 869–902.

Besley, T. & Case, A. (1994), Diffusion as a Learning Process: Evidence from HYV Cotton. Research Program in Development Studies, Center for International Studies, Woodrow Wilson School of Public and International Afffairs, Princeton University. Discussion Paper 174.

Beuermann, D. (2011), Information and Communication Technologies, Agricultural Profitability and Child Labor in Rural Peru. Inter American Development Bank, OVE Working Paper 02/11.

Binswanger, H. (1980), 'Attitudes Toward Risk: Experimental Measurement in Rural India', *American Journal of Agricultural Economics* **62**, 395–407.

Bobonis, G. & Finan, F. (2009), 'Neighborhood Peer Effects in Secondary School Enrollment Decisions', *The Review of Economics and Statistics* **91**(4), 695–716.

Camacho, A. & Conover, E. (2011), The impact of receiving price and climate information in the agricultural sector. Inter-American Development Bank, Working Paper Series No. IDB-WP-220.

Castillo, M., Cotla, C. R., Petrie, R. & Torero, M. (2013), On Measuring the Preferences of the Poor. Mimeo. George Mason University and International Food Policy Research Institute.

Castillo, M., Petrie, R. & Torero, M. (2008), Rationality and the Nature of the Market. Mimeo. Georgia Institute of Technology, Georgia State University and International Food Policy Research Institute.

Castillo, M., Petrie, R. & Torero, M. (2010), 'On the Preferences of Principals and Agents', *Economic Inquiry* **48**(2), 266–273.

Chong, A., Galdo, V. & Torero, M. (2009), 'Access to Telephone Service and Household Income in Poor Rural Areas using a Quasi Natural Experiment for Peru', *Economica* **76**(304), 623–648.

Courtois, P. & Subervie, J. (2013), Farmer Bargaining Power and Market Information Services, Paris, France: Institut National de la Recherche Agronomique (INRA). Presented at the CSAE Conferences 2013: Economic Development in Africa. Oxford (March).

Duflo, E., Glennerster, R. & Kremer, M. (2007), Using Randomization in Development Economics Research A Toolkit, *in* T. P. Schultz & J. A. Straus, eds, 'Handbook of Development Economics', Vol. 4, Elsevier, chapter 61, pp. 3895–3962.

Eckel, C. C. & Grossman, P. J. (2008), 'Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices', *Journal of Economic Behavior and Organization* **68**(1), 1–17.

Eguren, F. (2008), 'An Interview with Carlos Leyton, Minister of Agriculture', *La Revista Agraria* (101), 4–5.

Escobal, J., Ponce, C. & Hernandez-Asensio, R. (2010), Nuevos estilos de articulación en entornos de creciente vulnerabilidad ambiental: el caso de la dinámica territorial rural en la Sierra de Jauja, Junín. Informe de Avance - Etapa 2. Mimeo. Lima, Peru: Grupo de Análisis para el Desarollo (GRADE) and RIMISP.

Fafchamps, M. & Hill, R. V. (2008), 'Price Transmission and Trader Entry in Domestic Commodity Markets', *Economic Development and Cultural Change* **56**(4), 729–766.

Fafchamps, M. & Minten, B. (2012), ‘Impact of SMS-Based Agricultural Information on Indian Farmers’, *World Bank Economic Review* **26**(3), 383–414.

Foster, A. & Rosenzweig, M. (1995), ‘Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture’, *Journal of Political Economy* **103**(6), 1176–1209.

Futch, M. & McIntosh, C. (2009), ‘Tracking the Introduction of the Village Phone Product in Rwanda’, *Information Technologies and International Development* **5**(3).

Giné, X. & Mansuri, G. (2011), Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan. The World Bank, Policy Research Working Paper 5692.

Goyal, A. (2010), ‘Information, Direct Access to Farmers, and Rural Market Performance in Central India’, *American Economic Journal: Applied Economics* **2**(July), 22–45.

Holt, C. A. & Laury, S. K. (2002), ‘Risk Aversion and Incentive Effects’, *The American Economic Review* **92**(5), 1644–1655.

Horowitz, J. L. & Manski, C. F. (2000), ‘Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data’, *Journal of the American Statistical Association* **95**(449), 77–84.

Jensen, R. (2007), ‘The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector’, *The Quarterly Journal of Economics* **122**(3), 879–924.

Magnan, N., Spielman, D., Lybbert, T. & Gulati, K. (2013), Leveling with Friends: Social Networks and Indican Farmers' Demand for Agricultural Custom Hire Services. Unpublished Manuscript. The University of Georgia, Department of Agricultural and Applied Economics.

Manski, C. (1993), 'Identification of Endogenous Social Effects: The Reflection Problem', *Review of Economic Studies* **60**(3), 531–542.

Mitra, S., Mookherjee, D., Torero, M. & Visaria, S. (2013), Asymmetric Information and Middleman Margins: An Experiment with West Bengal Potato Farmers. Mimeo. Boston, MA: Boston University.

Mitra, S. & Sarkar, A. (2003), 'Relative Profitability from Prodution and Trade: A Study of Selectd Potato Markets in West Bengal', *Economic and Political Weekly* **38**(44), 4694–4699.

Molony, T. (2008), 'Running out of Credit: the Limitations of Mobile Telephony in a Tanzanian Agricultural Marketing System', *The Journal of Modern African Studies* **46**(4), 637–658.

Morgan, K. & Rubin, D. (2012), 'Rerandomization to Improve Covariate Balance in Experiments', *The Annals of Statistics* **40**(2), 1263–1282.

Munshi, K. (2004), 'Social Learning in Hetereogeneous Population: Technology Diffusion in the Indian Green Revolution', *Journal of Development Economics* **73**(1), 185–213.

Muto, M. & Yamano, T. (2009), 'The Imact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda', *World Development* **37**(12), 1887–1896.

Nakasone, E., Torero, M. & Minten, B. (forthcoming), 'The Power of Information: The ICT Revolution in Agricultural Development', *Annual Review of Resource Economics* **6**(1).

National Statistics Institute (2013), Resultados Definitivos del IV Censo Nacional Agropecuario 2012, Technical report, National Statistics Institute, Lima, Peru.

Nyarko, Y., Hildebrandt, N., Romagnoli, G. & Soldani, E. (2013), Market Information Systems for Rural Farmers Evaluation of ESOKO MIS, Year 1 Resutls. New York University Abu Dhabi, Center for Technology and Economic Development. Available at: http://www.nyucted.org/archives/1108.

Oster, E. & Thornton, R. (2012), 'Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up', *Journal of the European Economic Association* **10**(6), 1263–1293.

Stigler, G. (1961), 'The Economics of Information', *The Journal of Political Economy* **69**(3), 213–225.

Svensson, J. & Yanagizawa, D. (2009), 'Getting Prices Right: The Impact of the Market Information Service in Uganda', *Journal of the European Economic Association* **7**(2-3), 435–445.

Vasilaky, K. & Leonard, K. (2013), As Good as the Neighbors They Keep? Improving Farmers' Social Networks via Randomized Information Exchange in Rural Uganda. Unpublished Manuscript. University of Maryland, Department of Agricultural and Resource Economics.