Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 13-13

Investment, awareness, supermarkets, and profits: heterogeneous chili farmers in Indonesia

Han-Hsin Chang
Charles van Marrewijk
Randy Stringer
Wendy Umberger

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Kriekenpitplein 21-22
3584 EC Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to J.M.vanDort@uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors

Please direct all correspondence to the second author.

Han-Hsin Chang~*
~Asian Development Bank
6 ADB Avenue,
Mandaluyong City 1550
Philippines
Charles van Marrewijk
*Utrecht University
Utrecht School of Economics
Kriekenpitplein 21-22
3584 TC Utrecht
The Netherlands.
E-mail: J.G.M.vanMarrewijk@uu.nl
Randy Stringer
Wendy Umberger
Global Food Studies
Roseworthy Campus
University of Adelaide
Roseworthy SA 5371
Australia

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 13-13

Investment, awareness, supermarkets, and profits: heterogeneous chili farmers in Indonesia

Han-Hsin Chang[ab]
Charles van Marrewijk[b]
Randy Stringer[c]
Wendy Umberger[c]

[a]Asian Development Bank
Philippines

[b]Utrecht School of Economics
Utrecht University

[c] Global Food Studies
University of Adelaide

October 28, 2013

Abstract

Changes in the distribution channels for food production in developing countries raises the demand for high quality products sold through supermarkets at higher prices. We model the willingness of farmers to invest in high quality production, taking the role of traders into consideration. We test our model using data for Indonesian chili producers and find that (i) there is self-selection into high cost investment for production quality which leads to higher profits, (ii) there is a positive and significant income effect for participating in the modern (supermarket) retail channel, and (iii) the awareness of farmers regarding their product participation in the retail channel significantly affects their profits. We thus find both a 'real access gap' and a 'market efficiency gap'.

Keywords: Supermarkets, Farmers, High-value Agriculture, Indonesia, Heckman self-selection

JEL classification: D2, I3, L1, O1, Q1

Acknowledgements
This study was made possible by funding from the Australian Centre for International Agriculture Research (ACIAR). The project researchers gratefully acknowledge the support and guidance provided by ACIAR project managers Simon Hearn and David Shearer. Special thanks to PhD students who participated and gave helpful comments and suggestions in the seminar where the idea of this paper was first presented. The paper was also presented at the 2013 annual Asia Pacific Trade Seminar (APTS) in Nanjing, China and at Utrecht University School of Economics. We thank the seminar participants for valuable feedback and comments. Lastly, we thank Christopher Findlay for introducing and connecting the authors of this work, as well as for all the comments and suggestions given along the way.

1. Introduction

The increase in market penetration of supermarkets continues to transform the food-retail sector in developing countries (see Reardon et al, 2009; Reardon et al, 2007; Neven et al., 2006; Reardon et al., 2005; 2003; Weatherspoon and Reardon, 2003). In much of Asia, more than a decade of rising incomes and urbanization has resulted in large shifts in food consumption patterns and shopping behavior. Consumers are eating many more calories on average, while shifting from staple foods consumption towards higher-value fruits, vegetables and processed foods in consumption. Shoppers are willing and able to pay price premiums for convenience, food quality, and food safety (Mergenthaler et al, 2009a,b).[1]

These ongoing trends also affect agricultural producers through foreign investments and the higher standards imposed by supermarkets (Schipmann and Qaim, 2011a,b).[2] A growing literature aims to understand how the transformation of modern food retail chains influences smallholder market decisions. In its broadest sense, research on evolving food chains addresses how market liberalization, trade and foreign investment policies impact food producers, traders, wholesale markets, processors and retailers. This literature analyzes how rapid income growth and urbanization transform traditional chains, exploring poverty reduction opportunities, distributional implications and efficiency effects of smallholder participation (Reardon, Barrett, Berdegué and Swinnen, 2009). Key research issues are whether small producers are able to participate in more modern food chains and when they do participate, what are the growth, equity and development consequences?

To date the evidence suggests that supermarkets provide challenges, threats and opportunities to small farmers (Reardon et al, 2009;). On the one hand, better prices and a steady relationship with the buyer potentially improves the welfare for those who participate in the supermarket retail channel.[3] On the other hand, the quantity and quality requirements set by supermarkets are a challenge for small farmers. A growing body of literature has been looking into the determinants of the modern channel participation and the welfare implications of the potential exclusion of small-scale producers from entering the modern

[1] See also Minot and Roy, 2006; Traill, 2006; Regmi and Dyck, 2001; Jaffee and Morton, 1994; Minot, 1986.
[2] See also Reardon, Berdegué et al., 2007; Berdegué et al., 2007; Reardon, Henson et al., 2007; Berdegué et al., 2005; Henson and Reardon, 2005; Reardon and Swinnen, 2004; Neven and Reardon, 2004; Reardon et al, 2003; Reardon et al., 2001; Reardon and Barrett, 2000.
[3] Sexton (2012) argues that the agricultural market is characterized by imperfect competition in view of product heterogeneity, consolidation, and the dominate role of a few large processing, trading, and retailing firms.

retail channel (Rao and Qaim, 2010; Miyata et al., 2009; Moustier et al., 2009; Neven et al., 2009).[4] Exclusion from potentially more profitable food markets implies small farmers may be stuck in lower income poverty traps, and unable to gain the new technologies and management practices that modern food markets require.

This paper contributes to this literature in three ways. First, we develop a simple model that determines the sorting of famers, based on their household characteristics, into high- or low-cost investment for production, rather than channel selection. The investment level correlates positively with the probability of producing high-quality output meeting the quality standards of the supermarket, thus increasing the likelihood of accessing the modern retail channel. The model thus provides a theoretical foundation for the self-selection of investment, which influences a farmer's modern channel participation, and the economic impact of that decision. The model provides a testable hypothesis regarding differences in profitability.

Second, we incorporate an aspect overlooked in the current literature, namely that not all modern channel producers are aware of the final marketing channel of their product. This unawareness may arise because farmers do not sell directly to retailers but to intermediate traders, who represent and collect the fresh products for either the traditional market wholesaler or the dedicated supermarket wholesalers (Gulati et al., 2007). Asymmetric information regarding the value and quality of the product and its marketing channel places the intermediate trader in a better bargaining position over farmers with less market information. Some farmers are therefore unable to receive the maximum price the traders are potentially willing to offer. We model this aspect as a factor that affects traders' search cost. Empirically, we show how awareness in the marketing channel is an important aspect influencing farmers' profitability.

Third, based on our model we apply the treatment effect model (Heckman self-selection) to analyze the heterogeneous welfare implications for three types of farmers: aware, unaware, and traditional. The treatment effect is implemented in a two-stage estimation procedure, with the first stage dealing with the self-selection issue (the probability of making a high-cost investment decision) and the second stage dealing with the welfare implications.

[4] See also Hernández et al., 2007; Reardon, et al., 2007; Reardon, et al., 2007; Natawidjaja et al., 2007.

The empirical analysis uses household data from a survey of 597 chili farmers in Java, Indonesia.[5] Chili producers are chosen because chilies are a high-value agricultural product in Indonesia, providing farm households a source of income in addition to traditional or staple crops. The chili is an important vegetable consumed daily by most Indonesian families and supermarkets are increasingly including chilies as part of their product offerings. Like many other fresh fruit and vegetables (FFV) production, the production of chili is labor intensive, providing additional employment opportunities (Weinberger and Lunpkin, 2005). Previous studies suggest that FFV sold in the supermarket are usually marketed as high quality produce and that traders selling directly to the supermarkets usually receive a premium for their produce. However, production of high-quality FFV products that meet the stringent specifications of the supermarket channel usually comes at a higher input cost, resulting from the use of high-quality seeds, hired labor, capital investments required to ensure product quality standards (in terms of safety and appearance), and the packaging, storage and delivery requirements (Schipmann and Qaim, 2011b; Neven et al., 2009). These costs are challenging for the credit-constrained farmers. In addition, production uncertainty implies that even though investments are made and additional costs incurred, not all output will be of sufficient quality to meet the modern retail standards. Thus, investing in high-quality output is not necessarily rewarded with higher prices.

We find that aware and unaware farmers have similar basic household characteristics and production attributes. Our production sorting and welfare optimization model suggests that farmers who self-select into making high-cost investments for production earn higher profits. The treatment effect model confirms that endogenous sorting takes place based on farmers' productivity and other household characteristics. More interestingly, among modern channel producers, aware farmers benefit much more from participating in the modern channel than unaware farmers, after taking into account their investment level. This suggests that farmers benefit unequally even among supermarket suppliers. Our research thus shows that (i) farmers have to overcome a production and investment barrier in order to produce high-quality products suitable for the modern food retail channel and (ii) farmers need to be aware of the marketing channel of their product to obtain the full benefit from their investments as suppliers of the modern food retail channel.

[5] See Chowdhury el al. (2005) for an overview of structural change in the agricultural sector in Indonesia. See Awfaw et al (2010), Awfaw (2007), Minten et al (2009), and Miyata et al (2009) for treatment effect models.

2. Production sorting model

Let F be the total number of farmers, each producing one unit of output. Output is heterogeneous in terms of quality and can be sold in two separate channels: the modern channel and the traditional channel. The product quality is evaluated partially according to the appearance of a product, while other unobservable characteristics are evaluated through close monitoring during the production stage. An exogenous quality threshold set by the supermarket leads to sorting of outputs into two levels: high-quality and low-quality. Consumers willing and able to pay a premium for guaranteed quality and food safety shop in the supermarket.

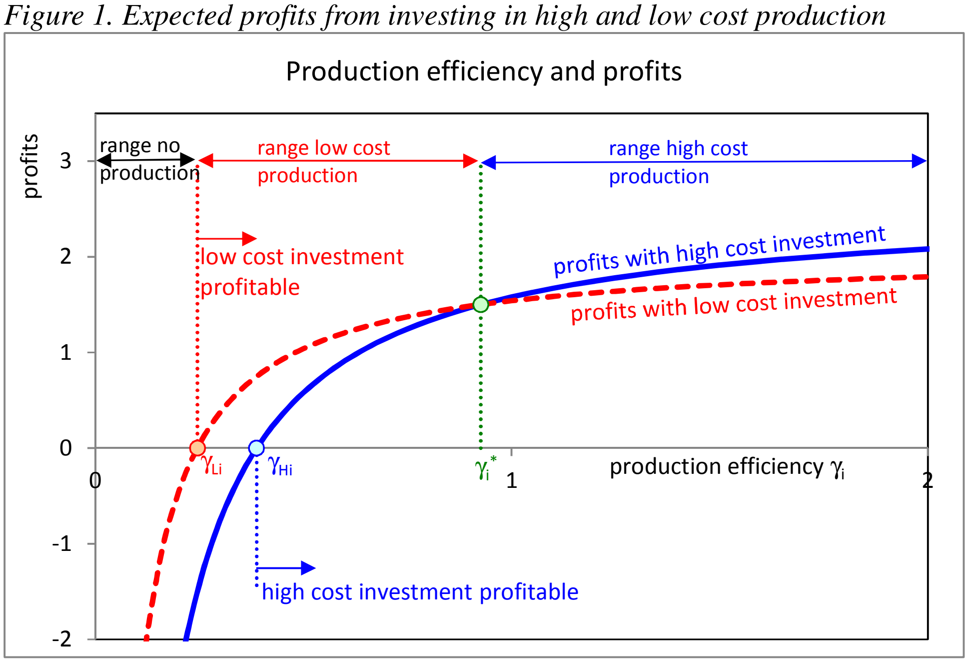
There are different prices offered to the famers in the two channels. Traders who procure for the modern channel buy only high-quality products. Although quality is partially observable, traders have to search for the farmers that produce high-quality output and incur search cost, S_i. These search cost are farmer-specific and depend on the ease with which the supermarket traders can find the farmers with high-quality products. The easier it is for the traders to find or contact the farmers with the output they require the lower the search cost for the trader to reach a specific farmer. The maximum price (or reservation price) the modern channel traders are willing to pay for the high-quality output is the difference between the price supermarkets pay to the trader P_S and the search costs supermarket traders incur to find a specific farmer: $P_{M,i} = P_S - S_i$, where the subscript M indicates Modern, S indicates Supermarket, and i is a farmer index. Traders who procure for the traditional channel do not care about quality and buy whatever output is available in the market. The search cost for the traditional traders is therefore negligible. The traditional channel traders offer a maximum price P_T.

Farmers are heterogeneous in terms of production efficiency γ. The efficiency is related to household characteristics, such as age, education, experience, and endowments. More educated farmers, for example, are expected to have better understanding of the specific requirements and details specified in the contract or agreement with the traders compared to the less educated farmers. Since the maximum price the supermarket pays to the traders is fixed, the price these traders are willing to offer to each farmer is a function of the search cost that differs between farmers: $P_M(S_i)$. Let $g(\gamma, S)$ denote the joint probability distribution of farmers' production efficiency and the search cost incurred to reach them. Farmers choose between making a high- or low-cost production decision, with investment costs C_H and C_L respectively ($C_H > C_L$). Production involves uncertainty. Farmers with high-cost investment produce high-quality output with probability β_H and low-quality output with probability $1\text{-}\beta_H$. Similarly, farmers with low-cost investment produce high-quality output with probability β_L and low-quality output with probability $1\text{-}\beta_L$. We assume that high-cost investment leads to higher expected quality: $\beta_H > \beta_L$.

Let $\delta \in (0,1]$ denote the probability that high-quality output is sold to the supermarket trader. The rejection rate $1-\delta$ is exogenous and independent of individual farmers' search cost.[6] Based on the exogenous market condition parameters $(\delta, \beta_H, \beta_L, P_S, P_T, C_H, C_L)$, farmers have three options: investing in high-cost production, low-cost production, or exiting the market (not entering into production). The expected profit π for farmer i with production efficiency parameter γ_i for whom the trader incurs search costs S_i is :

$$(1)\ E(\pi_i) = \begin{cases} \delta\beta_H(P_S - S_i) + (1-\delta\beta_H)P_T - C_H/\gamma_i & \\ \delta\beta_L(P_S - S_i) + (1-\delta\beta_L)P_T - C_L/\gamma_i & for \\ 0 & \end{cases} \begin{matrix} high\ cost\ investment \\ low\ cost\ investment \\ no\ production \end{matrix}$$

Figure 1. Expected profits from investing in high and low cost production



Parameters: $P_S = 4;\ S_i = 1; P_T = 1.5; \delta = 0.9; \beta_H = 0.8; \beta_L = 0.4; C_H = 1; C_L = 0.5.$

[6] The parameter captures the exogenous product rejection shock for farmers.

Figure 1 illustrates the production choice facing a farmer as a function of his production efficiency, given a particular set of other parameters. If production efficiency is below some critical value γ_{Li} even low cost production is not profitable. Similarly, high cost production with efficiency below some critical value γ_{Hi} is not profitable.[7] Note that these critical values are firm-specific as they depend on the search costs S_i. Figure 1 is drawn under the assumption that $\gamma_{Li} < \gamma_{Hi}$, in which case we can identify three ranges of production for the firm in question. If production efficiency is too low, namely $\gamma_i \in (0, \gamma_{Li})$, the firm does not produce. If production efficiency is intermediate, namely $\gamma_i \in [\gamma_{Li}, \gamma_i^*)$, the firm produces with low cost investment. If production efficiency is high, namely $\gamma_i \in [\gamma_i^*, \infty)$, the firm produces with high cost investment. The critical value γ_i^* which determines the borderline between low cost and high cost investment is determined by the switching condition, leading to equal profits for low and high cost investment.

(2) $\gamma_i^* = \frac{(C_H - C_L)}{\delta(\beta_H - \beta_L)(P_S - S_i - P_T)}$ (switching condition)

Clearly, for this switching condition to be operative two conditions must be fulfilled. First, the solution should be positive and finite. This implies that the term $(P_S - S_i - P_T)$ must be positive, or the search costs must be less than the supermarket price premium: $S_i < (P_S - P_T)$. In other words: firms with too high search costs will never switch to high cost investments. Second, for a switch to occur it must be worthwhile to produce at low cost for intermediately efficient firms, as illustrated in Figure 1. However, if $\gamma_{Hi} < \gamma_{Li}$ high cost investment profits always dominate low cost investment and a switch does not occur. As can be easily verified, this condition is equivalent to assuming $\gamma_{Hi} > \gamma_i^*$, that is switching occurs to the left of the point where high cost investment becomes profitable. Note that the conditions imposed so far do not exclude this possibility from occuring, as $\gamma_{Hi} < \gamma_{Li}$ is equivalent to assuming $(C_H - C_L)P_T < \delta(\beta_H C_L - \beta_L C_H)(P_S - S_i - P_T)$. Other things equal, it is therefore more likely to hold if (i) the probability of high quality production with high cost investment β_H rises, (ii) the probability of high quality production with low cost investment β_L falls, and (iii) the search costs S_i fall. We can combine the various possibilities by defining the *switching* efficiency level $\tilde{\gamma}_i \equiv max\{\gamma_i^*, \gamma_{Hi}\}$. High cost investment production then occurs if efficiency exceeds the switching level: $\gamma_i \geq \tilde{\gamma}_i$. Similarly, we can define the *minimum* efficiency for

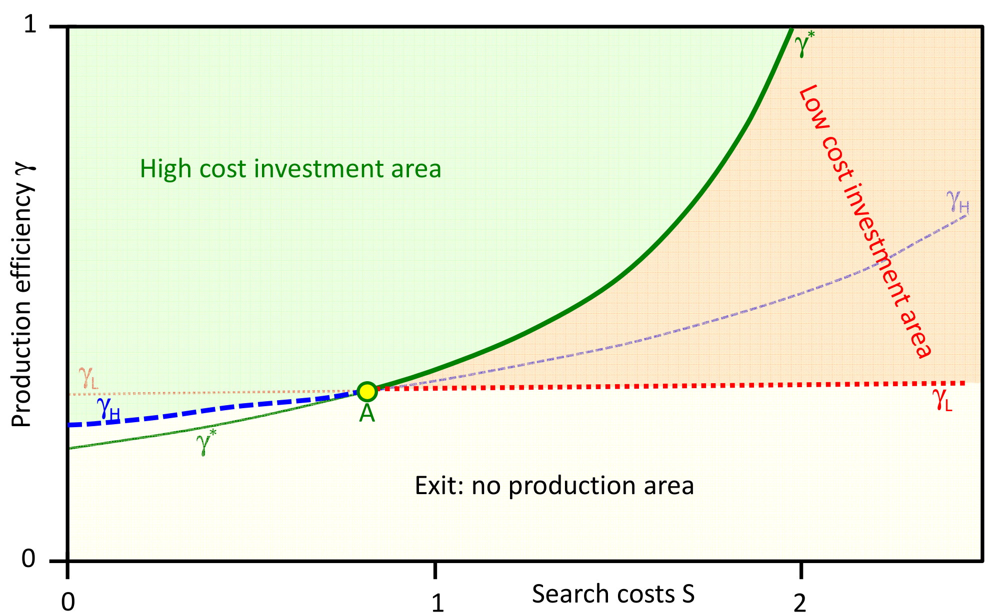
[7] Note that $\gamma_{Li} = C_L/[P_T + \delta\beta_L(P_S - S_i - P_T)]$ and $\gamma_{Hi} = C_H/[P_T + \delta\beta_H(P_S - S_i - P_T)]$.

positive production level $\underline{\gamma_i} \equiv min\{\gamma_{Li}, \gamma_{Hi}\}$. Low cost investment production then occurs if efficiency is in between the minimum efficiency and the switching level: $\underline{\gamma_i} \leq \gamma_i < \tilde{\gamma}_i$. As noted above, depending on the specific parameters this range may be empty for certain levels of low searching costs.

2.1 Equilibrium

The market is in equilibrium when high quality output sold in the supermarket matches the demand for high quality output that pays a high price: $Q^M = D_H$, while the rest of the output is sold though the traditional channel regardless of quality: $Q^T = Q - Q^M$. To determine the share of farmers producing with high and low investment costs and the share exiting the market we need to take both the search costs and the production efficiency into consideration. This is illustrated in Figure 2 in case there is no range of low cost investment for sufficiently low search costs. The figure shows γ_H and γ_L as defined in footnote 7and γ^* as defined in equation (2). All three curves intersect at point A. To the left of this point there is no low cost investment area, such that : $\underline{\gamma} = \gamma_H = \tilde{\gamma}$. To the right of point A we have: $\underline{\gamma} = \gamma_L < \tilde{\gamma}$. The range to the left of point A does not occur, of course, if point A is to the left of the vertical axis. A sufficient condition for that to happen is: $\beta_L C_H > \beta_H C_L$, that is the appropriately weighted probability of high quality output with low cost investment is sufficiently high.

Figure 2 Sorting of farmers into high and low investment costs and exit



Parameters: $P_S = 4;\ P_T = 1.5; \delta = 0.99; \beta_H = 0.99; \beta_L = 0.04; C_H = 1; C_L = 0.5.$

The curves in Figure 2 separate the plain into three regions and depict the sorting of farmers into high cost production, low cost production, and exit depending on their production efficiency and the search cost. Taking the joint density $g(\gamma, S)$ over production efficiency and search costs into consideration, we can now determine the share μ_H of firms which produce with high costs investment, the share μ_L which produces with low cost investment, and the share $\mu_E = 1 - \mu_H - \mu_L$ which exits the market:

(3) $\mu_H = \int_0^\infty \int_{\tilde{\gamma}}^\infty g(\gamma, S) d\gamma dS$ high cost investment

(4) $\mu_L = \int_0^\infty \int_{\underline{\gamma}}^{\tilde{\gamma}} g(\gamma, S) d\gamma dS$ low cost investment

(5) $\mu_E = \int_0^\infty \int_0^{\underline{\gamma}} g(\gamma, S) d\gamma dS$ exit

Since there is a total of F farmers available, and the probability of producing high quality output is equal to β_H for farmers with high cost investment and equal to β_L for farmers with low cost investment we can now determine the total volume of high quality and low quality production.

(6) $Q_H = (\mu_H \beta_H + \mu_L \beta_L)F$ high quality output

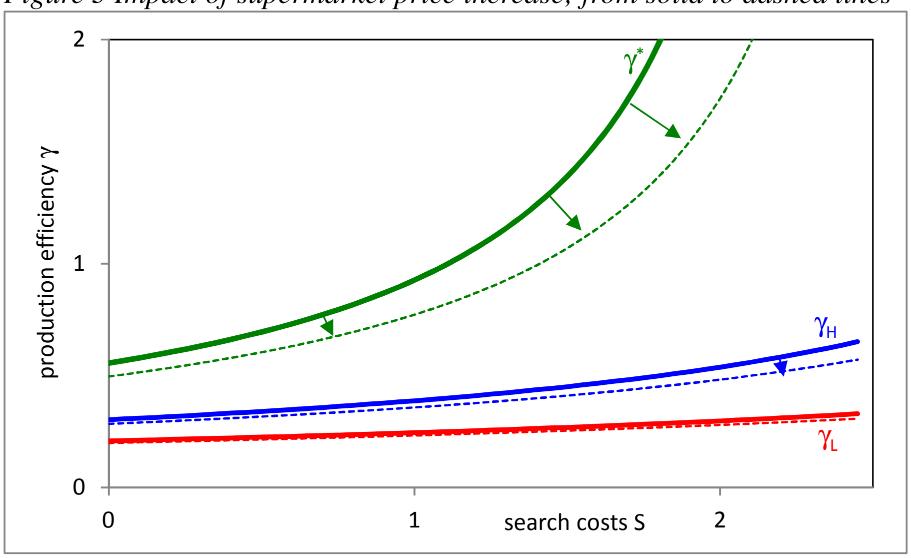
(7) $Q_L = (\mu_H(1 - \beta_H) + \mu_L(1 - \beta_L))F$ low quality output

The total output produced is the sum of high and low quality outputs: $Q = Q_H + Q_L$ and the total number of farmers that exit production is $F - Q$. Recall that only a fraction δ of high quality output will be sold through the modern supermarket channel: $Q^M = \delta Q_H$. The remaining output is sold through the traditional channel regardless of quality. As long as some high quality product is rejected by the supermarket traders ($\delta \neq 1$) the total high quality produced is greater than the quantity sold through the supermarket channel. Therefore, although only Q^M is priced and purchased as high quality product, the true amount of high quality food consumed is larger. This is an indirect welfare effect for the public often neglected. There will be a welfare loss as there will be less than the optimal number of farmers joining high quality output production when a proportion of high quality output is not rewarded with a high price. In this case there will be fewer farmers engaging in high cost investment with high quality production. At the same time, the rejected high quality product will be sold at a lower price to the consumers who shop at the traditional market either because they cannot afford standardized quality food or because they place a lower value on food quality.

2.2 Comparative static analysis

During the stage of economic transition in which the growing demand for high quality food exceeds the supply in the short-run, the excess demand drives up the procurement price in the modern channel. Recall that the maximum price the supermarket trader is willing to pay is positively related with the price P_S paid by supermarkets and negatively related to traders' search cost S_i, which differs per farmer. As economies develop and the price P_S supermarkets offer to traders rises, both the switching production efficiency and the minimum production efficiency fall.[8] This is illustrated in Figure 3. Since the high investment cost area increases and the exit area decreases, we can be sure that the share of high cost investment farmers rises ($\partial \mu_H / \partial P_S > 0$) and the share of exiting farmers falls ($\partial \mu_E / \partial P_S < 0$). Whether or not the share of low cost investment farmers rises or falls depends on the joint density $g(\gamma, S)$.

Figure 3 Impact of supermarket price increase; from solid to dashed lines



Parameters: $P_S = 4$ *and* 4.3 (*solid and dashed*); $P_T = 1.5$; $\delta = 0.9$; $\beta_H = 0.8$; $\beta_L = 0.4$; $C_H = 1$; $C_L = 0.5$.

Up to this point, we reflect little on supermarket traders' search cost and assume it is exogenously determined by the remoteness of the farmer. As mentioned, the easier it is for the traders to find or contact the farmers with the output they want, the lower the search cost for the trader to reach a specific farmer. Search costs are negatively correlated with the ease of communication traders experience with farmers and with the quality signals provided by the farmers. Although the farmers located further away from the market in principle have

[8] Note, for example, that $\partial \gamma_{Li} / \partial P_S = -\delta \beta_L \gamma_{Li}^2 / C_L$. Similarly for the other curves.

higher search costs, ownership of transportation or communication facilities, such as trucks and mobile phone can mediate the impact of physical distance. Other farm-related assets that signal producer's ability to preserve freshness or delay the product from perishing also decrease traders' search cost and are expected to increase profit. Note in particular that for an individual farmer it is the net premium paid by the supermarket channel over the traditional channel which determines its investment decision. The impact of a fall in search costs is therefore the same as the impact of a rise in the price paid by the supermarket.

2.3 Hypotheseis

To sum up the model and connect the conceptual model with the empirical work, following hypotheses are made. First, profit maximization ensures farmers sort into different investment levels based on their productivity. Thus, the most productive farmers self-select into high effort investment, and less productive farmers invest in low cost production, while the least productive farmers will not enter production.

- Hypothesis 1: More productive farmers self-select into high cost production and earn higher profits than farmers who invest in low cost production.

Second, although farmers make production decisions based on market price for high and low quality output, the search costs traders incur to find each farmer depends on how easy it is to find them. Thus, we expect that search cost will affect the maximum price the trader is willing to offer to each farmer.

- Hypothesis 2: The expected profits for farmers with lower search costs for traders are higher as these farmers are more likely to engage in high cost investment.

Third, awareness of the marketing channel explains the heterogeneity in profitability among farmers whose product enters the modern retail channel. On the one hand, we expect that aware farmers stand at a relatively better bargaining position as compared to their unaware neighbors, given the knowledge of their products' marketing channel. On the other hand, awareness can be thought of to lower search cost. As aware farmers are more likely to initiate contact or signal their ability to produce high quality output to the modern channel traders themselves, they lower traders search cost to find them.

- Hypothesis 3: Aware farmers earn higher profit than unaware farmers.

3. Empirical Methodology

Farm household income is determined by the income earned from agricultural production activities and by various socioeconomic factors. Since higher production capacity and productivity can lead to more investment and thus result in better output quality, we assume that the investment level directly affects the quality of the output produced, raising the probability that the farmer will receive a premium price for the product. High or low cost investment is thus a binary decision farm households make in order to maximize their net return from production (chili profit) and net household income.

Since the investment decision is endogenously determined, such that producers self-select into making high cost investment for production, the coefficient estimated based on OLS may be biased due to omitted variables in the specification. To eliminate this bias we use the treatment effect model, also known asthe Heckman self-selection model, to get an unbiased and consistent estimate of the effect of making high cost production investment on profit and income. The method is more common in the impact assessment literature, particularly when the treatment is not randomly applied (Greene, 2008; Maddala, 1986, 1983; Heckman, 1979, 1978). The strategy is to apply a two-stage estimation procedure. While the first stage evaluates the probability of making high cost investment (modern channel participation) with a probit model, the second stage uses the prediction from the first stage to evaluate the impact of the treatment on the welfare outcome. Maximum likelihood estimation is used, with all parameters estimated simultaneously.

The model estimates the effect of an endogenous binary treatment z_i on a continuous and fully observable variable y_i, conditioned on the independent variables x_i and w_i. The primary interest is in the welfare/output model:

(8) $y_i = x_i\psi + z_i\vartheta + v_i$ (2nd stage, output model)

where z_i is the endogenous dummy variable indicating whether the treatment is applied. The coefficient ϑ captures the effect of the endogenous decision, that is making high cost investment. The binary decision of whether the treatment is applied is modeled as the outcome of an unobserved latent variable z_i^*. It is assumed that z_i^* is a linear function of the exogenous covariates w_i and a random component u_i, that is: $z_i^* = w_i\theta + u_i$. The observed binary investment decision is:

(9) $z_i = \begin{cases} 1 & if\ z_i^* > 0 \\ 0 & otherwise. \end{cases}$ (1st stage selection model)

where u and v are bivariate normal with mean zero and covariance matrix $\begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}$.

Identification requires that at least one variable appearing in the selection equation is not included in the outcome equation. The identification variable often used in the literature is distance. Rao and Quim (2010), for example, use distance from a farmer's house to a tarmac road, while others used distance to a paved highway (Hernández et al., 2007) and to village head's land (Miyata et al., 2009). Distance from house to asphalt road is used as identification variable in this paper.

4. Data description

In this section, we first describe the survey and show how the three groups of producers (aware, unaware and traditional) are identified. The next sections provide summary statistics of the main variables and t-tests of difference in mean value between each group of farmers.

4.1 Survey design, awareness, and channel identification

Data for this study is collected by Sahara[9] in 2010 through a household survey of chili producers in West Java, Indonesia. There are in total 597 valid samples. Although chilies are one of the most important food crops consumed daily by Indonesians, the amount purchased through the modern channel is still limited. As a result, the share of chili supplying to the supermarket is small. The survey sampling is stratified based on the marketing channel to insure sufficient representation of producers from both the traditional and modern channel.

Producers who are aware of selling their chili to the supermarket are identified by their response from "experience with the modern channel". Those with a positive response of having sold chili to supermarkets are categorized into the aware farmers group. Identifying unaware modern channel producers is more difficult. This is done by an additional 'channel' identifier in the data, which identifies modern channel producers apart from those traditional channel producers. The channel identifier is available by asking the supermarket for a list of

[9] See Sahara (2012) for more information on survey design, data collection and identification of the modern channel supplier (see also Sahara et al., 2011).

suppliers in which they source their chili product from at the stage of the survey design to insure the modern channel producers are sufficiently represented in the sampling. With the channel identifier and famers response in their experience to the modern channel, unaware modern channel producers can be singled out as they denied selling chili product to the modern channel (have no experience), while they are at the same time identified as supplier by the supermarket. This gives us three separeate groups of chili producers: aware (A), unaware (U) and traditional (T) farmers. In our data there are 485 (81%) traditional farmers, 66 (11%) unaware farmers, and 46 (7.7%) aware farmers (Table 1).

Table 1. Numbers of observations in each group

	Channels		
Farmer type	Traditional	Modern	sum
Traditional (T)	485	0	485
Unaware (U)	0	66	66
Aware (A)	0	46	46
sum	485	112	597

4.2 Descriptive analysis

4.2.1 Welfare, dependent variables – Profit and Net income

Farmers' net income is derived by aggregating income from all of the income-earning activities for the household. The income from agricultural production related activities was calculated by subtracting the revenue earned from the costs associated with production. Aware farmers on average have significantly more net income than unaware farmers by 13 million Rupiah (Rp), approximately USD 1,350 a year. On average, modern channel producers' (aware and unaware combined) net income is significantly greater than traditional channel producers, but the difference between unaware farmers and traditional producers alone is not significantly different (Table 2, row 1). Disaggregating the net income by source, we see that net profit from chili production is on average 18.7 million for the aware farmers, followed by 8.8 million for the unaware and 6.1 million for the traditional farmers (Table 2, row 2). Net income from chili production thus accounts for a high percentage of income for most farmers, namely 26, 33, and 45 per cent for traditional, unaware, and aware farmers, respectively. Traditional channel producers have a greater proportion of income derived from other agricultural activities. Net income from other agricultural activities alone contributed to 30, 24 and 15 percent of the total net income for the traditional, unaware and aware farmers, respectively. Other significantly different income sources between the groups of farmers are:

agricultural trade, other trading activities, and net income from assistance program.[10]

*Table 2. Source of net income**

	mean			test statistics (p-value)			
Unit: million Rp	Traditional	Unaware	Aware	AvU	AvT	UvT	Channel
Net income	23.71	26.61	41.49	**0.10**	**0.00**	0.43	**0.01**
Net income from various income activities							
Chili	6.12	8.79	18.74	**0.03**	**0.00**	0.15	**0.00**
other Agri.	7.06	6.45	6.38	0.98	0.83	0.78	0.74
Livestock	0.72	0.47	0.65	0.63	0.88	0.45	0.52
Aquaculture	0.08	0.16	0.38	0.16	**0.00**	0.32	**0.01**
Agri. Trade	1.34	0.00	3.84	0.15	**0.06**	**0.09**	0.75
other trading	2.16	2.82	4.88	0.31	**0.09**	0.60	0.15
Rice / milling	0.28	0.00	0.00	na	0.60	0.53	0.41
Food processing	0.14	0.32	0.20	0.71	0.73	0.29	0.32
Other business	1.71	2.92	1.43	0.49	0.83	0.27	0.49
Agri. Wage labor	0.81	0.59	1.00	0.28	0.63	0.55	0.89
non-agi. Employment	2.48	3.32	2.47	0.66	0.97	0.39	0.50
pension	0.21	0.02	0.91	0.10	**0.07**	0.51	0.47
Remittances	0.38	0.62	0.42	0.57	0.86	0.28	0.37
Assistance program	0.01	0.10	0.05	0.37	**0.08**	**0.00**	**0.00**
Other.	0.20	0.06	0.15	0.44	0.87	0.54	0.57

* Excluding 17 obs. with negative net income, of which 15 traditional farmers and 2 unaware farmers group.

4.2.2 Welfare, independent variables

Five broad classes of explanatory factors are used in the models explaining farmers' welfare: household characteristics, land size & ownership attributes, non-land assets ownerships, chili production attributes, and habit of record keeping. Summary statistics of the variables used in the analysis are provided in Table 3, with cross group comparison of the mean value reported.

Household characteristics

As already pointed out by Sahara (2012), and similar to other studies, the data show that the modern channel participants are younger and receive more education as compared to the traditional channel producers (Rao and Qaim, 2010; Hernández et al., 2007). Among modern channel producers, household characteristics of unaware famers are not significantly different from aware farmers (Table 3, column 4: AvU). The significantly higher education received among supermarket producers suggests that these households have higher human capital than traditional channel producers. Boselie et al. (2003) have shown that farmers with higher human capital are more likely to improve their negotiation skills in dealing with the buyer.

[10] We use total net income and net income from chili production (chili profit) in logarithm form as dependent variables for our analysis. This reduces the number of observation due to negative net income values by 17 (2 from unaware and 15 from traditional) and by 120 due to negative profit from chili production (101 from traditional, 14 from unawary, and 5 from aware).

Table 3. Summary statistics of variables used, across groups comparison

	mean			Test statistics (p-value)			
	Traditional	Unaware	Aware	AvU	AvT	UvT	Channel
Household characteristics variables							
Household size (number)	4.58	4.23	4.46	0.40	0.61	0.09	0.12
Age (years)	46.29	44.03	42.78	0.50	**0.04**	0.12	**0.02**
Age spouse	40.13	38.88	36.09	0.19	0.01	0.37	**0.03**
Education (years)	6.43	8.02	7.96	0.92	**0.00**	**0.00**	**0.00**
Education spouse	6.63	7.66	8.05	0.49	**0.00**	**0.00**	**0.00**
Age between15-65 (%)	65.95	64.05	66.43	0.52	0.87	0.47	0.66
Age above 65 (%)	2.42	3.66	3.91	0.90	0.31	0.34	0.19
distance to road (km)	0.30	0.11	0.12	0.70	**0.07**	**0.02**	**0.00**
--- to sub-district market (km)	6.05	3.95	7.80	**0.00**	**0.03**	**0.00**	0.34
Total land size owned and cultivating land ownership and types							
Land owned							
Year 2005	4,079.00	4,876.41	3,195.34	0.40	0.35	0.40	0.89
Year 2010	4,544.28	5,145.24	3,432.92	0.40	0.29	0.56	0.90
Land cultivated							
Total (m^2)	6,433.41	8,089.61	8,097.52	1.00	0.18	0.11	**0.05**
% owned	51.65	44.72	36.22	0.23	**0.01**	0.20	**0.01**
% rented	24.35	32.39	50.75	**0.01**	**0.00**	**0.10**	**0.00**
% others	24.00	22.89	13.03	**0.10**	**0.05**	0.81	0.16
% irrigated	46.28	45.32	65.04	**0.01**	**0.01**	0.87	0.13
% Rain fed	11.51	7.14	5.90	0.75	0.21	0.24	**0.09**
% Dry land	42.21	47.54	29.06	**0.01**	**0.05**	0.36	0.62
Non-land assets variables (numbers of following asset)							
Assets in 2010							
mobile phone	1.19	1.47	1.61	0.57	**0.02**	**0.06**	**0.00**
motorbike	0.64	0.67	0.76	0.45	0.28	0.75	0.37
water pump	0.28	0.35	0.24	0.26	0.58	0.32	0.69
mist blower	1.12	1.30	1.41	0.52	**0.02**	**0.10**	**0.01**
power tiller	0.01	0.00	0.07	0.11	0.04	0.39	0.44
storage house	0.19	0.41	0.39	0.88	**0.00**	**0.00**	**0.00**
Assets in 2005							
mobile phone	0.53	0.45	0.87	**0.03**	**0.02**	0.52	0.31
motorbike	0.42	0.52	0.61	0.49	**0.06**	0.25	**0.05**
water pump	0.25	0.29	0.26	0.77	0.92	0.67	0.70
mist blower	0.96	0.79	1.17	**0.02**	0.17	0.18	0.87
power tiller	0.01	0.00	0.07	0.11	**0.03**	0.33	0.40
storage house	0.15	0.20	0.35	0.11	**0.00**	0.35	**0.01**
Chili production							
Area (m2)	3,137.45	4,465.80	5,304.97	0.55	**0.01**	**0.02**	**0.00**
Cost (million Rp)	11.28	20.81	22.56	0.79	**0.00**	**0.00**	**0.00**
Output(kg)	4,408.72	4,927.61	6,366.65	0.42	**0.08**	0.53	0.13
Types of chili planted	1.15	1.41	1.41	0.97	**0.00**	**0.00**	**0.00**
Seasons	1.24	1.62	1.67	0.68	**0.00**	**0.00**	**0.00**
Productivity (kg/ m^2)*	2.40	1.09	1.13	0.73	**0.07**	**0.02**	**0.00**
Cost per m^2 (Rp/ m^2) *	6408.72	4862.84	3583.01	**0.03**	0.14	0.34	**0.10**
Grading+	2.98	3.22	3.53	**0.02**	**0.00**	**0.06**	**0.00**
Price (Rp/kg)	6,236.60	7,637.88	8,976.27	**0.07**	**0.00**	**0.01**	**0.00**
Cost per kg (Rp/kg)*	3,052.92	4,992.44	3,900.03	**0.02**	**0.00**	**0.00**	**0.00**
Record keeping (Bivariate dummy, yes=1, no=0)							
Pesticide	0.12	0.47	0.46	0.89	**0.00**	**0.00**	**0.00**
date pesticide	0.06	0.15	0.15	0.99	**0.01**	**0.00**	**0.00**
price received	0.22	0.76	0.89	**0.08**	**0.00**	**0.00**	**0.00**
quantity output	0.21	0.76	0.85	0.25	**0.00**	**0.00**	**0.00**
For over 1 year	0.40	0.29	0.37	0.43	0.69	0.16	0.24

* outliers dropped. + See footnote 5 for detail of the grading scale. Significant t-test statistics are marked in bold.

Young producers are more willing to adapt to new technologies and production methods. The percentage of family members in different age categories (young / working / retired) gives an indication of the dependency ratio and the percentage of family member that can contribute to the labor force. The difference in age categories is, however, not significant.

Distance from the house to the nearest asphalt road is on average 200 meters shorter for modern channel producers as compared to traditional channel producers, and not statistically different between aware and unaware farmers. Distance from the house to a district center, in contrast, is significantly different between the groups. Aware farmers tend to live further away from the district market than traditional farmers, while unaware farmers live closer by.

Land size and ownership attributes

Farmers own on average 4,525 m^2 of land. The size is similar for the three groups of farmers. Ten percent have more than 1 hectare, while around 70 percent have less than 0.5 hectare, which makes it difficult to benefit from scale economies. Modern channel farmers cultivate a greater land area compared to traditional farmers (by 1,500 m^2) because they rent a larger area: aware farmers rent 3,458 m^2 on average (1.8 times greater than unaware farmers and 3.16 times greater than traditional farmers). The rented areas contribute to 51 percent of the total land area cultivated by aware farmers (compared to 32 and 24 percent for unaware and traditional farmers). The difference between each group is statistically different. Considering land quality, aware famers have a greater share of cultivated land irrigated (65%) and a much smaller share of dry land (29%) compared to both unaware and traditional farmers. Unaware farmers also have less irrigated and more dry land than the traditional farmers. Overall, 17 percent of all farmers cultivate land area over 1 hectare.

Non-land assets

Non-land assets are often pointed out in the literature as important determinants for modern channel participation. In 2005 aware farmers have on average significantly more mobile phones than unaware farmers or traditional farmers, which facilitates them to obtain market information and communicate with the traders. In 2010, a lot more farmers have mobile phones for communication, such that the difference between aware and unaware farmers disappeared, but the difference between modern and traditional farmers is significant. Transportation assets have been suggested to be critical for becoming a modern channel participants since supermarkets desire more frequent and timely delivery (Reardon et al., 2005). In our data, modern channel participants have more motor bikes in 2005 than traditional farmers. As for farming equipment, modern channel producers have more storage available. Aware farmers also have more mist blowers in 2005. Ownership of water pump and power tiller, however, do not differ across groups. As product appearance and quality consistency are important aspects emphasized by supermarkets better storage and equipment facilitating quality control can help farmers to meet supermarket standards.

Production attributes

Farmers differ in production efficiency. On average, modern channel farmers show a higher level of specialization in chili production. Aware and unaware farmers on average cultivate a larger area of chili, invest more in chili production and produce more outputs (kilos) than traditional channel farmers. The difference in output between aware and traditional farmers is nearly 2,000 kilos in 2010. As an additional indication for specialization, modern channel producers also plant more types of chili and are active in more seasons. The production costs also differ significantly between the three groups of farmers. Although modern channel farmers produce more than traditional farmers they also invest more. The total production cost is significantly higher for modern channel producers, but is insignificantly different between aware and unaware farmers.

Considering the differences in cost and land invested for chili production, the modern channel farmers seem to be the least productive (output per square meter), yet the most cost efficient ones (cost per square meters). The modern channel farmers produce on average 1.1 kilos of chili per square meter of cultivated land, while traditional producers more than double that amount. In line with the procurement quality requirements of the supermarket, chili produced by the modern producers receives better grading[11] and is sold at a higher price. The aware farmers received the highest average grading (3.53) for all three seasons, followed by the unaware farmers (3.22) and the traditional farmers with lowest grading (2.98).[12] In line with this, the average price (over three seasons) is highest for the aware farmers (8,976 Rp/kg) and lowest for traditional farmers (6,237 Rp/kg), with unaware farmers in between. The costs

---

[11] The grading inquiry in the survey was of 9 grading level, starting from the highest (1. *Superfull*), followed by: (2. *Super*), (3. *Medium*), (4. *Small*), (5. mix of 1 and 2), (6. mix of 2 and 3), (7. mixed of 3 and 4), (8. Other) and (9. no grading), consecutively. The grading reported in table 6 is based on a converted 5-level grading scale, which marked *superfull* as class A, combined 2 and 5 into class B, combined 3and 6 into class C, combined 3 and 6 into class D, and lastly the original 8 and 9 combined into lowest class E. The new A to E scale is than coded into 1 to 5 for quantitative comparison, with 5 as the highest grade and 1 as the lowest.

[12] We report only the average grading received over the three seasons. Detail of the separate grading for each season is available upon request. All producers receive higher chili grading in the early season (1.April) as compared to chili output in the later seasons (2.July and 3.September). Out of three seasons, the mean difference in grading is statistically significant between aware and unaware farmers in season 1 and 2, while the mean grading differences is statistically different between farmers of different channel in all three seasons.

required to produce better chili with better quality and higher price translates into higher investment per kilo of output. Unaware farmers invest most per kilo of output, followed by the aware and the traditional farmers. The difference between the three groups is statistically significant. Overall, the profit margin per kilo of chili is highest for the aware farmers, followed by the traditional and the unaware farmers.

Table 4. Summary statistics

Variables	Definition (units)	mean	s.d.
Dependent variables			
Net income	Net household income from various source in log (Rp.)	16.548	1.040
Chili profit	Net profit (revenue less cost) from chili production in log (Rp.)	15.228	1.413
Variables of interest			
Chili cost	Total cost invested for chili production in log (Rp.)	15.758	1.141
High Cost	Total cost invested for chili production above the median of all sampled farmers (dummy)	0.500	0.500
High cost per m^2	Total cost invested for chili production per square meter above the median of all sampled farmers (dummy)	0.499	0.500
Traditional	Whether farmers supply chili through the traditional channel (dummy)	0.812	0.391
Unaware	Farmers who supply to the supermarket but thought they are traditional channel producers (dummy=1 if unaware)	0.111	0.314
Aware	Farmers who know they are supplying to the supermarket (=1 if aware)	0.078	0.267
Productivity (γ)	Chili output produced per unit of land in log (KG/m^2)	0.155	0.996
Control variables			
Grading	Average chili grading received. Range between 1 to 5, with 5 as the best grading and 1 for the chili with the worst or without grading	3.047	0.985
Household size	Numbers of family members in the household (persons)	4.531	1.570
Age	Age of the household head (years)	45.769	10.979
Education	Years of education of the head of the household (years)	6.725	3.019
% working age	Proportion of family member between the age 15 to 65 (%)	0.658	0.197
% retired age	Proportion of family member above age 65 (%)	0.027	0.098
Experience	Years of experience in chili production (years)	8.920	6.750
Mobile phone	Mobile phone ownership in 2005 (5 years ago from the time of the survey (dummy)	0.745	0.436
Motor bike	Motor bike ownership in 2005 (dummy)	0.529	0.500
Water pump	Water pump ownership in 2005 (dummy)	0.265	0.442
Mist blower	Mist blower ownership in 2005 (dummy)	0.848	0.360
Power tiller	Power tiller ownership in 2005 (dummy)	0.013	0.115
Storage house	Storage house ownership in 2005 (dummy)	0.221	0.415
Chili land	Total land size devote to chili production (m^2) in 2010 (current asset)	8.280	1.144
Land owned	Land size owned (m^2) in 2005	6.055	3.663
% rented	Proportion of cultivated land rented (%)	0.273	0.378
Prod. record	Whether farmers keep production record, include: pesticide use and date of pesticide application on chili (dummy)	0.128	0.285
Market. record	Whether farmers keep marketing record, include: chili price, quantity sold (dummy)	0.327	0.464
Exogenous variable			
Distance	Distance to the nearest asphalt road	0.262	0.595

Record keeping

The percentage of record keeping among modern channel producers is double that of traditional producers. A significantly larger share keeps records on pesticide use, date of pesticide applied, price received, and quantities sold. The percentages vary a lot. Better educated farmers may not only have the ability to keep written records but also benefit from the insights delivered through better use of knowledge. We merge the 5 record keeping items into 2 record behavior categories, those related to production and those related to marketing.

Bivariate independent variables—high cost production dummy

We set an arbitrary cutoff investment level to distinguish between high cost and low cost producers because the true critical investment level required to produce high quality chili is unknown. In the treatment effect analysis, the critical bivariate investment decision variable is defined as the farmers who made investment above the median level among all sampled farmers. Thus by definition half of the households invest in high cost production. Similarly, we generate a relatively high cost per square meter dummy variable as a robustness check. Farmers whose investment in chili production is above the median level per square meter of land cultivated are considered high cost producers. Table 4 gives the summary statistics of all variables used.

5. Regression analysis

This section reports the empirical analysis of the self-selection phenomenon and the effect of investment choice, channel choice and awareness on the profitability of chili production and overall income.

5.1 Investment choice and effect on profit model

The output from the treatment effect model is split into two parts, with the estimation from the first stage selection equation reported in the lower panel and the second stage welfare output reported in the upper panel of Table 5. In the upper panel, the result from the ordinary least square (OLS) estimation is reported in columns 1 and 3 and of the treatment effect model in columns 2 and 4. In the lower panel, we show independent probit estimation in columns 1 and 3 next to the first stage outcome from the treatment effect model in columns 2 and 4. We demonstrate two model specifications, with and without controlling for record keeping habits (columns 3 and 4, respectively 1 and 2). The following models are estimated:

$$\text{high cost} = \psi(\text{ productivity, controls}) + \theta(\text{distance}) + u \qquad (1^{\text{st}} \text{ stage})$$

$$\text{chili profit} = \vartheta(\text{highCost}) + \psi(\text{awareness, productivity, controls}) + v \qquad (2^{\text{nd}} \text{ stage})$$

Table 5. Treatment effect, OLS and Probit regression output – chili profit

	(1)		(2)		(3)		(4)	
a. Outcome model: OLS and 2nd stage of the treatment effect model; dependent variable: chili profit.								
Model:	OLS		Treatment 2nd stage		OLS		Treatment 2nd stage	
High Cost	0.742***	(6.306)	1.483***	(4.184)	0.698***	(5.938)	1.405***	(4.143)
Unaware	0.424**	(2.445)	0.405**	(2.381)	0.229	(1.256)	0.223*	(1.700)
Aware	1.035***	(5.422)	0.996***	(5.306)	0.807***	(3.964)	0.772***	(4.554)
Productivity	0.392***	(5.761)	0.363***	(5.145)	0.396***	(5.866)	0.366***	(4.525)
Grading	0.148**	(2.517)	0.140**	(2.418)	0.143**	(2.447)	0.135**	(2.200)
Household size	-0.047	(-1.19)	-0.051	(-1.255)	-0.041	(-1.04)	-0.047	(-1.165)
Age	-0.013**	(-2.12)	-0.013**	(-2.036)	-0.012*	(-1.91)	-0.012**	(-1.979)
Education	-0.010	(-0.55)	-0.023	(-1.136)	-0.017	(-0.91)	-0.026	(-1.315)
% working age	0.359	(1.056)	0.431	(1.230)	0.445	(1.317)	0.481	(1.396)
% retired age	-0.891	(-1.39)	-0.807	(-1.225)	-0.912	(-1.44)	-0.819	(-1.367)
Mobile phone	0.013	(0.103)	-0.042	(-0.314)	-0.020	(-0.16)	-0.063	(-0.491)
Motor bike	0.236**	(2.031)	0.181	(1.484)	0.251**	(2.171)	0.192*	(1.689)
Water pump	0.086	(0.675)	0.004	(0.029)	0.060	(0.475)	-0.010	(-0.078)
Mist blower	0.278*	(1.759)	0.137	(0.786)	0.258	(1.646)	0.135	(0.827)
Power tiller	0.431	(0.912)	0.498	(1.024)	0.503	(1.074)	0.558	(1.255)
Storage house	0.255*	(1.910)	0.196	(1.405)	0.209	(1.569)	0.174	(1.425)
Experience	0.014	(1.619)	0.011	(1.252)	0.016*	(1.845)	0.012	(1.280)
Chili land	0.362***	(5.656)	0.348***	(5.486)	0.360***	(5.653)	0.350***	(4.975)
Prod. record					0.163	(0.775)	-0.004	(-0.017)
Market. record					0.358**	(2.432)	0.319**	(2.345)
Constant	11.16***	(17.99)	11.23***	(18.042)	11.04***	(17.91)	11.13***	(18.026)
adj. R^2	0.383				0.395			
b. Selection model: bivariate investment decision; dependent variable: high cost (dummy)								
Model:	Probit		Treatment 1st stage		Probit		Treatment 1st stage	
Productivity	0.158***	(2.695)	0.115	(1.527)	0.171***	(2.888)	0.132*	(1.745)
Household size	-0.008	(-0.207)	0.007	(0.150)	-0.005	(-0.12)	0.011	(0.245)
Age	-0.003	(-0.493)	-0.003	(-0.381)	-0.001	(-0.11)	0.001	(0.090)
Education	0.050**	(2.301)	0.039*	(1.676)	0.042*	(1.852)	0.027	(1.192)
% working age	-0.335	(-0.933)	-0.315	(-0.784)	-0.277	(-0.77)	-0.225	(-0.543)
% retired age	-0.373	(-0.549)	-0.325	(-0.434)	-0.445	(-0.65)	-0.411	(-0.482)
Mobile phone	0.329**	(2.437)	0.262*	(1.747)	0.312**	(2.275)	0.216	(1.438)
Truck	0.195	(1.622)	0.268**	(1.998)	0.220*	(1.797)	0.316**	(2.275)
Motor bike	0.282**	(2.096)	0.262*	(1.726)	0.260*	(1.916)	0.223	(1.485)
Water pump	0.47***	(2.859)	0.491***	(2.590)	0.46***	(2.763)	0.456**	(2.366)
Mist blower	-0.222	(-0.450)	-0.206	(-0.350)	-0.232	(-0.47)	-0.230	(-0.418)
Power tiller	0.125	(0.888)	0.189	(1.246)	0.076	(0.527)	0.138	(0.907)
Storage house	0.016*	(1.758)	0.013	(1.311)	0.019**	(2.071)	0.016	(1.502)
Experience	-0.008	(-0.207)	0.058***	(2.883)	-0.005	(-0.12)	0.132*	(1.745)
Land owned	0.07***	(3.702)	0.739***	(4.049)	0.07***	(3.433)	0.05***	(2.654)
% land rented	0.71***	(4.082)	0.115	(1.527)	0.73***	(4.088)	0.76***	(4.108)
Prod. record					0.76***	(3.031)	0.81***	(3.003)
Market. record					-0.003	(-0.02)	0.049	(0.306)
Distance to road	-0.169	(-1.612)	-0.239**	(-2.183)	-0.157	(-1.49)	-0.215**	(-2.189)
Constant	-1.5***	(-3.446)	-1.39***	(-2.817)	-1.69***	(-3.83)	-1.63***	(-3.244)
Pseudo R^2	0.121				0.137			
N	462		462		462		462	
athrho			-0.436**	(-2.172)			-0.419**	(-2.063)
lnsigma			0.132**	(2.425)			0.116**	(2.168)
Wald chi^2			247.054	p=0.00^			273.568	p=0.00^
Independent test			3.536	p=0.00^			4.255	p=0.04^

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$; ^ are p-value for Prob. > chi2

The first stage decision decides between entering into relatively high and relatively low cost production, which relates to higher and lower probability of supplying to the supermarket channel. This is significantly correlated with farmers' productivity (Table 5, lower panel). More productive farmers are more likely to invest in high cost for chili production, even after controlling for farmers' household characteristics, land ownership and other assets. Among the control factors, we see that more educated and more experienced farmers are more likely to invest in high cost production. In addition, non-land assets such as motor bikes and water pumps positively correlate with the endogenous decision in making high cost investment. The instrument variable, distance to road, is statistically significant. The negative coefficients suggest that farmers who live at more remote locations are less likely to invest in high cost production. This is in line with what we expected: lower search cost for traders makes it more likely for farmers to engage in high cost production.

In the second stage, the *high cost investment* decision dummy variable significantly affects farmers' profit from chili production in all specifications (Table 5, upper panel). The estimates from the treatment effect model indicates that other things being equal, farmers who invest in high cost production earn on average 140 percent higher profits than farmers who invest in relatively low cost production. The difference is net of the observed selection bias concerning farmers' investment choice in the first stage. This is much higher than the 70 percent suggested by the OLS estimation. In addition, the estimated coefficient regarding farmers' productivity is a statistically significant factor affecting both the investment decision (1^{st} stage) and the profitability of chili production (2^{nd} stage). The results show that more productive farmers are not only more likely to invest in high cost production; their profits are also higher as compared to less productive farmers. The result mirrors our earlier prediction from the theoretical model regarding the productivity sorting of farmers into different investment level, and how it affects profitability.

The coefficient for the aware and unaware farmers dummy variables are both positive and statistically significant, suggesting that modern channel participation gives producers additional profits as compared to the traditional farmers, even after controlling for the investment differences. More interestingly, the profit premium is significantly lower for unaware farmers as compared to the aware farmers. The aware farmers earn on average 72 percent higher profits than traditional channel farmers, while the profit premium for unaware farmers is only 22 percent. Since both aware and unaware farmers are modern channel producers, the difference in this premium has important implications as it suggests that participation in the modern retail channel provides additional profits which depend on the awareness of participation. The lower profit premium for the unaware farmers can potentially be explained by the difference in bargaining power or farmers' ability in assessing traders' willingness to pay. Farmers have to be aware of the marketing channel of their products to capture a greater proportion of the profit margin from supplying to the modern channel. In all specifications, we controlled for the quality of the output (grading), land area for chili production, other household characteristics and non-land assets owned.

The sign of estimated coefficients from the OLS and probit model as compared to the treatment effect model is the same for all the variables with significant estimates. The standard deviations of the estimated coefficients from the treatment effect model are much smaller for the key variables. The coefficients for most variables are not very different, with the exception of the estimated effect of the high-cost bivariate dummy, which is much higher for the treatment effect model than the outcome for the OLS estimation.

An interesting result shows up when record keeping behavior is distinguished between production and marketing records. We see that production record keeping (pesticide application dates and amount of pesticide used) is significant in determining the *first stage* investment decision, while marketing record keeping (price and quantity) is significant in affecting the *second stage* profit outcome. This result reinforces the link we find between productivity and the investment decision. Production records as a knowledge base for self-evaluation of productivity can support investment decisions, while marketing records provide a base for assessing the market conditions and rewards farmers with higher profit.

Lastly, to justify our model choice, the reported Wald test of all coefficients in the output (second stage) model being zero is rejected in both of the treatment effect specifications. The test statistics are $\chi^2 = 152$ and $\chi^2 = 282$, with $p < 0.01$, indicating the covariates used in the model are appropriate and at least one of the covariates used in the regression is not equal to zero. In addition, if $\rho = 0$, the errors u and v are independent and there is no endogeneity problem. The Wald test for independence reported at the bottom of Table 5 compares the joint likelihood of an independent probit model for the selection equation and an OLS regression model on the observed data against the treatment effect model likelihood. The likelihood test for $H_0 : \rho = 0$ gives $\chi^2 = 3.96$ and $\chi^2 = 5.68$ with both $p < 0.01$ rejects the null hypothesis at a statistically significant level. This suggests that using OLS regression results in biased estimation of the effect of high cost investment, and instead applying the treatment effect model is indeed appropriate since it allows the correlation between the two error terms to be non-zero ($\rho \neq 0$).

5.2 Channel choice and effect on net income

In this section, we explore the influence of participating in the modern food retail channel on income. Knowing the importance of investment on quality which may further influence modern channel participation, we estimate the following two stage treatment effect model:

$$\text{channel choice} = \theta(\text{distance}) + \psi(\text{cost, productivity, controls}) + u \qquad (1^{st}\text{ stage})$$

$$\text{net income} = \vartheta(\text{channel choice}) + \psi(\text{awareness, productivity, controls}) + v \qquad (2^{nd}\text{ stage})$$

Notice that we included the awareness indicator into the specification. The results are reported in Table 6, which is organized in the same way as Table 5, with record keeping behavior in the second specification, the probit and OLS regression reported in columns 1 and 3 and the first and second stage of the treatment effect model reported in columns 2 and 4.

From the outcome model, the positive and significant coefficients suggest that participating in the modern food retail channel increases farmers' net income by 75 percent as compared to non-participating farmers. This corroborates the earlier finding of a positive correlation between participation and income. In contrast, using the OLS model instead gives (negative) insignificant coefficients for the modern channel variable. In addition, our awareness indicator captures the additional welfare loss for the unaware farmers. After controlling for the first stage modern channel participation choice, awareness increases farmers' net income by an additional 37 percent. The effect of channel choice is stronger when we control for record keeping behavior (column 4).

From the selection model, the positive and significant coefficient of the chili costs suggests that farmers who invest more are more likely to become modern channel producers. The coefficient for the identification variable, distance to asphalt road, is negative and statistically significant in both treatment effect specifications. In the probit model, the distance coefficient instead become insignificant when record keeping behavior is included in the model.

Table 6. Treatment effect, OLS and Probit regression output -- Income

	(1)		(2)		(3)		(4)	
a. Outcome model: total net income.								
Model:	OLS		2nd stage treatment		OLS		2nd stage treatment	
Channel (0/1)	-0.179	(-1.42)	0.693***	(4.12)	-0.216	(-1.60)	0.750***	(4.26)
Aware (0/1)	0.365**	(2.01)	0.408**	(2.39)	0.367**	(2.01)	0.367**	(2.16)
Household size	0.025	(0.99)	0.029	(1.08)	0.026	(1.04)	0.029	(1.11)
Age	0.003	(0.61)	0.005	(1.05)	0.003	(0.69)	0.004	(0.89)
Education	0.033**	(2.30)	0.017	(1.15)	0.030**	(2.14)	0.023	(1.56)
% working age	0.007	(0.03)	-0.043	(-0.16)	0.036	(0.14)	-0.033	(-0.12)
% retired age	-0.260	(-0.64)	-0.588	(-1.27)	-0.266	(-0.66)	-0.515	(-1.13)
Mobile phone	0.30***	(3.22)	0.290***	(2.97)	0.296***	(3.14)	0.318***	(3.23)
Motor bike	0.44***	(5.21)	0.413***	(4.75)	0.440***	(5.24)	0.423***	(4.88)
Water pump	0.184**	(2.09)	0.222**	(2.40)	0.178**	(2.03)	0.230**	(2.50)
Mist blower	0.398***	(3.52)	0.337***	(2.88)	0.398***	(3.51)	0.352***	(3.05)
Power tiller	0.478*	(1.70)	0.507**	(2.12)	0.477*	(1.76)	0.443*	(1.91)
Storage house	0.513***	(5.75)	0.416***	(4.30)	0.498***	(5.59)	0.462***	(4.90)
Experience	0.016***	(2.81)	0.024***	(3.80)	0.017***	(2.92)	0.022***	(3.58)
Land owned	0.025*	(1.92)	0.017	(1.33)	0.023*	(1.77)	0.016	(1.28)
% land rented	0.307***	(2.60)	0.253**	(2.21)	0.312***	(2.66)	0.303***	(2.68)
Prod. record					0.240	(1.62)	0.215	(1.34)
Market. record					-0.017	(-0.15)	-0.385***	(-3.09)
Constant	14.75***	(45.03)	14.684***	(43.810)	14.714***	(44.52)	14.736***	(44.16)
Adj. R2	0.261				0.262			
b. Selection model on bivariate channel choice decision								
Model:	Probit		1st stage treatment		Probit		1st stage treatment	
Chili cost	0.255***	(3.54)	0.323***	(4.70)	0.181**	(2.28)	0.251***	(3.59)
Household size	-0.027	(-0.55)	-0.019	(-0.42)	-0.036	(-0.67)	-0.032	(-0.63)
Age	-0.008	(-1.05)	-0.008	(-1.20)	-0.001	(-0.17)	-0.005	(-0.567)
Education	0.057**	(2.39)	0.054**	(2.48)	0.033	(1.25)	0.020	(0.75)
% working age	0.244	(0.58)	0.216	(0.52)	0.334	(0.71)	0.297	(0.65)
% retired age	1.199	(1.63)	1.396*	(1.87)	1.018	(1.30)	1.291	(1.62)
Mobile phone	0.031	(0.19)	-0.063	(-0.40)	-0.088	(-0.48)	-0.178	(-1.05)
Motor bike	0.027	(0.19)	0.004	(0.03)	0.035	(0.23)	-0.020	(-0.14)
Water pump	-0.255	(-1.62)	-0.313**	(-2.11)	-0.292*	(-1.70)	-0.333**	(-1.96)
Mist blower	0.206	(1.02)	0.199	(0.99)	0.151	(0.67)	0.227	(1.17)
Power tiller	-0.290	(-0.52)	-0.149	(-0.40)	-0.004	(-0.01)	0.142	(0.43)
Storage house	0.325**	(2.06)	0.234	(1.47)	0.116	(0.66)	-0.039	(-0.22)
Experience	-0.05***	(-4.02)	-0.047***	(-4.27)	-0.040***	(-2.87)	-0.035***	(-3.02)
Chili land	0.072	(0.94)	0.125*	(1.65)	0.107	(1.178)	0.213**	(2.33)
Prod. record					-0.050	(-0.21)	-0.028	(-0.12)
Market. record					1.389***	(8.14)	1.426***	(7.96)
Distance to road	-0.447**	(-2.14)	-0.484***	(-3.552)	-0.305	(-1.22)	-0.452**	(-2.50)
Constant	-5.40***	(-4.81)	-6.810***	(-6.348)	-5.260***	(-4.25)	-6.974***	(-6.24)
PseudoR2	0.146				0.299			
N	462		462		462		462	
athrho			-0.658***	(-6.815)			-0.734***	(-6.735)
lnsigma			-0.061	(-1.475)			-0.066	(-1.603)
Wald_chi2			263.449	p=0.000			45.366	p=0.000
Independent test			46.443	p=0.000			41.815	p=0.000

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$

Among the control variables, we see that most household characteristics do not have a strong correlation with farmers' channel choice, with the exception of experience and water pump ownership. Similar to other studies, we see that more experienced farmers are less likely to participate in the modern channel. It is argued that the longer the farmers have been growing chili, the stronger the relationship they formed with the buyers they are familiar with. These farmers are then less likely to seize new market opportunities, especially when it requires making adjustments in the production method for better output quality. The independent probit model explains slightly less than 14 percent of the variation, but is significantly improved to nearly 30 percent when record keeping behavior is included in the specification.

The results also suggest that farmers' net income correlates positively with farm households' non-land assets. Farm households with more experience in chili production have a higher income level. Household who have a higher percentage of farm land rented also have a higher income level. This suggests that these farmers are likely to have better knowledge and capability to mobilize their production capacity when necessary. Notice that, unlike the previous section, keeping marketing records is significantly affecting farmers' channel choice in the first stage. The result suggests that better documentation of price and quantity increases farmers' knowledge about the market. Record keeping behavior is costly, however, and shows up as a significantly negative factor for net income. Overall, the Wald test and independence test again confirm the appropriateness of using the treatment effect model.

5.3 Robustness check

As a robustness check for the chili profit model, we use cost per square meter invested instead of total investment (Table 7). Again the treatment effect model suggests that more productive farmers engage in relatively high cost (per unit land) production, which is rewarded with higher profits. The profit premium for participating in the modern channel is still positive and significant. Aware farmers are again enjoying higher profit margins for participating in the modern channel as compared to the unaware farmers. The OLS and probit model shows similar results for most of the variables, but with an insignificant coefficient for the unaware farmers and investment level in affecting chili profits. The profit premium for the aware farmers as compared to the traditional farmers is significantly positive and is double the premium enjoyed by the unaware farmers as compared to the traditional farmers. Overall, the Wald-test is again in support of the treatment effect model.

Table 7. Robustness check for chili profit

Dependent var.	(1) High cost /m2		(2) High cost /m2		(3) Chili profit		(4) Chili profit	
model	Probit		Treatment 1st stage		OLS		Treatment 2nd stage	
Productivity	0.883***	(9.38)	1.022***	(8.05)	0.449***	(5.98)	0.256**	(2.15)
High Cost					0.134	(1.06)	1.008***	(2.64)
Unaware					0.305	(1.61)	0.311**	(2.27)
Aware					0.78***	(3.67)	0.78***	(4.68)
Grading					0.144**	(2.38)	0.118*	(1.87)
Distance to road	-0.092	(-0.74)	-0.223	(-1.43)				
Production record	0.305	(1.23)	0.330	(1.29)	0.261	(1.21)	0.216	(1.03)
Marketing record	0.272*	(1.71)	0.401**	(2.30)	0.386**	(2.52)	0.240	(1.55)
Constant	-1.50***	(-3.16)	-1.92***	(-3.54)	10.2***	(16.35)	10.2***	(16.74)
Pseudo /adj. $R2$	0.258				0.348			
N	577		459		562		459	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$;
Wald chi2(21) = 280.82 (prob>chi^2=0.000), independence test = 4.543 (prob>chi^2=0.033); athrho = -0.511(-2.131)

So far, we have treated investment as a dummy variable, indicating the relatively high or low cost investment farmers engage in for chili production. To expose the full information we have from the data, we run two separate OLS regressions. The first regression is the same as our probit specification in estimating the total costs farmers invested for chili production in section 5.1. We then use the predicted cost from the first regression to explain farmers' profit from chili production for the second OLS regression. The variables used are the same as before except for the cost variable, which is now continuous. Table 8 reports the coefficients for the key variables of interest. The complete output is provided in appendix A1 and A2.

Table 8. Robustness check for chili profit with continuous cost[+]

Dependent variable	(1) Linear total Cost		(2) Chili profit	
Productivity	0.249***	(5.679)	0.363***	(3.948)
Predicted total cost			0.455**	(1.992)
Unaware			0.329*	(1.764)
Aware			0.745***	(3.531)
Grading			0.134**	(2.199)
Distance to road	0.031	(0.375)		
adj. $R2$	0.2339		0.3499	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$. [+]Full specification is the same as the model in column 4 of Table 5 with exception of different cost variable used. Coefficients for the control variables are omitted in the Table

The robustness check for the effect of the investment decision on profits again supports our earlier findings (Table 8). The coefficient for the predicted costs indicates that a 1 percent increase in costs results in a 45 percent higher chili profits, while a 1 percent increase in productivity increases farmers' total cost invested in chili production by 25 percent. In addition, supplying to the modern channel increases farmers' profits even after controlling for the investment level. Unaware modern channel producers have on average 33 percent higher income than the traditional channel producers, while aware modern channel producers have an even higher profit margin (75 percent higher) as compared to the traditional channel producers. The difference in profit margin between the aware and the unaware farmers is statistically significant

6. Conclusion

The dynamics in the food retail sector which led to the rising importance of high-value supermarkets in developing countries provides both challenges and opportunities for farmers. Marginalization of small farmers has been one of the main concerns. While a growing number of studies looks into the determinants of the modern channel participation a theoretical foundation is often lacking. In this article, we provide a simple model illustrating the self-selection of farmers into different levels of investment for production, and how the result of the self-selection affects farmers' welfare. We thus provide a theoretical foundation for analyzing the increasing demand for high quality food supplied by the supermarket. We link investment differentials to the likelihood of modern channel participation as private standards require higher input investments to insure output quality and consistency.

Empirical studies support the connection between the investment level and modern channel participation. Based on evidence for micro-farm households in Indonesia, we arrive at three main conclusions.

First, farmers who invest in high cost production are rewarded with high profits. There is clearly self-selection based on farmers' productivity for the investment level chosen. Land ownership, and ownership of other non-land assets, such as trucks and water pumps, are also important factors that endogenously determines farmers' investment choice.

Second, we find a positive and significant income effect for modern channel participation. Moreover, self-selection of participation is correlated with farmers' investment level. In other words, more productive farmers are more likely to make high levels of investment for production. More investment increases farmers' chance of producing high quality output, which raises the likelihood to become suppliers for the modern food retail channel. These results confirm the previous concerns that the requirements of supermarket in output quality and consistency pose a challenge to the resource-poor and credit constrained farmers in participating in the modern food retail channel.

Third, depending on the farmers' awareness of their participation in the modern food retail channel, farmers who are unaware of their participation can be deprived from obtaining the full profit margin as compared to their neighbors who are aware of their participation. Controlling for investment choice, the profit premium enjoyed by the aware farmers can be up to four times larger than the profit premium enjoyed by the unaware farmers. Our results mirror the two gaps pointed out by Torero and Gulati (2004), namely the 'real access gap' and the 'market efficiency gap', which need to be solved by institutional change and infrastructure (see also Torero, 2011, and Gulati et al., 2007). We show that even after farmers overcome the 'real access gap' of becoming an efficient and capable modern channel suppliers, those who fail to address the 'market efficiency gap' regarding farmers' ability to utilize market information will be unable to reap the full benefits of engaging in the modern food retail value chain.

7. References

Asfaw, A. (2007), "Supermarket purchases and the dietary patterns of households in Guatemala", IFPRI discussion paper 696. Washington, DC.

Asfawa, S., D. Mithöfer and H. Waibel (2010), Agrifood supply chain, private-sector standards, and farmers' health: evidence from Kenya", *Agricultural Economics* 40: 251-263.

Berdegué, J.A, F. Balsevich, L. Flores and T. Reardon (2005), "Central American supermarkets' private standards of quality and safety in procurement of fresh fruits and vegetables", *Food Policy*, 30(3), 254–269.

Berdegué , J. A., R. Hernández, J. Ortega, and T. Reardon (2007). Strawberry growers and modern market channels in Mexico. Micro Report module 3 of component 1. Regoverning Market Program, IIED, London.

Boselie, D., S. Henson and D. Weatherspoon (2003), "Supermarket procurement practices in developing countries: redefining the roles of the public and private sectors", *American Journal Agricultural Economics*, 85(5): 1155–1161.

Chowdhury, S., A. Gulati, and E.Gumbira-Said, (2005), "High value products, supermarkets and vertical arrangements in Indonesia", MTID Discussion Paper No.83. International Food Policy Research Institute, Washington, pp 1-83.

Christiaensen, L., L. Demery and K. Jesper (2011), "The (evolving) role of agriculture in poverty reduction – An empirical perspective", *Journal of Development Economics*, 96: 239-254.

Dethier, J.J and A. Effenerger (2012), "Agricultural and development: a brief review of the literature", *Economic System*, 36: 175-205.

FAO. (2011), "The state of food and agriculture", FAO, Rome

Greene, W. H. (2008), Econometric analysis. Upper Saddle River, NJ: *Prentice Hall.*

Gulati, A., N. Minot, C. Delgado and S. Bora (2007), "Growth in high-value agriculture in Asia and the emergence of vertical links with farmers, " p91-108 in Swinnen, J. (Eds.), *Global Supply Chains, Standards and the Poor: How the Globalization of Food Systems and Standards Affects Rural Development and Poverty*. London: CABI Press.

Heckman, J. (1978), "Dummy endogenous variables in a simultaneous equation system", *Econometrica*, 46(4): 931-959.

Heckman, J. (1979), "Sample selection bias as a specification error", *Econometrica*, 47(1): 153-162.

Henson, S. and T. Reardon (2005), "Private agri-food standards: implications for food policy and the agri-food system", *Food Policy*, 30: 241–253.

Hernández, R., T. Reardon and J.A. Berdegué (2007), "Supermarkets, wholesalers, and tomato growers in Guatemala", *Agricultural Economics*, 36(3), 281–290.

Jaffee, S., and J. Morton (1994), *Marketing Africa's high value foods: Comparative experiences of an emergent private sector*. Dubuque, Iowa: Kendall/Hunt Publishing Co..

Maddala, G. S. (1983), Limited-dependent and qualitative variables in econometrics. Cambridge: Cambridge University Press.

Maddala, G. S. (1986), "Disequilibrium, self-selection, and switching models", p.1633-1682 in Z. Griliches, and D.I. Michael (Ed.), *Handbook of Econometrics*. North-Holland: Elsevier.

Minten, B., L. Randrianarison and J.F.M. Swinnen (2009) "Global retail chains and poor farmers: Evidence from Madagascar", *World Development*, 37 (11): 1728–1741.

Minot, N. (1986), Contract farming and its effect on small farmers in less developed countries, Working paper no. 31. Lansing: Michigan State University.

Minot, N. and D. Roy (2006), "Impact of high-value agriculture and modern marketing channels on poverty: an analytical framework", Mimeo. Washington, DC: Markets, Trade and Institutions Division paper no. 120 and MTID discussion paper no 73. Washington, DC: International Food Policy Research Institute.

Miyata, S., N. Minot and D. Hu (2009), "Impact of contract farming on income: Linking small farmers, packers, and supermarkets in China", *World Development*, 37 (11): 1781–1790.

Mergenthaler, M., K. Weinberger and M. Qaim. (2009a), "Consumer valuation of food quality and good safety attributes in Vietnam", *Review of Agricultural Economics,* 31(2): 266-283.

Mergenthaler, M., K. Weinberger and M. Qaim. (2009b), "The food system transformation in developing countries: a disaggregate demand analysis for fruits and vegetables in Vietnam", *Food Policy* 34 (5): 426-436.

Moustier, P., Phan Thi Giac Tam, Dao The Anh, Vu Trong Binh and Nguyen Thi Tan Loc (2009), "The role of farmer organizations in supplying supermarkets with quality food in Vietnam", *Food Policy*, 35 (1): 69-78.

Natawidjaja, R., T. Reardon, S. Shetty, T.I. Noor, T. Perdana, E. Rasmikayati, S. Bachri, and R. Hernandez (2007), "Horticultural Producers and Supermarket Development in Indonesia", UNPAD/MSU Report published by the World Bank/Indonesia, June.

Neven, D. and T. Reardon (2004), "The rise of Kenyan supermarkets and the evolution of their horticulture product procurement systems", *Development Policy Review*, 22(6): 669-699.

Neven, D., T. Reardon, J. Chege and H. Wang (2006). "Supermarkets and consumers in Africa: the case of Nairobi, Kenya", *Journal of International Food and Agribusiness Marketing*, 18(1/2): 103-123.

Neven, D., M. M. Odera, T. Reardon, and H. Wang (2009), "Kenyan supermarkets, emerging middle-class horticultural farmers, and employment impacts on the rural poor", *World Development*, 37(11), 1802–1811.

Natawidjaja, R., T. Reardon, S. Shetty, T.I. Noor, T. Perdana, E. Rasmikayati, S. Bachri, and R. Hernandez (2007), "Horticultural producers and supermarket development in Indonesia", UNPAD /MSU/World Bank. *World Bank report* no. 38543.

Rao, E.J.O. and M. Qaim (2010), "Supermarkets, farm household income, and poverty: insights from Kenya", *World Development*, 39(5):784-796.

Reardon, T., C.B. Barrett, J.A. Berdegué and J.F.M Swinnen (2009), "Agrifood industry transformation and small farmers in developing countries", *World Development*, 37(11): 1717-1729.

Reardon, T., and C.B. Barrett (2000), "Agro-industrialisation, globalization and international development: an overview of issues, patterns and determinants", *Agricultural Economics*, 23: 195–205.

Reardon, T., J.A. Berdegué and P. Timmer (2005), "Supermarketization of the 'emerging markets' of the Pacific Rim: development and trade implications", *Journal of Food Distribution Research* 36(1):3-12.

Reardon, T., J. A. Berdegué, R. Medina, M. Aguirre, F. Echánove, A. Martínez, R. Cook, N. Tucker and F. Balsevich (2007). "The Rise of Supermarkets in Mexico: Implications for the Domestic Horticultural Supply Chain and Development", *USDA Report*, East Lansing: Michigan State University.

Reardon, T., J.M. Codron, L. Busch, J. Bingen, and C. Harris (2001), "Global change in agrifood grades and standards: agribusiness strategic responses in developing countries", *International Food and Agribusiness Management Review*, 2: 195–205.

Reardon, T., S. Henson and J. Berdegué, (2007) "'Proactive fast-tracking' diffusion of supermarkets in developing countries: implications for market institutions and trade", *Journal of Economic Geography* 7: 399-431.

Reardon, T., and J. Swinnen (2004), "Agrifood sector liberalization and the rise of supermarkets in former state-controlled economies: Comparison with other developing countries", *Development Policy Review*, 22(5): 515–524.

Reardon, T., P. Timmer, C.B. Barrett and J. Berdegué (2003), "The rise of supermarkets in Africa, Asia and Latin America", *American Journal of Agricultural Economics*, 85(5): 1140-1146

Reardon, T., P. Timmer and J.A. Berdegué (2004), "The rapid rise of supermarkets in developing countries: induced organizational, institutional, and technological change in agrifood system", *Journal of Agricultural and Development Economics*, 1(2): 15-30

Regmi, A., and J. Dyck (2001), "Effects of urbanization on global food demand". In: Regmi, A. (Ed.), Changing Structures of Global Food Consumption and Trade. ERS WRS 01-1. Economic Research Service, United States Department of Agriculture, Washington, DC.

Sahara (2012), "Channel choice decisions and market relationships between farmers and buyers in the Indonesian chili industry", University of Adelaide dissertation series; Adelaide, Australia.

Sahara, W. Umberger, N. Minot, and R. Stringer (2011), "Determinants of market channel choice Among chili farmers in Indonesia." Paper prepared for the Australian Agricultural and Resource Economics Meetings, 7-10 February 2012, Fremantle, Western Australia.

Schipmann C. and M. Qaim (2011a) "Supply chain differentiation, contract agriculture, and farmers' marketing preferences: The case of sweet pepper in Thailand", *Food Policy*, 36: 667–677.

Schipmann C. and M. Qaim (2011b), "Modern Food Retailers and Traditional Markets in Developing Countries: Comparing Quality, Prices, and Competition Strategies in Thailand", *Applied Economic Perspectives and Policy*, 33(3): 345–362. doi:10.1093/aepp/ppr018.

Sexton R.J. (2013), " Market power, misconceptions, and modern agricultural markets", *American*

Journal of Agricultural Economics, 95 (2): 209-219. doi: 10.1093/ajae/aas102.

Torero, M. and A. Gulati (2004), "Connecting small holders to market: role of infrastructure and institutions", USAID Policy brief, International Food Policy Research Institute (IFPRI).

Torero, M. (2011), "A framework for linking small farmers to markets", Paper presented at the IFAD Conference on New Directions for Smallholder Agriculture 24-25 January, 2011.

Traill, W.B. (2006), "The rapid rise of supermarkets?", *Development Policy Review*, 24(2): 163-174.

Weatherspoon, D.D. and T. Reardon, (2003) "The Rise of Supermarkets in Africa: Implications for Agrifood Systems and the Rural Poor", *Development Policy Review,* 21(3): 333-55.

Weinberger, K. and T. Lumpkin (2005), "Horticulture for Poverty Alleviation - The Unfunded Revolution", AVRDC Working Paper No. 15. Available at SSRN: http://ssrn.com/abstract=781784 or http://dx.doi.org/10.2139/ssrn.781784

8. Appendix

Appendix A1 Full result of robustness check Table 7.

Dependent var.	(1) High cost /m2		(2) High cost /m2		(3) Chili profit		(4) Chili profit	
model	Probit		Treatment 1st stage		OLS		Treatment 2nd stage	
High Cost					0.134	(1.06)	1.008***	(2.64)
Unaware					0.305	(1.61)	0.311**	(2.27)
Aware					0.775***	(3.66)	0.782***	(4.67)
Productivity	0.883***	(9.37)	1.022***	(8.05)	0.449***	(5.98)	0.256**	(2.15)
Grading					0.144**	(2.37)	0.118*	(1.87)
Household size	0.025	(0.57)	0.046	(0.92)	-0.037	(-0.90)	-0.053	(-1.22)
Age	0.004	(0.49)	0.003	(0.36)	-0.014**	(-2.16)	-0.014**	(-2.22)
Education	0.031	(1.35)	0.022	(0.94)	-0.009	(-0.44)	-0.021	(-0.97)
% working age	0.810**	(2.06)	1.297***	(2.84)	0.417	(1.18)	0.136	(0.36)
% retired age	-0.190	(-0.25)	0.233	(0.30)	-0.966	(-1.47)	-1.100*	(-1.88)
Mobile phone	0.121	(0.82)	0.013	(0.08)	0.007	(0.06)	0.036	(0.26)
Motor bike	-0.070	(-0.52)	-0.022	(-0.15)	0.268**	(2.23)	0.260**	(2.17)
Water pump	0.270*	(1.86)	0.274*	(1.77)	0.114	(0.87)	0.048	(0.37)
Mist blower	0.403**	(2.29)	0.342*	(1.66)	0.325**	(2.00)	0.175	(1.10)
Power tiller	0.225	(0.40)	0.162	(0.26)	0.494	(1.02)	0.456	(1.05)
Storage house	-0.105	(-0.69)	0.011	(0.07)	0.178	(1.29)	0.210*	(1.66)
Experience	-0.013	(-1.26)	-0.027**	(-2.25)	0.020**	(2.25)	0.029***	(3.08)
Land owned 05	-0.046**	(-2.28)	-0.032	(-1.47)				
% land rented	0.368*	(1.96)	0.618***	(2.87)				
Chili land					0.484***	(7.75)	0.498***	(7.55)
Production record	0.305	(1.23)	0.330	(1.29)	0.261	(1.21)	0.216	(1.03)
Marketing record	0.272*	(1.71)	0.401**	(2.30)	0.386**	(2.52)	0.240	(1.55)
Distance to road	-0.092	(-0.73)	-0.223	(-1.43)				
Constant	-1.50***	(-3.16)	-1.91***	(-3.54)	10.16***	(16.34)	10.17***	(16.73)
Pseudo /adj. $R2$	0.258				0.348			
N	577		459		562		459	

t statistics in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$; Wald chi2(21) = 280.82 (prob>chi^2=0.000), independence test = 4.543 (prob>chi^2=0.033); athrho = -0.511(-2.131)

Appendix A2 Full result of robustness check table 8.

Dependent variable	(1) Linear total Cost		(2) Chili income	
Predicted total cost			0.455**	(1.992)
Unaware			0.329*	(1.764)
Aware			0.745***	(3.531)
Productivity	0.249***	(5.679)	0.363***	(3.948)
Grading			0.134**	(2.199)
Distance to road	0.031	(0.375)		
Household size	0.011	(0.364)	-0.046	(-1.127)
Age	-0.005	(-0.976)	-0.013*	(-1.949)
Education	0.015	(0.921)	-0.015	(-0.740)
% working age	-0.154	(-0.560)	0.479	(1.372)
% retired age	-0.265	(-0.515)	-0.823	(-1.252)
Mobile phone	0.263**	(2.552)	-0.093	(-0.650)
Motor bike	0.217**	(2.321)	0.174	(1.381)
Water pump	0.184*	(1.801)	0.016	(0.113)
Mist blower	0.357***	(2.939)	0.097	(0.522)
Power tiller	0.231	(0.631)	0.417	(0.862)
Storage house	0.184*	(1.691)	0.083	(0.577)
Experience	0.015**	(2.140)	0.014	(1.504)
Chili land			0.437***	(6.688)
Production record	0.444**	(2.444)	0.067	(0.283)
Marketing record	0.216*	(1.861)	0.294*	(1.837)
Land owned 05	0.067***	(4.825)		
% land rented	0.634***	(4.827)		
Constant	14.366***	(43.847)	3.959	(1.215)
adj. $R2$	0.234		0.350	