Doctoral Thesis

UNIVERSITY OF TRENTO - Italy

CIFREM

INTERDEPARTMENTAL CENTRE FOR RESEARCH TRAINING IN ECONOMICS AND MANAGEMENT

DOCTORAL SCHOOL IN ECONOMICS AND MANAGEMENT

SHOCKS, COPING STRATEGIES AND THEIR CONSEQUENCES. AN APPLICATION TO INDONESIAN DATA

A DISSERTATION SUBMITTED TO THE DOCTORAL SCHOOL OF ECONOMICS AND MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DOCTORAL DEGREE (PH.D.) IN ECONOMICS AND MANAGEMENT

Francesca Modena

April 2008

ADVISORS

Prof. Gabriella Berloffa
University of Trento

Prof. Christopher Leslie Gilbert
University of Trento

DOCTORAL COMMITTEE

Prof. Guido Ortona
University of Piemonte Orientale
“Amedeo Avogadro”

Prof. Carlo Federico Perali
University of Verona

Prof. Maria Luigia Segnana
University of Trento

Acknowledgements

Writing a Ph.D. thesis indirectly involves many people and can be considered the final step of the “schooling” track or as the first towards the world of scientific research.

I would like to thank many people. Starting from the very beginning, I am grateful to Prof. Giuseppe Folloni for introducing me to economic research at the time of the undergraduate courses and for encouraging me to embark on the Ph.D. The present work was mainly carried out at CIFREM, University of Trento. I wish to express my gratitude to all the staff and students, in particular my colleagues on the programme: Francesca B., Francesca M., Giulia and Oxana. I would like to thank also the Department of Economics, and in particular the Grade group who helped and supported me during these years.

A special thank goes to Prof. Basu and the Department of Economics at Cornell University, for the opportunity to spend a few months in Ithaca, for the stimulating discussions and the friendly atmosphere. I would like to express my gratitude to my supervisors, Profs. Christopher Gilbert and Gabriella Berloffa, their continued presence and support added greatly to my educational experience. I could not have wished for better teachers.

In the end, I would like to express my gratitude to my family, who encouraged me to strive for the best I could do in everything I did and to always remember why I am doing it.

Contents

Introduction

The analysis of the way in which risk and/or uncertainty affect household economic choices is a key issue in developing countries, where the economic environment is riskier, households may be more vulnerable to risk and shocks may have dramatic consequences on household living conditions (Morduch, 1994; Dercon , 2005c). There are two components of risk, to which different aspects of households decisions are associated: the possibility of unforeseeable future events, and the actual occurrence of shocks. The former leads people to try and reduce future risk and/or its consequences (ex ante risk management strategies), whereas the latter forces households to cope with the consequences of shocks when they occur (ex post risk coping strategies). This dissertation refers mainly to the effects of shocks faced by poor people in poor countries, and in rural areas in particular, while the problem of ex ante responses to risk is touched upon only in the last chapter.

Banerjee (2005) posits the question "are the poor just like you and me except in that they have less money? Or is it useful to think of them as being subject to different pressure from the rest of the population and therefore sometimes making choices that are very different?" (Banerjee, 2005, p.2). In discussing the topic of risk, the differences between developed and developing countries, and between poor and non-poor people, become relevant. The sources of risk that affect the two types of economies are very different, as different is the ability to cope with shocks and the risk-management strategies adopted. Some types of risk appear to be much more prevalent, less protected, and to lead to more serious consequences in poor countries than in other part of the world (Fafchamps, 1999), and for poor people in these economies, especially for those living in rural areas. Rural villages have in general not only lower income levels, but also a higher income variability compared to non-rural areas (Fitzsimons, 2007).

One of the most relevant sources of uncertainty is variation in business income. In poor economies incomes are much more dependent on business sources than on safe wage. Business activities (mainly farms in rural areas) are typically small, the production technologies adopted are quite simple and cannot provide insurance against economic or natural events. For example, most of the farms in poor rural countries are highly dependent on weather variability (mainly rainfall shocks), and plant disease and pests may cause crop losses. Another important source of uncertainty related to business risk is the fluctuation in prices. Moreover, poor people are exposed to health problems caused by their scarce nutrition and by the environment in which they live or work (World Bank, 2000).

Not only poor rural economies face a high incidence of risk, but poor people are also less protected against shocks (Fafchamps, 1999; Dercon, 2005b). The failure of insurance and credit markets (particularly in rural areas, Fitzsimons, 2007), and the fragility of social security, make poor people vulnerable against hardships. Low incomes limit the possibility to save and accumulate assets, hence reducing the effectiveness of self protection strategies, and the use of productive assets to smooth consumption is costly in terms of future incomes. Moreover, poor households may invest in safe but low return activities to reduce their exposure to risk, and this affects their long-term income. Nevertheless, evidence suggests that in some cases risk management strategies (both ex ante and ex post) offer at least some protection against shocks. These strategies involve mainly informal mechanisms and mutual support networks between households and villages. When these strategies fail, the lack of insurance and protection may lead to poverty traps or households may require several years to recover from the shock (Dercon, 2005b).

To summarize, shocks are costly to households in terms of lost income, reduced consumption, and sale or destruction of assets. Poor households developed several mechanisms to both smooth income (ex ante and ex post), and protect consumption (consumption smoothing), and these mechanisms may be costly in their turns. This dissertation investigates the way in which potential and actual shocks affect household economic conditions and choices. We will focus on three main issues. First, we explore which strategies are chosen for different types of shocks. Second, we investigate which are the specific mechanisms adopted in the face of the most common shock (crop loss), and the consequences on consumption. Finally, we analyze the consequences of risk and shocks on a particularly important household decision: how much to invest in children education. The source of data we employ is the Indonesian Family Life Survey (1993 round). A large amount of information on household economic conditions (income sources, consumption, types of assets), and household and individual characteristics were collected, as well as a detailed section on shocks experienced by the household during the five years prior to the interview. These data provide a very useful source to analyze shocks, and to document the relationship between types of shocks, household characteristics, and coping strategies.

We start with the review of the theoretical and empirical literature of households' behaviour in the face of risk and shocks in developing countries. The theoretical framework that underlines the analysis is an intertemporal model of consumption and asset investment choices. We discuss the specificities of these choices in poor economies, where insurance and credit markets are less developed, and households have typically low incomes. We present models that account for these issues, in particular models with precautionary savings or liquidity constraints, and the asset smoothing hypothesis. We next discuss the main findings of the empirical studies that investigate consumption and asset investment choices in developing countries, focusing in particular on the different behaviour between poor and non-poor households. A descriptive analysis of the data follows. The suggestion of the literature are exploited to investigate the sources of shocks faced by rural Indonesian households, and the mechanisms they develop to deal with negative events. The focus is on the extent to which the effects of shocks and the coping mechanisms differ for households with different characteristics, and in particular between poor and non-poor.

The second chapter models households responses to shocks. Most of the literature examines the role of specific responses to hardships, without examining how households choose between different strategies. We investigate households decisions with regard to coping mechanisms focusing on how households' wealth and the type of shocks experienced (whether demographic or economic, idiosyncratic or common) affect the strategies used. Data show that the majority of shocks elicit only a single response, and some shocks elicit multiple responses. We thus consider a situation of non-exclusive multinomial choice, and propose two models which require at least one response, and account for non-exclusive and dependent multiple responses. In the first and simpler specification, the Poisson-multinomial, households first choose the number of responses to a specific shock, and then the specific choices are identified to maximize household utility conditional on the former choice. The second specification, the threshold-multinomial, generalizes the standard multinomial logit model by supposing that agents will choose more than one response if the utility they derive from other choices is "close" to that of the utility-maximizing choice. We compare these models with the widely used Marginal Logit Model.

The third chapter investigates whether rural Indonesian households smooth income following a crop loss (i.e. whether they adopt strategies, labour supply adjustments in particular, to recover the income reduction due to the shock), and how much consumption smoothing is achieved through income smoothing strategies (i.e. how much of the increase in income is transferred onto consumption). We extend the approach of Cameron and Worswick (2003) and construct quantitative measures of income shocks and of households' responses to the shock, analyzing more than one response type, and differentiating between poor and non-poor farmers. We find that the model that drives household consumption behaviour is different for poor and non-poor households, and this leads shocks to have different impacts on households consumption and savings choices.

The last chapter extends the analysis and explores the role of both ex ante income risk and income shock occurrence on child education, and, in particular, on school progression from a school level to the next one. The main original feature of the chapter is the assumption of irreversibility of the decision to withdraw a child from school: once the child drops our, he/she cannot return to school. In this way, temporary interruptions in schooling in the face of risk or short-term shocks have long term impacts on the child human capital. We underline that the impact of uncertainty on schooling decisions is more subtle that what suggested by much of the existing development literature.

Chapter 1

Literature Review and Data Overview

1.1 Introduction

Many studies document the high income variability in developing countries related to various sources of risk (Townsend, 1994; Morduch, 1995; Kinsey *et al.*, 1998). Several surveys review the relationship between risk, poverty, and household responses to shocks in poor countries. Alderman and Paxson (1992) review the literature on household ex ante and ex post strategies to insure consumption against shocks, examining the effectiveness of these mechanisms. Morduch (1995) investigates how risk affects household production and employment decisions. Townsend (1995) focuses on the ability of households to insure their consumption against idiosyncratic or common shocks, and on the role of various mechanisms in providing insurance. Fafchamps (1999) summarizes the literature on rural poverty and risk in developing countries, examining the limits of coping strategies and discussing the policy implications. Dercon (2005a) focuses on the constraints households face in using insurance mechanisms, and on the role of economic policies.

This chapter summarizes the main findings regarding the way in which potential and actual shocks affect household economic conditions and choices, with a particular focus on the different coping strategies adopted and on their consequences. The theoretical framework that underlines the analysis is an intertemporal model of consumption and asset investment choices. Under the assumptions that there are no borrowing constraints, that agents are rational forward looking, and that utility function is separable over consumption and leisure, households seek to keep the marginal utility of consumption constant over time. This means that households spread transitory income changes over the remaining periods of their life-cycle, and therefore consumption is substantially unaffected by these shocks (consumption smoothing). The life-cycle models and the permanent income hypothesis capture this behaviour. However, poor households, especially in developing countries, may have no access to perfect credit markets, and they may find it hard to save. Moreover, the vast majority of people living in poor areas do not rely on wage income, rather on farm profits, for which the availability of some productive assets becomes crucial. This makes them cautious in running down assets to smooth consumption. To account for these issues, alternative models are developed that relax some of the assumptions of the permanent income hypothesis. In particular, this chapter reviews models with precautionary savings or liquidity constraints, and the asset smoothing hypothesis.

The theoretical predictions of these models form the basis for several empirical studies that investigate consumption and asset investment choices in developing countries. Many of them test whether consumption is smoothed, both across space, by relying on social networks, and over time, by saving and borrowing, or by accumulating and selling non-financial assets. A growing empirical literature considers a framework in which assets generate income and focuses on the trade off between the use of assets as a buffer against shocks (consumption smoothing behaviour) and their maintenance for future income and consumption (asset smoothing behaviour). This chapter summarizes the main empirical findings related to consumption and asset choices in poor countries and in the presence of shocks.

A descriptive analysis of the data follows the literature review. The main objective is to document the sources of shocks that affect rural Indonesian households, and how types of shocks and household characteristics affect the use of different coping strategies, focusing on the differences between poor and non-poor households. Data analysis is based on the 1993 round of the Indonesia Family Life Survey data (IFLS)[1], and the focus is on the rural sample. An interesting feature of this database is that respondents were asked whether the household had experienced an economic hardship in the five years prior to the interview, the type of shock, when it happened (year and month), what measures were taken and the costs of overcoming it[2]. For these reasons this database is particularly suitable to examine the relationship between shocks and household decisions in poor countries.

---

[1] IFLS is conducted by RAND in collaboration with UCLA and Lembaga Demografi, University of Indonesia.

[2] Appendix A shows in detail the questionnaire for household economic shocks.

1.2 Types of shock and their consequences on household living conditions

The extent to which shocks are costly to households (in terms of lost income, reduced consumption, or the sale and destruction of assets), and the ability to cope with them depend on the nature of the shock, on environmental and economic circumstances, and on specific household characteristics. As regard the nature of the hardship, shocks that affect single individuals or households have different consequences compared to shocks that occur at the village level. Health and economic shocks may be associated with different costs. Moreover, the frequency and intensity of shocks are important factors in determining the extent of the consequences. It is therefore useful to first classify the sources of risks, and then present their consequences, and the means households can rely on to cope with them.

1.2.1 Types of shock

Two broad classifications of hardships are proposed in the literature. The first is between demographic and economic shocks. The former includes mainly health shocks, the latter directly affect economic variables (i.e. income, assets). The second distinction, correlated with but not implied by the demographic-economic one, is between idiosyncratic and common shocks. Idiosyncratic shocks are household specific, while common shocks affect group of households, or the entire community. Some shocks are more likely to be idiosyncratic (for example illness or unemployment), others to be common (for example droughts or epidemics), but few risks purely belong to one category (Dercon, 2005a). Moreover, even commonly experienced events may produce idiosyncratic effects according to specific household characteristics.

Distinguishing between economic and demographic shocks, and between idiosyncratic and common ones, is important because they may have different consequences on households' living conditions and they may allow or require different strategies. Several papers document that demographic shocks are more costly than economics ones. For example, Kenjiro (2005) find that in rural Cambodia, the economic damage caused by sickness is more severe than that caused by a crop loss. Gertler and Gruber (2002) find evidence that in Indonesia the economic costs associated with major illness are high and cause a severe reduction in household consumption. Moreover, demographic shocks may prevent households from using some coping strategies, for example households are less likely to use the labour supply response to cope with health shocks because they may affect the ability of the households to provide labour (Kochar, 1995).

Aggregate variations may be more difficult to insure and to cope with because households cannot rely on community assistance and social networks (Dercon, 2002; Alderman and Paxson, 1992). Even the use of assets as a buffer against shocks may differ between idiosyncratic and common risks. When the majority of households try to sell assets, their prices fall and the gain from the asset sale decreases (Morduch, 1994; Frankenberg, Smith and Thomas, 2003).

These two classifications however, do not exhaust the characteristics of the shocks that are relevant for the analysis of their consequences. For example, the frequency and intensity of shocks, and the autocorrelation over time are important features that may exacerbate the effect of shocks to income. High intensity and low frequency risks are in general more difficult to deal with than small and frequent shocks (Dercon, 2005a). Fafchamps (1999) points out that shocks may not be independently distributed over time: experience one type of risk may raise vulnerability to other shocks, and in this case coping is more difficult (Alderman, 1996). Moreover, some shocks may be non-stationary since they lead to permanent consequences.

1.2.2 Direct and indirect effects of shocks, and long-run consequences

When analysing the consequences of shocks, it is useful to distinguishing between direct and indirect effects, and between short-run and long-run consequences. Shocks may have direct impacts on incomes (income shocks) or assets (asset shocks), and indirect effects as consequences of the strategies used by households to deal with risk. Income shocks are more likely to have transitory effects, whereas asset shocks are more likely to have long term consequences: a shock that causes the loss of human or physical assets that are crucial for income production can have a negative impact not only on current well-being, but on future income and human capital as well (McPeak, 2004). From this point of view, the direct effects of demographic shocks may be severe when they affect the human capital formation of children. Dercon and Hoddinott (2003) find that shocks affecting child health in Zimbabwe and Ethiopia lead to persistence effects on the future adult human capital and earnings. Alderman *et al*. (2006) and Hoddinott and Kinsey (2001) show that the health of Zimbabwean children, and in particular that of children living in poor households, is affected by drought, and this has long term consequences on subsequent human capital formation.

Indirect effects of shocks are consequences of both ex ante and ex post risk-management strategies. Ex ante strategies involve diversification across crops, the use of a variety of production techniques, portfolio diversification, migration, etc. They may have negative consequences when, for example, the fear of risk force poor households to take safe but less profitable decisions (Morduch, 1990; Alderman and Paxson, 1992). Dercon (1996) finds that poor households in Tanzania who choose low risk activities have an income per unit of land that is 20 percent lower than that of rich households who choose riskier but more profitable strategies. Morduch (1990) suggests that Indian farmers with a low level of assets cultivate mostly safe varieties of rice instead of risky but high-return varieties. As suggested by Rosenzweig and Wolpin (1993), when risk averse farmers face borrowing constraints and low incomes in a risky environment they could underinvest in key productive assets. Using Indian data, Rosenzweig and Wolpin find that poor farmers invest more in bullocks than in irrigation equipment even if the return of the latter is higher. This is due to the fact that poor farmers prefer to invest in liquid assets (bullock), than to finance irriversible investents (pumps) even if the latter are more profitable.

Ex post coping strategies may have negative consequences when they destroy or reduce the physical, financial, human or social capital of the household (Dercon, 2005c). This could increase the risk of entering poverty and could make households vulnerable to future hardships. In this sense, maintaining income in the short-run could come at the expense of long-term well-being. For example, recent empirical and theoretical papers find evidence that, in the presence of credit constraints and lack of assets, income shocks force children to be withdrawn from school or to work while remaining in school, with permanent consequences on their human capital (Beegle, Dehejia and Gatti, 2006; Fitzsimons, 2007; de Janvry *et al*., 2006; Guarcello *et al*. 2003; Jacoby and Skoufias, 1997; de Janvry *et al*., 2006).

1.3 Risk coping strategies

An essential part of the literature on risk is the study of the strategies developed by households to deal with shocks and their consequences. There are different ways to characterize these strategies. The World Bank (WDR 2000/01) classifies the ex post mechanisms according to whether they involve formal insurance mechanisms or informal arrangements between individuals and communities (see table 1.1). Informal strategies are developed at the individual and household level, or are group based (communities, villages). Formal mechanisms may be market based or provided by governments/NGOs. In many developing countries informal mechanisms appear to significantly contribute in coping with risk, while formal credit and insurance markets are less developed and asset poor households face constraints in access to these institutions.

Table 1.1
Risk coping strategies

	Informal mechanisms		Formal mechanisms	
	Individual and household	*Group based*	*Market based*	*Publicly provided*
Coping with shocks	▪ sale of assets ▪ loans from money lenders ▪ child labour reduce food consumption ▪ seasonal or temporary migration	▪ transfers from networks of mutual support	▪ sale of financial assets ▪ loans from financial institutions	▪ social assistance ▪ workfare ▪ subsidies ▪ social funds ▪ cash transfers

Source: World Bank, *World Development Report, 2000/01: Attacking Poverty*

Alternatively, one can choose as a classifying criterion the relationship between these strategies and a specific objective that the household may try to achieve, such as consumption or income smoothing. Households can ex ante deal with risk "making conservative production or employment choices and diversifying economic activities" (Morduch, 1995, p. 104) (income smoothing), or may perform strategies that protect consumption (consumption smoothing). Alderman and Paxson (1992) propose a classification according to whether risk coping strategies smooth consumption across households (risk-sharing strategies), or over time (risk-protection strategies). The former are collective strategies that involve the mutual support of social networks, the latter are individual strategies based on self-protection. Since the effects of shocks on consumption is one of the issues we are interested in, we follow this classification to discuss coping mechanisms.

1.3.1 Risk sharing strategies

Risk can be shared between household members (intra-household risk-sharing), or can be spread across different households (inter-households risk sharing). In the first case the household is viewed as a risk-sharing institution, in which individual-specific shocks in income can be insured by the other household members (Mazzocco, 2004; Mazzocco and Saini, 2006; Dercon and Krishnan, 2000).

Many theoretical and empirical studies on risk-sharing characterize the household as a single unit, ignoring the intra-household decisions, and focusing on the inter-households component of risk-sharing. According to this literature, units for risk-pooling are extended families (Foster, 1993; Witoelar, 2005), friends and relatives (Fafchamps and Lund, 2003), ethnic groups (Grimard, 1997), and villages (Townsend, 1994). Inter-household risk-sharing is more efficient within small communities, where the problems of asymmetric information are relatively unimportant with respect to large groups. The ways in which the risk is spread may be formal and informal credit transactions, social transfers and social assistance (Milanovic, 2000), generalized reciprocity (Sahlins, 1972), and gift exchange (Cashdan, 1985).

Perfect risk-sharing within a community (inter-households) implies that "total village resources in any time period are distributed so as to equate the weighted marginal utility of consumption across households" (Alderman and Paxson, 1992, p.16). Formally, consider a social planner who maximizes the sum of weighted utilities across households (Townsend, 1994; Fafchamps, 1999; Morduch, 2005):

$$\begin{aligned} &\max_{c_{it}} \sum_{i=1}^{N} \omega_i \sum_{t=0}^{\infty} \beta^t \sum_{s_t \in S_t} \Pr ob(s_t) U_i \left[c_{it}(s_t) \right] \\ &s.t. \quad \sum_i c_{it}(s_t) = \sum_i y_{it}(s_t), \end{aligned} \tag{1.1}$$

where ω_i is the welfare weight associated with individual *i*, β is the discount factor, s_t denotes the state of the world at time t, and $y_{it}(s_t)$ is the individual stochastic income. Defining $\lambda(s_t)$ the Lagrange multiplier associated with each constraint, the first order condition is given by:

$$\omega_i \beta^t \Pr ob(s_t) U_i^{'} \left[c_{it}(s_t) \right] = \lambda(s_t)$$

where $\lambda(s_t)$ depends only on aggregate income. This implies that individual consumption varies only with aggregate income and not with individual income. From the first order conditions we can derive:

$$\frac{\omega_j}{\omega_i} = \frac{U_i^{'}(s_t)}{U_j^{'}(s_t)} = \frac{U_i^{'}(s_t^{'})}{U_j^{'}(s_t^{'})}, \quad \forall \ s, \ s^{'}, \ i \ and \ j. \tag{1.2}$$

If two individuals have equal welfare weights and the same utility function, they have the same consumption levels at any time and for any state of the world. If $\omega_j > \omega_i$, it follows that $c_j > c_i$ for every period, since perfect risk sharing implies that the ratio of consumption is fixed across state of nature and over time.

The finding that individual consumption varies only with aggregate income is the basis for testing risk sharing. The main form for the test is that employed by Cochrane (1991), Mace (1991), and Townsend (1994). Following Townsend[3], constant relative risk aversion utility functions are assumed (CRRA):

$$U(C_{it}) = \frac{C_{it}^{(1-\gamma)}}{1-\gamma}, \tag{1.3}$$

where C_{it} is the consumption of household *i*, and γ is the measure of relative risk aversion.

Pareto efficiency implies that:

$$\log c_{it} = \log \overline{C}_t + \frac{1}{\gamma}(\log \omega_i - \log \overline{\omega}), \tag{1.4}$$

where $\overline{C}_t$ and $\overline{\omega}$ are average values. This equation says that under the hypothesis of perfect risk sharing, individual consumption depends on average consumption and not on individual income. Based on this implication, risk sharing is tested regressing household consumption on average group consumption, household income and shocks variables. Household should be affected only by common shocks, whereas idiosyncratic shocks should be insured within the community. Most studies rejects the hypothesis of full insurance between communities, but supports a partial risk-sharing. Townsend (1994) finds that the marginal propensity to consume out of household income is no greater than 0.14 in Indian villages; using the same data Ravallion and Chaudhuri (1997) estimate a marginal propensity to consume between 0.12 and 0.46.

Morduch (2005) extends the test presented above and regresses the growth of household consumption on time-specific fixed effects and on household income growth. If income shocks translate into consumption variations, the coefficient on income growth is expected to be one, otherwise if there is consumption smoothing, income growth should have no effects on consumption. Using Indian data, the author finds estimates of the coefficients on income growth between 0.19 and 0.32. The test used by Morduch does not allow to distinguish if consumption is smoothed inter-temporally or the risk is shared within a community or a village. Udry (1994) tests for risk pooling in northern Nigeria focusing on the role played by credit contracts in sharing the risk between households. The author develops two models: a competitive model of the credit market, and a bilateral model of loan contracting that permits imperfect information and equilibrium default. Results of both models reject the hypothesis

[3] While Cochrane (1991) and Mace (1991) test consumption insurance on US data, Townsend (1994) tests this property for a developing country (India).

of complete risk sharing, but find evidence that state-contingent loan transactions are important mechanisms for risk pooling within a community.

The degree of risk sharing may differ for poor and non-poor households. Dercon and Krishnan (2000) find evidence that non-poor households are better able to insure themselves within the community. Similarly, Santos and Barrett (2006) find that in southern Ethiopia the poor are excluded from informal insurance networks, and do not receive transfers in response to shocks; De Weerdt (2004) shows that poor households in a village in Tanzania have less social networks than the rich, and often they can rely only on other poor households. This makes them “vulnerable in the face of idiosyncratic risk” (De Weerdt, 2004, p.31).

In summary, empirical findings show that the effects of shocks are partly spread across households, but that there is not complete risk sharing. The degree of risk pooling depends on the type of shocks. Households can rely on family, or community, assistance to cope with idiosyncratic, but not with covariant shocks. Moreover, poor households may be excluded from social networks, or have the possibility to rely only on other poor households, thus reducing their ability to smooth consumption through risk-sharing, and increasing their need to rely on self-protection strategies.

1.3.2 Self-protection strategies: savings and assets

Households can smooth consumption not only across space, but also over time, by saving and borrowing, or by accumulating and selling non-financial assets (Bardhan and Udry, 1999, Dercon, 2005a). Under the assumptions that there are no borrowing constraints and that agents are rational forward looking, households seek to keep the marginal utility of consumption constant over time. This means that households spread transitory income changes over the remaining periods of their life-cycle, and therefore consumption is substantially unaffected by these shocks (consumption smoothing). The life-cycle models and the permanent income hypothesis capture this behaviour. However, poor households, especially in developing countries, may have no access to perfect credit markets, they may find it hard to save or be cautious in running down assets to smooth consumption. To account for these issues, alternative models are developed that relax some of the assumptions of the permanent income hypothesis: models with precautionary savings and/or liquidity constraints, and the asset smoothing hypothesis.

The permanent income hypothesis

Consider a consumer who has a total value of financial assets at time t (A_t), and receives a labour income y_t. The consumer maximizes his lifetime utility defined as

$$U_\tau = u(c_\tau) + E_\tau\left(\sum_{t=\tau+1}^{T} \beta^{t-\tau} u(c_t)\right). \tag{1.5}$$

With complete markets, in each period agents may borrow or lend at the interest rate *r*, so that the stock of financial assets evolves according to $A_{t+1} = (A_t + y_t - c_t)(1+r)$. The value function for the agent's problem is the following:

$$V_t(A_t + y_t) = \underset{c_t}{Max}\{u(c_t) + \beta E_t V_{t+1}[(1+r_t)(A_t + y_t - c_t) + y_{t+1}]\} \tag{1.6}$$

The first order conditions imply that:

$$u_t^{'}(c_t) = \beta(1+r)Eu^{'}(c_{t+1}) \tag{1.7}$$

Assuming a quadratic utility function (so that the marginal utility is linear) and $\beta(1+r) = 1$, this equation simplifies to

$$c_t = Ec_{t+1} \tag{1.8}$$

Equation (1.8) means that the consumption at time t is equal to the expected value of future consumption, that is household makes consumption plans such that expected consumption is constant (Bardhan and Udry, 1999).

Since assets must be zero at the end of the last period ($A_{T+1} = 0$), we have also that $c_T = A_T + y_T$; therefore we obtain the intertemporal budget constraint:

$$\sum_{t=1}^{T} \frac{c_t}{(1+r)^t} = A_t + \sum_{t=1}^{T} \frac{y_t}{(1+r)^t}. \tag{1.9}$$

The discounted value of consumption from t to T equals the value of assets at time *t* plus the expected discounted value of income stream form *t* to T (Bardhan and Udry, 1999; Deaton, 1992). From equations (1.8) and (1.9), and setting T to infinity, we have:

$$c_t = \frac{r}{1+r}\left(A_t + E_t \sum_{\tau=t}^{\infty} \frac{y_\tau}{(1+r)^{(\tau-t)}}\right). \tag{1.10}$$

This equation describes the main implication of the permanent income hypothesis (Deaton, 1992), according to which current consumption is equal to the value of current assets plus the expected actual value of the income stream from *t* to infinity. Equation (1.10) shows that only shocks that affect the household's expectations about future income (permanent shocks) change current consumption, while transitory shocks lead primarily to additions to assets or to the use of previously accumulated assets rather than to changes in consumption (Friedman, 1957).

To summarize, the assumptions required by this model are (Browning and Lusardi, 1996; Deaton, 1992): agents have intertemporally additive utility functions and maximize expected utility over an infinite time horizon, they have quadratic preferences, and a constant real interest rate equal to the rate of time preference; finally, agents face perfect capital markets. Some of these strong assumptions cannot reasonably expect to hold for poor households in developing countries (Deaton, 1997), where people face high income risks and credit markets are imperfect. Therefore, we now discuss two models that rule out some of the assumptions of the permanent income hypothesis, and that may be more appropriate to analyze household's decisions in these contexts.

Precautionary savings

The precautionary saving model relaxes the assumption that agents have quadratic utility functions and assumes that marginal utility is convex. Considering the marginal utility of consumption as a "shadow-price" for consumption; "because there is diminishing marginal utility of consumption, this price is higher when consumption is low than when it is high. […] It is also reasonable to suppose that the marginal utility of consumption rises more rapidly when consumption is low than when it is high", making then possible to assume the convexity of the marginal utility of consumption (Deaton, 1997, p. 361). In the case where $u_t^{'}(c_t)$ is convex, an increase in the uncertainty of future consumption (i.e. the mean remains unchanged but the spread around the mean increases) will increase the expected future value of the marginal utility. To keep equality (1.7), the current marginal utility of consumption must increase, and hence current consumption will decrease and savings increase. Prudent individuals aim to protect themselves against both long term and short-term misfortunes. The more individuals are prudent, i.e. the more $u_t^{'}(c_t)$ is convex, and the more risky is the environment, the more agents save.

Consider the example of an individual whose expected future income is higher than the current one, for example students in medical school (Browning and Lusardi, 1996). As shown above, under the permanent income hypothesis, and with quadratic utility functions, students will keep their consumption level constant over time, thus borrowing today and saving in the future when their earnings will increase. In the precautionary savings model, with convex marginal utility of consumption, prudent students account for the uncertainty of future income: they will not borrow today just because future income is expected to be higher, but they may increase today savings to deal with future income variability.

Liquidity constraints

The precautionary saving model maintains the assumption of the permanent income hypothesis that capital markets are perfect. Models with liquidity constraints relax this hypothesis. Deaton (1991) investigates how households with borrowing constraints are able to smooth consumption through the accumulation and use of financial and physical assets. Agents are considered impatient, facing liquidity constraints and living for an infinitely horizon (a dynasty). Liquidity constraints are accounted for by assuming that borrowing is impossible and asset levels must be non-negative in each period. Labour income is uncertain, and a precautionary motive for saving is assumed. In this model, the "cash on hand" hold by the household at time *t* is relevant in determining period *t* consumption. Define x_t as "cash on hand" (assets plus current income, i.e. $x_t = A_t + y_t$); the Euler equation for intertemporal allocation is given by

$$u_t^{'}(c_t) = \max\left(u_t^{'}(x_t), \frac{1+r}{1+\delta} E_t u^{'}(c_{t+1}) \right), \qquad (1.11)$$

where r is the interest rate and δ is the rate of time preference. This equation implies that x_t is the maximum the household can consume in period t, so that $c_t = f(x_t) \le x_t$. This means that if the household has no assets accumulated at time *t*, current consumption cannot be higher than current income.

The main finding of Deaton's basic model is the following

$$\begin{cases} c_t = f(x_t) = x_t, & x_t \le x^* \\ c_t = f(x_t) \le x_t, & x_t \ge x^* \end{cases} \qquad (1.12)$$

where, with i.i.d. income, x^* is a little lower than mean income. Consumption is equal to cash-on- hand if the latter is below the critical value x^*; given the borrowing constraint, this implies that household will have no assets in the next period. If $x \ge x^*$, household saves some fraction of cash-on-hand in this way increasing the future level of assets. This is the "buffer stock" behaviour: assets are only used to buffer fluctuations in income, and even small levels of assets allow households to considerably smooth consumption.

It is important to notice that, without uncertainty and borrowing constraints, impatient consumers will borrow to increase current consumption. Assuming constraints and prudence in the face of uncertain incomes, agents have two motives for accumulating assets in good years, even if they are impatient (Deaton, 1991). But they do not save if their wealth is below a critical value, becoming vulnerable against negative shocks. Moreover, consumption can not be smoothed as a consequence of a series of negative hardships.

The behaviour resulting from credit constraint and from precautionary saving are similar and may not be easily distinguished (Browning and Lusardi, 1996). Liquidity constraints reinforce prudent behaviour, since the possibility of binding borrowing constraints in the future is an additional reason to accumulate precautionary savings (Deaton, 1992).

Since credit markets may be underdeveloped in poor economies, especially in rural areas, this model may be the relevant one for many households in these countries. However, the framework described above does not take into account that the vast majority of people living in poor areas do not rely on wage income, rather on farm profits, for which the availability of some productive assets becomes crucial. Therefore, we need to consider in more details the role of assets in this type of environment.

The role of liquidating productive assets and asset smoothing

In many developing countries, and in rural areas in particular, assets have more than one function: they store wealth, and contribute directly to the income generation process. A growing theoretical and empirical literature focuses on the role of assets as buffer stocks in a framework in which assets generate income. This leads to a trade off, in the sense that selling assets to smooth consumption today could have important implications for future income and hence for future consumption. This may lead households, and especially poor households, to be more cautious in running down assets in the face of transitory shocks. The use of productive assets as buffer stocks was introduced by Rosenzweig and Wolpin (1993). They assume that farmers cannot borrow to smooth consumption, that assets (i.e. bullocks, land, irrigation equipment) contribute to agricultural production and income, and that earnings are stochastic (shocks are both idiosyncratic and common). Farmers decide how many assets to purchase maximizing the expected lifetime utility under the constraint that consumption plus the purchase of assets must equal farm profits. Moreover, a minimum subsistence consumption level is required. Rosenzweig and Wolpin focus on productive assets and distinguish them according to their degree of liquidity. Zimmermann and Carter (2003) introduce the trade off between productive (see the Rosenzweig and Wolpin’s approach) and unproductive assets (see the Deaton's model), and focus on their safety and profitability. Productive assets (for example land) are risky but high return assets, unproductive ones are safe but low-return (for example grain). Households face two choice problems: they have to decide how to allocate resources between consumption and assets, and they have to solve the portfolio choice problem.

The key issue of these models is that they capture the trade off between current consumption and asset accumulation for future income, and hence for future consumption. We present a simplified version of the Zimmermann and Carter's model to describe the asset smoothing approach.

The production function is defined as:

$$F(T_{it}, A_{it}, \theta_{it}, \theta_{vt}) = \theta_{it}\theta_{vt} D \cdot (T_{it})^{\sigma} + \mu A_{it} \tag{1.13}$$

where θ_{it} and θ_{vt} are individual and common shocks, A and T denote nonproductive and productive assets respectively (grain and land), D and σ are productivity parameters, and μ is the rate of return on grain. The budget constraint is given by:

$$c_t \leq F(T_{it}, A_{it}, \theta_{it}, \theta_{vt}) - P_{Tt}(T_{it+1} - T_{it}) - (A_{it+1} - A_{it}) \tag{1.14}$$

where P_{Tt} is the relative price of the productive asset and the price of A is the numeraire.

The household maximization problem can be defined in terms of the value function as:

$$V_t(T_t, A_t) = \max_{T_{t+1}, A_{t+1}} u(c_t) + \beta E_t V_{t+1}(T_{t+1}, A_{t+1}) \tag{1.15}$$

where β is the discount factor. Solving this problem households decide how to allocate resources between consumption and assets, and between safe and risky assets. The two basic strategies identified by Zimmerman and Carter are: accumulate safe and less profitable assets to minimize the fluctuations on assets levels and aim to maintain these levels (asset smoothing), choose risky and profitable assets, and use them as buffer to smooth consumption (consumption smoothing). Results are supportive of the idea that poor households acquire a safe and less remunerative portfolio, and aim to maintain their level of assets to preserve their current and future income. They pursue asset smoothing instead of consumption smoothing. Consumption and investment strategies are very different for rich households. Rich households acquire a high-return portfolio which is used to smooth consumption.

Many empirical papers focus on the trade off between the use of productive assets as a buffer against shocks and their maintenance for future income and consumption. Fafchamps, Udry and Czukas (1998) find little evidence of livestock sales as a buffer against severe rainfall shocks in West Africa: livestock sales compensate only for at most thirty percent of village income shocks. Kazianga and Udry (2004) find that households in rural Burkina Faso "intentionally destabilized consumption in order to conserve livestock through the drought period" (Kazianga and Udry, 2004, p.24); the little consumption smoothing that households can do is achieved through savings in grain stocks (non-productive assets). A different result comes from Rosenzweig and Wolpin (1993). They show that bullock sales and purchases are important in allowing consumption smoothing for Indian households. This may have implications for the efficiency of agricultural production, in fact results suggest that poor farmers prefer to invest in liquid assets (bullock), and use them to smooth consumpion, than to finance irriversible investments (irrigation equipement) even if the latter are more profitable.

Some papers suggest that consumption and asset accumulation strategies vary according to the initial level of assets owned by the household, and test whether there exists a threshold above and below which asset management strategies bifurcate. Hoddinott (2006) find that farmers in Zimbabwe sell livestock to protect consumption (consumption smoothing) only if the initial level of livestock is above a given threshold. Otherwise, below that threshold households do not sell productive assets and reduce consumption in the face of a shock (asset smoothing). Barrett and Carter (2005) find that below a given assets threshold, households reduce consumption in order to preserve their stock of assets, while above that threshold assets are sold to protect consumption[4].

This literature suggests that consumption behaviour may be different for households with different endowment of assets. This issue will be explored in detail in chapter three.

1.3.3 Self-protection strategies: income smoothing via labour market

In a framework in which assets generate income, an implication of the models described above is that households may cope with risk by adopting strategies that smooth income (Morduch, 1995; Dercon, 2005a). For example they may choose inputs and production techniques that reduce variability, and/or diversify income sources (as noted above). These

[4] A further issue related to the use of productive assets as a buffer against shocks is the correlation between asset prices and shocks. Productive assets may be sold to a lower price because they are less liquid (Fafchamps, 1999), but even if productive assets were perfectly liquid, there could be in any case a negative correlation between asset prices and shocks (Zimmerman, 1993). Moreover, when assets, livestock for example, are sold in a hurry, the selling price could underestimates their real value, especially if markets are not well developed (Fafchamps, Udry and Czukas, 1998).

strategies could have long-term consequences, since they emphasize the gap between rich and poor and may push poor households into a poverty trap.

However, households may rely on income-based strategies when a shock occurs (ex post), through income earning activities, and in particular labour supply adjustments. There is empirical evidence of the use of labour supply as a response to shocks. Kochar (1999) finds evidence that farm households in India are able to cope with idiosyncratic crop shocks shifting from own-farm to off-farm work. Similarly, Cameron and Worswick (2003) find that the extra income generated by the labour supply response to a crop loss is important in allowing Indonesian households to avoid reducing consumption expenditure.

There is a general finding that poor households rely more than rich ones on the labour supply response. Maitra (2001) shows that Indian farmers differ in their response to similar shocks and in their ability to smooth consumption according to whether they are constrained or unconstrained. Farmers with unrestricted access to credit (medium and large farms) deal with a shocks using state contingent transfers (for example credit) and without changing their leisure and consumption behaviour. Constrained farmers (small farms) with restricted access to credit are able to insure consumption against unanticipated income changes only if they adjust their market participation in response to the shock (shifting from own-farm to off-farm work, hence leaving leisure unaffected).

This literature suggests that labour supply adjustments may be an important means to overcome income shocks, and that its use may differ between poor and non-poor households. Chapter three will explore these issues analyzing which variables influence the adoption of labour supply responses, and what is the role of this strategy in compensating the income reduction due to a specific shock (crop loss).

1.4 The limit to risk coping strategies and policy interventions

Before turning to data description we want to underline that coping strategies may be insufficient in protecting households or they may involve socially inefficient choices, thus justifying the need for policy interventions. These issues will not be examined in the thesis, but it is important to be aware of their existence and to keep them in mind when reading the results of the following chapters.

The previous paragraphs reviewed several coping strategies available to households. We underlined that the effectiveness of these strategies depends on the type of shocks, as well as on households' characteristics. In particular, we pointed out that poor households are typically less insured against shocks. Poor households have limited access to credit, they have a low level of savings, and they may be excluded from informal insurance networks. The inability to cope with shocks, and the insurance and credit market failure, may result in poverty traps (Dercon, 2005b): the poor may have no resources to overcome shocks and the shock-induced welfare loss may push households below some poverty threshold. With respect to the type of shock, we highlighted that some types of shock may be more difficult to insure than others, because of their frequency and intensity, and because they may have persistent effects. Asset shocks for example are more likely to lead to long run consequences when they affect productive assets or human capital accumulation; a growing literature shows that health shocks may have permanent consequences, in particular for children.

Even when households are able to deal with risk, the risk-management strategies they adopt may lead to inefficient choices. This is particularly true for ex ante strategies, and for poor households. We discussed this issue with regard to the income smoothing strategies, and to the need for precautionary savings. Poor households have to reduce risk (because of the lack of means to cope ex post with the shock), and hence they invest in low risk and low return activities. The need to use assets to smooth consumption may lead households to invest more in liquid assets, even if they have low returns, than in irreversible investments with high returns. This is costly in terms of low returns and low future incomes[5], and it may lead to risk-induced poverty persistence.

However, we discussed how the problem of shock-induced inefficient choices can arise also with ex post strategies; for example requiring children to drop out of school or to work may have long-term consequence and be socially inefficient. Moreover, productive assets may be sold to insure consumption, and this leads to future lower incomes.

The cost of risk in terms of persistent poverty, long term consequences and social inefficient choices justifies the need for public interventions. Dercon (2005b) points out that "a key objective for the provision of insurance against poverty is to make sure that risk is not a reason for persistent poverty, by ensuring that uninsured risk does not force the poor to invest in activities and assets of low profitability to limit their own exposure to risk" (Dercon, 2005b, p. 9). Dercon classifies the policy interventions between ex ante and ex post measures. The former aims to support the households' ability to protect themselves against shocks. They may be accomplished for example through measures that support saving behaviour and improve the access to credit. Ex post policies are safety net programs, like

[5] As suggested by Fafchamps (1999), precautionary savings are often liquid and not very productive.

emergency food assistance, employment guarantee schemes, and publicly funded health care for the indigent (Barrett and McPeak, 2006)[6].

1.6 Evidence about shocks and coping strategies in the Indonesia Family Life Survey

1.5.1 General description

This dissertation uses the Indonesia Family Life Survey data, in particular the 1993 round (IFLS1). In 1993, about 7200 households were interviewed in 13 provinces representing 83% of the Indonesian population. Indonesia is a country with a high variation in culture, geography and economic conditions. The relevant feature of the IFLS sample is that the 13 provinces included in the survey are able to capture this heterogeneity. The IFLS is extremely rich and contains detailed information on a variety of topics at the individual, household, and community level. For each IFLS1 household, representative members (typically the household heads) provided household-level demographic and economic information, such as household composition, incomes, assets, savings, expenditures, and the value of household consumables that were self-produced. In addition, several household members aged 15 and older (the household's head and the spouse if present, and up to three other members randomly selected) were asked to provide detailed individual information on a variety of topics, including education, employment, time allocation, marital history, migration, health, individual assets and non labour income. Another section provides child information (education and health) and was administered to selected children aged 0-14[7].

In the 1993 questionnaire a specific section on household economic shocks was presented to collect information on whether the household had experienced an economic hardship in the past five years, the type of shock, when it happened (year and month), what measures were taken and the costs of overcoming it. Six types of shock are analyzed in the IFLS dataset:

---

[6] Safety net must be designed in order to avoid negative externalities; these programs may incentive participants to leave their informal risk sharing arrangements, with negative consequences for the other households in the network (Dercon, 2005a).

[7] The second wave of IFLS was fielded in 1997 (IFL2), and the third wave in 2000 (IFLS3). About 90% of all the IFSL1 individuals were re-interviewed in 1997, and all household members aged 15 or older were interviewed individually. In each wave of IFLS, a community survey is added to the individual and household questionnaires (see Frankenberg and Karoly (1995) and Frankenberg and Thomas (2000) for a description of the surveys). We prefer to use the 1993 round because it gives more complete information on shocks and individual labour income, and to compare our results with that of other studies that use the 1993 round (for example Cameron and Worswick, 2003)

- death of a household member
- sickness of a household member
- crop loss
- household or business loss due to a disaster
- unemployment of a household member
- fall in the price of crop.

We noted in paragraph 1.2 that different types of shock involve different coping strategies. From this point of view it is useful to note that death and sickness are clearly demographic shocks, while the remaining four categories represent economic shocks. Crop loss and price falls are expected to be common shocks, death and sickness are usually considered as idiosyncratic hardships. This is only a general distinction: as mentioned above each shock may have both a common and idiosyncratic component, and households characteristics may influence the probability of experiencing the shock.

The 1993 round provides information also on the measures adopted to cope with the shocks. The survey allows us to distinguish six possible risk-coping mechanisms:

- extra job
- loan (including a loan from families or friends)
- asset sale (sale of next harvest, food supply, cattle/poultry, jewelry, other assets)
- family assistance
- use savings
- cuts expenditure.

This dissertation focuses on rural households and the database used in this chapter includes only those households that supplied a complete set of income and demographic data. After dropping income outliers (about 1% of the total sample) the sample reduces to 3601 rural households interviewed in 1993. Description of survey questionnaires, selected variables definition and the construction of household income and other main variables are shown in appendix A. In table 1.2 we show some descriptive statistics for the sample that we will use in the following sections of this chapter.

In the rural Indonesia, more than half of households (58%) own a farm, 30% own a non farm business, and the majority of households' heads work as self employed (60%) (see table 1.2)[8]. Nearly a third of heads have a second job, and nearly half of Indonesian rural households have at least one adult income earner other than the head. The level of education

[8] The households in which the head is private worker/employee (about one out of three) have an household income that is statistically higher than that of households with a self-employed head.

is in general low: the average years of school completed by the household head are about 4 years, where the length of the primary school in Indonesia is 6 years; the average education is even lower for female heads. Households have on average more than four members, two of whom have less than 18 years of age.

Table 1.2
Means of Household Characteristics

	Weighted means	*Standard errors*
Farm ownership	0.58	0.01
Business ownership	0.30	0.01
1992 value of farm assets (excluding zeroes)	5785.46	398.86
1992 value of business assets (excluding zeroes)	1497.77	389.53
1992 value of non-business assets (excluding zeroes)	5112.58	520.86
Household income	1012.98	28.66
Age household head	46.47	0.27
Head inactive	0.11	0.01
Head employee	0.28	0.01
Head self employed	0.60	0.01
Head family worker	0.01	0.00
Head has second job	0.27	0.008
Head years of school completed	3.73	0.07
Household size	4.36	0.04
# of household members aged 0 to 5	0.61	0.01
# of household members aged 6 to 17	1.25	0.02
# of household members aged 18 to 64	2.26	0.20
# of household members aged over 64	0.24	0.01
# of income earners (other than head)	0.68	0.02
# of male adults unschooled	0.17	0.01
# of female adults unschooled	0.39	0.01
# of male adults with primary education	0.65	0.01
# of female adults with primary education	0.67	0.01
# of male adults with secondary education	0.26	0.01
# of female adults with secondary education	0.16	0.01
# of male adults with high education	0.02	0.00
# of female adults with high education	0.01	0.00

The table reports the main characteristics of Indonesian rural households. Income, assets and expenditure are in thousands of rupiah.

The following paragraphs will examine whether the type of shock and household's characteristics (wealth indicators in particular) affect the propensity to experience and report the shocks, the households' ability and the cost to overcome them, and the different measures adopted. As pointed out in the previous sections, poor and rich households may differ in their exposure to risk and in their ability to insure against its effects; moreover different types of shock may be associated with different responses and different economic damages.

1.5.2 Self-reported shocks and household characteristics

To understand the extent to which shocks occur in the Indonesian rural context, table 1.3 shows the number (and percentages) of households reporting different number of shocks. It must be noted that when a household experienced the same type of shock more than once over the last five years, only one event is reported in the dataset; this allows us to capture only the cross-sectional frequencies. As pointed out by Newhouse (2003), there are no information on which shock is reported in the case of multiple occurrences, but the frequencies increase as we get close to the survey's year (1993); this suggests that recent shocks are more likely to be reported.

About one household out of three (1216 households) experienced at least one shock in the five years reporting period (table 1.3). Not all the shocks are equally frequent (see table 1.4): the most frequent ones are sickness and crop loss (9.4% and 15.4% respectively), whereas business loss due to a disaster and unemployment affect only a few households (1.6% and 2% respectively). It is worth noting that if we restrict the attention on farm households (defined as those who reported to own a farm business and at least a farm asset) about a fourth of them experienced a crop loss in the last five years.

Table 1.3
Number of shocks reported (previous 5 years)

# shocks		*# households*		*Weighted percentages*	
0		2,385		67.39	
1	at least one shock	913	1216	24.92	32.61
2		244		6.21	
3		52		1.34	
more than 3		7		0.08	

The table reports the number of shocks experienced by Indonesian rural households during the five years prior to the interview

Table 1.4
Reported household shocks, by type of shocks (previous 5 years)[9]

Type of shock	*# households*	*Weighted percentages*	*Commonality of the shock[10]*	
			weighted mean	*weighted median*
Death	284	7.02	4.74 (0.33)	3.57
Sickness	376	9.40	5.99 (0.44)	3.85
Crop loss	561	15.40	8.86 (0.82)	6.89
Disaster	63	1.60	4.38 (0.38)	3.70
Unemployment	65	1.99	3.96 (0.24)	3.57
Fall in price of crop	239	6.57	6.10 (0.62)	3.85

The table reports the number of households, and the percentage of all households sampled, reporting shocks of each type over the five year period 1989-93. The commonality of the shock is defined as the percentage of households reporting the same shock in the same village in 1993. Standard errors are in parenthesis

As noted above, an important characteristic of the shock is the extent to which it is common or idiosyncratic. Crop loss and price falls are expected to be more common than demographic shocks, even if each shock may have an idiosyncratic component, both in the probability of experiencing the shock, and in the extent of its effects. Looking at the percentage of households that experienced the same shocks in the same village in 1993 (fourth column of table 1.4), we can see that crop loss is the most common shock, even if the percentage is quite small. However, it has to be underlined that this commonality variable may be biased because only one occurrence of the same shocks is reported.

The occurrence and report of shocks may depend on household's characteristics. Some households may be more vulnerable against certain hardships, or less likely to report the shocks they experience. Looking at the frequencies of various shocks by household's wealth (per capita expenditure tercile), shown in table 1.5, we can observe however that the weighted percentages do not differ markedly except for sickness (for poor households) and disaster. With respect to the incidence of the former, the finding that poor households are less likely to experience/report a sickness may be unexpected if we believe that the probability of *experiencing* a sickness depends on household's economic/sanitary conditions,

[9] The final sample consists of 3601 rural households.
[10] Villages with no households reporting shocks are excluded from the median and the mean.

which are typically worse for poor households. To explore this idea we estimated a logit model to test whether these characteristics (the main flooring type used in the house, the main water source for drinking and cooking, where the household drain its sewage) affect the probability of having a sickness in 1993 (controlling for other household and village variables). Results are not supportive of this hypothesis. Moreover, the probability of *experiencing* this type of shock may be related to the number of children and old members in the household. The Pearson chi-squared tests reject the hypothesis of independence between having a sickness in 1993, and the number of children in the house (aged 0-5 and 6-11), as well as having a sickness and the number of members older than 64. These are higher in poor than in rich households, and this should lead to the opposite finding of the one reported in table 1.5.

There may be however other elements behind this result. First, rich households may be more likely than poor households to *report* a sickness for several reasons. Sickness is associated with direct and indirect costs, mainly medical expenses and the reduction in earned incomes (Gertler and Gruber, 2002). This reduction may occur because the sick householder is an income earner, or because there is need for a care-giver at home, and the care-giver is an income earner. The higher is the earned income that is lost, the higher is the indirect cost of sickness (opportunity cost), and the higher could be the reporting probability. As suggested by Newhouse (2003), another possible explanation for the finding that the frequency of reported sickness is higher for rich households than for the poor is that the IFLS survey asked to report a health shocks that requires hospitalization or continuous treatment. Higher income households are clearly more likely to report such a shocks.

Moreover, there is consistent evidence in the literature that not only wealthier but also more educated households are more likely to *report* sickness (Newhouse, 2003; Schultz and Tansel, 1997). IFLS data show that the estimated proportion of households that report a sickness is significantly lower if the head is unschooled or if he/she did not complete the primary school. Since the number of households with head illiterate is higher for poor than for non-poor households, the former have a lower probability to report sickness, thus confirming the result in tale 1.5.

With respect to the incidence of business loss due to a disaster, the probability of reporting this type of shock (the questionnaire asked for events that caused an economic damage) may be related to the economic value of the business: the higher is the value of business, the higher is the economic damage and the higher could be the probability of reporting the shock. Households in the top third of the expenditure distribution have a significant higher weighted mean of business profit than households in the bottom or in the middle third, and this could explain why the number of households reporting a business loss ("disaster") is higher for rich than for poor households.

Similarly, large farms or households whose income strongly depends on the farm business may have a higher propensity to report a crop loss. Data do not support this hypothesis. Focusing on the farm sample, there are not significant differences in the frequencies of reported crop loss by the size of the farm (tercile of 1992 farm assets) and by the share of farm profit on total income.

Table 1.5
Reported Household Shock Experience, by Per capita consumption (PCC) tercile, 1989-93

Type of shock	*Weighted percentages*		
	Bottom third	Middle third	Top third
Death	6.9	6.8	7.3
Sickness	6.8*	9.9	11.8
Crop loss	16.5	15.1	14.5
Disaster	0.3*	1.7*	2.9*
Unemployment	2.1	2.2	1.7
Price falls	7.1	5.7	6.9

The table reports the number of shocks experienced by Indonesian rural households during the five years prior to the interview, by per capita consumption tercile
* the difference in the mean is statistically significant at the 5% level

1.5.3 The economic cost of the shocks

As noted in paragraph 1.2, beside the frequency, the intensity is another important feature of shocks. We discussed above that some shocks are more frequent than others. We now present some data on the economic cost of the shock. The 1993 survey of IFLS reports the cost of overcoming each type of shock, estimated by the household's head. The variable is summarized in Table 1.6, for the five year period 1989-93. Leaving aside business loss due to a disaster and unemployment (which affect only a few households), the most costly shocks are the demographic ones: weighted means are 460 and 560 thousands of rupiah for death and sickness respectively, whereas the cost of crop loss and price falls are about 300 and 260. The difference in the medians is less marked, but the contrast persists.

Since the same absolute costs may have a completely different incidence on households' economic situation and since we have seen that wealthier households are more likely to report a sickness, in the last column of table 1.6 we present also the medians of the ratio between the economic costs and the 1993 household income. In terms of the percentage of household income spent to overcome shocks, crop loss is actually costlier than sickness. The implication is a greater proportional incidence of crop loss for poor households.

Table 1.6
Cost of Overcoming Shocks (shocks previous 5yrs)

	Thousand rupiahs			1993 cost/1993 household income (weighted median)
	# households	Weighted mean	Weighted median	
Death	258	464.8	230	63.5
Sickness	354	561.6	200	17.8
Crop loss	500	309.5	125	29.4
Disaster	58	972.4	200	30.0
Unemployment	55	584.0	225	18.9
Price falls	200	258.0	100	17.1

The table reports the self-reported cost of overcoming different types of shocks experienced during the five years prior to the interview

In interpreting these findings, one should keep in mind that survey respondents may not correctly report the costs of the shock, and there are missing values (the percentage of missing values varies from about 6% for sickness to about 16% for price falls). Moreover, the interpretation of this cost variable is not simple. As noted in paragraph 1.2, shocks may have both direct effects (in terms of lost income or destruction of assets) and indirect ones, as a result of the coping measures adopted[11]. It is not clear whether households report only the direct costs or also the indirect ones.

Because of the problems in interpreting the self-reported cost of the shocks and possible measurement errors in the declared values, we will not use this variable in the dissertation. We will estimate in each chapter the appropriate cost of the shock which is relevant for the decisions we will explore.

[11] Direct costs may be, for example, the medical expenditures related to health shocks. Data provide evidence that households that report a sickness in 1993 have a mean value of 1993 medical expenditures about five times larger than the mean value of those who did not report a sickness.

1.5.4 Risk coping strategies

As noted in the literature review, risk coping strategies vary with the type of shock (demographic or economic, idiosyncratic or common) and with household's characteristics (poor households may differ from rich households in their responses to shocks). This section examines data on households responses to shocks, differentiating for types of shocks and household's wealth.

Tables 1.7 and 1.8 report the weighted frequencies of responses to shocks by types of shock and by household's per capita consumption tercile. Percentages do not add up to 100 because households may choose more than one measure to cope with the same shock. Table 1.8 does not report disaster and unemployment shocks because there are only few observations.

Table 1.7
Shock Responses, 1989-1993

	Death	Sickness	Crop loss	Disaster	Unemployment	Price falls
Extra job	14.4%	13.9%	43.5%	19.4%	37.2%	39.0%
Take loan	24.3%	30.9%	21.6%	23.2%	27.4%	16.4%
Sell assets	32.5%	36.9%	21.0%	37.2%	8.7%	19.0%
Family assistance	33.4%	22.0%	8.0%	7.9%	29.0%	4.6%
Use savings	13.3%	14.6%	3.4%	5.2%	4.1%	3.3%
Cut down on household expenses	4.9%	7.9%	19.9%	22.3%	15.7%	32.0%

The table summarizes the percentages of those households which experienced each type of shock who identify each response mode. Because of multiple responses, percentages sum to more than 100%

Looking at table 1.7 some interesting facts emerge. The first regards the use of labour as an insurance mechanism in the face of economic shocks (crop loss and price falls in particular): nearly half of households that experience a crop loss respond with the labour supply, in line with the findings of other papers (Cameron and Worswick, 2003). This strategy is less well used to cope with demographic shocks: table 1.7 shows that only about 14% of households reporting sickness or death, change their labour supply choices. As noted in paragraph 1.2, the reason could be that these shocks may reduce the labour force both directly (when the sick household member is the income earner) and indirectly (through the need for domestic labour to take care of sick household members).

Households may cope with shocks by taking loans: data suggest that this measure is important for Indonesian rural household and its use does not vary according to the type of the shock. Family assistance (which includes receiving money and/or assistance in the form of goods from family or friends) is an important mean to overcome death, sickness and unemployment of a household member, while it is marginal as a response to crop loss, business loss due to a disaster, and price falls. A similar pattern emerges for the sale of assets, even if in this case its role in coping with economic shocks is not so marginal[12]. A possible explanation, as suggested by the literature and as pointed out in the previous paragraphs, could be that informal risk-sharing strategies and asset sales are able to protect against idiosyncratic shocks, but provide little insurance against common shocks (when all households in the same village try to sell assets, their prices fall).

The last issue regards the role of savings and expenditure cut. Savings are more used to cope with demographic than with common shocks, even if the percentage of households that use this measure is in general low. Cut down on household expenses follows the opposite pattern: 20% and 30% of households reporting a crop loss and a price fall respectively reduce consumption in the face of such shocks. The percentage is much lower for demographic shocks (between 5% and 8%), suggesting that consumption is better protected against these hardships.

The coping strategies adopted depend not only of the type of shocks, but on households' characteristics as well. Focusing on the distinction between poor and non-poor (defined on the basis on per-capita consumption), table 1.8 shows that poor households are more likely to use the labour supply response to cope with economic hardships (as pointed out by other authors, Cameron and Worswick, 2003, Newhouse, 2003, Maitra, 2001), whereas rich households are more likely to use savings, sell assets and take a loan.

[12] The variable "selling assets" is constructed aggregating the following measures reported in the questionnaire: selling next harvest in advance (below market value), food (rice) supply, cattle/poultry, jewelry, and other assets. Future research may consider to disaggregate the single measures, and to examine the different behaviour of productive and non productive assets, liquid and non-liquid assets, as suggested by the literature reported in the previous sections.

Table 1.8
Shock Responses, 1989-1993, by PCC tercile

PCC in bottom/middle/top third		Death	Sickness	Crop loss	Price falls
Extra job	%	7.7	20.0	48.3	45.6
		20.7*	14.6	48.6	42.4
		15.1	10.1	32.0**	28.7**
Take loan	%	24.0	35.2	18.4	19.7
		25.9	35.9	19.5	13.5
		23.1	23.9*	27.8*	15.3
Sell assets	%	30.8	28.0	22.3	20.8
		30.1	35.8	20.2	19.8
		36.3	43.5**	20.3	16.4
Family assistance	%	40.8	23.8	7.8	1.6
		30.0	20.7	7.1	2.7
		29.0	21.9	9.2	9.6*
Use savings	%	13.8	10.4	1.7	0.0
		6.5**	11.3	1.2	2.2
		19.3	20.2**	8.0**	7.8**
Cut down on household expenses	%	5.1	6.4	15.8	28.3
		6.9	7.7	22.8	32.1
		2.7	9	21.8	35.9

The table summarizes the percentages of those households which experienced each type of shock who identify each response mode, by per capita consumption tercile.
** the difference in the means is statistically significant at the 5% level
* the difference in the means is statistically significant at the 10% level

The definition of poor and non poor households is a crucial aspect in the context of the thesis' theme. Rural households may be poor in terms of income, but relatively rich in terms of wealth (wealth is defined as the difference between the value of assets and liabilities owned by the household, UNECE *et al*., 2007). Current income is often a poor-proxy for long-run resources, while wealth measures are important to understand household economic well-being dimension (Juster *et al*, 1999; Mishra *et al*., 2004; UNECE *et al*., 2007). For example, some households may have low incomes and relatively high wealth because their incomes are only temporarily low. In the presence of sufficiently developed capital and insurance markets, assets owned by the household significantly affect the stringency of the households' liquidity constraints (as suggested by Deaton, 1992). In the context of developing economies, this may not happen because credit markets often fail. Moreover, many rural households own assets not only to accumulate wealth, but also to provide the basis for generating income. In this framework, households with a low level of assets may aim to maintain that level, and their consumption behaviour may be more sensitive to current income than to wealth.

In the descriptive analysis of this chapter we used per capita consumption to identify poor and non poor households. In chapter two and four we estimate permanent income as a proxy for household's wealth, while chapter three focuses on farm households, and defines poor and non poor households on the basis of their level of farm assets (as a measure of wealth and a proxy for credit constraints)[13].

1.7 Conclusions

Indonesian households, as in many developing countries, have volatile incomes (Newhouse, 2003), especially in rural areas where household economy is highly dependent on family business (mainly farm business). In a framework in which insurance and credit markets are less developed, and incomes are typically low, risk may lead to inefficient economic choices. This chapter discussed households' behaviour in the face of risk, on the basis of both the review of the literature and the analysis of Indonesian data.

The theoretical framework that underlines the analysis is an intertemporal model of consumption and asset investment choices, that accounts for the fact that poor households, especially in developing countries, may have no access to perfect credit markets, and they may find it hard to save. Moreover, the vast majority of people living in poor areas do not rely on wage income, rather on farm profits, for which the availability of some productive assets becomes crucial.

Both the literature review and the descriptive analysis of Indonesian data, presented in the second part of this chapter, highlight that household responses to shocks depend on the nature of the shock (whether demographic or economic, idiosyncratic or common,…), and on household characteristics. For example, the most frequent shocks reported by Indonesian households are sickness of a household member and crop loss; crop loss is also the most common one, in terms of the percentage of households that experienced the same shocks in the same year in the same village. Households responses to these hardships are quite different: for example, households are less likely to use the labour supply response to cope with health shocks because they may affect the ability of the households to provide labour. More generally, family assistance and asset sale are important means to overcome

[13] Since poverty is a multi-dimensional phenomenon, another dimension that is really important in identifying poor versus non poor households is human capital owned by members of the household. However, in this thesis we adopt a more restrictive definition of poverty which focuses on economic conditions.

demographic shocks, while they are marginal as a response to economic hardships. On the other hand, households are more likely to take an extra job to cope with economic shocks. Relatively few households draw down savings as a coping strategy, and the percentage is even lower for economic shocks, in the face of which a higher percentage of households cut expenditure. The link between types of hardships, households' characteristics and coping strategies will be explored in detail in chapter two.

A characteristic that has important consequences on the way in which households respond to shocks is clearly household's wealth (and the presence of liquidity constraints that may be related to it). Indonesian data shows that poor and non-poor households differ in the choices of coping strategies: the former are more likely to use the labour supply response, while rich households are more likely to sell assets (for demographic shocks) and to use savings. This result is in line with the finding of the theoretical and empirical literature that when liquidity constraints are binding, households need to rely on autarchic savings, both to build a buffer stock of assets and to self-finance profitable investments. This implies that households below a certain threshold of assets reduce consumption in order to preserve their stock of assets, while above that threshold assets are sold to protect consumption. Chapter three will focus on consumption/saving behaviour of poor and non-poor households in the face of a specific shock (crop loss), investigating whether households follow a consumption smoothing strategy or the need to accumulate assets prevails.

The final point highlighted in this chapter is that the need to reduce risk and to build a buffer stock may be costly in terms of low returns and low future incomes (safe assets have typically low returns, and precautionary savings are often liquid and not very productive, as suggested by Fafchamps, 1999). This addresses another important issue highlighted in the literature: whether shocks and the coping strategies adopted to overcome them cause long-run effects. This may happen also when children are withdrawn from school in the face of shocks, since they may not be able to restart school or to recover the educational gap: in this way temporary schooling interruptions have lasting impacts. This issue will be explored in chapter four, looking at the effects of both ex ante risk and ex post shocks on child education.

Chapter 2

Models for Non-Exclusive Multinomial Choice, with Application to Indonesian Rural Households

2.1 Introduction

Textbook discussions of discrete choice modelling focus on binomial and multinomial choice models in which agents select a single response. We consider the situation of non-exclusive multinomial choice. One possibility, extensively used in the sociology literature, is to adopt the so-called Marginal Logit Model (MLM) which posits an independent binomial model for each choice (Agresti and Liu, 1999). The MLM is simple to compute but it has two disadvantages: it allows the possibility of null response which may not always be realistic, and, in a way which we will make clear below, it models choice outcomes rather than the choice process. We propose two alternative models which require at least one response and give rise to estimates which are interpretable within the standard stochastic utility framework.

The analysis of the way uncertainty affects poor households, and their responses to this uncertainty, is a key issue in developing countries. We apply the choice models that we develop to the responses of Indonesian rural households to demographic and economic shocks. The structure of the interviews from which we take our data requires a shock to have a response. While the majority of shocks elicit only a single response, some shock instances elicit multiple responses. It appears that multiple responses are to a large extent associated with particular interviewers employed in the survey process.

We develop two models. In the first and simpler specification, choices are modelled as sequential: a household first chooses the number of responses to a specific shock, and then the specific choices are identified to maximize household utility conditional on the former choice. In our particular case, we may think of the interviewer as selecting the number of responses and the interviewee identifying the particular responses. In the second specification, we generalize the standard multinomial logit model by supposing that agents will choose more than one response if the utility they derive from other choices is “close” to that of the utility-maximizing choice. In effect, this supposes selection of a utility maximizing band, which will contain at least one choice but may contain more than one. This specification makes choice of the number of responses joint with choice of the particular responses. We believe this will generally be the most realistic specification.

The chapter is organized into eight sections. The second reviews how the literature treats non-exclusive multinomial choice and presents the models we propose. The third section reviews how households cope with shocks and why coping strategies may have long run negative consequences. Section four and five discuss the data and the empirical methodology. Section six presents empirical results, and section seven tests the model specification. The final section concludes.

2.2 Non-exclusive multinomial choice

The standard multiple choice model is posed in terms of maximization of a random utility function. Our application is to adjustment to shocks. Any such adjustment imposes costs. We adapt the random utility framework by modelling choice as resulting from minimization of a random cost function. We suppose that households may experience one of a number $S \geq 1$ of shocks and respond to each shock experienced from a choice set comprising $M \geq 2$ of adjustment modes. Our data, in which all responses are available for all shock types, therefore allow a total of MS shock-response pairs.

Interviewees are asked to report which responses did they make, not by how much they responded for each possible response. In a neoclassical framework it is natural to suppose that all “consumptions” (i.e. all goods that enter the utility function) will adjust to shocks since these affect the marginal utility of money and hence disturb the entire set of first order conditions. However, in our data many interviewees report one or only a few response modes. The survey results can be interpreted in either of two ways:

a) agents report the single or the few most important responses because this is how the interview is structured or interpreted by the interviewer, even though other responses are non zero

b) discreteness of choices or fixed costs of adjustment imply that some responses are not made.

Let us interpret the adjustment cost in terms of fixed costs. In the extreme case in which there are high fixed costs, respondents would choose only a single response mode, the one that minimize such costs. If fixed costs are not high, and marginal utilities change substantially when households adjust consumptions (i.e. the second derivatives of the utility function are large), it may make sense to make multiple adjustments (making one large adjustment may be worse than making two small adjustments because of the curvature of the utility function).

The resulting optimization problem is complicated. There are two sets of optimality conditions. The first requires marginal utilities to be proportional to prices for each mode in which a positive adjustment is made. This determines the size of the adjustment conditional on the adjustment mode being selected. The second condition relates to the choice of adjustment modes, given that the scale of the adjustments determined by the first condition. This defines a discrete choice problem with the complication that the characteristics of each choice depend on the other choices made.

We do not observe the size of adjustments and so we are unable to investigate the first optimality condition. We therefore simply look at the second condition. Write c_{hms} for the cost of adjustment mode *m* to shock *s* for household *h*. These costs, which may be interpreted as the fixed costs of adjustment, will depend on a vector x_h of household characteristics. There are *H* households in the sample. Following the random utility approach, we assume that adjustment costs have a deterministic and a stochastic component and write

$$c_{hms} = f_{hms} + \varepsilon_{hms}. \tag{2.1}$$

The household chooses its adjustment mode(s) to minimize adjustment costs. Satisfaction of the budget constraint forces at least one response. Standard microeconomic theory suggests that it will be optimal for the household to make multiple responses such that marginal adjustment cost is equalized across modes. Either discreteness (for example, in taking an extra job) or fixed costs may result in zero adjustment in one or more modes.

Multinomial logit

This model is appropriate for data which only permit a single response. Write $r_{hms} = 1$ if household *h* chooses response $m \in \{1, 2, ..., M\}$ in response to a shock of type *s*. Define

$p_{hms} = \Pr\left(r_{hms} = 1 \mid x_h\right)$ as the probability that response *m* is the cost-minimizing response to shock *s*. For simplicity, focus on the first response p_{h1s}. Henceforth, we omit the shock subscript *s* where this does not result in ambiguity. Ignoring the possibility of ties

$$p_{h1} = \Pr\left(r_{h1} = 1\right) = \Pr\left(c_{h1} < c_{hm}, m = 2,..,M \mid x_h\right) \tag{2.2}$$

and similarly for the remaining *M*-1 choices. Following Domencich and McFadden (1975) assume that the stochastic cost components ε_{hm} follows an extreme value (Gnedenko) distribution. Then $c_{hm} = f_{hm} + \varepsilon_{hm}$ also has the same distribution as does the cost of the minimizing choice $c_h^* = \min_{m \in \{1,..,M\}} c_{hm}$. Hence the probabilities p_{hm} are logistic:

$$p_{hm} = \frac{e^{-f_{hm}}}{\sum_{j=1}^{M} e^{-f_{hj}}} = \frac{a_{hm}}{\sum_{j=1}^{M} a_{hj}} \quad \left(m = 1, 2, .., M\right) \tag{2.3}$$

where $a_{hm} = e^{-f_{hm}} \quad \left(m = 1, 2, .., M\right)$. (The minus signs reflect the fact that we are minimizing costs rather than maximizing utilities as in the standard random utility model).

The Marginal Logit Model

In many circumstances it will be possible for agents to have multiple responses. The simplest approach is to consider the marginal probability of choosing each response:

$$p_{hm} = \frac{e^{-f_{hm}}}{1 + e^{-f_{hm}}} = \frac{a_{hm}}{1 + a_{hm}} \tag{2.4}$$

This is the Marginal Logit Model (MLM) of Agresti and Liu (1999) – see also Loughin and Scherer (1998), Agresti (2003, chapter 11) and Liu and Agresti (2005). Marginalization allows us to estimate the probability of each response separately.

Although authors are not explicit in this regard, the MLM is best seen as a reduced form model responding to the question "Which responses are selected?" rather than a structural model responding to the question "Which responses do respondents select?" To see the difference between these two questions, note that there are 2^M-1 distinct sets of non-null choices from *M* elements. The answer to the former question is obtained from an *H* by *M* matrix of zeros and ones for each of the household-response pairs, while the answer to the latter question derives from a matrix of *H* by 2^M-1 matrix of zeros and ones, with a single one in each row corresponding to each of the possible response sets available to each household. Because the questions are different, the likelihoods are also different – the MLM likelihood is the likelihood of the vectors of responses while the structural likelihood is that of the household choices.

Despite the fact that the standard interpretation of the MLM is in terms of marginal probabilities, it is possible to give the model a structural interpretation in terms of stochastic utility theory. To see this, suppose that respondents compare the cost c_{hm} of adopting response *m* with a benchmark cost $c_{h0} = f_{h0} + \varepsilon_{h0}$ where the stochastic element ε_{h0} also follows an extreme value distribution. The MLM specification (2.4) results if the fixed cost component f_{hm} is identically zero.

In what follows, we develop two models that account for non-exclusive and dependent multiple responses: a Poisson model and a threshold model, which both generalize the random utility approach. The Poisson model supposes that the household makes a sequential decision, first choosing the number $m_h = \#(\Omega_h)$ of responses to a shock, and then identifying the best (i.e. cost-minimizing) m_h response, i.e. the set Ω_h conditional on this choice m_h. By contrast, the threshold model supposes that the number m_h of responses is an outcome of the response identification decision.

The Poisson-Multinomial Model[14]

The survey design obliges households to identify at least one response to any shock. Hence the number m_h of responses is a non-zero integer: $m_h \in \{1, 2, .., M\}$. If $m_h - 1$ follows a Poisson process with mean $\mu_h = \mu(x_h)$, where the unit displacement reflects the impossibility of a null response, we may write

$$\Pr(m_h \mid x_h) = \frac{e^{-\mu(x_h)} \mu(x_h)^{m_h - 1}}{(m_h - 1)!} \tag{2.5}$$

Consider first the case in which $m_h = 1$ and response *j* is selected. In the logit framework

---

14 There exists a literature on so-called multinomial-Poisson models in which individuals make multiple responses across a range of response modes. An example is transport mode frequencies for different transport modes in which households may use different modes on different occasions – see Terza and Wilson (1990). These models replace the multinomial response probabilities with Poisson frequencies. Our models adopt the polar opposite case in which responses remain categorical but the number of responses is variable and is modelled as Poisson.

$$\Pr\left(\Omega_h=\{j\}\mid m_h=1,x_h\right)=p_{hj}^1=\frac{e^{-f_{hj}}}{\sum_{m=1}^{M}e^{-f_{hm}}}=\frac{a_{hj}}{\sum_{m=1}^{M}a_{hm}} \tag{2.6}$$

Combining equations (2.5) and (2.6)

$$\Pr\left(\Omega_h=\{j\}\mid x_h\right)=p_{hj}=\frac{\theta_h a_{hj}}{\sum_{m=1}^{M}a_{hm}} \tag{2.7}$$

where $\theta_h=e^{-\mu_h}$.

Turning to the case in which $m_h = 2$ with responses *i* and *j* selected, we need to consider the probability that *j* is the overall cost minimizing choice and that *i* is the next best, and the converse situation in which *i* is the overall cost minimizing choice and that *j* is the next best. Following the derivation shown in Appendix 2B, the probability of choosing responses *i* and *j* given that $m_h = 2$ is given by:

$$\Pr\left(\Omega_h=\{i,j\}\mid x_h\right)=p_{hij}=\frac{a_{hi}a_{hj}\theta_h\mu_h}{\sum_{m=1}^{M}a_{hm}}.\left(\frac{1}{\left(\sum_{m=1}^{M}a_{hm}-a_{hj}\right)}+\frac{1}{\left(\sum_{m=1}^{M}a_{hm}-a_{hi}\right)}\right). \tag{2.8}$$

The argument is similar in the case that three responses are selected - see Appendix 2B.

The Threshold Multinomial Model

As previously, let Ω_h be the set of responses made by household *h* to a particular shock. We generalize the random utility framework by introducing a household-specific threshold $t_h=t\left(x_h\right)\geq 0$. Within this framework, the household may choose the first response to shock *s* either because this is the cost-minimizing choice or because one of the other choices is cost-minimizing but the cost of the first choice is sufficiently close.

First consider households which make a single response. Define

$$p_{h1}=\Pr\left(\Omega_h=\{1\}\right)$$

To fix ideas, consider the case in which $c_{h1}\leq c_{h2}\leq\ldots\leq c_{hM}$. If $\Omega_h=\{1\}$, it must be the case that response 2 is not "close" to response 1, i.e. $c_{h1}<c_{h2}-t_h$. It follows that

$$\begin{aligned} p_{h1} &= \Pr\left(c_{h1} \le c_{hm} - t_h, m = 2,...,M\right) \\ &= \Pr\left(f_{h1} + \varepsilon_{h1} \le f_{hm} + \varepsilon_{hm} - t_h, m = 2,...,M\right) \\ &= \Pr\left(\varepsilon_{h1} - \varepsilon_{hm} \le f_{hm} - f_{h1} - t_h, m = 2,...,M\right) \end{aligned} \tag{2.9}$$

If the errors have the extreme value distribution, and generalizing to the case in which response *j* is chosen:

$$p_{hj} = \frac{e^{-f_{hj}}}{e^{-f_{hj}} + \sum_{m=1,m\neq j}^{M} e^{-(f_{hm}-t_h)}} = \frac{a_{hj}}{a_{hj} + \lambda_h \sum_{m=1,m\neq j}^{M} a_{hm}} \tag{2.10}$$

where $\lambda_h = e^{t_h} \ge 1$.

Now consider a household which responds using two modes, say 1 and 2:

$$p_{h12} = \Pr\left(\Omega_h = \{1,2\}\right)$$

We need to consider two cases, that in which choice mode 1 is cost minimizing while mode 2 is sufficiently close to be also chosen, and the converse case in which 2 is cost minimizing and 1 is also chosen. Using the same notation

$$\begin{aligned} p_{h12} &= \Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 3,...,M\right) \\ &+ \Pr\left(c_{h2} \le c_{h1} \le c_{h2} + t_h \,\&\, c_{h2} \le c_{hm} - t_h, m = 3,...,M\right) \end{aligned} \tag{2.11}$$

In the general case for two choices in which $\Omega_h = \{j,k\}$, this probability is given by

$$p_{hjk} = \frac{(\lambda_h - 1)a_{hj}a_{hk}}{\left(a_{hj} + a_{hk} + \lambda_h \sum_{m\notin\Omega_h}^{M} a_{hm}\right)} \cdot \left[\frac{1}{\left(a_{hj} + \lambda_h \sum_{m=1,m\neq j}^{M} a_{hm}\right)} + \frac{1}{\left(a_{hk} + \lambda_h \sum_{m=1,m\neq k}^{M} a_{hm}\right)}\right] \tag{2.12}$$

See Appendix 2B for derivation of this result and also for the case in which three responses are selected.

Independence from Irrelevant Alternatives

As is well known, multinomial logit suffers from the independence from irrelevant alternatives (IIA) property. From an algebraic standpoint, IIA states the relative odds of any two response choices *m* and *n* are independent of the characteristics of other possible responses. This can be seen from the multinomial logit model in which equation (2.3) implies

$$\frac{p_{hm}}{p_{hn}} = \frac{e^{-f_{hm}}}{e^{-f_{hn}}} = \frac{a_{hm}}{a_{hn}} \quad (m, n = 1, 2, .., M) \qquad (2.13)$$

Within the multinomial logit framework, IIA can be resolved by allowing for unobserved heterogeneity in household utility (here cost) functions, reflected in an unobserved error component η_{hm} so that $c_{hm} = f_{hm} + \eta_{hm} + \varepsilon_{hm}$. On this specification, equation (2.3) becomes

$$p_{hm} = \frac{e^{-(f_{hm}+\eta_{hm})}}{\sum_{j=1}^{M} e^{-(f_{hj}+\eta_{hj})}} \quad (m = 1, 2, .., M) \qquad (2.14)$$

This defines the Mixed Multinomial Logit (MMNL) model – see Brownstone and Train (1999) and McFadden and Train (2000). The MMNL model can be estimated by simulated maximum likelihood once a distribution is specified for the error η_{hm}.

Moving away from the algebra, the IIA "thought experiment" asks us to conceive of an additional response, say a new mode of transport. In multinomial logit, the new response will attract customers from the existing modes such as to leave the relative probabilities of choosing the existing modes unaffected. The MMNL model allows the new mode to be correlated with existing modes, breaking this prediction. However, IIA ceases to be so problematic once we translate this thought experiment to the case of non-exclusive choice since choice of the new mode does not preclude choice of the existing modes. Indeed, if the characteristics of the new and existing modes are similar, it is likely that both will be chosen even without allowing for an explicit correlation structure as permitted by the MMNL.

This argument applies directly to the MLM even though equation (2.4) implies that the ratio of the marginal probabilities of any two responses are independent of the properties of the other responses. By contrast, the Poisson-multinomial model is closer to that of the standard multinomial since the number of responses is independent of the response characteristics. IIA continues to apply to a respondent who chooses to make a single response. Although equation (2.8) shows that, for respondents who make multiple responses, the odds of any two choices do depend on the properties of the entire response set, this is in a manner which makes it difficult for the model to reflect correlations among responses. Finally, by endogenizing the number of responses chosen, the threshold multinomial model becomes closer to the MLM in this regard – if a new mode is similar to an existing mode, it is likely that their expected costs (or utilities) will differ by less than the threshold. This is reflected in equation (2.12) which makes the probability of each response depend on the characteristics of the entire choice set.

The IIA issue does not arise in the application we consider in the following sections since we lack variables which vary both by household and response mode. For this reason, we do not further pursue mixed models.

2.3 Shocks, household responses and their consequences

As noted in chapter one, the analysis of the uncertainty affecting households, and their responses to this uncertainty, is a key issue in developing countries, where poor people are exposed to risks that affect household living conditions (Morduch, 1994; Dercon, 2005c). Shocks can have a major impact on the possibility of the household escaping poverty or may induce a non-poor household to enter poverty. Uncertainty is therefore central to our understanding of poverty. World Bank (2000) notes the importance of policies that help poor people to manage the risks they face.

A growing theoretical and empirical literature focuses on the analysis of income variability and on the ability of households to overcome income risks. Poor people have developed mechanisms to deal with hardships. Often these involve informal insurance arrangements between individuals and entire communities. Although these strategies offer some cushion against shocks, they are not always sufficient with the consequence that shocks may push households into poverty or exacerbate their existing poverty status. Even when households are able to deal with risk, the strategies they adopt may have negative long-run consequences (Dercon, 2005c). This is particularly true for ex ante risk-management mechanisms. These strategies may involve diversification across crops, the use of a variety of production techniques, portfolio diversification or migration. Negative long term consequences may arise when, for example, risks leads poor households to choose safe but less profitable choices (Morduch, 1990; Alderman and Paxson, 1992; Rosenzweig and Wolpin, 1993). However, the problem of shock-induced inefficient choices can arise also with ex post strategies that deal with the consequences of shock; for example requiring children to drop out of school or to work may have long-term consequences and be socially inefficient (de Janvry *et al.*, 2006). Moreover, productive assets may be sold to maintain consumption, resulting in lower incomes in the future. In this sense, short run income maintenance may be at the expense of longer-term well-being. Hence, if we are to design appropriate income protection frameworks, it is important to understand how households cope with actual shocks and the possibility of future shocks, and to evaluate which responses are costlier for households. Our application is a contribution to this literature.

2.4 Data

The data used for this study are from the 1993 Indonesia Family Life Survey data (IFLS1). 7224 households were interviewed over a wide range of issues. Our focus is on the section of the survey relating to demographic and economic shocks. Respondents were asked whether their household had experienced an economic shock in the past five years, the type of the shock, when it happened (year and month), what measures were taken and the costs of overcoming the shocks.

As noted in chapter one, six types of shock are analyzed in the IFLS dataset:

- i) death of a household member
- ii) sickness of a household member
- iii) crop loss
- iv) household or business loss due to a disaster
- v) unemployment of a household member
- vi) fall in the price of a crop.

We distinguish between demographic and economic shocks – demographic shocks are death and sickness, while economic shocks are the remaining four categories. The nature of the shock is important because it has implications for the ability to cope with its consequences (see Dercon, 2002), and influences the response adopted. A related distinction is between idiosyncratic and common shocks which is correlated with but not implied by the demographic-economic distinction.

Turning to the measures adopted to cope with the shocks, the survey allowed us to distinguish six possible responses:

- i) extra job
- ii) loan (including a loan from families or friends)
- iii) asset sale (sale of next harvest, food, cattle or poultry, jewellery or other assets)
- iv) family assistance
- v) use of savings

vi) reduction of expenditures.

The survey questionnaire was not explicit as to whether single or multiple responses to a shock were sought.[15]

The responses identified in the survey are all ex post risk-coping strategies.[16] They can be divided into two categories: risk sharing strategies that smooth consumption across households, and intertemporally smoothing strategies that smooth consumption over time. Risk sharing responses involve either formal institutions, such as formal credit transactions, or informal mechanisms, such as transfers between families or friends. Alternatively, households can smooth consumption intertemporally by saving and borrowing, or by accumulating and selling non-financial assets (Alderman and Paxson, 1992; Bardhan and Udry, 1999; Dercon, 2002).

Table 2.2
Reported Household Shock Experience

Type of shock	1989-93		1992-93	
	# rural households	per cent	# rural households	per cent
Death	254	7.8%	111	3.5%
Sickness	325	10.0%	169	5.2%
Crop loss	544	16.8%	340	10.5%
Business loss	60	1.8%	35	1.0%
Unemployment	54	1.7%	28	0.9%
Price falls	231	7.0%	147	4.5%

The table reports the number of households, and the percentage of all households sampled, reporting shocks of each type over the five year period 1989-93 and the two year sub-period 1992-93 used in the subsequent analysis.

We include only those households in our dataset that supplied a complete set of income and demographic data. After dropping income outliers (1% of the total sample), and considering only rural households, the sample reduces to 3246 households. 1116 households (34.4% of the total sample) experienced at least one shock in the five year reporting period, 697 of them (21,5% of the total sample) experienced at least one shock in the final two years, 1992-93. Table 2.1 reports the number of households that experienced each type of shock over the five years 1989-93 and the two years 1992-93.

[15] The survey also includes an explicit question on the costs associated with each shock. We do not use the answers to this question in this chapter for two reasons. First, response is partial. Second, it is unclear whether this variable measures the pre- or post-response (i.e. gross or net) costs associated with shocks. (The variable we have defined in equation (2.1) is on a gross basis).

[16] Ex ante risk-management strategies shape the risks households face by choosing particular activities (diversification across crops, the use of a variety of production techniques, etc).

The most frequent shocks are sickness and crop loss. Business loss and unemployment affect only a few households. In view of the low incidence of these shocks in our data, we aggregate these into a single category reducing the number *M* of shock types to five for the purposes of econometric analysis.

Table 2.2 shows the percentage of multiple responses for each shock. The majority of households report a single response. This is consistent with the view either that responses are interdependent - the fact of having chosen (or reported) one response mode reduces the probability of choosing (or reporting) others – or that many interviewers interpreted the survey question as requiring a single response.

Table 3.2
Percentage of Multiple Responses Reported

	1989-93	1992-93
Death	19.7%	22.5%
Sickness	19.7%	16.6%
Crop loss	18.6%	17.4%
Business loss due to a disaster	15.0%	8.6%
Unemployment	24.0%	21.5%
Price falls	15.0%	17.7%

The table reports the percentage of those households which experienced each type of shock who reported multiple responses over the five year period 1989-93 and the two year sub-period 1992-93 used in the subsequent analysis.

Table 2.3 shows the percentage of households that responded in each manner for each shock. These statistics suggest that household responses differ between demographic shocks (death, illness) and economic shocks (crop loss, business loss or unemployment, price falls). The data suggest an important role for family and community assistance in the case of demographic shocks, while this measure appears relatively less important as a response to crop loss and price falls.

Cameron and Worswick (2003) have argued that labour supply responses help Indonesian households to smooth consumption in the face of a crop loss. This response appears particularly important for economic shocks. It is also apparent that economic shocks are more likely to lead to a decline in consumption than are demographic shocks.

Table 2.4
Shock Responses by Household (1992-93)

	Demographic		Economic		
	Death	Sickness	Crop loss	Business loss or unemployment	Price falls
Extra job	13.5%	7.7%	47.9%	34.4%	42.2%
Take loan	28.8%	36.0%	18.8%	29.5%	17.0%
Sell assets	27.0%	27.2%	15.6%	23.0%	17.7%
Family assistance	36.0%	21.9%	7.0%	13.0%	4.0%
Use savings	15.3%	17.7%	5.0%	3.3%	4.0%
Cut down on household expenses	4.5%	4.7%	22.6%	14.7%	32.6%

The table summarizes the percentages of those households which experienced each type of shock who identify each response mode the two year sub-period 1992-93 used in the subsequent analysis. Because of multiple responses, percentages sum to more than 100%.

We turn now to the explanatory variables we will use in the econometric model. Microeconomic theory indicates that two variables are potentially important in explaining shock responses. The first is the extent to which shocks are common across households. Standard discussions indicate that informal insurance mechanisms are better able to cope with idiosyncratic shocks than with common shocks. The dataset allows us to identify the village in which a household is resident. We define a commonality variable as the weighted percentage of households (other than the household in question) that experienced the same shock in the same village. Let Z_{hs} be the percentage of families that experience shock *s* in the village in which household *h* is resident. The commonality variable, modifies to exclude the household in question, is defined as $z_{hs} = Z_{hs}\frac{n_h}{n_h - 1} - \frac{1}{n_h - 1}\delta_{hs}$, where n_h is the number of households surveyed in the village and δ_{hs} is a dummy variable equal to one if the household *h* has experienced shock. The modification is important because the unmodified variable Z_{hs} will not be independent of δ_{hs} in villages in which there is a small number of reporting households giving rise to potential simultaneity bias.

Some shocks are more likely to be idiosyncratic (for example illness or unemployment) or common (for example droughts or epidemics), but few risks purely belong to one category (Dercon, 2005a). Each shock may have an idiosyncratic component, both in the probability of experiencing the shock, and in the extent of its effects. The data shows that crop loss is the most common shock, followed by price fall and sickness of a household member.

Commonality is on average lower for demographic than for economic shocks (the median of commonality is 4.5 for economic and 4.1 for demographic shocks, averages are respectively 8.8 and 5.4, and the difference is statistically significant).

The second potentially important variable is household permanent income. Certain shock responses are more easily available to rich households than to poor households. We should therefore expect that the probability of choosing a specific mode will be affected by the household's wealth. For example, poor people are less able to save and accumulate assets, and they have restricted access to credit because of lack of collaterals. Poor households are thus more vulnerable and have limited means to deal with risk. We measure household wealth through estimated permanent income. Construction of the permanent income variable is discussed in Appendix 2A.

In addition to these economic variables, the survey design (in particular, incomplete instructions) may allow the identity of the interviewer to play a role in determining the number of responses chosen by the household. Since we are able to identify the interviewer for each respondent, we relate the number of responses chosen by each household to the average number of responses elicited by the same interviewer, excluding responses given by the household in question.

2.5 The empirical model

We define the cost associated with a certain measure and a certain shock as a random cost with a stochastic and a deterministic component – see section 2. We assume that households choose the response mode that minimizes this adjustment cost or, in the case of multiple responses, that has a cost sufficiently close to the best choice. To implement the models we need to specify the deterministic cost component (f_{hms} in equation (2.1)), the household specific threshold (t_{hs}), and the Poisson parameter μ_{hs}. Hence the empirical strategy has two components: the first is the definition and estimation of the deterministic cost, and the second involves the adjustment of the multinomial model to account for non-exclusivity of responses.

The three variables we use vary across households and shocks but not across response modes. This implies that the response probabilities to a particular shock all depend on the same three variables. Alternative models imply different nonlinear mappings of these variables into the unit interval. In principles, each such mapping might depend on a large number of parameters, but with a limited number of observations it is difficult to identify all parameters. (In particular, a number of shock-response pairs are poorly represented in the data). We impose structure on these mappings by jointly estimating the response probabilities across shocks and by imposing a degree of response homogeneity. Agresti and Liu (2001) refer to this approach in the context of the MLM as a Simultaneous MLM.

We adopt a linear specification for the deterministic cost component:

$$f_{hms} = \kappa_{ms} + \gamma_m z_{hs} + \alpha_{ms} y_h^P \qquad (2.15)$$

z_{hs} is the variable defined in section 4 that captures the commonality of the shock in each village, and y_h^P is the estimated household permanent income. We impose homogeneity on the intercepts for the demographic and economic shocks respectively

$$\begin{aligned} \kappa_{ms} &= \kappa_m^d \quad (s = 1,2) \\ \kappa_{ms} &= \kappa_m^e \quad (s = 3,4,5) \end{aligned} \qquad (2.16)$$

In the Poisson model we posit

$$\mu_{hs} = \exp(\mu_s + \beta_s v_{hs}) \qquad (2.17)$$

where v_{hs} is the average number of responses elicited by household *h*'s interviewer to shock type *s*. Similarly, for the threshold model we suppose that the threshold, beneath which response costs are regarded as indistinguishable, is influenced by the identity of the interviewer, motivating the specification

$$t_{hs} = \exp(\tau_s + \varphi_s v_{hs}), \qquad (2.18)$$

where $s = \{d,e\}$ (demographic and economic).

The estimated cost functions are latent and therefore have an arbitrary zero. This implies that we need to normalize the parameters. We do this by setting the adjustment costs c_{h1s} to zero for each shock type *s*. The implied parameter restrictions are

$$\kappa_{1s} = \gamma_1 = \alpha_{1s} = 0 \quad (s = 1,..,S). \qquad (2.19)$$

To evaluate the importance of modelling multiple responses as interdependent, we compare the results obtained from these models with those from the MLM which treats each response as an independent decision. Both the threshold and the Poisson-multinomial models use the identical adjustment costs expression (2.15) but substitutes the MLM probability

(2.3) for the threshold and the Poisson probabilities. Note that parameter normalization is not required in the MLM case – the alternative to responding in a particular manner is not making that response. The implication is that it is only the intra-response mode differences that are comparable across the two models. To obtain comparability between the MLM, the threshold and Poisson probabilities, we can re-normalize equation (2.15) for the MLM model, as

$$f_{hms} = (k_{1s} + \kappa_{ms}) + (g_1 + \gamma_m) z_{hs} + (a_{1s} + \alpha_{ms}) y_h^P + \beta_s v_{hs} \tag{2.20}$$

in conjunction with the restrictions given by equations (2.19).

2.6 Results

We have estimated the parameters of the deterministic component of the cost function (2.15) using the MLM, the Poisson-multinomial and the threshold multinomial models. Results are reported in Table 2.4 (see end of the chapter).

- There are marked differences between the reactions to demographic and economic shocks. In the case of economic shocks, the labour supply response is associated with the lowest costs, holding other variables in the model constant. Relative to this, demographic shocks increase the probability of taking a loan, sale of assets and family assistance with respect to the labour supply response. Use of savings and expenditure reduction remain the least favoured responses . These results reinforce the findings of the literature that poor households use the labour market to smooth income more than they use financial assets to smooth consumption (Maitra, 2001; Kochar, 1999; Cameron and Worswick, 2003). However, the results are qualified by the fact that adjustment of labour supply is less likely in the face of demographic shocks. As suggested by Kochar (1995), demographic shocks, such as loss of family members and sickness, may affect the ability of the household to provide labour, making labour income ineffective as a source of insurance.
- The estimated models provide clear evidence that the probability of responding to shocks through use of savings increases with permanent income (i.e. the cost of responding through the use of savings is significantly negatively related to permanent income)[17]. Only the richest households are able to use savings to smooth consumption.

[17] We ran the Poisson and threshold models with the full set of α coefficients. All apart from α_5 ("use savings") were close to zero. Even in the MLM model we cannot reject the hypothesis that all the coefficients on permanent income except α_5 are equal to zero ($\chi^2(5) = 3.81$, tail probability 0.58).

- Shock commonality increases the probability of a labour supply response. In the MLM specification the estimated coefficients on commonality are only significant for the responses "extra job" and "asset sales". These results provide strong evidence of both the importance and limitations of informal insurance arrangements for poor communities. Shock commonality is likely to be lower for demographic than for economic shocks accounting for the lower probability of adjusting labour supply in the face of demographic shocks.
- The average number of responses for the interviewer is highly significant for economic shocks, and fairly significant for demographic shocks irrespective of model specification confirming that interviewer identity plays an important role in determining the number of responses. In the Poisson-multinomial model, an increase in the average number of responses elicited by the same interviewer, excluding responses given by the household in question, increases the expected counts, holding other variables constant. Similarly, in the threshold-multinomial specification the interviewer variables has a positive effect on the threshold, thus increasing the number of responses chosen by the household.

These results are all in line with the qualitative conclusions drawn from Tables 1-3 in section 4.

2.7 Testing the model specification

As noted in section 2, the MLM and the multinomial specifications answer two different questions and maximize different likelihood functions. The MLM likelihood is defined in terms of the probability of each response being selected. If household *h* selects response *m* to shock *s*, $r_{hms} = 1$. This outcome occurs with probability p_{hms}. Similarly, the outcome and $r_{hms} = 0$ occurs with probability $(1-p_{hms})$. The overall probability can be written in the binomial form $p_{hms}^{r_{hms}}(1-p_{hms})^{1-r_{hms}}$ and the log-likelihood is $r_{hms}\ln p_{hms} + (1-r_{hms})\ln(1-p_{hms})$. The overall log-likelihood function is

$$LB = \sum_{h=1}^{H}\sum_{s=1}^{S}\sum_{m=1}^{M}\left[r_{hms}\ln p_{hms} + (1-r_{hms})\ln(1-p_{hms})\right] \qquad (2.21)$$

In the multinomial specification the likelihood function is maximized over the entire set of all the possible combinations of responses. The entire set of possible combinations up to three choices is given by Q = 41 possibilities and we index these by q such that $\Omega_{hs}^{1} = \{1\}$,

…, $\Omega_{hs}^{6}=\{6\}$, $\Omega_{hs}^{7}=\{1,2\}$ …, $\Omega_{hs}^{11}=\{1,6\}$ etc. Define $\tilde{r}_{hqs}=1$ if the combination Ω^{q} is chosen, 0 otherwise, with $q=1,\ldots,Q$ and let $\tilde{p}_{hqs}$ and $1-\tilde{p}_{hqs}$ be the associated probabilities. The log-likelihood for the Poisson and the threshold models is defined as:

$$LJ=\sum_{h=1}^{H}\sum_{s=1}^{S}\sum_{q=1}^{Q}\left[\tilde{r}_{hqs}\ln\tilde{p}_{hqs}+(1-\tilde{r}_{hqs})\ln(1-\tilde{p}_{hqs})\right] \tag{2.22}$$

The MLM and multinomial likelihood functions are not directly comparable although either can be transformed into the other. Given the MLM choice probabilities p_{hms} estimated from the MLM model, the corresponding multinomial probabilities $\tilde{p}_{hs}^{q}$ may be computed as

$$\tilde{p}_{hqs}=\prod_{m\in\Omega^{q}}p_{hms}\prod_{m\notin\Omega^{q}}\left(1-p_{hms}\right) \tag{2.23}$$

Equivalently, given the multinomial probabilities $\tilde{p}_{hs}^{q}$ we may compute the associated probabilities p_{hms} associated with each choice as

$$p_{hms}=\sum_{q=1}^{Q}1\left(m\in\Omega^{q}\right)\tilde{p}_{hqs} \tag{2.24}$$

where the function $1(v)$ returns the value unity if v is true and zero if false.

We use expressions (2.23) and (2.24) to calculate the structural (choice-based) likelihood based on the estimated MLM probabilities and the MLM (response-based) probabilities for the two structural multinomial models. Table 2.5 lists the maximized log-likelihoods on both bases for all three specifications. The two multinomial models have higher log-likelihoods irrespective of the choice basis. On the MLM choice basis, the threshold-multinomial model slightly out-performs the Poisson-multinomial model, but the ranking is reversed on the choice basis.

Table 2.5
Log-likelihoods

	MLM basis *LB*	Structural basis *LJ*
Marginal logit model	-2125.75	-2518.34
Poisson-multinomial	-2124.66	-2454.87
Threshold-multinomial	-2123.93	-2456.67

The table records the results maximized log-likelihoods from equations (2.21) and (2.22) for the three models considered .

The three models we have considered are not nested and comparison of likelihoods is therefore at best a criterion for good fit and not a test. In what follows, we first use a version of the paired *J* test introduced by Davidson and Mackinnon (1981).[18]

Index the three specifications as (*b*,*p*,*t*) for the MLM, Poisson-multinomial and threshold-multinomial models respectively. Write the estimated individual choice-based probabilities as $p_{hms}^{j} \quad (j = b, p, t)$ and the estimated joint choice-based probabilities as $\tilde{p}_{hqs}^{j} \quad (j = b, p, t)$. Construct the two sets of differences

$$d_{hms}^{j} = p_{hms}^{j} - p_{hms}^{b} \quad (j = p, t) \tag{2.25}$$

To perform the *J*-type test we include these differences d_{hms}^{j} additively in the augmented MLM model. The *J* test statistic for the MLM null against alternative *j* is the one-sided *t* statistic on the coefficient λ_j of the variable d_{hms}^{j}.

The procedure for testing the two multinomial models is identical. We construct the four set of differences[19]

$$\tilde{d}_{hqs}^{jk} = \tilde{p}_{hqs}^{j} - \tilde{p}_{hqs}^{k} \quad \left((j,k) = (p,b),(t,b),(p,t),(t,p)\right) \tag{2.26}$$

Regarding model *j* as the null, we re-estimate the model including the difference $\tilde{d}_{hqs}^{jk}$ as an additive regressor. The *J* test statistic for null *j* against alternative *k* is the *t* statistic on the coefficient λ_{jk}. Monte Carlo evidence has established that *J* tests have a pronounced tendency to over-reject in finite samples – see McAleer and Pesaran (1986) and McAleer (1987).

Test outcomes are listed in Table 2.6. At the 5% level, neither of the multinomial models rejects the MLM model whereas the two multinomial models reject each other with the threshold model also being rejected by the MLM model. These outcomes are not easy to reconcile with the likelihood values reported in Table 2.5. We note that the properties of the *J*-style test have not been established for nonlinear environments and it also seems possible that our sample, although large, is insufficient to give reliable results.

---

[18] A "paired" non-nested test is a test between a pair of two hypotheses from a larger set of hypotheses (McAleer, 1995). Our test should be thought of as a *J*-type test rather than a pure *J* test since our models do not fall within the linear regression class.

[19] In effect three since $\tilde{d}_{hqs}^{tp} = -\tilde{d}_{hqs}^{pt}$.

Table 2.6
***J* Test Results**

		Alternative hypothesis		
		Marginal logit model	Poisson -multinomial	Threshold -multinomial
Null hypothesis	Marginal logit model	-	1.80 [3.60%]	1.41 [7.93%]
	Poisson-multinomial	1.20 [11.5%]	-	2.94 [0.16%]
	Threshold-multinomial	2.23 [1.29%]	2.58 [0.50%]	-

The table records the results of the *J* tests for each pair of models. The test statistics are distributed as Student *t*. Tail probabilities are given in parentheses. The hypotheses tests are all one-sided so rejection of the null at the conventional 95% level is appropriate if the tail probability is inferior to 5%. The tests are calculated using equations (2.25) and (2.26).

We obtain clearer results from an alternative approach using a linear probability (LPM) framework. To test the MLM null, consider the six regressions

$$r_{hms} = \delta_j p^j_{hms} + \delta_k p^k_{hms} + u^{jk}_{hms} \quad \left(j,k = b,p,t; k \neq j \right) \tag{2.27}$$

If the MLM model (H_j say) is valid, we should find $\delta_j = 1$ and $\delta_k = 0$. Conversely, if hypothesis k (H_k) is valid we should find $\delta_j = 0$ and $\delta_k = 1$. Similarly, in the multinomial framework, we consider the six regressions

$$\tilde{r}^j_{hqs} = \delta_j \tilde{p}^j_{hqs} + \delta_k \tilde{p}^k_{hqs} + \tilde{u}^{jk}_{hqs} \quad \left(j,k = b,p,t; k \neq j \right) \tag{2.28}$$

The tests have the same form. As is well-known, the LPM suffers from heteroscedasticity and so in all cases we use a heteroscedasticity-robust estimate of the variance-covariance matrix.

Results are reported in Table 2.7. The upper block of tests relates to the MLM repose basis. The tests fail to discriminate between the alternative models even though the estimated coefficients give a greater weight to the Poisson and threshold probabilities than to those from the MLM itself. By contrast, using the joint choice basis (Table 2.7, lower block) the two multinomial models decisively reject the MLM model. Although it remains true that neither multinomial model is able to reject the other, the estimated coefficients give a higher weight to the threshold model in line with the log-likelihoods reported in Table 2.5.

Table 2.7
Linear Probability Model Test Results

Individual choice basis		δ_b	δ_p	δ_t	H_j versus H_k	H_k versus H_j
H_j	Marginal logit model	0.398	0.603	-	1.21	0.53
H_k	Poisson-multinomial	(1.03)	(1.55)		[29.9%]	[59.0%]
H_j	Marginal logit model	0.315	-	0.692	1.77	0.42
H_k	Threshold-multinomial	(0.86)		(1.88)	[17.0%]	[65.9%]
H_j	Poisson-multinomial	-	0.334	0.672	0.83	0.24
H_k	Threshold-multinomial		(0.64)	(1.29)	[43.7%]	[78.3%]
Joint choice basis						
H_j	Poisson-multinomial	0.011	1.005	-	0.10	32.7
H_k	Marginal logit model	(0.09)	(8.09)		[90.8%]	[0.00%]
H_j	Threshold-multinomial	0.049	-	0.975	0.26	36.8
H_k	Marginal logit model	(0.42)		(8.58)	[77.0%]	[0.00%]
H_j	Poisson-multinomial	-	0.287	0.733	2.18	0.43
H_k	Threshold-multinomial		(0.78)	(2.01)	[11.3%]	[65.0%]

The table reports the results of the test based using the linear probability (LPM) framework, described in equations (2.27) and (2.28). Hypothesis test statistics (column 4 and 5) are heteroscedasticity-corrected F tests. Heteroscedasticity-robust t-statistics are in "(.)" parentheses and tail probabilities in "[.]" parentheses.

In summary, the test outcomes depend on the way the model is framed. If the question is, "Which response modes will be adopted?", this motivates an individual choice approach. In this case, the standard MLM model appears adequate. If, instead, the question is, "How will households respond?", a joint choice is required. In this second context, it is important to explicitly acknowledge the joint nature of multiple response choices and the MLM model is clearly inadequate. The evidence is less decisive in relation to the choice between alternative multinomial specifications although there is some suggestion that the threshold-multinomial model is slightly superior to the Poisson-multinomial model. Since the threshold model is also to be preferred in relation to the IIA critique, this is the model we recommend.

2.8 Conclusions

Multinomial choice models have traditionally focussed on exclusive choice. Survey design may however permit multiple responses. One possibility is to model such responses using the MLM, but this allows the possibility of null response and, in general, fails to fully reflect the structure of the choice process. We have developed two models which generalize the McFadden's now standard random utility framework to allow for the possibility of multiple response. In the first of these models, the respondent first decides on the number of responses and then chooses the actual responses to maximize utility conditional on that prior choice. In the second, threshold, model, the two decisions are made jointly, with the agent choosing multiple responses if utility outcomes are sufficiently close.

These models are both relatively straightforward from a computational standpoint provided the number of responses selected remains small. From a theoretical standpoint, the models are an advance over the standard MLM approach in that they see respondents as comparing response with each other rather than with a common benchmark. A disadvantage of the Poisson multinomial model is that it inherits the IIA property from multinomial logit. Both the MLM and threshold multinomial models, which make the number of responses endogenous to the respondent's decision process, are less vulnerable to this problem.

We apply this framework to modelling the responses of households in rural Indonesia to demographic and economic shocks. The survey design obliges respondents to nominate at least one response to any such shock. A minority of households nominate multiple responses. The incidence of multiple responses appears to be primarily a function of the identity of the interviewer, and it appears that interviewers may have interpreted the survey instructions differently. Both the Poisson and threshold multinomial models outperform the MLM model. Choice between the two multinomial models is less clear but the data appear to be marginally more favourable to the threshold model.

There are also substantive conclusions. Macroeconomic theory emphasizes the role of individual household savings as a device for smoothing consumption in the face of income shocks. Our data for rural Indonesian households demonstrates the importance of this mechanism but only for the richest households. By contrast, the theoretical literature on shock response in development economics has emphasized the role of labour supply and informal insurance arrangements at the family and village level, but has noted that informal arrangements only work well when shocks are idiosyncratic. We develop a measure of the commonality of shocks and show that response choice does indeed depend on commonality. Shock commonality increases the probability of a labour supply response with respect to all other coping mechanisms. Results provide confirmation of the importance of labour supply as a strategy to cope with economic and common shocks, and of the role of informal arrangements in coping with demographic and idiosyncratic hardships.

Table 2.4
Estimated Models

	Marginal logit model		Poisson-multinomial model		Threshold-multinomial model	
	Coeff.	*t*	*Coeff.*	*t*	*Coeff.*	*t*
κ *(demographic)*						
extra job	2.81	6.91				
loan	0.93	2.51	-1.54	-6.89	-1.49	-6.75
sell assets	1.13	2.99	-1.40	-6.05	-1.34	-5.89
family assistance	1.17	3.04	-1.36	-5.56	-1.33	-5.46
use savings	2.94	6.27	0.24	0.71	0.15	0.41
cut expenses	3.41	7.53	0.64	1.83	0.62	1.82
κ *(economic)*						
extra job	1.54	5.46				
loan	2.05	6.88	0.40	2.13	0.34	1.92
sell assets	2.05	6.73	0.38	1.93	0.34	1.83
family assistance	3.22	8.91	1.40	5.10	1.28	4.84
use savings	4.71	10.26	2.82	7.45	2.63	7.26
cut expenses	1.94	7.09	0.31	1.78	0.24	1.44
γ *(commonality)*						
extra job	-0.29	-4.18				
loan	0.08	1.17	0.27	3.32	0.26	3.40
sell assets	0.20	2.22	0.37	4.04	0.36	4.06
family assistance	0.12	0.97	0.31	2.53	0.31	2.63
use savings	0.11	0.72	0.31	2.00	0.30	2.01
cut expenses	0.01	0.15	0.20	2.72	0.19	2.79
α *(permanent income)*						
use savings	-0.55	-5.03	-0.52	-5.52	-0.47	-5.43
β *interviewer*						
demographic	-0.30	-1.05				
economic	-0.69	-3.46				
μ, τ *(intercept)*						
demographic			-2.84	-4.74	-3.00	-3.04
economic			-5.12	-10.05	-6.29	-9.27
μ,τ *(interviewer)*						
demographic			1.09	2.31	1.56	1.90
economic			2.73	7.29	4.01	7.84
log likelihood	-2125.7444		-2454.87		-2456.67	

The table reports the estimated parameters of the deterministic component of the cost function (2.15) using respectively the MLM, the Poisson-multinomial and the threshold multinomial models. In the Poisson and threshold models, parameters are normalized setting the adjustment cost c_{h1s} to zero for each shock type s ($\kappa_{1s} = \gamma_1 = \alpha_{1s} = 0 \quad (s = 1,..,S)$). The estimated threshold t_{hs} and Poisson parameters μ_{hs} are also reported.

Appendix 2A – Income Equation Estimation

We adapt the methodology used by Paxson (1992) and Cameron and Worswick (2003) to decompose household income[20] into permanent and transitory components. We estimate the following equation:

$$Y_h = \alpha_0 + \alpha_1 X_h^P + \alpha_2 X_h^T + \nu_h \tag{2A.1}$$

where Y_h is the income of household h and X_h^P and X_h^T are variables viewed as determinants of permanent and transitory income respectively. This allows us to decompose income as:

$$\hat{Y}_h^P = \hat{\alpha}_0 + \hat{\alpha}_1 X_h^P \tag{2A.2}$$

$$\hat{Y}_h^T = \hat{\alpha}_2 \left(X_h^T \right). \tag{2A.3}$$

The fitted residual includes both a permanent component not captured by X_h^P and transitory income shocks not captured by X_h^T (Cameron and Worswick, 2003).

The variables included in X_h^P are the number of household members in each age categories, the number of adult members (age 18-64) in each education/gender category, dummies variables that indicate the occupation of the household head[21], a dummy that identifies if there is a householder who has a non-farm business, the value of land and provincial dummies. X_h^T includes dummy variables for the shocks experienced in the previous two years. There are two complications. First, not all shocks can be treated as transitory. For example, death of a household member may affect income in a permanent way. Hence, deaths occurred in the previous five years are included in the estimation of the permanent component[22]. Second, households with a non-farm business are more likely to experience a household or business loss due to a disaster. This motivates the inclusion of an interaction term between the dummy business loss and a dummy that equals one if the household owns a non-farm business.

---

[20] Household income (Y_h) is calculated as the sum of the following variables: wages earned by each household member, net profit generated by the farm, net profit generated by the household business, household income other than from business or employment (pension, scholarship loan, insurance claim, winnings, gift from family or friends, other), total income from household assets (other that farm and business assets). Appendix A shows the survey questions used to construct income.

[21] Self employed workers, employees or family workers.

[22] IFLS survey asked to report a sickness that requires hospitalization or continuous treatment. This may suggest that sickness of a household member can be treated as a permanent shocks. I constructed two measures of permanent income, with and without sickness, but results do not change.

Table 2A.1
Descriptive Statistics

Variable	Mean	s.d.	Min	Max
Household income	1080	1598	- 104	20130
Death	0.08	0.27	0	1
Sickness	0.05	0.22	0	1
Crop loss	0.11	0.31	0	1
Business loss	0.01	0.10	0	1
Business loss*non farm business	0.01	0.08	0	1
Unemployment	0.01	0.09	0	1
Price falls	0.05	0.21	0	1
Household owns a non-farm business	0.32	0.47	0	1
Land value	3062	11077	0	200000
# household members aged 0 to 5	0.65	0.81	0	5
# household members aged 6 to 11	0.71	0.84	0	4
# household members aged 12 to 17	0.64	0.83	0	5
# household members aged 18 to 64	2.32	1.06	0	12
# household members aged over 64	0.20	0.50	0	8
# males 18-64 without education	0.17	0.40	0	3
# females 18-64 without education	0.37	0.54	0	3
# males 18-64 – primary school only	0.63	0.65	0	5
# females 18-64 – primary school only	0.64	0.62	0	5
# males 18-64 up to secondary school	0.29	0.55	0	4
# females 18-64 up to secondary school	0.19	0.44	0	4
# males 18-64 high school	0.03	0.17	0	3
# females 18-64 high school	0.01	0.11	0	2
Head employee	0.29	0.45	0	1
Head self-employed	0.69	0.46	0	1
Head family worker	0.02	0.12	0	1

The table summarizes the descriptive statistics of the variables used in the income equation. “Death” refers to the death of a household member occurred in the five years prior to the interview. All other negative shocks refer to the two years 1992-93. Household income and land value are in thousands of rupiah.

Table 2A.2
Income decomposition equation estimates
Dependent variable: Household income

Permanent components	**Coefficient**	***t***
Death	-267.32	-3.41
Household owns a non-farm business	587.00	10.11
Land value	0.02	4.17
# household members aged 6 to 11	83.00	2.31
# household members aged 12 to 17	97.73	2.95
# household members aged over 64	-37.40	-0.57
# males 18-64 up to secondary school	591.78	9.00
# females 18-64 up to secondary school	570.48	5.72
# males 18-64 high school	1737.66	6.23
# females 18-64 high school	1913.40	4.63
Head employee	736.84	4.5
Head is self-employed	11.90	0.08
Intercept	242.10	1.21
Transitory components	**Coefficient**	***t***
Sickness	-52.54	-0.51
Crop loss	-88.20	-1.21
Business loss	340.17	0.98
Dummy own business*business loss	-351.90	-0.76
Unemployment	-424.54	-2.32
Price falls	-57.82	-0.51
N	3246	
$F_{34,3211}$	25.61	
R-squared	0.317	

The table reports the OLS estimates from equation (2A.1). The household income is regressed on a set of variables that determines the permanent and the transitory income components. The dummies "Head employee" and "Head is self-employed" refer to the work status of the household head, head does not work or is a family worker is the omitted category. "Death" refers to the death of a household member occurred in the previous five years. All other negative shocks refer to the two years 1992-93. Income equation also contains controls for the number of children in the house aged 0-5, the number of adult males and females (aged 18-64) unschooled and graduated only from the elementary school, and provincial dummies.
Robust *t* statistics in parentheses.

.

Appendix 2B – Calculation of multinomial probabilities

Poisson-multinomial model

In the case in which $m_h = 2$ with responses i and j selected, we need to consider the probability that j is the overall cost minimizing choice and that i is the next best, and the converse situation in which i is the overall cost minimizing choice and that j is the next best. Hence, using the notation already established:

$$\Pr\left(\Omega_h = \{i, j\} \mid m_h = 2, x_h\right) = p_{hij}^2 = \frac{a_{hj}}{\sum_{m=1}^{M} a_{hm}} \cdot \frac{a_{hi}}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{a_{hi}}{\sum_{m=1}^{M} a_{hm}} \cdot \frac{a_{hj}}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)}$$

$$= \frac{a_{hi} a_{hj}}{\sum_{m=1}^{M} a_{hm}} \cdot \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)} \right) \tag{2B.1}$$

Combining equations (2.5) and (2B.1),

$$\Pr\left(\Omega_h = \{i, j\} \mid x_h\right) = p_{hij} = \frac{a_{hi} a_{hj} \theta_h \mu_h}{\sum_{m=1}^{M} a_{hm}} \cdot \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)} \right) \tag{2B.2}$$

The argument is similar in the case that three responses are selected. We obtain

$$\Pr\left(\Omega_h = \{i, j, \ell\} \mid x_h\right) = p_{hij\ell} = \frac{a_{hi} a_{hj} a_{h\ell} \theta_h \mu_h^2}{2\sum_{m=1}^{M} a_{hm}}.$$

$$\left[\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)}\left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj} - a_{h\ell}\right)}\right)\right.$$

$$+\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)}\left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{h\ell}\right)}\right)$$

$$\left.+\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{h\ell}\right)}\left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{h\ell}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj} - a_{h\ell}\right)}\right)\right]$$

(2B.3)

Threshold Multinomial Model

In the case of double responses, equation (2.11) gives the probability of choosing modes 1 and 2:

$$p_{h12} = \Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \;\&\; c_{h1} \le c_{hm} - t_h, m = 3,...,M\right)$$
$$+\Pr\left(c_{h2} \le c_{h1} \le c_{h2} + t_h \;\&\; c_{h2} \le c_{hm} - t_h, m = 3,...,M\right)$$

Consider the first term in this expression. We may split this into two further components as

$$\Pr(c_{h1} < c_{h2} \le c_{h1} + t \;\&\; c_{h1} \le c_{hm} - t_h, m = 3,...,M)$$
$$= \Pr(c_{h1} < c_{hm} - (1-\delta_{m2})t_h, m = 2,...,M) - \Pr(c_{h1} < c_{hm} - t_h, m = 2,...,M)$$

where δ_{ij} is the Kronecker delta, $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$

This is illustrated in terms of the distribution function of c_2 for the case of M =3 in Figure 2.1.

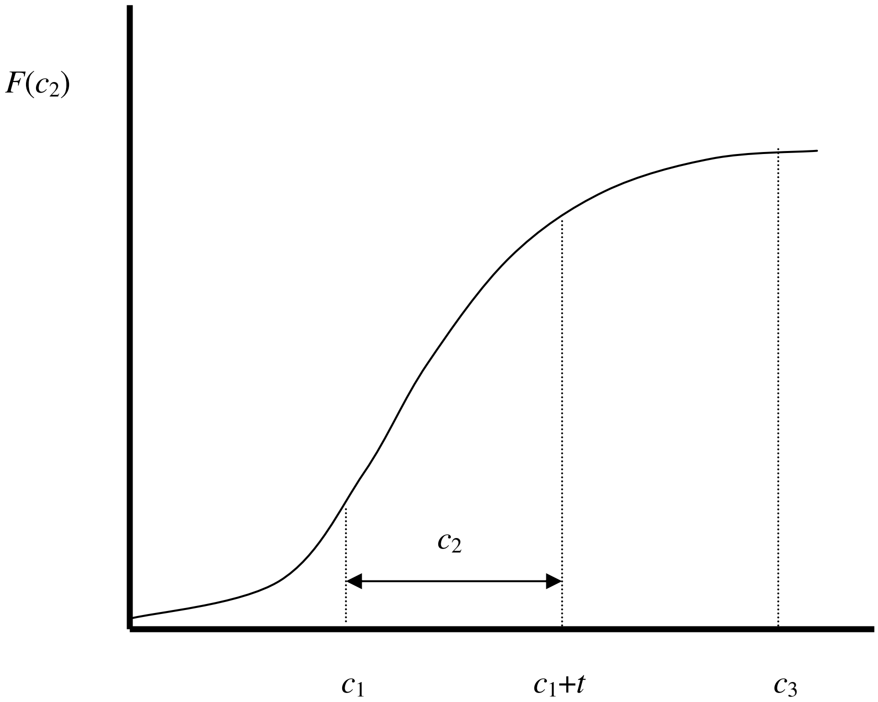


Figure 2.1: Distribution function of c2

Using equation (2.10), this probability becomes

$$\Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 3,\ldots,M\right)$$
$$= \frac{a_{h1}}{a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}} - \frac{a_{h1}}{a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}} \qquad (2B.4)$$
$$= \frac{\left(\lambda_h - 1\right) a_{h1} a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}\right)}$$

The second component of equation (2.11) follows directly as

$$\Pr(c_{h2} < c_{h1} \le c_{h2} + t_h \,\&\, c_{h2} \le c_{hm} - t_h, m = 3,\ldots,M)$$
$$= \frac{a_{h2}}{a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}} - \frac{a_{h2}}{a_{h2} + \lambda_h \sum_{m=1,m\neq 2}^{M} a_{hm}} \qquad (2B.5)$$
$$= \frac{(\lambda_h - 1) a_{h1} a_{h2}}{(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm})(a_{h2} + \lambda_h \sum_{m=1,m\neq 2}^{M} a_{hm})}$$

Combining equations (2B.4) and (2B.5), we obtain

$$p_{h12} = \frac{(\lambda_h - 1) a_{h1} a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)} \cdot \left[\frac{1}{\left(a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}\right)} + \frac{1}{\left(a_{h2} + \lambda_h \sum_{m=1, m\neq 2}^{M} a_{hm}\right)} \right]. \quad (2B.6)$$

Generalization to the case in which three modes are chosen is yet more complicated. For notational simplicity, let $\Omega_h = \{1, 2, 3\}$. Then

$$\begin{aligned} p_{h123} &= \Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right) \\ &+ \Pr\left(c_{h2} \le c_{h1}, c_{h3} \le c_{h2} + t_h \ \& \ c_{h2} \le c_{hm} - t_h, m = 4, ..., M\right) \\ &+ \Pr\left(c_{h3} \le c_{h1}, c_{h2} \le c_{h3} + t_h \ \& \ c_{h3} \le c_{hm} - t_h, m = 4, ..., M\right). \end{aligned} \quad (2B.7)$$

As previously, we analyze the three components separately. Consider the first component:

$$\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right).$$

This probability depends on the values of both c_2 and c_3 which are taken as following independent and identical Gumbel distributions.

Hence

$$\begin{aligned} &\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right) \\ &= \Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right) \\ &. \Pr\left(c_{h1} < c_{h3} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right). \end{aligned}$$

Using equation (2.8), we may write this joint probability as

$$\begin{aligned} &\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \ \& \ c_{h1} \le c_{hm} - t_h, m = 4, ..., M\right) \\ &= \frac{(\lambda_h - 1)^2 a_{h1}^2 a_{h2} a_{h3}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=2, m\neq 3}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)}. \end{aligned}$$

(2B.8)

It follows that:

$$p_{h123} = (\lambda_h - 1)^2 a_{h1} a_{h2} a_{h3} \cdot \left[\frac{a_{h1}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)} \right.$$

$$+ \frac{a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + \lambda_h a_{h1} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + \lambda_h a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)}$$

$$\left. + \frac{a_{h3}}{\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h3} + \lambda_h a_{h1} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h3} + \lambda_h a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)} \right].$$

Chapter 3

Income Shocks, Coping Strategies, and Consumption Smoothing

3.1 Introduction and literature review

As reviewed in chapter one, a growing theoretical and empirical literature analyzes the effects of shocks on households' living conditions in developing countries, and on the coping strategies adopted to overcome them. Ex post coping mechanisms include risk-pooling strategies, that spread the effects of shocks across households in a community, and self-protection strategies. The latter involve taking loans (Udry, 1990, 1994; Fafchamps and Lund, 2003), selling financial or non-financial assets (Deaton, 1991; Rosenzweig and Wolpin, 1993; Zimmerman and Carter, 2003), and increasing the labour supply to cope with shocks (Kochar, 1999; Maitra, 2001).

This chapter focuses on self-protection strategies[23]. The use of assets as a buffer to cope with shocks has been widely explored in the literature. Recent studies give primary attention to asset choices in the face of shocks in a framework in which assets contribute directly to the income generation process. The use of assets both to generate income and to smooth consumption leads to a trade off, in the sense that selling assets to smooth consumption today could have important implications for future income and hence for future consumption. This could lead households, and especially poor households, to be more cautious in running down assets to smooth consumption. Barrett and Carter (2005) suggest that below a given assets threshold, households reduce consumption in order to preserve

[23] This chapter focuses on the most frequent shock in rural Indonesia, crop loss. Since Indonesian data suggest that risk sharing strategies are less used than self-protection strategies in coping with crop loss, we will underline the role of the latter mechanisms, labour supply response in particular, and we will not take into account inter-household transfers and network.

their stock of assets (asset smoothing), while above that threshold assets are sold to protect consumption (consumption smoothing). Zimmerman and Carter (2003) find that rich households acquire high-return portfolios and run down assets to smooth consumption. On the other hand, poor households choose safe but low-return portfolios and reduce consumption in the face of shocks in order to defend their asset base (asset smoothing). Hoddinott (2006) shows that the probability of selling assets (animals) in the face of a negative income shock depends on the prior level of assets: households with more than two animals are considerably more likely to sell them than households with only one or two animals. This literature points out that households may respond differently to income shocks depending on the level of their asset ownership.

Self-protection strategies may also involve the use of labour markets as insurance institutions. In this way consumption smoothing is achieved through ex post income smoothing (Morduch, 1995; Dercon, 2002). Kochar (1999) finds evidence that farm households in India are able to cope with idiosyncratic crop income shocks shifting from own-farm to off-farm work. In this way households reduce their income variability and smooth consumption in response to shocks without relying on savings or assets. Similarly, Cameron and Worswick (2003) find that the extra income generated by the labour supply response to a crop loss is important in allowing Indonesian households to avoid reducing consumption expenditure. Maitra (2001) finds that Indian farmers differ in their response to shocks and in their ability to smooth consumption according to whether they are constrained or unconstrained. Farmers with unrestricted access to credit (medium and large farms) deal with shocks using state contingent transfers (for example credit) and without changing their leisure and consumption behaviour. Constrained farmers (small farms) with restricted access to credit are able to insure consumption against unanticipated income changes only if they adjust their market participation in response to the shock, shifting from own farm work to the labour market.

Most of the empirical studies on risk coping through self-protection strategies consider two questions. The first explores the role played by coping strategies in mitigating the effects of shocks. Many studies investigate how responsive the risk-coping strategies are to shocks, without estimating how much of the income shock is compensated (by regressing a variable which indicates a specific risk coping strategy on measures of shocks (Pan, 2007)). For example, Udry (1995) regresses savings (overall saving, grain, livestock and cash saving) on a measure of shocks to income defined as a weighted average of the number of self-reported negative events; Hoddinott (2004) regresses the net number of sales of oxen on a measure of rainfall shocks; Rosenzweig and Wolpin (1993) estimate an ordered probit of purchase and sale of productive assets on income variations; McPeak (2004) estimates a tobit model to analyze how rainfall shocks influence livestock sales; Kochar (1999) analyzes how market labour hours respond to income shocks.

The second question explores whether consumption can be smoothed against transitory income changes. This literature refers to the permanent income hypothesis (PIH) according to which the marginal utility of current consumption is equal to the discounted expected marginal utility of future consumption (Deaton, 1992). In this model only permanent income innovations are completely reflected in consumption changes, whereas transitory shocks are smoothed. Empirically this may be tested by regressing household consumption, or savings, on the permanent and transitory components of income. If households are able to smooth consumption, the coefficient on transitory income should be near zero, and the coefficient on permanent income should be near one in the consumption equation. Paxson (1992) tests for consumption smoothing in Thailand. She finds that the propensity to save out of transitory income is quite high, suggesting that savings are used to buffer consumption from income shocks. Gertler and Gruber (2002) test whether Indonesian households insure consumption against health shocks regressing the growth in log per-capita consumption on change in health. They do not find evidence of full consumption insurance against illness. Kochar (1999), Maitra (2001) and Cameron and Worswick (2003) show that households can smooth consumption by adjusting labour supply. Using data from rural Burkina Faso, Kazianga and Udry (2004) find that about 50% of changes in transitory income are passed onto consumption, with no significant differences for poor and rich households. Jalan and Ravallion (1997), who regress changes in consumption on changes in per capita income, show that poor households are less well insured against shocks: 40% of an income shock is being passed onto current consumption for the poorest households, while rich households are protected from almost 90% of an income shock.

These results suggest that, for some households, consumption choices may not be based on the permanent income model, but rather be driven by current income and the need to accumulate savings. In order to understand which model drives household consumption behaviour, we should estimate the income variation due to coping strategies and examine how much of this income gain is passed onto consumption. Few papers estimate a quantitative measure of the increase in income due to ex post responses to shocks. Fafchamps *et al*. (1998) estimate how much of the income fluctuation is compensated by sales of livestock by multiplying the predicted number of livestock sold when facing with shocks by their median price. However, since the actual price may be different for different households, this measure is only an approximation of the actual income recovery. Cameron and Worswick (2003) base their measure on self-reported income. More precisely, they estimate the size of the crop loss and of the income generated from the labour supply response by including controls for the labour supply response in the income equation. However, none of these papers examine how much of the increase in income due to coping strategies is passed onto consumption.

This work investigates whether rural Indonesian households smooth income following a crop loss (i.e. whether they adopt strategies, labour supply adjustments in particular, to recover the income reduction due to the shock)[24], and how much consumption smoothing is achieved through income smoothing strategies (i.e. how much of the increase in income is transferred onto consumption)[25]. We extend Cameron and Worswick (2003) approach and construct quantitative measures of income shocks and of households' responses to the shock, analyzing more than one response type, and differentiating between poor and non-poor farmers. Indeed, as several papers argue, the cost and ability to deal with risk may differ between poor and non-poor households (Zimmerman and Carter, 2003). Non-poor households may have access to several coping mechanisms, such as borrowing or selling assets, whereas poor households are more likely to be constrained, and hence they may have to rely on the labour supply response to mitigate the effects of shocks.

In summary, there are two questions this chapter seeks to answer. First, which variables influence the adoption of income smoothing strategies and whether these strategies completely or partly recover the income reduction due to the shock. Quantitative measures of

[24] "Employment and wages are likely to be more flexible in largely agricultural societies in which a high proportion of the workforce is self-employed or works in the informal sector" (Manning, 2000, p. 130). The case of Indonesia is consistent with this framework. The flexibility of Indonesian labour markets and the availability of alternative employment opportunities for those who lose their jobs, mostly in small-scale enterprises and the informal sector, supported the adjustments in labour supply as one important aspect of the response to shocks, even in the face of the economic crisis of 1997-98 (Manning, 2000).

[25] The question of how much consumption smoothing is achieved trough income smoothing strategies may depend on whether income is pooled within the household and on how resources are allocated within the household (Mazzocco, 2004; Witoelar, 2005). This chapter does not examine the intra-household behaviour. This aspect may be taken into account for future model developments.

income shocks and household's ability to cope with the shock are estimated for poor and non-poor households, and using these measures we can explore whether the increase in income due to income-smoothing strategies partially offsets, or exceeds, the income loss due to the shock. Second, this chapter investigates whether households smooth consumption and the role played by different coping strategies in mitigating consumption reductions. The analysis is conducted distinguishing between consumption behaviour of poor and non-poor households. These issues are explored using the Indonesian Family Life Survey data.

The chapter is organized as follows. The theoretical model is presented in section II. Section III discusses the data, and section IV summarizes the methodology and main results of Cameron and Worswick (2003), establishing the motivation for the extensions proposed in this chapter. Section V and VI presents respectively the empirical methodology and results. Section VII concludes.

3.2 Theoretical Framework

The model developed in this section is a simple intertemporal model with a household farming production function subject to exogenous income shocks. Leisure and asset investment decisions are included in the household optimization problem. Assets are defined as productive (farm assets) and non-productive (financial assets). They have direct effects on income levels, and can also serve as a buffer to smooth consumption against shocks (Rosenzweig and Wolpin, 1993; Zimmerman and Carter, 2003; Newhouse, 2005). Household members may work in the family farm and/or as salary workers.

The farm profit function is defined as $\Pi_{ft} = \pi(\Phi_{t-1}, h_t^f, s_t)$, where h_t^f is the labour input, Φ_{t-1} is the level of productive and unproductive assets owned by the household at the end of the previous year, and s is a transitory random shock. Shocks are assumed exogenous, and uncorrelated over time. The farm profit increases with positive shocks, and decreases as a consequence of negative shocks, such that $\partial\Pi / \partial s > 0$. Total income of the household comes from the farm and off-farm labour. There is evidence that in the face of a shock household members reduce the hours worked in the farm and increase the hours of work in the wage employment sector (Kochar, 1999; Maitra, 2001; Cameron and Worswick, 2003). In this chapter, we are not interested in examining the trade-off between farm and non-farm labour, and hence we assume that household members work a fixed amount of hours in the family farm, so that h_t^f is exogenous, and it varies with s (negative shocks reduce h_t^f). The remaining time endowment ($T_s = T - h_t^f$) can be allocated to either leisure

(l_t) or off-farm work. Let y_t^w be the income earned by family members on wage employment. It is defined as $y_t^w = w_t(T_s - l_t)$.

Total household income can be written as:

$$I_t = \Pi_{ft}(\Phi_{t-1}, h_t^f, s_t) + w_t(T_s - l_t). \quad (3.1)$$

Assets[26] evolve according to:

$$\Phi_t = \Phi_{t-1} + \phi_t \quad (3.2)$$

where ϕ_t is the amount of assets purchased, or sold when $\phi_t < 0$, by the household at time t (for simplicity we assume no depreciation and no interest rate).

The budget constraint the household faces is given by:

$$c_t + p_\Phi(\phi_t) = I_t \quad (3.3)$$

where the price of the consumption good is normalized to one, and p_Φ is the price of assets. Households can either sell productive assets or decrease financial assets to increase consumption. However, we assume that households face a constraint on assets defined as:

$$\phi_t \geq g(Z, \Phi_{t-1}) \quad (3.4)$$

i.e., there is a borrowing limit that depends on the level of assets previously accumulated and on household characteristics (Z)[27] (Newhouse, 2005).

The farm profit equation written above shows that future productivity is a function of current asset accumulation strategies. Today's sale of assets has important implications for future income and hence for future consumption. This form of non separability between current and future consumption leads households, and especially poor households, to be more cautious in running down assets in the face of transitory shocks. Hence the trade off captured in this model is not only between consumption and off-farm labour, but also

[26] As mentioned above, assets are a broad definition and include both productive and non productive assets. However, considering a sample of farm households it is reasonable to suppose that productive assets constitute the majority of total assets owned by the households.

[27] The assumption is that $\Phi_t \geq \overline{\Phi}(Z)$, the level of assets at time t must be greater than or equal to a household specific threshold which is a function of household characteristics. For productive assets this threshold is assumed to be positive, $\overline{\Phi}(Z) > 0$. Subtracting Φ_{t-1} from both sides, the constraint becomes $\Phi_t - \Phi_{t-1} \geq \overline{\Phi}(Z) - \Phi_{t-1}$, that is $\phi_t \geq \overline{\Phi}(Z) - \Phi_{t-1}$. The reduced form becomes $\phi_t \geq g(Z, \Phi_{t-1})$ (equation (3.4)).

between current consumption and asset accumulation for future consumption (Zimmerman and Carter, 2003).

Each period utility function is defined as $u_t(c_t, l_t)$, where we allow for consumption and leisure choices to be non-separable (Kochar, 1999; Kazianga and Udry, 2004); more precisely we assume that $\partial^2 u_t / \partial c \partial l > 0$.

The household's Bellman equation is defined as:

$$V_t(\Phi_{t-1}, s_t) = \max_{l_t, \varphi_t} \left\{ u(\Pi_{ft}(\Phi_{t-1}, s_t) + w_t(T_s - l_t) - p_\Phi \varphi_t \, , \, l_t) + \beta E_t V_{t+1}(\Phi_t, s_{t+1}) \right\} \quad (3.5)$$

The first order conditions for households for which the constraint is not binding are the following[28]:

$$\begin{cases} \dfrac{\partial u}{\partial c_t} = \beta E_t \dfrac{1}{p_\Phi} \dfrac{\partial V_{t+1}}{\partial \Phi_t} & (a) \\ \dfrac{\partial u}{\partial c_t} = \dfrac{1}{w_t} \dfrac{\partial u}{\partial l_t} & (b) \end{cases} \quad (3.6)$$

Equations (3.6a) and (3.6b) solve respectively the trade off between consumption and assets purchase, and between consumption and leisure. Looking at the equation (3.6a), a negative shock that decreases the farm profit will increase the marginal utility of income all else equal. Assuming the value function as concave in assets[29], to keep equality household must increase consumption and decrease Φ_t, i.e. the household will choose a lower level of ϕ_t. A similar result comes from equation (3.6b). A negative shock increases the marginal utility of income all else equal, and decreases the marginal dis-utility of off-farm work[30]. To

[28] We do not consider for simplicity the time constraint.

[29] The first derivative of the value function with respect to Φ_t is:

$$\frac{\partial V_t}{\partial \Phi_{t-1}} = \frac{\partial u}{\partial c_t} \cdot \frac{\partial c_t}{\partial \Pi_t} \cdot \frac{\partial \Pi_t}{\partial \Phi_{t-1}} > 0,$$

The second derivative is given by:

$$\frac{\partial^2 V_t}{\partial \Phi_{t-1}{}^2} = \frac{\partial^2 u}{\partial c_t{}^2} \cdot \left(\frac{\partial c_t}{\partial \Pi_t} \right)^2 \cdot \left(\frac{\partial \Pi_t}{\partial \Phi_{t-1}} \right)^2 + \frac{\partial u}{\partial c_t} \cdot \frac{\partial^2 c_t}{\partial \Pi_t{}^2} \cdot \left(\frac{\partial \Pi_t}{\partial \Phi_{t-1}} \right)^2 + \frac{\partial u}{\partial c_t} \cdot \frac{\partial c_t}{\partial \Pi_t} \cdot \frac{\partial^2 \Pi_t}{\partial \Phi_{t-1}{}^2}$$

Assuming $\dfrac{\partial^2 u}{\partial c_t{}^2} < 0$, $\dfrac{\partial^2 c_t}{\partial \Pi_t{}^2} < 0$, and $\dfrac{\partial^2 \Pi_t}{\partial \Phi_{t-1}{}^2} < 0$, the value function is concave in assets.

[30] This derives from $\partial^2 u_t / \partial c \partial l > 0$ and from the effect of a negative shock on T_s.

keep equality (3.6b) household reduces l_t. Hence, in the face of a negative shock households reduce the amount of assets (by either buying less or selling productive assets or by reducing financial assets), and/or increase the labour market participation to overcome the hardship.

Equation (3.6) holds only if the household is not constrained in period t, i.e. if $\phi_t > g(Z_t, \Phi_{t-1})$[31]. If the constraint is binding, $\phi_t = g(Z_t, \Phi_{t-1})$, the first order conditions take the form:

$$\begin{cases} \frac{\partial u}{\partial c_t} = \beta E_t \frac{1}{p_\Phi} \frac{\partial V_{t+1}}{\partial \Phi_t} + \gamma_t & (a) \\ \frac{\partial u}{\partial c_t} = \frac{1}{w_t} \frac{\partial u}{\partial l_t} & (b) \end{cases} \tag{3.7}$$

where γ_t is the multiplier for the constraint. Equation (3.7a) means that the marginal utility of consumption for constrained households is greater than the marginal utility that would be optimal without constraints.

Substituting (3.7b) into (3.7a), we have that $\frac{\partial u}{\partial l_t} \frac{1}{w} = \beta E_t \frac{1}{p_\Phi} \frac{\partial V_{t+1}}{\partial \Phi_t} + \gamma_t$. This implies that $\left(\frac{\partial u}{\partial l_t}\right)_c > \left(\frac{\partial u}{\partial l_t}\right)_u$, where the subscripts *c* and *u* identify respectively the constrained and unconstrained framework. The marginal utility of leisure and of consumption are both greater when households face constraints. This implies that in general constrained households consume less and work more than if they were unconstrained, and these effects are even more pronounced in the face of a negative shock.

In the empirical analysis we will not estimate the structural model presented in this section, rather we will use the theoretical prediction of this model to guide the specification and interpretation of the reduced form equations that will be estimated.

[31] Applying the Kuhn-Tucker conditions, the multiplier for the constraint has to be equal to zero when the constraint is not binding.

3.3 The Data

The data used for this study are from the 1993 Indonesian Family Life Survey (IFLS1). 7224 households were interviewed over a wide range of issues. Only those households that supplied a complete set of income and demographic data are included in the dataset. After dropping income and assets outliers (about 1% of the total sample), and focusing on the rural area, the sample includes 3601 rural households; of these, 2183 are farm households, defined as those who reported to own a farm and at least one farm asset in the year of the survey. Respondents were asked whether their household had experienced an economic shock in the past five years, the type of the shock, when it happened (year and month), what measures were taken and the costs of overcoming the shock. The survey permits only one occurrence of the same shock in the period 1989-93 to be reported by the same household, and there is evidence that the most recent shocks are more likely to be reported[32]. Nearly 34% of the total rural sample has experienced at least one shock in the past 5 years. The incidence of the different types of shocks is reported in table 3.1. The most frequent shocks are sickness and crop loss, whereas business loss and unemployment affect only a few households. Focusing on the farm sample, the percentage of households that suffered a crop loss is nearly 24%. Column six of table 3.1 reports the medians of the percentage of farmers that experienced the same shock in the same village in 1993, considering only villages in which there is at least one household reporting the shock. As expected, crop loss is the most common shock, with a median percentage of 6.7 (and a maximum of 40%).

Table 3.1
Number of households reporting shocks by type of shock (1988-93)

Type of shock	Rural sample		Farm sample		
	# households	per cent	# households	per cent	Commonality - medians[33]
Death	284	7.9	174	7.9	3.7
Sickness	376	10.4	232	10.5	3.9
Crop loss	560	15.6	538	24.3	6.7
Disaster	63	1.75	41	1.9	3.7
Unemployment	65	1.81	25	1.1	3.9
Price falls	239	6.64	215	9.7	4.2

The table reports the number of rural and farm households, and the percentage of all households, reporting shocks of each type over the five year period 1989-93. The commonality of shocks is reported for shocks occurred in 1993 and for the farm sample, where the commonality is the percentage of households reporting the same shock in the same village in 1993.

[32] For example, 31% of the crop loss experienced in the period 1988-93 are reported to occur in 1993, and 63% in 1992-93.

[33] Villages with no households reporting shocks are excluded from the median.

Since crop loss is the most frequent shock in rural Indonesia, and one of the major sources of risk in poor rural areas, in the empirical analysis we will focus on this type of shock. This choice clearly raises some issues about which sample to use. Cameron and Worswick (2003) use the entire sample of rural households. This sample selection poses some problems since the crop loss is a shock that should affect only those who have a farm production. Table 3.1 shows that only a few rural non-farm households report a crop loss. In order to avoid to include households that cannot report a crop loss because they do not own a farm, we prefer to drop some observations and to restrict the sample to farm households.

Table 3.2 shows the percentage of farm households that use different measures in response to crop losses reported in the period 1988-93 and in 1993 only. Nearly 40% of the total respondents report taking an extra job to overcome a crop loss. Other important responses are “cut down on household expenses”, “take a loan” and “sell assets”. Indonesian data confirm the suggestion in the literature that informal insurance mechanisms, such as family and community assistance, may be used less in the face of common shocks, like for example crop loss (Alderman and Paxson, 1992). As the percentage of households that experience the same shock in the same village increases, the community may provide less insurance against it.

The importance of different responses may vary according to the wealth/size of the farm and hence table 3.3 reports the percentages of responses by 1992 farm assets quartiles, where the bottom 25% of the asset distribution identifies small farms. The entire set of the previous five years crop losses is used to increase the number of observations. As pointed out by other authors (Kochar, 1999; Newhouse, 2005, Maitra, 2001), labour supply adjustment is a measure used particularly by poor farmers. Indeed, the percentage of households that take an extra job decreases as we move from poor to rich farmers. Owners of large farms (measured in terms of farm assets) are more likely to run down assets and to use savings than owners of small farms, even if the percentage of households that use savings remains low.

Table 3.2
Responses to a crop loss

	1988-93	1993
Measure Taken	% of households	% of households
Extra job	45.0	39.2
Loan	21.2	20.0
Sell assets	20.0	17.5
Family assistance	6.9	7.2
Savings	4.5	5.4
Cut down on household expenses	20.8	29.0

The table reports the percentage of farm households that adopted each type of coping strategy in the face of a crop loss experienced in the period 1988-93 and 1993. Because of multiple responses, percentages sum to more than 100%

Table 3.3
Responses to crop loss experienced in the past 5 years, by 1992 farm assets percentiles

	crop loss responses – 1988-93			
Measure Taken	Bottom 25%	25-50%	50-75%	Top 25%
Extra job	54.6	50.0	43.9	30.7
Loan	24.8	18.0	13.0	29.2
Sell assets	14.2	15.6	25.2	25.4
Family assistance	3.5	9.4	8.6	6.2
Savings	0.7	2.3	3.6	11.4
Cut down on household expenses	22.0	23.4	22.3	15.4

The table summarizes the percentages of those households which experienced a crop loss over the five year period 1989-93 who adopted each response mode. Because of multiple responses, percentages sum to more than 100%

By examining the main differences between small, medium and large farms reported in table 3.4, it is possible to observe that the higher propensity of small farms to use labour supply is not correlated with a significant larger household size or other demographic characteristics. Indeed, the only significant difference (except the obvious one related with assets and income) is in the number of household members with secondary and high education. Therefore, it seems that the adoption of different strategies is more related with the values of assets (maybe because of borrowing constraints) than with other household characteristics. This descriptive evidence suggests that it is important to distinguish small and large farms in the analysis of income and consumption smoothing behaviour.

Table 3.4
Sample means of household characteristics by 1992 value of farm assets

	Bottom 25%	25-75%	Top 25%
1993 Crop loss	0.08 (0.27)	0.07 (0.26)	0.07 (0.26)
Labour supply as a response to a 1993 crop loss	0.51 (0.51)	0.37 (0.46)	0.30 (0.46)
Own farm land	0.45 (0.50)*	0.97 (0.17)	0.99 (0.07)
Business ownership	0.23 (0.42)	0.26 (0.44)	0.28 (0.45)
’92 value of business assets (excluding zeroes)	489.70 (1467.00)	482.20 (1674.23)	4460.96 (29347.50)*
’92 value of non-business assets	1907.59 (4198.26)*	3050.02 (6488.73)*	5819.95 (16482.29)*
Household income	676.29 (1049.87)*	793.12 (1174.43)*	1624.23 (2159.15)*
Head inactive	0.04 (0.20)	0.06 (0.23)	0.04 (0.19)
Head employee	0.17 (0.38)	0.13 (0.33)	0.11 (0.31)
Head years of education	3.53 (3.40)	3.82 (3.58)	4.62 (4.19)*
Household size	4.57 (2.03)	4.48 (1.96)	4.87 (2.09)
Number of income earners (other than head)	0.67 (0.91)	0.68 (0.91)	0.75 (0.92)
Number of male adults with secondary education	0.20 (0.40)	0.26 (0.51)	0.42 (0.67)*
Number of female adults with secondary education	0.10 (0.40)	0.16 (0.40)	0.27 (0.50)*
Number of male adults with high education	0.01 (0.09)	0.02 (0.13)	0.05 (0.23)*
Number of female adults with high education	0.003 (0.06)	0.01 (0.10)	0.02 (0.15)*

The table reports mean characteristics of farm households, by the size of the farm (quartiles of 1992 farm assets distribution). Standard errors are in parentheses. * the difference in the means is statistically significant at the 5-percent level. Income and assets are in thousands of rupiah.

Since the role of labour supply in the face of a crop loss in rural Indonesia has been analyzed also by Cameron and Worswick (2003), the next paragraph reviews their approach in order to highlight the main differences with the aim and methodology of this chapter.

3.4 The approach of Cameron and Worswick (2003)

Cameron and Worswick (2003) analyze the role of the labour supply response in allowing Indonesian rural households to smooth consumption in the face of a transitory shock. They consider two cases: the constant labour supply case (CLS), and the variable labour supply case (VLS) in which they permit changes in labour market behaviour following crop loss. In the CLS case, quantitative measures of permanent and transitory income components are constructed in two steps. First they estimate the following equation (Paxson's (1992) method):

$$Y_h = \alpha_0 + \alpha_1 X_h^P + \alpha_2 X_h^T + \varepsilon_h \quad (3.8)$$

where Y_h is household income, X_h^P and X_h^T are vectors of variables that determine permanent and transitory income respectively. The estimated coefficients are then used to construct the permanent ($\hat{Y}_h^P$) and transitory ($\hat{Y}_h^T$) income components as follows:

$$\hat{Y}_h^P = \hat{\alpha}_0 + \hat{\alpha}_1 X_h^P \quad (3.9)$$

$$\hat{Y}_h^T = \hat{\alpha}_2 X_h^T \quad (3.10)$$

The variables used to identify permanent income are the number of adults in several education/gender categories, the occupation of the household head, the value of farm land, and provincial dummies. The variables used to identify transitory income are a dummy for reported crop loss, and the same dummy interacted with the value of the farm land, to account for the fact that larger farms may have larger transitory income. Note that this measure of transitory income includes both the income loss due to the crop loss and the increase in income due to coping strategies, labour supply response in particular. To estimate the size of the crop loss net of any labour supply income (VLS), equation (3.8) becomes:

$$Y_h = \gamma_0 + \gamma_1 X_h^P + \gamma_2 X_h^T + \gamma_3 X_h^{LS} + u_h \quad (3.11)$$

where X_h^{LS} captures the increase in income due to the labour market changes following crop loss. This is a dummy that equals one if the household had a labour supply response, interacted with the number of household members aged 12-64. Because of the endogeneity of labour supply response, equation 3.11 is estimated including the selection terms for those who respond with the labour supply and those who do not[34]. The predicted income components constructed from equations 3.8 and 3.11 are then used to estimate the

[34] Selection terms are computed from a probit regression of labour supply response to crop losses over the period 1989-93, on the number of members aged 18 to 64 with secondary education, the number of household members aged 65 and over, and the land value.

households' marginal propensity to save out of permanent and transitory income. The saving equation is

$$S_h = \beta_0 + \beta_1 \hat{Y}_h^P + \beta_2 \hat{Y}_h^T + \beta_3 \hat{\varepsilon}_h + \beta_4 Z_h + e_h \qquad (3.12)$$

where S_h are household savings, $\hat{Y}_h^T$ is the measure of transitory income derived either from the CLS case or from the VLS case, $\hat{\varepsilon}_h$ is the fitted residual from income equation (3.8) in the CLS case, and from (3.11) in the VLS case ($\hat{u}_h$); Z is a vector of variables that control for the lifecycle characteristics of the household (Paxson, 1992).

Results show that in the CLS case, the marginal propensity to save out of transitory income is not statistically different from one (the estimated coefficient is 1.90 with t = 2.4), and the marginal propensity to save out of permanent income is statistically different from one (the estimated coefficient is 0.16 with t = 2.8). These findings are supportive of the consumption smoothing behaviour. In the VLS case, where $\hat{Y}_h^T$ is defined as the crop loss net of any labour supply income ($\hat{\gamma}_2 X_h^T$), the marginal propensity to save out of transitory income is about 0.21, statistically significant, and statistically different from one. Cameron and Worswick conclude that the increase in income due to the labour supply response is important in allowing rural households to smooth consumption: households that do not change their labour supply when facing the shock reduce their expenditure by about 79% of the loss in income due to the crop loss.

Note that these results imply that for all households, whether poor or non-poor, consumption is closely related with permanent income, and that the transitory income generated by the labour-supply response is partly transferred onto consumption (the amount needed to compensate for the crop loss) and the remaining part (if positive) is saved.

However, as suggested by the previous descriptive analysis and by various papers in the literature, both the importance of labour supply as a response to shocks and the link between consumption and permanent or transitory income may be quite different for poor and non-poor households. Rich households are less likely to adjust their labour supply because they can insure consumption through alternative coping strategies (taking out a loan, selling assets, use of savings, etc). on the other hand, poor households may face borrowing constraints and they may need to build a stock of assets to protect themselves from future risk and/or to finance investments. As a consequence, their marginal propensity to consume out of both permanent and transitory income may be quite different from that of rich farmers.

Indeed, using the same methodology of Cameron and Worswick, on farm sample instead of the entire rural sample, and distinguishing between small and medium/large farms, we find that non-poor farmers insure consumption without having to rely on the labour market, whereas owners of small farms are found to reduce consumption both in the constant labour supply case (CLS) and in the variable labour supply case (VLS). In the CLS case the marginal propensity to consume out of transitory income is 1.7 (t=3.2) for poor farmers and 0.2 (t=0.5) for owners of medium and large farms, and we cannot reject the joint hypothesis that the former is equal to one and the latter equals zero. The same results can be obtained for the VLS case (see appendix 3.A for details). Furthermore, poor farmers have a lower marginal propensity to consume out of *permanent* income (estimated coefficients are respectively 0.5 and 0.9 for poor and non-poor farmers, and are both statistically significant).

Given these findings, it seems important to examine in more detail the differences in labour supply and consumption behaviour between small and large farms, trying to identify whether households follow a consumption smoothing strategy or the need to accumulate assets prevails. We propose a different empirical methodology to estimate a quantitative measure of the income reduction produced by the crop loss, and of the household's ability to recover from the shock. In particular, the impact of the shock is allowed to differ for poor and non-poor farmers, and more than one coping strategy is considered in evaluating the income recovery from the shock. Then we use these estimates to obtain more information of the consumption behaviour of different households.

3.5 Empirical Methodology

The aim of this section is twofold. The first is to construct quantitative measures of income shock and of the household's ability to cope with the shock, in order to estimate how much of the income loss is compensated by different coping strategies. Second, we explore whether households actually follow a consumption smoothing strategy in the face of shock and the role played by different coping mechanisms.

3.5.1 Income Equation and Measures of Shock

This section estimates a quantitative measure of the income loss produced by the crop shock, and of the household's ability to recover from the shock. Several methodologies have been used to measure income shocks. Rosenzweig (1988) uses the difference between a household current income and its mean income over a nine-year panel. Jacoby and Skoufias

(1997) define the idiosyncratic shock as the deviation of the change in log full income from the change in the village-season-year mean and the aggregate shock as the mean change itself. Beegle, Dehejia, and Gatti (2003) measure transitory crop shocks using the reported values of crop loss (due to insects, rodents, and other calamities). Kochar (1999) measures income shock as the residual from a regression of crop profits on variables determining the household's expectations of profits (a set of household dummy variables, reflecting all time-invariant factors, and a set of time-varying demographic variables). Paxson (1992) measures the transitory income component regressing total household income on a set of variables that affect transitory income (in her study this set consists of deviations of rainfall from its average level). Similarly, Cameron and Worswick (2003) estimate transitory income regressing household income on self-reported shocks (crop loss).

This section describes the empirical methodology we propose, which leads to consistent estimates under the assumption that the crop loss is exogenous. This issue will be explained in detail below. The predicted income for households that do not report a crop loss in 1993 is estimated first. With no crop loss, this predicted income would be an appropriate estimate of the household's income for all households[35].

For households that experienced a crop loss in 1993 the difference between the observed income and the predicted income is constructed. This difference is regressed on a set of variables that affect the magnitude of the income shock (e.g. farm assets) and the household's ability to cope ex post with the hardship.

Income for households that do not report a crop loss is defined as:

$$Y_h = \alpha_0 + \alpha_1 X_h^P + \alpha_2 X_h^T + \varepsilon_h \tag{3.13}$$

where Y_h is the 1993 household income, X_h^P is a vector of variables that determine permanent income, and X_h^T is a set of other variables that may affect household income in a transitory way. In this analysis, X_h^P are demographic characteristics, location dummies, and wealth indicators. The latter includes the value of household's assets held at the end of the year prior to the interview, distinguishing between farm, business non farm, and non business real and financial assets. To account for the non-linearity of the income function, the coefficient on farm assets is interacted with dummies that indicate whether the farm is small, medium and large (defined according to quartiles of 1992 farm assets distribution).

[35] It is worth noting, however, that estimating permanent income using cross-sectional data instead of panel data does not allow to model the dynamics of predicted income, and to solve the problem of unobserved heterogeneity (Abul Naga and Bolzani, 2000).

Other variables that may contribute to permanent income are head's characteristics (education, occupation type), the family size, and village-province specific information. Dummies for non labour income sources (pension, winnings, and gift from family/friends) are also included in the income equation. Winnings and gifts (X_h^T) are omitted from the estimation of permanent income and included in a positive transitory income component. The parameters in (3.13) can be consistently estimated by applying OLS on the sub-sample of households with no crop loss under the assumption that the crop loss is exogenous.

For households that reported a crop loss in 1993, the difference between actual and predicted income is constructed:

$$\Delta Y_h = Y_h^{CL} - \hat{Y}_h \tag{3.14}$$

where Y_h^{CL} is current income for households that reported a crop loss, and $\hat{Y}_h$ is the predicted income for these households, on the basis of the parameters estimates from equation (3.13). This difference can be explained by the sum of the loss produced by the shock and the gains from the ex post coping strategies that are reflected in the income measure (plus the effects of unobservables). To estimate the size of the crop loss and of the increase in income given by coping strategies, the following regression is estimated:

$$\Delta Y_h = \beta_0 + \beta_1 X_h^S + \beta_2 X_h^{LS} + \beta_3 X_h^L + \beta_4 X_h^A + u_h \tag{3.15}$$

where X_h^S are shock variables that will be explained below, X_h^{LS} and X_h^L are vectors of variables that determine the size of the increase in income due to the following strategies: "labour supply" and "sell assets or take a loan" respectively. X_h^A is the value of 1992 non-productive assets owned by the household[36].

The income shock caused by a crop loss has two components: the constant term in (3.15), and the farm specific component estimated using the vector X_h^S, which includes the value of 1992 farm assets. This second component allows us to link the income loss with the size of the farm: the larger is the farm, the larger may be the crop loss. To account for possible non-linearity in the functional form, the coefficient on farm assets is interacted with dummies that

[36] As a check for the validity of using the estimates of equation (3.13) to construct the dependent variable in (3.15), the following equation is estimated: $Y_h^{CL} = \beta_0 + \beta_1 X_h^S + \beta_2 X_h^A + \beta_3 X_h^{LS} + \beta_4 X_h^L + \beta_5 \hat{Y} + u_h$; $\hat{\beta}_5$ is found to be non statistically different from one (F(1,148)=2.46, Prob>F=0.12).

indicate whether the farm is small, medium and large (defined according to quartiles of 1992 farm assets distribution).

The extra labour income given by the labour supply response is estimated using the dummy labour supply (self-reported strategy) interacted with the number of household members aged 13-64 (following Cameron and Worswick (2003) and Kochar (1995), households with more people of working age may increase their labour supply by more). In order to control for possible effects on reported income of other strategies, (e.g. "take loan" or "sell assets"), we include dummy variables for whether the household used at least one of these coping mechanisms, and the value of 1992 non-business assets (X_h^A)[37].

Least squares estimation of (3.15) may lead to biased estimates of the parameters because of the endogeneity of the labour supply response. Let us assume that the decision to respond with the labour supply to the crop loss can be described by a probit equation

$$LS_h^* = \delta Z_h + \eta_h \tag{3.16}$$

with $LS_h=1$ if $LS_h^* > 0$ and zero otherwise. LS_h is a dummy that equals one if the household reported labour supply as a response to the crop loss and zero otherwise, and Z is the set of independent variables that explains the probability of taking an extra job. If we believe that the outcome equation depends upon the regime (LS=0 or LS=1), then we should split the sample into "labour users" (LS=1) and "non-labour users" (LS=0) and estimate an income equation corresponding to (3.15) for the two sub-samples (switching regression model). If it is the case, the problem with the labour supply response can be treated as a sample selection problem. The error terms on the income equations for the two sub-samples would have variances σ_0^2 and σ_1^2 and covariances σ_{02} and σ_{12} with η_h. From the probit equation, we can derive the selection terms defined as:

$$h_h = \begin{cases} \dfrac{f(\hat{\delta} Z_h)}{F(\hat{\delta} Z_h)} & if\ LS_h = 1 \\ \dfrac{f(\hat{\delta} Z_h)}{1 - F(\hat{\delta} Z_h)} & if\ LS_h = 0\ and\ CL_h = 1 \end{cases} \tag{3.17}$$

where f is the standard normal density function and F is the cumulative distribution function. CL_h is a dummy that equals one if the household reported a crop loss in 1993, zero

[37] Non business and non farm assets are added as additional regressors because the measure of household income includes the income from the rent/lease/profit-sharing of non-business assets.

otherwise. To have consistent estimates, each income regression should be augmented with the corresponding selection term.

If all coefficients but the one on X_h^{LS} are the same for LS=0 and LS=1, it follows that we can pool the entire sample of "labour users" (LS=1) and "non-labour users" (LS=0) and augment this regression equation with the selection terms. Equation (3.15) thus becomes:

$$\Delta Y_h = \beta_0 + \beta_1 X_h^S + \beta_2 X_h^A + \beta_3 X_h^{LS} + \beta_4 X_h^L + \sigma_{12} \frac{f(\hat{\delta} Z_h)}{F(\hat{\delta} Z_h)} \; LS_h$$
$$+\sigma_{02} \frac{f(\hat{\delta} Z_h)}{1 - F(\hat{\delta} Z_h)} (1 - LS_h) CL_h + u_h \qquad (3.18)$$

We conducted a Chow test to check whether all coefficients are the same for LS=0 and LS=1, and we cannot reject this hypothesis (F(6,149)=0.84, Prob>F=0.54). Therefore we use (3.18) to estimate the total effect of the shock (the sum of income loss and income gains), the contribution of labour supply response, and the income loss net of any recovery. Shock measures are defined as:

$${}_1\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_2 X_h^A + \hat{\beta}_3 X_h^{LS} + \hat{\beta}_4 X_h^L \qquad (3.19)$$

$${}_2\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_3 X_h^{LS} \qquad (3.20)$$

$${}_3\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S \qquad (3.21)$$

where ${}_1\hat{Y}_h^S$ represents the total effect of the income shock, and it may be positive or negative according to whether the increase in income generated from coping mechanisms is higher or smaller than the income loss caused by the shock. ${}_2\hat{Y}_h^S$ is the effect of the income shock when only labour supply response is considered, and ${}_3\hat{Y}_h^S$ is the measure of income loss net of any recovery. These measures allow us to answer to the first question of the chapter: whether households partly or completely offset the income loss by using labour supply and/or different coping strategies.

3.5.2 Consumption equation

The second question this chapter seeks to answer is the extent to which farmers smooth consumption in the face of short-term shocks (like for example the crop loss) or allow consumption to vary in order to build a stock of assets. Consumption is measured by non durable annual household expenses[38]. As suggested by Deaton (1997), a way to test for consumption smoothing is to write consumption as a function of permanent and transitory income components, together with other variables, for example the demographic structure of the household:

$$C_{ht} = \alpha + \beta Y_{ht}^{P} + \gamma Y_{ht}^{T} + \theta Z_{ht} + u_{ht}$$

In a "strict" version of the permanent income hypothesis, the parameters β and γ would be one and zero, respectively, but even the evidence that $\beta > \gamma$ would suggest that households smooth their consumption relative to income (Deaton, 1997)[39]. Following this approach, the consumption equation can be written as:

$$\begin{aligned} C_h = {} & \gamma_0 + \gamma_1 \hat{Y}_h^P + \gamma_2 \cdot ({}_3\hat{Y}_h^S) CL_h + \gamma_3 \hat{Y}_h^{LS} CL_h + \gamma_4 \hat{Y}_h^A CL_h + \gamma_5 \hat{Y}_h^L CL_h \\ & + \gamma_6 \hat{u}_h \cdot CL_h + \gamma_7 \hat{\varepsilon}_h \cdot (1 - CL_h) + \gamma_8 Z_h + \nu_h \end{aligned} \tag{3.22}$$

where $\hat{Y}_h^P$ is the permanent income component, ${}_3\hat{Y}_h^S$ is the measure of the income shock net of any recovery (as reported in equation 3.21), $\hat{Y}_h^{LS}$ is the extra labour income, and $\hat{Y}_h^A, \hat{Y}_h^L$ are the predicted income gains due to other coping strategies (non-business assets and "take a loan or sell assets" respectively). $\hat{u}_h$ and $\hat{\varepsilon}_h$ are the fitted residuals from income equations (3.18) and (3.13) respectively, and Z_h is a set of variables that measure the life-cycle stage of the household. Following Paxson (1992), the variables included in Z_h are the number of household members in each age categories[40].

---

[38] The expenditure variable used in this chapter includes the total value of goods self-produced by the household. Durable goods are not included because it is difficult to impute the appropriate measure of the service flow derived from that.

[39] This approach has been used by Paxson (1992) and Cameron and Worswick (2003). They estimate the level of household savings as a linear function of permanent income, transitory income, the residual from the income equation (unexplained income), and a set of variables that measure the life-cycle stage of the households. Paxson includes also the variability of the household's income.

[40] All the regressors in (3.22) but Z_h are estimates, consequently coefficient standard errors may be biased (Davidson and MacKinnon, 2004).

The inclusion of income measures of the shock (net of any recovery) and of the coping strategies allows us to analyze the impact of the crop loss and of different responses to the shock on consumption. To examine the different behaviour of constrained and unconstrained households, both the permanent income and the measure of the crop loss are interacted with dummies to identify small, medium and large farms.

3.6 Results

3.6.1 Income equation estimates

Income equation for households that did not report the crop loss

Estimates of the income equation (3.13) are reported in table 3.5[41]. The R^2 (0.37) indicates that independent variables explain a fairly high proportion of income variability. Results are in line with standard income equation estimates. It should first be noted that coefficients on 1992 farm assets confirm the non-linearity of the income function. Coefficients on all the other 1992 assets are highly significant[42]. Households with head employed in the private and government sector are expected to have a higher income than all the other households, all else equal. Households with the head self-employed have a higher income than households with head inactive or family worker. Head's education is positively and significantly related to income only for secondary and high levels, whereas the coefficients on head's primary education (both complete and incomplete) are not statistically different from zero. Other variables that have a significant effect on household income are the number of income earners, other than the head, and non labour income sources (such as gifts and winnings, and the presence of a household member that receives a pension). Finally, provincial dummies are also included in the income equation.

[41] Sample characteristics of the main variables included in the income regressions are shown in table 3B.1.

[42] Recall that the profit function is expressed in the theoretical model as a function of assets at time t-1, therefore we use the previous year's level of assets. Non business and non farm assets are added as additional regressors because the measure of household income includes the income from the rent/lease/profit-sharing of non-business assets. Appendix A explains in detail the survey questions used to construct household income.

Table 3.5
Income equation estimates

Variables	*Coeff.*	*t*
Permanent income variables		
1992 farm assets*dummy small farm	-0.44	-1.99
1992 farm assets*dummy medium farm	0.02	0.94
1992 farm assets*dummy large farm	0.02	4.41
1992 business non-farm assets	0.01	3.87
1992 non business assets	0.03	3.14
Head employee	1207.49	8.71
Head self employed	171.95	1.93
Head complete primary educ	84.78	1.44
Head secondary educ	819.31	6.20
Head high educ	1899.87	4.20
Nr. of income earner	151.55	4.32
Pension (if someone receives a pension)	1279.46	3.60
Household size	140.95	3.32
Household size^(2)	-10.91	-2.80
electricity in the village	83.07	1.48
Intercept	-262.86	-1.51
Positive transitory income variables		
Winnings	381.71	3.99
Gift	146.27	2.01

Number of obs= 2020
F(28, 1991) = 16.34
R-squared= 0.37

The table records the results from equation (3.13) and estimates the predicted income for households that did not report a crop loss in 1993. Dependent variable is 1993 household income. This regression includes also provincial dummies. Both income and assets are measured in thousands of rupiah. Standard errors are robust.

Income equation for households that reported a crop loss

The difference between the observed income and the estimated income (ΔY_h) is calculated for households that report a crop loss in 1993. Table 3.6 shows some descriptive statistics of ΔY_h for the sub-sample with CL=1. The mean value is -45 (thousands of rupiah), but there is a high variability around it. Differentiating for those that report labour supply response and those who do not, the former have a positive mean value of ΔY_h (175), while the mean is negative for the group with LS=0 (-187). The difference in the means of ΔY_h for the two groups (LS=0 and LS=1) is statistically significant at the 1% level.

Table 3.6
Descriptive statistics of the difference in incomes

Variable	*Obs.*	*Mean*	*p25*	*p50*	*p75*
ΔY_h^{CL}	163	-44.57	-627.51	-219.32	221.84
ΔY_h^{CL} *LS	64	175.01	-653.88	-125.87	374.12
ΔY_h^{CL} *(1-LS)	99	-186.52	-600.46	-302.18	136.30

The table presents the descriptive statistics of the difference in incomes for households that reported a crop loss in 1993. The difference in the means is statistically significant at the 1% level

Equation 3.18 allows us to identify the different determinants of ΔY_h. In order to control for the endogeneity of labour supply we have added selection terms obtained from a probit equation that estimates the probability of using the labour supply to deal with the shock[43]. Before presenting the results for equation (3.18), we discuss the probit estimates (table 3.7).

A first important finding is that proxies for the household's access to the credit market (such as the value of the land and the presence of at least one financial institution in the village) have a negative and significant effect on the probability of adopting a labour supply response, whereas the probability increases significantly with the commonality of the shock, measured by the proportion of households that experienced a crop loss in the same village[44]. The probit equation controls also for the soil quality in the village. Households living in villages with a poor or average soil quality are more likely to respond with the labour supply

[43] Following Cameron and Worswick (2003), the probit model is estimated over the sample of households who reported a crop loss over the past five years instead of in 1993 only to increase the number of observations. There is no universally accepted goodness of fit measure for probit (Kennedy, 2003). Table 3.7 reports McFadden's pseudo-R^2, and the percentage correctly predicted. The McFadden's R-squared is a transformation of the log-likelihood. Define ℓ^* the maximized log-likelihood using probit, and ℓ_0 the log-likelihood using the same model but with only the intercept. R^2 is defined as $R^2 = 1 - \frac{\ell^*}{\ell_0}$. Another goodness of fit measure is the percentage of correct predictions. According to this measure, an observation is predicted as y=1 if the estimated probability is greater or equal than one-half. Otherwise the observation is predicted as y=0. The fit measure is given by $\frac{1}{n}\sum_{i=1}^{n}\left[y_i \hat{y}_i + (1-y_i)(1-\hat{y}_i)\right]$. The percentage of correct predictions is 64.85% and pseudo R-squared = 0.11. Cameron and Worswick (2003) reported a percentage of correct prediction of 59.14.

[44] Several studies link the possibility of relying on coping strategies to the shock commonality. For example, informal insurance mechanisms and asset sales are less useful as risk coping instruments when the extent to which shocks are common across households in the same area is high. This may force households to use other coping strategies, such as increasing labour supply.

to a crop loss than households living in villages in which the soil quality is high. A possible explanation is that farmers that cultivate a high soil quality have higher farm profits[45] and hence they may rely on strategies other than extra job to cope with the crop loss.

As regards demographic variables, the probability of varying the labour supply decreases with the age of the household head (older heads are expected to be less likely to take an extra job) and with the number of adult members with secondary education (educated adults are more likely to be active in the formal job market, hence they are less likely to change their work time allocation). Households in which the spouse of the head is inactive are less likely to adopt a labour supply response. This evidence would suggest that labour supply responses affect the leisure choices of working people, but do not induce non-working individuals to enter the labour market. Another significant variable is the number of female household members aged 13-17, suggesting a possible incidence of transitory shocks on girls' labour supply. The coefficient relating the presence of boys in the same age category is not significantly different from zero.

Table 3.7
Probit equation for the labour supply response

Variables	*Coef.*	*z*
Land value	-0.00004	-3.44
Dummy if credit in village	-0.30	-2.43
Proportion of other households experiencing a crop loss in the same village	1.17	3.64
Age household head	-0.02	-3.43
# of adult members with secondary education	-0.14	-1.81
# of female with age 13-17	0.22	1.95
Spouse is inactive	-0.36	-2.69
Poor soil quality in village	0.38	2.03
Average soil quality in village	0.41	2.79
intercept	0.25	0.91

Number of obs=532
Pseudo R-squared = 0.11
Percentage correctly predicted = 64.85

The table records the results from the probit regression that estimates the probability of responding with labour supply to a crop loss. Dependent variable is a dummy that equals one if the household had a labour supply response in the face of a crop loss over the period 1989-93

[45] Farmers living in villages with a high soil quality have farm profits that are statistically higher (at 0.01% level) than other farmers.

Table 3.8 reports the results of estimating equation (3.18). The income loss produced by the crop loss is the sum of the intercept and of a household specific term related to the value of 1992 farm assets, i.e. to the size of the farm. The estimated coefficient on farm assets decreases, in absolute value, as we move from small to large farms. This finding may be explained with the nonlinearity of the profit function. With decreasing returns on farm assets, the marginal effect of an increase in assets on the income loss is larger for low than for high levels of assets[46].

The size of the increase in income as a result of labour supply responses is estimated by interacting the labour supply dummy with the number of household members aged 13-64. The income gain due to the extra job as a response to 1993 crop loss is found to be statistically significant and to have a high impact on household income. In particular, each member aged 13-64 allows households to gain about 433 thousands of rupiah of extra labour income after a crop loss[47]. In order to control for possible effects on household income of other coping strategies, we include the dummy "take a loan/sell assets", which has a positive effect on ΔY_h, but it is not statistically significant (t=1.29) (when entered separately, the dummies "take a loan" and "sell assets" had statistically similar coefficients, hence we pool them to increase the number of observations). Finally, the 1992 value of non business and non farm assets has a positive coefficient, meaning that rich households have a higher ability to recover from the shock (even if the t statistic is not very high; t=1.52)

[46] Estimating equation (3.18) on the value of farm assets and the value of farm assets squared, the latter has a negative coefficient confirming the non linear relationship between income and farm assets.

[47] Kochar (1995) suggest that "the segmentation of labour markets by gender may make a household's vulnerability to crop income shocks a function of its demographic composition" (Kochar, 1995, p.159), so that households with a higher number of males in the working age with respect to females are more likely to smooth crop shocks. This hypothesis seems to be confirmed by our data: the dummy labour supply interacted with the number of males aged 13-64 is significant (coefficient=449.07, t=2.07), while the interaction with the number of females 13-64 is non statistically significant (coefficient=412.24, t=1.14).

Table 3.8
Income equation estimates
Dependent variable: $\Delta Y_h = Y_h^{CL} + \hat{Y}_h$

Variables	*Coeff.*	*t*
Measure of income loss		
1992 farm assets*small farm	-2.03	-2.11
1992 farm assets*medium farm	-0.12	-2.13
1992 farm assets*large farm	-0.02	-2.62
Recovery's measures		
1992 non-business assets	0.05	1.52
LS*N_1364	432.71	2.34
Dummy sell assets or take a loan	353.97	1.29
Other variables		
1st selection term* LS	-121.35	-0.23
2nd selection term* (1-LS)	-1130.89	-2.40
Dummy cut expenditure	140.04	0.72
Intercept	-1001.04	-2.66
Number of obs= 163		
F(9, 153) = 3.05		
R-squared= 0.21		

The table records the results from equation (3.18), and estimates the size of the income reduction due to the crop loss and of the increase in income due to coping strategies. The sample is households that had a crop loss in 1993. Both income and assets are measured in thousands of rupiah. Standard errors are robust.

From this regression we construct the measures of the income reduction caused by the crop loss and of the income gains due to labour supply response and other coping strategies. Predicted measures are summarized in tables 3.9a and 3.9b. The income reduction caused by the crop loss does not vary significantly with the size of the farm, and the mean is about 1260 thousands rupiah. If we add the extra labour income, the average value of income variation ($_3\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_3 X_h^{LS}$) is about -750, suggesting a significant impact of the labour supply response. Indeed, when only households that used extra labour are considered, the mean value of the income loss plus the extra labour income is 87.61, i.e. on average they recover all the income loss (the mean value of the extra labour income is about 1298, somewhat larger than the crop loss)[48]. The income recovery given by holding a higher initial level of non-business assets is small, and, as expected, higher for large farms.

Table 3.9a
Descriptive statistics of predicted variables

Variable	*Obs.*	*P25*	*Mean*	*p50*	*p75*
$Y_h^{CL} * CL$	163	165.00	924.27	442.00	1006.00
$\hat{Y}_h^{CL} * CL$	163	323.21	924.27	754.64	1361.52
$(\hat{\beta}_0 + \hat{\beta}_1 X_h^S) * CL$	163	-1378.13	-1261.42	-1169.49	-1104.70
$\hat{\beta}_2 X_h^A * CL$	163	23.65	152.99	74.82	186.02
$\hat{\beta}_3 X_h^{LS} * CL * LS$	64	865.41	1298.12	12.98.12	1730.83
$\hat{\beta}_4 X_h^L * CL * L$ [a]	62		353.97		

The table presents the descriptive statistics of the predicted income for households that had a crop loss in 1993, and the estimated measures of income loss caused by a crop loss and of income gains due to coping strategies.
[a] L is a dummy that equals one if the household reported sell assets or take a loan as a repose to the 1993 crop loss

Table 3.9b
Means of predicted measures of shocks, by size of the farm

	Averages			
Variable	*Small farms*	*Medium farms*	*Large farms*	*All farms*
${}_1\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_2 X_h^A + \hat{\beta}_3 X_h^{LS} + \hat{\beta}_4 X_h^L$	-368.87	-533.83	-370.30	-461.68
$\hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_2 X_h^A$	-1122.65	-1160.09	-986.90	-1109.98
${}_2\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S + \hat{\beta}_3 X_h^{LS}$	-639.84	-771.47	-832.83	-750.28
${}_3\hat{Y}_h^S = \hat{\beta}_0 + \hat{\beta}_1 X_h^S$	-1253.68	-1260.26	-1280.63	-1261.42

N=163

The table presents the mean values of the predicted measures of income loss and gains (for households that reported a crop loss in 1993), considering each coping strategy separately and distinguishing for poor and non-poor farmers.

[48] The mean values of the income reduction caused by the crop loss and of the income gain due to the labour supply response estimated by Cameron and Worswick (2003) are respectively -1235 and 838 thousands rupiah.

3.6.2 Consumption equation estimates

On the basis of the theoretical model presented in section 3.2, constrained households consume less and work more than if they were unconstrained, and these effects are even more pronounced in the face of a negative shock. Moreover, as previously suggested, for some households, consumption choices may not be based on the permanent income model, but rather be driven by current income and the need to accumulate savings. Consumption equation (3.22) is then analyzed focusing on the differences between constrained (small farms) and unconstrained (medium and large farms) households. Results are reported in table 3.10[49]. The first difference between poor and non-poor farmers is related to the income measure which appears to be relevant for consumption choices. According to the permanent income hypothesis, consumption is determined by permanent income and not by current income: this implies that consumption should be unaffected by transitory income changes[50].

Our estimates suggest that consumption of medium and large farms is indeed determined by permanent income, while the crop loss has no impact on non-durable expenditures[51]. Consumption of small farms is instead influenced by both permanent and transitory income. The implications of this result for consumption smoothing are better grasped if we consider the difference in the estimated marginal propensity to consume out of permanent and transitory income. As suggested by Deaton (1997), a statistically positive difference between these two coefficients would represent evidence that households are willing/able to smooth consumption relative to income. This result is confirmed for medium and large farms (p-value=0.000), but not for small farms. Indeed, the estimated coefficients on permanent income, crop loss, and extra labour income on consumption for small farms are statistically equal (test for the equality of the three coefficients: F(2,2162)=0.84, Prob>F=0.43)[52].

---

[49] Five outliers which belong to the top percentile of the expenditure distribution are excluded from the expenditure regression. OLS estimates of (3.22) may be biased because households choose simultaneously consumption levels and coping strategies in the face of a shock. To control for the endogeneity of labour supply response, we included the selection terms, calculated from the probit for the labour supply response, in the consumption equation. They are not statistically significant and estimates do not significantly change.

[50] Precisely, consumption smoothing means that short-term fluctuations in income are spread over a long time horizon, which can be defined as the lifetime (life-cycle model), or as an infinite horizon (which would be appropriated for a dynasty, according to the permanent income hypothesis).

[51] We cannot reject the hypothesis that the effect of the shock on consumption is the same for medium and large farms (F(1,2161)=0.80, Prob>F=0.373). Hence we pool the two groups.

[52] Flavin (1985) explores whether the empirical rejection of the permanent income hypothesis occurs because agents are myopic, or because some agents face liquidity constraints. She finds that the observed excess sensitivity of consumption to current income is due to liquidity constraints. The extent to which consumption is affected by the presence of borrowing constraints is examined also by Zeldes (1989).

A second distinction between poor and non-poor households is the magnitude of the marginal propensity to consume out of the relevant income measure. Rich farmers and owners of medium farms consume more than 70% of their permanent income[53]. The marginal propensity to consume out of current income for poor households is about 0.5[54], with the consequence that about one half of the current income is transferred into savings. This seems to confirm what suggested by the literature: in the face of exclusion from financial markets, poor households have to perform an autarchic saving strategy, to build a buffer stock of assets and to self-finance profitable investments (Carter and Barrett, 2005; Fafchamps, 1999).

The third result that emerges from table 3.10 is that different coping strategies that change current income have different impacts on consumption for poor households. The income generated by the measures "non-business assets" and "take a loan or sell assets" is entirely used to mitigate the consumption reduction due to the crop loss, even if, as reported in tables 3.9a and 3.9b, these measures have only a marginal role in compensating the income loss[55]. As noted above, the marginal propensity to consume out of extra labour income is about 0.5, and statistically lower than the one estimated for the other measures.

[53] The marginal propensities to consume out of permanent income are 70% and 90% for medium and large farms respectively, and they are statistically different (F(1,2162)=3.69, Prob>F=0.055).

[54] Test on coefficients:

a) extra labour income=crop loss=permanent income=0.5: F(3,2162)=0.69, Prob>F=0.559
b) extra labour income=crop loss=permanent income=0.7: F(3,2162)=4.80, Prob>F=0.002.

[55] Coefficients on "non-business assets" and "take a loan or sell assets" are non statistically different in consumption equation, hence we consider both measures together.

Table 3.10
Expenditure equation estimates

Variables	*Coef.*	*t*
$\hat{Y}_h^P$ -small farms	0.45	6.28
$\hat{Y}_h^P$ -medium farms	0.71	7.71
$\hat{Y}_h^P$ -large farms	0.91	10.19
${}_3\hat{Y}_h^S$ - small farms	0.63	3.03
${}_3\hat{Y}_h^S$ - medium/large farms	0.08	0.87
$(\hat{\beta}_2 X_h^A + \hat{\beta}_4 X_h^L)$ -small farms	1.13	1.80
$\hat{\beta}_3 X_h^{LS}$ - small farms	0.46	1.88
Transitory positive income	1.20	5.70
$\hat{\varepsilon}_h * (1 - CL)$	0.13	1.22
$\hat{u}_h * CL$	0.36	6.13
members age 0-5	1.67	0.05
members age 6-11	185.31	4.72
members age 12-17	267.16	6.10
members age 18-64	225.93	6.74
members 65 year or over	126.16	2.02
intercept	367.80	3.81

R-squared= 0.33
Number of obs= 2178

The table records the results from equation (3.22). Dependent variable is 1993 non-durable household expenditure. Standard errors are robust

3.7 Conclusions

This chapter uses the 1993 round of the Indonesian Family Life Survey to explore two issues. First, we seek to explore which variables influence the adoption of ex-post income smoothing strategies and whether these strategies completely or partly recover the income reduction due to the shock. Quantitative measures of income shocks and households' ability to cope with the shock are estimated for poor and non-poor households, and using these measures we can explore whether the increase in income due to income-smoothing strategies partially offsets, or exceeds, the income loss due to the shock. Our focus is on the labour supply response in the face of a crop loss. The second issue we explore is whether households smooth consumption relative to income and the role played by coping strategies (labour supply response in particular) in mitigating consumption reductions.

The descriptive analysis and the literature presented in the chapter suggest that the coping strategies adopted to overcome a crop loss, and the link between consumption and permanent or transitory income may be quite different between poor and non-poor households. Non-poor households may have access to several coping mechanisms, such as borrowing or selling assets, and they may have enough resources to leave consumption unaffected in the face of a transitory shock. On the other hand, poor households have low endowment of assets and they may face borrowing constraints, thus needing to accumulate savings. In order to examine these issues, the analysis is conducted distinguishing between consumption behaviour of poor and non-poor households.

The theoretical framework that underlines the analysis is a life-cycle model in which income is generated by farm profits and by off-farm labour income. Productive and unproductive assets, together with an exogenously determined amount of labour, enter as a determinant of farm profits, and the remaining amount of time can be allocated to either leisure or wage market. A negative shock reduces farm profits (and the amount of farm labour) and increases the marginal utility of off-farm labour income. The model predicts that the marginal utility of leisure and of consumption are both greater when households face credit constraints. This implies that in general constrained households consume less and work more than if they were unconstrained, and these effects are even more pronounced in the face of a negative shock.

The empirical methodology follows a two-stage procedure. The predicted income for households that do not report a crop loss is estimated first. If the crop loss is exogenous, this predicted income is an appropriate estimate of the household's income for all households. For households that experienced a crop loss in 1993, the difference between the observed income and the predicted income is constructed. This difference is regressed on a set of variables that affect the magnitude of the income shock (e.g. farm assets), and the size of the income gains due to the labour supply response and other coping strategies. From that regression we construct measures of the income reduction caused by a crop loss and of the income gains due to coping strategies, and we can explore how much of the income loss the labour supply response and other strategies contribute to compensate. These measures are also used to examine the consumption choices of different households.

In order to explore whether households smooth consumption relative to income, non-durable household consumption is regressed on permanent income, the measures of crop loss and income gains, and other household characteristics. We focus on the differences between constrained (small farms) and unconstrained (medium and large farms) households.

Our results suggest that the income gain given by the extra labour supply response completely compensate the income reduction caused by the crop loss, while the role of non-business assets and “take a loan or sell assets” is in general marginal in recovering the income loss. As regard consumption behaviour, there are two main differences between poor and non-poor households. First, while medium and large farms smooth consumption relative to income, this is not so for small farms: for the latter, the main components of transitory income (crop loss and the extra labour income) have an effect on consumption that is statistically significant and equal to the one associated with permanent income. This is not so for the income gain due to coping strategies other than labour supply: what poor households receive from taking a loan or selling assets is entirely used to mitigate the consumption reduction due to the crop loss, whereas the extra labour income due to the labour supply response is partly transferred into savings.

The second distinction between poor and non-poor households concerns the marginal propensity to consume out of the relevant income measure: the former save about a half of their current income, whereas the latter consume a fraction of their permanent income close to one. This result may confirm the need for poor households, that are more likely to be excluded from financial markets, to rely on autarchic saving strategies.

Appendix 3A

Following the Cameron and Worswick (2003)'s approach, household income is regressed on a set of variables that determine the permanent and transitory income components, where the transitory income is defined as the income reduction due to self-reported crop loss. The first column of table 3A.1 (CLS) reports the results from equation (3.8). In this framework the predicted measure of income loss includes also the income gains due to coping strategies. The second column (VLS) of table 3A.1 estimates the effect of the crop loss net of labour supply response. From these regressions we predict the household permanent income, and the size of the crop loss in the constant labour supply case (CLS) and in the variable labour supply case (VLS). Predicted income components are then used as regressors in the consumption equation, as reported in table 3A.2.

Table 3A.1
Income equation estimates

	CLS case		*VLS case*	
Variables	*Coef.*	*t*	*Coef.*	*t*
Income shock				
CL*farm assets*farm assets bottom25%	-1.13	-1.69	-1.81	-2.40
CL* farm assets* farm assets 25-100%	-0.01	-2.14	-0.02	-2.48
LS*# members age 13-64	-	-	158.60	1.86
Positive transitory income variables				
winnings	353.43	3.91	355.12	3.93
gift	133.23	1.96	125.67	1.84
Permanent income variables				
1992 farm assets	0.02	3.93	0.02	3.96
1992 business non-farm assets	0.01	4.64	0.01	4.66
1992 non-business assets	0.02	2.30	0.02	2.30
pension	1182.01	3.84	1192.43	3.89
Head employee	1126.65	8.18	1133.38	8.31
Head self employed	141.64	1.49	145.86	1.56
Head complete primary educ	64.57	1.16	64.50	1.16
Head secondary educ	754.05	6.00	744.17	5.93
Head high educ	1431.93	3.92	1421.86	3.88
Nr. of income earner	176.24	4.52	173.64	4.68
Household size	145.71	3.70	137.19	3.48
Household size square	-11.21	-3.08	-10.55	-2.91
electricity in the village	81.03	1.55	82.57	1.58
Intercept	-318.46	-1.99	-307.86	-1.93
Number of obs= 2183				
R-squared= 0.38				

The table records the results from income equations (3.8) and (3.11). Dependent variable is 1993 household income. Regressions include also provincial dummies. The selection terms that control for the endogeneity of the labour supply response are not statistically significant (they are jointly not statistically different from zero). Standard errors are robust. Household income and assets are in thousands of rupiah.

Table 3A.2
1993 non durable expenditure equation estimates.

	CLS case				VLS case			
Variables	*Coef.*	*t*	*Coef.*	*t*	*Coef.*	*t*	*Coef.*	*t*
Y_h^P	0.84	12.27	0.84	12.22	0.84	12.25	0.84	12.19
$\hat{Y}_h^S$	0.62	1.69	-	-	0.56	2.05	-	-
$\hat{Y}_h^S$ - small farms	-	-	1.70	3.22	-	-	1.05	3.33
$\hat{Y}_h^S$ - medium/large farms	-	-	0.21	0.45	-	-	0.18	0.41
Transitory positive income	1.24	4.93	1.25	4.98	1.24	4.93	1.25	4.96
ε	0.32	6.54	0.32	6.53	0.32	6.54	0.32	6.53
intercept	285.64	2.86	283.56	2.84	291.37	2.92	288.96	2.90
F-Tests: $\hat{Y}_h^S - non\,small = 0$ $and\ \hat{Y}_h^S - small = 1$			F(2,2167)=0.96 Prob>F=0.38				F(2,2167)=0.10 Prob>F =0.91	
R-squared= 0.33								
Number of obs= 2178								

This table reports the results of the expenditure regression. Regression includes also the number of household members in each age category. Standard errors are robust

Appendix 3B

Table 3B.1 summarizes the sample characteristics for the main variables included in the income regressions presented in tables 3.5 and 3.8. The mean for dummy variables represents the proportion of that group: for example, 81% of the household heads work as self-employed, and 13% as private or government workers. 26% of the heads are illiterate and 35% did not complete the elementary level.

Table 3B.1
Sample means of variables included in the income regression

Variable	*Mean*	*Std. Dev.*	*Min*	*Max*
Shocks				
1993 Crop loss	0.08	0.26	0	1
1993 Labour supply	0.03	0.17	0	1
"take a loan or sell assets" (1993)	0.03	0.16	0	1
Household economy				
Household income	970.97	1506.37	-100	18464
1992 farm asset	5551.32	15204.49	0	268025
1992 business assets	399.60	7811.01	0	356650
1992 illiquid non-business assets	3936.08	9829.80	0	275430
1992 liquid non-business assets	190.53	628.64	0	9500
whether a household member receives a pension	0.02	0.13	0	1
whether the household receives winnings	0.17	0.38	0	1
whether the household receives gifts	0.16	0.37	0	1
# income earner other than the head	0.69	0.92	0	6
Work status of the household head				
Head is inactive	0.05	0.21	0	1
Head is employed	0.13	0.34	0	1
Head is self employed	0.81	0.39	0	1
Head is family worker	0.01	0.11	0	1
Education of the household head				
Head unschooled	0.26	0.44	0	1
Head incomplete primary	0.35	0.48	0	1
Head completed primary	0.23	0.42	0	1
Head secondary education	0.14	0.34	0	1
Head high education	0.02	0.12	0	1
Household size	4.59	2.01	1	16
Electricity in the village	0.76	0.43	0	1

The table reports the sample means of key variables used in the income equations reported in tables 3.5 and 3.8. Household income, expenditure and assets are in thousands of rupiah

Chapter 4

The Effects of Risk and Shocks on School Progression in Rural Indonesia

4.1 Introduction and literature review

In developing countries where incomes are volatile and on average low, and financial and insurance markets are incomplete (Townsend, 1994; Morduch, 1990), ex ante risk and ex post shocks may have a large impact on child labour and child education. Many papers support the idea that, in poor countries, child time allocation may be used as a strategy to both insure ex ante against risk and to cope ex post with negative hardships in the face of incomplete insurance markets and credit constraints (Jacoby and Skoufias, 1997; Beegle at al., 2006; Baland and Robinson, 2000; Ranjan, 2001). Children may be used as means to amass a buffer stock, and/or they may be sent to work to earn money and support the family. Children may work as part-time family workers, helping their parents at home and in the fields. Moreover, children may drop out to save the costs of schooling.

Focusing on the ex post responses to shock, many empirical and theoretical studies analyze the extent to which child time allocation provides a means to cope with transitory income shocks. Sawada and Lokshin (2001) find evidence of a positive effect of unanticipated shocks on child labour in rural Pakistan. Sawada (2003) shows that children's propensity to drop out of school in rural Pakistan is related to transitory shocks. Beegle, Dehejia, and Gatti (2003; 2006) find that in Tanzania, crop shocks lead to a significant increase in the level of child labour, but that households with durable assets can mitigate these effects. Jacoby and Skoufias (1997) examine the extent to which income shocks affect school participation of children in rural India, and the role played by incomplete financial markets in determining child schooling decisions. Guarcello et al. (2003) analyze the effects of collective and individual shocks on household decisions concerning children's school attendance and labour supply in Guatemala. Their results suggest that credit rationing is a key determinant of the household's decision to invest in the human capital of children. Jensen (2000) shows that school enrolment rates substantially decline in the presence of adverse agricultural conditions in Cote d'Ivoire. Thomas *et al*. (2004) find evidence that the 1997 Indonesian crisis had negative effects on education of young children in poor households. Duryea *et al*. (2007) analyze the impact of household short-term economic shocks (unemployment of the household head) on schooling and employment transitions of young people in Brazil.

Another strand of literature examines how ex ante risk affects education and child labour decisions. Kazianga (2005) shows that child time may be used as an ex ante measure to diversify risk in an uncertain environment. He finds that the less well developed are other formal or informal insurance mechanisms, the more is child labour used. Maitra *et al*. (2006) show that, in India, children serve as a buffer to deal with the uncertainty created by the labour market. Households may prefer neither to send the children to work nor to send them to school, instead keeping them idle to deal with uncertainty. Fitzsimons (2007) finds evidence that, in Indonesia, aggregate village components of risk negatively affect children education, whereas households are able to insure themselves against idiosyncratic risk without reducing child schooling. Jalan and Ravallion (2001) find no evidence that risk decreases child schooling in Rural China; indeed, they find that the coefficient on income risk is positive and significant at the 5% level.

The use of child time as a risk management strategy, both ex ante and ex post, may have long run consequences. Children who are withdrawn from school may not be able to restart school or to recover the educational gap: in this way temporary schooling interruptions have lasting impacts. This issue is less well established in either the theoretical or the empirical literature. Jacoby and Skoufias (1997) find that eliminating the negative effect of shocks on child schooling would have only modest effects on children's human capital attainment. Guarcello *et al*. (2003) observe that child labour has a high degree of persistence, because it is difficult to return to school. Duryea *et al*. (2007) point out the potential long-term consequences of short term shocks for children's human capital.

Few papers propose a theoretical treatment of child schooling choices in the presence of risk taking into account the irreversibility or state dependence of school attendance. De Janvry et al. (2006) develop a model of child schooling decisions in the face of income shocks with a re-entry cost as an additional cost of schooling. This allows them to consider the dependence of child enrollment on previous schooling decisions. Using Mexican data, they find evidence that shocks force children to drop out of school with long run effects on children's human capital given the state dependence of school attendance. They find that this state dependence is higher in secondary than in primary school, suggesting that school attendance is more flexible for children at the elementary than at the secondary level.

The first contribution of this chapter is that it analyzes the role that both ex ante risk and ex post shocks play in determining parental decisions with regard to child schooling, differentiating by school levels. The key assumption of the model is that withdrawal from school is an absorbing state, that is children cannot re-enroll once they stop going to school. In this way, temporary interruptions in child schooling in the face of risk or short-term shocks have long term impacts on the child human capital. Given irreversibility of withdrawal from school, risk and uncertainty enter the schooling decision in different ways from that proposed by the literature (which suggests that income risk decreases child education, see Fitzsimons, 2007): in the face of household future income variability parents are more likely to send children to school in the current period to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults). When income is revealed the household may withdraw the child from school if this should prove necessary (but cannot re-enrol the child if he/she has been previously withdrawn).

This chapter analyzes the progression decisions from one education level to another (in particular, from elementary to junior secondary school) using the Indonesian Family Life Survey. The focus on school transitions has been suggested by the low continuation rates in Indonesia after graduation, and in particular after the completion of the primary level. Moreover, differentiating the decision to attend school by school levels allows us to better capture the effects of the state dependence.

The chapter is organized as follows. Section II presents some evidence on child schooling in Indonesia and descriptive analyses of the data. The theoretical model is presented in section III. Section IV discusses the empirical methodology and section V summarizes the results. Section VI concludes.

4.2 Evidence on child schooling in Indonesia

This section first presents a brief picture of the educational system in Indonesia, showing how it is organized, and outlining the main problems and the main changes over recent years. Then data on child education are discussed, with reference to the 1993 round of the Indonesian Family Life Survey (IFLS1). The second part of this paragraph examines the main differences in enrolment rates with regard to gender, poverty status and level of schooling. Finally, data on school infrastructure and on household educational expenditure are briefly discussed.

4.2.3 The Indonesian education system

Over past decades, Indonesia has invested a large amount in the education system, improving both the quantity and quality of education. By 1984 the Indonesian government had fully implemented the six year compulsory education for primary school age children, and the 1994 law increased the compulsory basic education to nine years: six years of primary education (for ages 7-12) and three years of junior secondary education (ages 13-15). Currently, the formal school system consists of the following levels:

Table 4.1
The Indonesian schooling system

Level		*Year*	*Age*	*Cost of schooling*
Elementary school	compulsory education (since 1984)	1-6	7-12[a]	Public schools are free; tuition fees for private schools
Junior secondary school	compulsory education (since 1994)	7-9	13-15	Public schools are free; tuition fees for private schools
Senior secondary school		10-12	16-18	No free education. Partial government support for the senior secondary education
Under graduate	Higher education	13-16	19-22	No-free education. Government does not support attendance of public universities
Post graduate		17-18	23-24	
Doctorate		19-21	25-27	

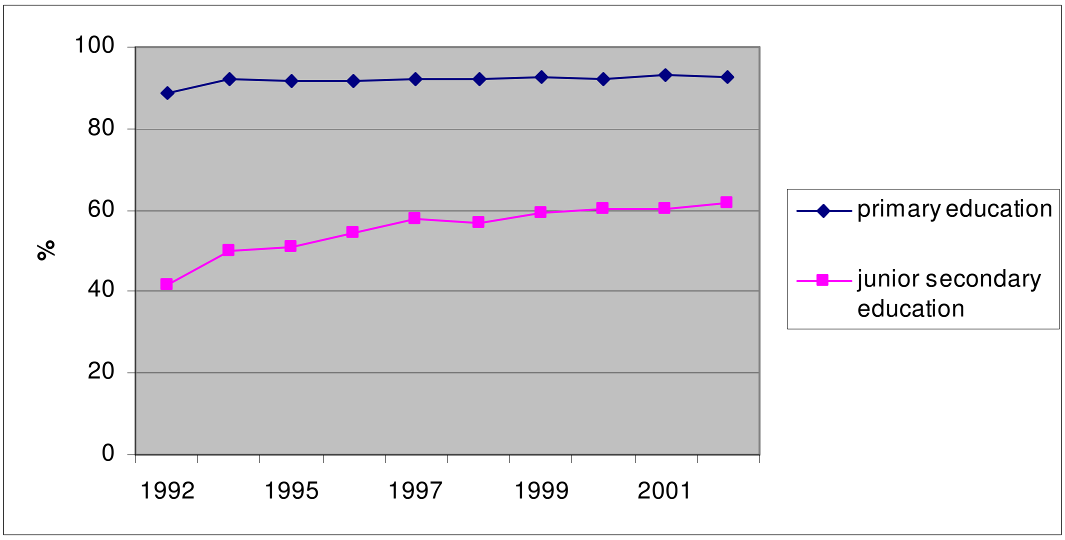
The table presents the Indonesian schooling system.
[a] 6 years olds can also be admitted

Source: American Institutes for Research (2002). On-line source: http://www.ieq.org/pdf/2nd_Ed_Casestudy_paper.pdf

At the junior and senior secondary levels there are two curricula: academic and vocational. The public sector provides the greater part of the education services, and compulsory education is free, but there is a strong private education sector as well.

The survey utilized in the empirical analysis of this chapter is from the first round of the Indonesian Family Life Survey (IFLS1), conducted in 1993. In that year, only the primary schooling level was compulsory, and since the early 1980's government has made a major effort to achieve the goal of universal elementary education. These policies were largely successful in improving primary educational levels, but two aspects of the Indonesian educational system remained particularly problematic (Manning, 2000; UNDP, 2004): high drop out rates from both primary and secondary school (the drop out rate was nearly 20% in 1993/94 for primary school, and 8% in junior secondary school in 1990/91), and low continuation rates from primary to junior secondary school (the continuation rate was close to 62% in 1993/94). These high drop out and low continuation rates have been seen as causes of adult illiteracy and reliance on child labour (Manning, 2000). A new policy was thus required to achieve universal education, and in 1994 compulsory basic education was increased to nine years. Currently, elementary education is almost universal: primary net enrolment ratios reached 93% in 2002 (figure 4.1), with no significant differences between rural and urban areas, between girls and boys, or across poverty quintiles (UNDP, 2004). As a result of the 1994 law, continuation rates from primary to junior secondary rose to nearly 72% in 1999-2000, and the net enrolment ratio at junior secondary school increased from 42% in 1992 to 62% in 2002 (figure 4.1), although with significant differences between rural and urban areas (lower in rural areas) and between poor and rich households. Recent policy interventions have significantly reduced dropout rates (only 3% at elementary level, and 4% at junior secondary level in 2000/01). In spite of this, a substantial proportion of children still fail to complete the nine year basic education cycle (UNDP, 2004), and children from poor families are more likely to drop out of school early and to be called upon to support their parents (Acedo, 2002). Manning (2000) reports that approximately 15%-20% of children aged 12-13 and 20%-30% of children aged 14-15 are not enrolled in school and may potentially be involved in the work force.

Figure 4.2. Net enrollment ratios, primary and junior secondary school



Source: Susenas, UNDP (2004). Net enrollment ratio is the share of children of official age that are enrolled in school

4.2.4 Data – Child education

This sub-section discusses the main findings on child education in rural Indonesia, based on IFLS1 (1993) data and using individual, household and village sampling weights for the descriptive analysis. Most children start school at age 7, and only a few individuals are at school after age 17. This suggests a focus on the sample of children aged 7-17 years. Approximately 75% of Indonesian children in the 7-17 age group were attending school in 1993. The number of years in education averaged four, with marginally statistically significant differences between girls and boys (slightly higher for boys). Table 4.2 provides further information on average values of child and household characteristics. To briefly summarize, the average values of schooling years were 3.3 years for the mother and 4.2 years for the father (the length of the primary level was six years). Nearly all siblings aged 7-12 years were attending school, while 60% of siblings aged 13-15 years were in school and less than one out of three of young adults aged 16-18 years were in school.

As regards the economy of the family, farm earnings were the main income source for rural households. Farms were typically owned by the household head and in many cases were cultivated with the help of family members (children may have been involved in the family farm). On the basis of the survey, nearly half the households did not have an income earner other than the head.

Table 4.2
Mean characteristics of households and children

	Weighted mean[a]	*Standard Error*
Age of child	11.74	0.05
Female child	0.48	0.01
Child attending school	0.75	0.01
Child years of schooling (completed)	3.93	0.05
Years of schooling mother (if there is the mother)	3.24	0.08
Years of schooling father (if there is the father)	4.39	0.10
# sibling attending school aged 7-12 years/ # sibling aged 7-12 years	0.90	0.01
# sibling attending school aged 13-15 years/ # sibling aged 13-15 years	0.61	0.02
# sibling attending school aged 16-18 years/ # sibling aged 16-18 years	0.28	0.02
Household owns house	0.93	0.01
Log value of mthly per capita expenditure	10.29	0.02
Household owns farm	0.63	0.01
Household owns non farm business	0.32	0.01
Head is self-employed	0.65	0.01
Head is self-employed with the help of family worker	0.34	0.01
Head is employee	0.27	0.01
Number of income earners other than the head	0.79	0.02
Number of children	*4498*	
Number of households	*2239*	

The table presents the main characteristics of the households and of the children. Per capita expenditure is in rupiah.
[a] Household and individual sampling weights are used.

Looking at the primary activity of children during the week prior to the interview, the percentage of those that reported attending school is nearly 66% with significant differences between boys and girls (table 4.3; this information was only sought from children of ten or more years old[56]). A substantial proportion of children reported work or job search as their primary activity, and a similar percentage of girls reported doing housework. In Indonesia there is a considerable jump in work participation after 14 years (Fitzsimons, 2007) and hence the same data are presented for the 10-14 age group (table 4.4). The percentage of children that reported attending school as primary activity increases to about 79% but is slightly lower for girls. Among 10-14 years olds, approximately 9% reported work or looking for a job as their primary activity, with no significant differences between males and females. The proportion of girls that did the housework decreases compared with the age 10-17 group, but is still substantial.

Table 4.3
Primary activity during the week prior to the interview, children aged 10-17 years[a]

	Male		*Female*		*Male and female*	
	Freq.	*Weighted %*	*Freq.*	*Weighted %*	*Freq.*	*Weighted %*
Working	227	14.0	189	12.6	416	13.3
Job searching	77	5.5*	39	3.1*	116	4.4
Attending school	1,149	66.9*	1,039	62.7*	2,188	64.9
Housekeeping	13	0.9*	153	10.1*	167	5.3
Other	184	12.4	159	11.6	342	10.0
Total # of children	*1,650*	*100*	*1,579*	*100*	*3,229*	*100*

The table reports the primary activity of children aged 10-17 years during the week prior to the interview.
[a] this question was only asked of children 10+ years old
* the difference in the weighted proportions are statistically significant

Table 4.4
Primary activity during the week prior to the interview, children aged 10-14 years[a]

	Male		*Female*		*Male and female*	
	Freq.	*Weighted %*	*Freq.*	*Weighted %*	*Freq.*	*Weighted %*
Working	65	6.2	64	5.7	129	5.9
Job searching	27	3.1	18	2.3	45	2.7
Attending school	913	80.6	861	77.2	1,774	79.0
Housekeeping	7	0.6*	46	4.0*	53	2.3
Other	97	9.5	97	10.6	194	10.1
Total # of children	*1,084*	*100*	*1,086*	*100*	*2,195*	*100*

The table reports the primary activity of children aged 10-14 years during the week prior to the interview.
* the difference in the weighted proportions is statistically significant

[56] A survey by Asra (1993, 1996) suggests that the work of children under 10 years old is very small in Indonesia compared with those 10 plus years old.

We examine now in detail the school participation rates by gender, poverty status and school levels.

Primary education

Net enrolment ratios. Data from the IFLS show that in Indonesia the 1993 net enrolment ratio (NER) was approximately 87% (table 4.5), confirming the findings of other studies (UNDP, 2004). Achieving universal primary education has been a national goal since the early 1980's, and these data show that ten years on, Indonesia achieved high levels of access to basic education for children aged 7 to 13 years. Looking at table 4.5, we can see that there is gender equity in access to primary education (the participation rate is only slightly higher for girls than for boys). Enrollment rates increase as per capita consumption expenditure (PCE) rises, both for males and females, even if the disparities between poverty terciles are not so marked.

Gross enrolment ratios. Students may enroll late (over 7 years of age) or may repeat grades, with the consequence that they complete primary school when they are older than twelve. To account for that, Gross Enrolment Ratios (GER) for primary education are reported (table 4.5). If many children outside the official primary school age range are enrolled in primary school, the GER will exceed the NER by a large amount. Data show that in Indonesia there are many under-age and over-age children who attend elementary level, in particular for boys and among poor households.

Table 4.5
Net and gross enrollment ratios in primary schools by PCC tercile and gender

	NER (%)			*GER (%)*		
	Male	*Female*	*Male and Female*	*Male*	*Female*	*Male and Female*
Lower third	80.7	82.3	81.5	94.1	91.9	93.0
Middle third	91.1	89.4	90.2	99.5	99.3	99.4
Upper third	89.3	90.7	90.0	99.2	98.1	98.7
All	*87.0*	*87.4*	*87.1*	*97.6*	*96.4*	*96.9*

The table reports the net and gross enrollment ratios for elementary school.
Net enrollment ratio (NER)= enrolled children in the age group 7-12 / total # of children in the age group 7-12
Gross enrollment ratio (GER)= enrolled children / total # of children in the age group 7-12

Junior secondary education

Net and gross enrolment ratios. In 1993 junior secondary education was not compulsory, and there was a low continuation rate after the graduation from primary school. Table 4.6 reports NER and GER for the junior secondary level. In 1993 only approaching 40% of the age group was enrolled in junior secondary school, with a lower participation rate for girls than for boys (37% and 42% respectively). Data show considerable disparities between poverty terciles, with an NER for the top third of the PCC distribution more than twice than that of the lowest third. The gender gap was relatively low for poor and rich households, whereas it was high for the middle third. There is a significant jump in male junior secondary enrolment rates between the lower and the middle third of PCC distribution, while the corresponding jump for female enrolment is between the middle and the upper tercile[57]. The GER for junior secondary is about 50%. As for the NER, there are considerable differences by gender and household wealth.

Table 4.6
Net and gross enrollment ratios in junior secondary schools by PCC tercile and gender

	NER			*GER*		
	Male	*Female*	*Male and Female*	*Male*	*Female*	*Male and Female*
Lower third	26.6	24.5	25.6	36.4	31.8	34.1
Middle third	42.4	28.2	35.3	53.8	39.9	46.9
Upper third	58.3	56.8	57.6	78.9	71.4	75.2
All	*42.4*	*36.5*	*39.5*	*56.4*	*47.7*	*52.0*

The table reports the net and gross enrollment ratios for junior secondary school.
Net enrollment ratio (NER)= enrolled children in the age group 13-15/ total # of children in the age group 13-15
Gross enrollment ratio (GER)= enrolled children/total # of children in the age group 13-15

As discussed above, there is a small fraction of children not enrolled in primary school, and a considerable number of children not attending junior secondary level. Considering out-of-school children, we may distinguish between children who have never enrolled in school, and those who have at some time been to school. This second group may be further divided between children who drop out before the completion of the school level, and children who stop going to school after graduation. These decisions may have different implications for the child's human capital and the likelihood that he/she becomes literate when adult.

Tables 4.7 reports the weighted percentages of children and young adults who are not enrolled in school, but have attended at some time, by age and gender. Less then 5% of children aged 8 through 11 years are not enrolled in school, with a small gender gap. After the age of 12 the percentage of children not enrolled in school increases (at that age children should have graduated from the primary level): about 11% of 12 year old and 24% of 13 year old children are not enrolled in school, as is the case for more than half children aged 15 to 17 years. With increased age, the gender gap significantly rises, as girls are less likely to be

[57] A possible explanation is that households in the middle third of the PCC distribution were able and willing to send boys to junior secondary school, but were less willing to do so for girls. As in primary education, there was a large number of children aged younger than 13 or older than 15 attending junior secondary school.

enrolled in school. The distribution of out-of-school children by age suggests that the largest loss occurs at the transition points between school levels. To better describe this issue, table 4.8 shows the distribution of out-of-school children, who have at some time been to school, according to the number of schooling years completed. More than half of out-of-school children aged 7-17 years, and who have at some time been to school, left school after graduation from primary level. The proportion of out-of-school children who completed less than six schooling years is nearly 27% for both boys and girls. This percentage increases to 33% for poor households, compared with 21% of rich households. Rich households have a higher percentage of out-of-school children that completed junior secondary school.

School progression rates at different educational stage can be further analyzed estimating the conditional survival function (we follow the procedure of Sawada and Lokshin, 2007). Let n_k denote the number of children who have completed education at least at stage k-1 (the set is not right-censored at education level k-1). Among these n_k students, define as h_k the number of children who have completed education at least at level k, so that $h_k = n_{k+1}$. Then h_k/n_k is the fraction of students who progressed to a higher stage of education, conditional on having completed stage k-1; this is an estimate of the conditional survival probability at education level k. Results suggest that in 1993 the probability of continuation after the completion of the elementary school in rural Indonesia was 65% (as reported in section 4.2.1, Manning (2000) and UNDP (2004) reported a continuation rate close to 62% in 1993/94), and the probability of continuation to senior secondary school was 73%.

Table 4.7
School non-enrollment percentages, by gender and age

Child Age (yrs)	*Male (%)*	*Female (%)*	*Male and female (%)*
7	0.7	0.5	0.6
8	2.6	1.4	2.0
9	1.6	2.6	2.0
10	3.0	3.5	3.2
11	4.4	3.8	4.1
12	12.1	11.0	11.5
13	20.6	28.7	24.6
14	34.6	43.8	38.8
15	50.9	52.0	51.4
16	63.9	73.4	67.9
17	61.2	70.2	65.7

The table reports the percentages of children that are not enrolled in school by age and gender. Percentages are computed using individual sampling weights and on the sub-sample of children that have at some time been to school.

Table 4.8
Distribution of out-of-school children according to schooling years completed and gender

School level	*Number of years completed*	*Male (%)*	*Female (%)*	*Male and female (%)*
Elementary – not completed	<6	28.2	26.9	27.6
Elementary – completed	6	58.9	62.1	60.5
Junior secondary – not completed	7-8	4.3	2.3	3.4
Junior secondary - completed	9	7.9	8.2	8.0
Senior secondary	> 9	0.7	0.5	0.6
TOT		*100*	*100*	*100*
N		*878*		

The table reports the distribution of out-of-school children according to number of schooling years completed. Percentages are computed using individual sampling weights and on the sub-sample of children 7-17 years old that are not at school but have at some time been to school

Table 4.9 presents some data on the time utilization of out-of-school children who have previously attended school, focusing on the sub-sample of children aged 10-14 years. Data suggest that being out-of-school does not necessarily mean working: some children neither go to work nor go to school (this category of children is referred to as "idle children" in the literature). The proportion of children involved in the labour force is statistically greater for boys than for girls, whereas girls are more likely to do housework and be involved in childcare. In considering the link between being out-of-school and working, it is also important to bear in mind that the direction of causation is not necessarily form child labour to being out-of-school but could also be that children work because they are unable to go to school (Probe report, 1999)[58].

[58] IFLS2 (1997) asked parents why the child stopped going to school or has never been to school. Besides the motive "to help parents" and "could not afford", other reasons were: "child doesn't want to go", "child was not able to study", "child was not accepted". Similarly, the Probe report presents the situation of an eight-year old boy dropped out in class one because of being teased and beaten by other children; he spends the day working.

Table 4.9
Primary activity during the week prior to the interview, out of school children (aged 10-14 years) – *weighted percentages*

	Male	*Female*	*Male and female*
Working[59]	37.1*	26.6*	31.6
Job searching	16.5	11.5	13.9
Housekeeping	2.7*	20.3*	12.0
Other	43.7	41.6	42.5
Total	*100*	*100*	*100*
Number of children	*143*	*164*	*307*

The table shows the primary activity during the week prior to the interview of out of school children. Percentages are computed using individual sampling weights and on the sub-sample of children of 10-14 years old that are not at school but have at some time been to school.
* the difference in the weighted proportions is statistically significant

Considering children who have never enrolled in school, the 1993 Indonesian data show that nearly 4% of males and 5% of females aged 7-17 years had never been to school (the gender gap is significant at the 10% but not at the 5% level), with significant differences between poor and non-poor households (table 4.10). These findings may partly reflect the fact that children may start school after age seven and hence the same data are presented for the 10-17 age group. The proportion of "never-enrolled" boys substantially decreases, in particular for poor households, suggesting that many poor households send their boys to school late. The proportion of "never-enrolled" girls decreases less than that of boys, and the gender gap becomes larger. There are significant disparities between poor and non-poor households for both males and females.

Table 4.10
Proportion of "never enrolled" children, by PCC tercile and gender *(weighted %)*

	Age 7-17		*Age 10-17*	
	Male	*Female*	*Male*	*Female*
Lower third	8.4	10.1	3.8	8.1
	(0.010)	(0.010)	(0.009)	(0.015)
Middle third	2.7	3.9	1.5	2.3
	(0.007)	(0.008)	(0.006)	(0.008)
Upper third	1.6	2.5	0.3	1.6
	(0.005)	(0.006)	(0.003)	(0.006)
All	*4.3*	*5.5*	*1.8*	*4.0*
	(0.005)	*(0.006)*	*(0.004)*	*(0.006)*

The table reports the percentages of children that have never been to school. Percentages are computed using individual sampling weights. Standard errors in parenthesis.

[59] In Indonesia the minimum legal working age is 15, but the ILO convention was only ratified in 1999 (Fitzsimons, 2003).

"Never-enrolled" children, aged ten or more years, may be involved in the labour force (nearly 30%), may be looking for a job or doing housework (25%), or may have other activities (45% report "other" as primary activity during the week prior to the interview). For households in the lower third of the PCC distribution, the proportion of "never-enrolled" children that work increases to 40%, while adding those that are looking for a job or doing housework, the proportion rises to 63%. As noted in table 4.9, children that have never been to school are not necessarily actively involved in work, including domestic work, and the need for child labour appears higher for poor than for rich households.

4.2.3 Data - Infrastructural facilities: school availability and school quality

An important issue in determining school attendance is the availability of school facilities. Educational policy in Indonesia has focused on the provision of basic elementary education for all, implementation of efforts to increase access to primary education together with improvement of education quality. According to IFLS data, in 1993 almost all rural villages had at least one elementary school (99%), and many of them (83%) had more than one school. Lack of schooling infrastructure emerges as a problem at the secondary level: in 1993 nearly 65% and 85% of villages did not have a junior or a senior secondary school respectively, and very few villages had more than one such school. Poor infrastructure increases the cost of sending a child to secondary school, mainly as the result of transportation costs and the cost of room rental and meals. Moreover, parents may be reluctant to send their daughters to school outside the village (as suggested by the Probe report, 1999, for India), reducing the possibility for girls to increase their educational levels.

Table 4.11
Number of schools in the villages

	Weighted mean	*Standard Error*
# of primary schools	2.99	0.15
Ratio of good primary schools	0.49	0.04
# of junior secondary schools	0.50	0.08
Ratio of good junior secondary schools	0.40	0.06
# of senior secondary schools	0.19	0.05
Ratio of good senior secondary schools	0.42	0.11
# of villages with at least one elementary school	*130*	
# of villages with at least one junior secondary school	*54*	
# of villages with at least one senior secondary school	*24*	
# of villages	*132*	

The table shows the number of schools in the village. Descriptive statistics are computed using the village sampling weights.

Even when school facilities are available at a convenient distance, their quality may be inadequate. As reported in table 4.11, nearly half of the elementary schools are considered of good quality by the head of the village, while the proportion of good schools decreases to about 40% at the junior and senior secondary levels.

4.2.4 Data - Educational expenditure

Several studies indicate that the costs of education play a major role in determining the choice of poor families to withdraw children from school. In Indonesia, basic public school is provided free, but this does not imply that parents bear no educational expenditures. Moreover, as noted above, the cost of attending secondary school is greater than that of primary school because many children will be obliged to go to school outside the home village. Table 4.12 reports the weighted averages of educational expenditure, distinguishing between expenditures for children inside and outside the household. Households devote nearly 12% of their total expenditure to education, with significant differences according to PCC terciles (the share is 7% for poor households, and 14% for rich households).

Table 4.12
Educational expenditure

	Weighted mean	*Standard Error*
Total annual educational expenditure	289.7	17.5
Annual educational expenditure for children inside the household	220.5	11.7
Annual educational expenditure for children outside the household	69.8	12.2
Share on total annual expenditure (by PCC tercile) (%)		
Lower third	7.0	0.004
Middle third	8.0	0.004
Upper third	14.0	0.006
Average	12.0	0.003
# of households	*2339*	

The table reports weighted means and standard errors of educational expenditure. Household sampling weights are used. Annual educational expenditures are in thousands of rupiah.

In the empirical strategy we will analyze in detail educational expenditure, estimating the cost of attending each level of school.

This section emphasized the low continuation rate after the graduation from the elementary level. Enrollment rates in junior secondary school are low, with large differences between poor and rich households. The theoretical and empirical model will therefore focus on the schooling decisions made at the transition points, and in particular on school progression at the point of the graduation from the primary level.

4.3 The theoretical model

The model developed in this section aims to analyze the determinants of the school progression in rural Indonesia, and in particular the role played by ex ante risk and ex post shocks. Consider a household in which there are one parent and one child (Basu and Van, 1998).[60] The parent earns an exogenous income (y_t , with t=1,2,3), which is not known by the household prior to the period in question, and decides on the child's time use. We focus on the schooling decisions of children who have graduated from the elementary level. We can divide children's life time into three periods: t=1,2,3. The first two periods are the child schooling phase (after the completion of the primary level), where 1 and 2 are respectively the junior and senior secondary school (each of them lasts three years). In the third period the child has become an adult, and his human capital depends on the decisions made in the schooling phase. As suggested in the paragraph 4.2, many children stop attending school after the completion of the primary level, and this makes the analysis of progression choices an important issue. Hence, this model focuses on the decisions the household faces at the transition points: continuation from elementary to junior secondary school (decision point t=1, at the beginning of period 1), and, conditional on an affirmative choice in that case, continuation from junior to senior secondary school (decision point t=2, at the beginning of period 2)[61].

If the child is enrolled in school, the parent pays an annual cost of schooling defined as p for junior secondary school and $q > p$ for senior school. If the child does not attend school, he may work and earn an income which depends on the highest education level completed (as well as on experience): w_t is the income earned at time t if the highest level is elementary school, x_t and z_t are the incomes of junior and senior secondary graduates respectively. The key assumption of this model is the irreversibility of the decision to

[60] This may be simply a convention, we can think of the two parents as "one parent" and the children as "one child" (Basu and Van, 1998). In the empirical analysis, we will consider more than one child per household.

[61] Sawada and Lokshin (2007) analyze the school progression in rural Pakistan. They develop a theoretical framework to model the sequential schooling decisions.

withdraw the child from school: once the child drops out, he cannot return to school. This implies that if the child stops attending school at the elementary level, this will be the highest grade completed at the end of the schooling phase, and for the remainder of his working life (period three), and hence he will earn a lower income than that he would have earned by continuing to attend school ($x_3 > w_3$ if the child has stopped after the junior school and $z_3 > x_3 > w_3$ if the child has continued to senior school).

To summarize the outcomes (all discounted) are:

	Parent's income			Child's income			School Cost	
	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$	$t = 3$	$t = 1$	$t = 2$
only elementary school	y_1	$\tilde{y}_2$	$\tilde{y}_3$	w_1	w_2	w_3	0	0
elementary + junior secondary school	y_1	$\tilde{y}_2$	$\tilde{y}_3$	0	$x_2 > w_2$	$x_3 > w_3$	p	0
junior + senior secondary school	y_1	$\tilde{y}_2$	$\tilde{y}_3$	0	0	$(a + \tilde{z}_3) > x_3$	p	$q > p$

The tilde indicates a variable which is not known by the household prior to the period in question. We assume that adult incomes are random as is the child's adult income after graduating from junior secondary school. School costs and incomes for those who have not completed junior secondary school are modelled as fixed. The outcomes assume that the child's income and the cost of schooling are certain. These assumptions are made to simplify the analysis. The earnings of a senior secondary school graduate contains a term a which depends on ability. Ability is observed by the family when it makes education decisions but not by the econometrician, who is obliged to treat it as random, with distribution function F(a). Ability enters as a determinant of the adult income only if the child is graduated from the senior secondary school, i.e. we assume that ability affects productivity only at high levels of human capital.

We suppose that household decisions can be characterized by a standard strictly concave utility function $u(\cdot)$ which has total net household income as its argument. Following Jacoby and Skoufias (1997) and Fitzsimons (2007) we do not consider child leisure or child education directly in the utility function. Parents care about a child's schooling as it contributes to household income. We suppose that the child continues to be a member of the household even when he becomes an adult, implying that the sum of adult's (ex child) and parent's incomes enters in the utility function, and that the source of income does not matter. This assumes homogeneous preferences within the household (or that there is a benevolent dictator). A further assumption of the model is that there are no savings and borrowings (De Janvry *et al.*, 2006)[62]. This implies that utility depends only on current income and costs. Our assumption is more reasonable than the polar alternative of full equalization of expected marginal utility over time (Fitzsimons, 2007). Intermediate cases are more difficult to analyze.

The problem is solved as a dynamic programme. The household first evaluates the expected decision to attend senior secondary school and then moves back to the earlier decision with respect to junior secondary schooling conditional on the anticipated senior secondary school decision.

Consider first the decision to continue to senior secondary school conditional on having completed both elementary and junior secondary school. The household will do this if

$$u\left(y_2-q\right)+\delta Eu\left(\tilde{y}_3+\tilde{z}_3+a\right)\geq u\left(y_2+x_2\right)+\delta Eu\left(\tilde{y}_3+x_3\right) \qquad (4.1)$$

where δ is the discount rate from period 3 to period 2. Note that y_2 is known at the time this decision is made.

For simplicity, suppose that income is stationary[63] so $E_s y_t=\overline{y}$ for all $s < t$ and define $\overline{u}=u\left(\overline{y}\right)$. Similarly suppose $E_s\tilde{z}_3=0$ for s=1,2. Write the income deviations as $y_t=\left(1+\eta_t\right)\overline{y}$ and $z_3=\left(1+\zeta_3\right)a=\lambda(1+\zeta_3)\overline{y}$. Assume $\begin{pmatrix}\eta_3\\ \zeta_3\end{pmatrix}\sim N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & \rho\theta\\ \rho\theta & \theta^2\end{pmatrix}\sigma^2\right)$ and $\eta_2\sim N(0,\sigma^2)$, and suppose that η_3 and ζ_3 are independent of η_2. Note that $u'\left(y_2\right)\simeq\overline{u}'+\eta_2\overline{yu}''=\left(1-r\eta_2\right)\overline{u}'$ where $r=-\frac{\overline{yu}''}{\overline{u}'}$, the coefficient of relative risk aversion evaluated at $y=\overline{y}$. On this basis, we can expand the left hand side of equation (4.1) as

$$u\left(y_2-q\right)\simeq\overline{u}+\left[\eta_2\overline{y}-\left(1-r\eta_2\right)q\right]\overline{u}'$$

[62] As noted in chapter two, only the richest households are able to accumulate savings and use them to cope with shocks.

[63] This is consistent with the analysis conducted in chapters two and three, where we distinguish between permanent and transitory income. In a regression we will report in table 4C.4 (appendix 4C), the coefficient on 1993 income in a 1997 income regression is about 0.3, significantly smaller than one. This implies that there is some persistence in income shocks, but not a lot.

$$Eu\left(\tilde{y}_3+\tilde{z}_3+a\right)\simeq\overline{u}+\overline{u}'E\left[\left(\left(\tilde{y}_3-\overline{y}\right)+\tilde{z}_3+a\right)\right]+\tfrac{1}{2}\overline{u}''E\left[\left(\left(\tilde{y}_3-\overline{y}\right)+\tilde{z}_3+a\right)^2\right]$$

and

$$=\overline{u}+a\overline{u}'+\tfrac{1}{2}\left[Var\left(\tilde{y}_3\right)+2Cov\left(\tilde{y}_3,\tilde{z}_3\right)+Var\left(\tilde{z}_3\right)+a^2\right]\overline{u}''$$

$$=\overline{u}+a\overline{u}'-\frac{r}{2}\left[\left(1+2\rho\lambda\theta+\lambda^2\right)\sigma^2+\theta^2\right]\overline{yu}'$$

On this basis, we can expand the left hand side of equation (4.1) as

$$u\left(y_2-q\right)+\delta Eu\left(\tilde{y}_3+\tilde{z}_3+a\right)\simeq\overline{u}+\left[\eta_2\overline{y}-\left(1-r\eta_2\right)q\right]\overline{u}'+\delta\left[\overline{u}+a\overline{u}'\right]$$
$$-\frac{\delta r}{2}\left[\left(1+2\rho\lambda\theta+\lambda^2\right)\sigma^2+\theta^2\right]\overline{yu}'$$

Similarly, the right hand side of equation (4.1) is

$$u\left(y_2+x_2\right)+\delta Eu\left(\tilde{y}_3+x_3\right)\simeq\overline{u}+\overline{u}'\left[\eta_2\overline{y}+\left(1-r\eta_2\right)x_2\right]+\delta\left(\overline{u}+x_3\overline{u}'\right)-\tfrac{1}{2}\delta r\overline{yu}'\sigma^2.$$

It follows that the child will continue to senior school if

$$\delta\left(a-x_3\right)\geq\left(1-r\eta_2\right)\left(q+x_2\right)+\tfrac{1}{2}\delta r\left[\left(2\rho\theta+\lambda\right)\lambda\sigma^2+\theta^2\right]\overline{y}. \qquad (4.2)$$

The left hand side of equation (4.2) is the discounted anticipated benefit from attending senior secondary school, i.e. net additional income. The right hand side contains two terms: the first is the financial cost of attending senior secondary school, adjusted for any deviation of period 2 marginal utility from its average level. The second term is the money equivalent of the utility cost of the uncertainty associated with adult (period 3) income.

Suppose the population distribution of ability is distributed with mean zero and variance ω^2[64]. The probability Q that the child will attend senior secondary school, as evaluated by the econometrician, is then

$$Q=\Pr\left[\delta\left[\left(a-\overline{z}\right)+\left(\overline{z}-\overline{x}_3\right)\right]\geq\left(1-r\eta_2\right)\left(q+x_2\right)+\tfrac{1}{2}\delta r\left[\left(2\rho\theta+\lambda\right)\lambda\sigma^2+\theta^2\right]\overline{y}\right]$$
$$=F\left(\frac{\left(\overline{z}-\overline{x}_3\right)-\frac{1-r\eta_2}{\delta}\left(q+x_2\right)-\tfrac{1}{2}r\left[\left(2\rho\theta+\lambda\right)\lambda\sigma^2+\theta^2\right]\overline{y}}{\omega}\right). \qquad (4.3)$$

64 The distribution of abilities admits the possibility that *a<0*. This problem can be overcome by truncating the ability distribution at zero, or some fixed positive number. Truncation results in the introduction of an additional term in the denominator of the density function, required to ensure that the truncated distribution integrates to unity. The estimates reported in this chapter ignore this complication which in practice is likely to make very little difference.

where F(.) is the distribution function of $\left(\frac{a}{\omega}\right)$ and $\bar{z}$ is the mean of the ability variable a across the population. Equation (4.3) shows that the probability of a child attending senior secondary school, as evaluated by the econometrician, depends mainly on four factors:

a) Positively on the difference in expected earnings $(a - x_3)$ from attending senior secondary school.

b) Negatively on the full cost $(q + x_2)$ of attending senior secondary school.

c) Positively on the proportional deviation η_2 of the household's period 2 income from its normal level.

d) Negatively on the variance $\theta^2\sigma^2$ of adult earnings of senior secondary school graduates relative to their own expectations.

Now go back one period and consider the decision to attend junior secondary school. If the child is withdrawn from school at this stage, family utility is

$$\begin{aligned} &u(y_1 + w_1) + \gamma E_2 u(\tilde{y}_2 + w_2) + \gamma\delta E_2 u(\tilde{y}_3 + w_3) \\ &\quad \simeq (1+\gamma+\gamma\delta)\bar{u} + \left[\eta_1 \bar{y} + (1 - r\eta_1) w_1 + \gamma w_2 + \gamma\delta w_3 - \tfrac{1}{2} r(\gamma+\gamma\delta)\sigma^2 \bar{y}\right]\bar{u}' \end{aligned} \tag{4.4}$$

Instead, if the child continues at school, family utility is

$$\begin{aligned} &u(y_1 - p) + E_2 \max\left[\gamma u(\tilde{y}_2 + x_2) + \gamma\delta u(\tilde{y}_3 + x_3), \gamma u(\tilde{y}_2 - q) + \gamma\delta u(\tilde{y}_3 + \tilde{z}_3 + a)\right] \\ &\quad = u(y_1 - p) + E_2\left[\gamma u(\tilde{y}_2 + x_2) + \gamma\delta u(\tilde{y}_3 + x_3)\right] \\ &\quad + E_2 \max\left[\gamma\{u(\tilde{y}_2 - q) - u(\tilde{y}_2 + x_2)\} + \gamma\delta\{u(\tilde{y}_3 + \tilde{z}_3 + a) - u(\tilde{y}_3 + x_3)\}, 0\right] \end{aligned} \tag{4.5}$$

The first set of terms on the right hand side of equation (4.5) show the benefit of attending junior secondary school. The final term gives the value of the option of continuing to senior secondary school. We may approximate

$$\begin{aligned} &u(y_1 - p) + E_2\left[\gamma u(\tilde{y}_2 + x_2) + \gamma\delta u(\tilde{y}_3 + x_3)\right] \\ &\quad = (1+\gamma+\gamma\delta)\bar{u} + \left[\gamma x_2 + \gamma\delta x_3 + \eta_1 \bar{y} - (1 - r\eta_1) p - \tfrac{1}{2} r(\gamma+\gamma\delta)\sigma^2 \bar{y}\right]\bar{u}' \end{aligned} \tag{4.6}$$

Consider this final term:

$$\begin{aligned} &E_2 \max\left[\gamma\{u(\tilde{y}_2 - q) - u(\tilde{y}_2 + x_2)\} + \gamma\delta\{u(\tilde{y}_3 + \tilde{z}_3 + a) - u(\tilde{y}_3 + x_3)\}, 0\right] \\ &\quad \simeq \max\left[\gamma\delta(a - x_3) - \gamma(1 - r\eta_2)(q + x_2) - \tfrac{1}{2}\gamma\delta r\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}, 0\right]\bar{u}' \end{aligned}$$

The two expressions inside the square brackets are equal to zero when

$$\eta_2 = \frac{1}{r} - \frac{\delta(a - x_3)}{r(q + x_2)} + \frac{\delta\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}}{2(q + x_2)} = \eta^* \quad (4.7)$$

η^* is the level of period two income which would leave the household indifferent between allowing the child to continue to senior secondary school or withdrawing. It follows that

$$\begin{aligned}
&E_2 \max\left[\gamma\{u(\tilde{y}_2 - q) - u(\tilde{y}_2 + x_2)\} + \gamma\delta\{u(\tilde{y}_3 + \tilde{z}_3 + a) - u(\tilde{y}_3 + x_3)\}, 0\right] \\
&= \int_{\eta^*}^{\infty} \left\{\gamma\delta(a - x_3) - \gamma(1 - r\eta_2)(q + x_2) - \tfrac{1}{2}\gamma\delta r\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}\right\}\varphi(\eta_2)\, d\eta_2 \\
&= \left[\gamma\delta(a - x_3) - \gamma(q + x_2) - \tfrac{1}{2}\gamma\delta r\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}\right]\left(1 - \Phi\left(\frac{\eta^*}{\sigma}\right)\right) + \gamma r\sigma(q + x_2)\varphi\left(\frac{\eta^*}{\sigma}\right) \\
&= \left[\gamma\delta(a - x_3) - \gamma\left(1 - r\sigma\Lambda\left(\frac{\eta^*}{\sigma}\right)\right)(q + x_2) - \tfrac{1}{2}\gamma\delta r\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}\right]\left(1 - \Phi\left(\frac{\eta^*}{\sigma}\right)\right)
\end{aligned} \quad (4.8)$$

where $\Lambda(.) = \frac{\varphi(.)}{1 - \Phi(.)}$, the inverse Mills ratio, and $\varphi(.)$ and $\Phi(.)$ are the standard normal density and distribution functions respectively[65]. Combining equations (4.4), (4.5), (4.6), (4.7) and (4.8), the child will attend junior secondary school provided

$$\begin{aligned}
&\left[\gamma x_2 + \gamma\delta x_3 - (1 - r\eta_1)p\right] \\
&\quad + \left[\gamma\delta(a - x_3) - \gamma\left(1 - r\sigma\Lambda\left(\frac{\eta^*}{\sigma}\right)\right)(q + x_2) - \tfrac{1}{2}\gamma\delta r\left[(2\rho\theta + \lambda)\lambda\sigma^2 + \theta^2\right]\bar{y}\right]\left(1 - \Phi\left(\frac{\eta^*}{\sigma}\right)\right) \\
&\geq \left[(1 - r\eta_1)w_1 + \gamma w_2 + \gamma\delta w_3\right]
\end{aligned} \quad (4.9)$$

The left hand side of equation (4.9) contains two terms. The first is the net benefit from going to junior school if the child will not continue to senior school. Let us analyze the second term, which is multiplied by the probability $\left(1 - \Phi\left(\frac{\eta^*}{\sigma}\right)\right)$ of continuation, in greater detail. This represents the discounted net benefit of continuing to senior secondary school conditional of having completed junior secondary school. $\gamma\delta(a - x_3)$ and $\gamma(q + x_2)$ are respectively the benefit and the cost of attending senior secondary school.

[65] This uses the result that, for $\eta_2 \sim N(0, \sigma^2)$, $\int_{\eta^*}^{\infty} \eta\varphi\left(\frac{\eta}{\sigma}\right)d\sigma = \sigma\varphi\left(\frac{\eta^*}{\sigma}\right)$ (Maddala, 1983, p.365).

$\left(-\gamma r\sigma\Lambda\left(\frac{\eta^*}{\sigma}\right)(q+x_2)\right)$ is the expected reduction in the period two utility cost of sending the child to senior school in the event of a positive period two income shock (which reduces period two marginal utility)[66]. $\frac{1}{2}\gamma\delta r\left[(2\rho\theta+\lambda)\lambda\sigma^2+\theta^2\right]\bar{y}$ is the utility cost of the period three income uncertainty associated with adult income. The right hand side of equation (4.9) is the wage associated with having completed only the elementary level.

Rearranging equation (4.9) in terms of ability, we find that the child attends junior high school provided

$$\begin{aligned} a \geq \frac{1-r\eta_1}{\gamma\delta}(p+w_1) - \frac{1}{\delta}\left[\Phi\left(\frac{\eta^*}{\sigma}\right)x_2 - w_2 - \left(1-\Phi\left(\frac{\eta^*}{\sigma}\right)\right)q\right] - r\sigma\varphi\left(\frac{\eta^*}{\sigma}\right)(q+x_2) \\ -\left[\Phi\left(\frac{\eta^*}{\sigma}\right)x_3 - w_3\right] + \tfrac{1}{2}r\left(1-\Phi\left(\frac{\eta^*}{\sigma}\right)\right)\left[(2\rho\theta+\lambda)\lambda\sigma^2+\theta^2\right]\bar{y} = a^* \end{aligned} \tag{4.10}$$

It follows that the probability of attending junior high school is

$$P = \Pr\left(a \geq a^*\right) = 1 - F\left(\frac{a^*}{\omega}\right) \tag{4.11}$$

Note that equation (4.10) cannot be solved analytically since η^* depends on a through equation (4.7). Ignoring this dependency, we conclude that the probability of the child attending junior high school depends on eight factors:

a) Positively on the period 2 difference in earnings $(x_2 - w_2)$ from attending junior secondary school, multiplied by the probability $\Phi(\eta^*)$ of non-continuation.

b) Positively on the period 3 difference in earnings $(x_3 - w_3)$ from attending junior secondary school, multiplied by the probability $\Phi(\eta^*)$ of non-continuation.

c) Positively on the period 3 difference in expected earnings $(a - w_3)$ from attending both junior and senior secondary school, multiplied by the probability $1-\Phi(\eta^*)$ of continuation.

---

[66] Note that $\frac{\partial}{\partial\sigma}\sigma\varphi\left(\frac{\eta^*}{\sigma}\right) > 0$. Proof:

$$\frac{\partial}{\partial\sigma}\sigma\varphi\left(\frac{\eta^*}{\sigma}\right) = \varphi\left(\frac{\eta^*}{\sigma}\right) - \sigma\varphi'\left(\frac{\eta^*}{\sigma}\right)\frac{\eta^*}{\sigma^2} = \varphi\left(\frac{\eta^*}{\sigma}\right) + \left(\frac{\eta^*}{\sigma}\right)^2\varphi\left(\frac{\eta^*}{\sigma}\right) = \varphi\left(\frac{\eta^*}{\sigma}\right)\left[1+\left(\frac{\eta^*}{\sigma}\right)^2\right] > 0$$

d) Negatively on the full cost $\left(p+w_1\right)$ of attending junior secondary school.

e) Negatively on the full cost $\left(q+x_2\right)$ of attending senior secondary school multiplied by the probability $1-\Phi\left(\eta^*\right)$ of continuation to senior secondary school.

f) Positively on the proportional deviation η_1 of the household's period 1 income from its normal level.

g) Positively on the standard deviation of the family's period 2 earnings through the option of continuation to senior secondary school.

h) Negatively on the variance $\theta^2\sigma^2$ of adult earnings of senior secondary school graduates relative to their own expectations.

A trivial extension of the model allows the household permanent income to be a determinant of the earnings of a senior secondary school graduates. We suggest that households with high permanent income are likely to be well-placed to find well paid jobs for their educated children, but cannot do much for drop-out children. This allows us to replace $\tilde{z}_3$ with $\tilde{z}_3+\kappa\overline{y}_3$ for some $\kappa>0$. In this way the probability of attending senior secondary school depends positively on household permanent income, and the probability of attending junior secondary school is positively related to the permanent income interacted with the probability of continuation.

To summarize, the model considers two sources of income uncertainty. The first is the household's uncertainty with respect to the earnings of the child once graduated from the senior secondary school. What matters here is the variability of income of senior secondary graduates when they become adults (period three). Under the assumption of a concave utility function (i.e. risk aversion), an increase in uncertainty of future earnings decreases the expected adult utility (benefits) from senior secondary school. This uncertainty has hence a negative impact on the decision to continue to senior secondary school, and, consequently, to junior secondary school. The second source of income variability is the household's uncertainty with respect to its income over the time the child may remain at school, hence for period two. Under the assumption of irreversibility, the probability of continuing to junior secondary school depends positively on the variance of period two household income, as seen by the household, through the option of continuing to senior secondary school in the event of a positive period two income shock.

Finally, this model shows the impact of shocks ex post. Under the assumption that there are no savings and borrowings, a negative income shock increases current period marginal utility, and hence the utility cost of schooling rises. Since there is no implication for future parental income, future marginal utility stays the same so the expected utility benefit of education is unaffected. Thus a negative shock decreases the probability of attending school (and the opposite holds for positive shock). Shocks are the ex post counterpart of ex ante uncertainty. It is the possibility of positive shocks, making education more affordable, that drives the positive effect of the variance of household income over the time the child will remain at school. We analyze both shock responses ex ante, to the anticipation of shocks, and ex post, to their realizations.

4.4 Empirical strategy

This section presents the empirical methodology used to estimate the school progression of children, based on the predictions of the theoretical model. In order to implement the model, we need information on schooling costs, wage levels by educational attainment, a measure of income shocks, and estimates of income uncertainty. The following sections aim to construct these variables, showing the methodology used and outlining the main results. The empirical model for schooling decisions is then discussed.

4.4.1 Cost of schooling

Total educational expenditure can be written as $E_h = p_{h1}N_{h1} + p_{h2}N_{h2} + p_{h3}N_{i3}$, where p_{hj} is the cost of the school level *j* for the household h, and N_{hj} is the number of children in the household attending education level *j*. Expenditure for the different levels of schooling is not directly observable but may be inferred from regressing total expenditure for education on the number of children attending each school level. The regression also controls for household income and province. We use the following model which is additive in a set of three multiplicative components, one for each level of schooling:

$$E_h = y_h^{\eta} e^{\gamma' P_h} \sum_{j=1}^{3} \beta_j N_{hj} \tag{4.12}$$

where E_h is the total educational expenditure of family *h*, y_h is household income, *P* are province dummies, N_{hj} is the number of children attending school level *j*, and β_j is the coefficient associated with N_{hj}. The cost of sending an additional child to school level *j* is given by $y_h^{\hat{\eta}} e^{\hat{\gamma}' P_h} \hat{\beta}_j$. Assuming constant economies of scale, that is assuming linearity of educational expenditure on the number of children attending school[67], the marginal effect is also the average effect, and is our measure of the cost of schooling.

As already noted, equation (4.12) includes household income, thus allowing the cost of schooling to vary across households. According to this formulation, households estimate the cost of the various levels of schooling as their expected expenditures given their incomes. Alternatively, one might view income-related expenditures as discretionary, implying that these expenditure components should be omitted from the cost variables. If this were the case, the estimated cost would become $\overline{y}^{\hat{\eta}} e^{\hat{\gamma}' P_h} \hat{\beta}_j$, where $\overline{y}$ is the sample average income. We construct both estimates of the cost of schooling: the first conditional on household income, the second conditional on average income.

Household income and the number of children attending school may both be endogenous. Potential instruments for household income are the education of the head and the value of non-business assets owned by the household the year before the interview. The numbers of children attending each school level can be instrumented with the predicted count variables estimated using Poisson regression models on the sub-sample of households with at least one child in the reference age group (children aged 7-12 for the elementary school, 13-15 for junior secondary school, and 16-18 for senior secondary school). Table 4.13 reports the number (and percentages) of households reporting different number of children (boys and girls) attending elementary, junior and senior secondary school for each sub-sample. The majority of households with at least one child in the secondary school age group do not have children attending school (53% and 74% of households with at least one boy in the reference age group do not send their male children to junior and senior secondary school respectively; percentages of households with zero girls attending secondary school are 61% and 79%). This confirms what was pointed out in section 4.2, that there is a low continuation rate after the completion of the elementary school, and more generally, that the largest loss in the number of children attending school occurs after the elementary compulsory level.

[67] The square number of children attending school is not statistically significant in the educational expenditure equation, thus rejecting the hypothesis of a quadratic relationship between expenditure and the number of children.

Table 4.13
Number of children in the household attending school, by gender and school level

	Elementary school				*Junior secondary school*				*Senior secondary school*			
	# boys		# girls		# boys		# girls		# boys		# girls	
	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%
0	98	9.2	109	9.9	312	53.3	340	60.6	322	74.4	322	78.7
1	746	70.1	770	69.9	251	42.9	204	36.4	101	23.3	77	18.8
2	197	18.5	191	17.3	17	2.9	16	2.8	9	2.1	10	2.4
3	22	2.1	31	2.8	2	0.3	1	0.2	-	-	-	-
4	2	0.2	-	-	1	0.2	-	-	-	-	-	-
>4	-	-	-	-	2	0.3	-	-	1	0.2	-	-
N[a]	*1,065*	*100*	*1101*	*100*	*585*	*100*	*561*	*100*	*433*	*100*	*409*	*100*

The table reports the number (and percentages) of households reporting different number of children (boys and girls) that attend each school level.
[a] Number of households with at least one child (boys or girls) in the reference age group (i.e. there are 1065 households with at least one boy in the age group 7-12)

Because of the high percentage of zeroes in the number of children enrolled in secondary school, the count variables are estimated with the zero-inflated Poisson model (ZIP)[68]. The results of Poisson and ZIP models are reported in Appendix 4A (tables 4A.1 and 4A.2). The main results are summarized as follow.

Household socio-economic variables. Household permanent income is statistically significant only for the number of girls attending junior school[69]. Indicators for the status of the house (whether the household utilizes electricity, owns a private toilet, and whether the main water source is located inside the house), have a positive and significant effect on the number of young adults attending senior secondary school. Ownership of a farm or non farm business may increase child work participation (Cockburn and Dostie, 2007), and hence decrease the likelihood that a child is in school. To control for that, dummies for having a farm or non-farm business are included in the count models. Results show that having a farm has a negative and significant impact only on the number of girls attending elementary school.

[68] The ZIP model responds to the failure of the Poisson model to account for excess zeros by changing the mean structure to allow zeros to be generated by two distinct processes. The ZIP assumes that there are two latent groups. An individual in the Always-0 Group (group A) has an outcome of 0 with a probability of 1, while an individual in the Not Always-0 Group (group B) might have a zero count, but there is a nonzero probability of a positive outcome. The membership in group A is modeled through a logit or probit model (Long and Freese, 2001).
[69] Permanent income is estimated regressing household income on a set of variables that determine the permanent component, as discussed in chapters two and three. The wage of children in the reference age group are subtracted from household income.

Household characteristics. The role of household composition is analyzed including the number of children and young adults in the household; to capture possible different impacts, children are divided in four age categories, corresponding to the pre-school period (age 0-6) and to three school levels (age 7-12, age 13-15, and age 16-18). The number of males and females in the reference age group is clearly highly significant in the count model for males and females respectively; the number of older children in general has a positive and significant effect in the count models.

Finally, dummies for the religion of the household is included as additional regressors: living in a Christian family increases the number of household members attending the senior school.

Head and spouse education. The number of boys and girls attending junior secondary school are positively related to the education of the head and the spouse of the head respectively. The number of girls attending the elementary level negatively depends on having a household head that is illiterate or that completed only the elementary level. Educational dummies are not statistically significant in the count model for senior school.

As suggested in section 4.2, lack of school infrastructure emerges as a problem at the junior and senior secondary education levels. The numbers of elementary, junior and senior schools in the village are included in the count models. They are positive and significant in the count models for the female sample (for elementary and junior school), suggesting that parents may be more reluctant to send girls to school outside the home village (as pointed out by the Probe report, 1999, for India).

Estimates from the Poisson and zero inflated Poisson models are then used as instruments for the number of children attending school in the educational expenditure equation[70]. Table 4A.4 (Appendix 4A) reports the results of equation 4.12, and table 4A.5 summarizes the estimated costs of schooling, by level of school and gender. Both set of estimates, conditional on household and on average income, are presented. As stated in section 4.2, elementary school is provided free (there are no school fees), but this does not imply that parents bear no educational expenditures. As expected, schooling costs are higher for higher education levels (there are school fees at junior and senior levels and many children are

[70] Table 4A.3 reports the actual and predicted number of children attending school. Both the Poisson and the zero-inflated Poisson (ZIP) estimates are reported for junior and senior secondary school. The Vuong test of ZIP versus standard Poisson supports the ZIP model for senior secondary, but not for junior secondary school. Estimates from the ZIP models are used as instruments for the number of children attending junior and senior secondary.

obliged to go to school outside the home village, as noted in section 4.2). Estimates show a higher cost for girls than for boys, especially for secondary school.

4.4.2 Income levels by education

If the child does not go to school, he may work and earn an income which depends on his education and experience. As suggested in the theoretical section, incomes may be classified by level of school completed and by experience as follows:

Grade completed	*Junior secondary (3 years)*	*Senior secondary (3 years)*	*Rest of the life (37-42 years)*
Elementary Years of schooling >=6 & <9	w_1 Experience <= 3yrs	w_2 3<Experience <= 6yrs	w_3 Experience > 6yrs
Junior secondary Years of schooling >=9 & <12	-	x_2 Experience <= 3yrs	x_3 Experience > 3yrs
Senior secondary Years of schooling >=12 & <15	-	-	z_3

Children who have completed only the elementary school will earn w_1 and w_2 respectively for the three years corresponding to junior and senior secondary school, while the income earned for the rest of the life is w_3. Assuming 55-60 years as the retirement age (Leechor, 1996), and in view of the fact that the official age at which children should complete primary school is 12 years, the length of "rest of life" is approximately 37-42 years. Similarly, children who have graduated from junior secondary school will earn an income defined as x_2 and x_3 respectively for the first three years of work (the time they would otherwise have spent at senior school) and for the remainder of their working life.

Predicted child incomes follow from an income regression for all household members older than 10 years that report work information[71]. We run separate regressions for males and females, including educational dummies to estimate returns to schooling, and allowing the return to education to vary across provinces. The income equation can be written as:

[71] Adult household members (the household head and the spouse, if present, are selected, and up to three other members) were randomly selected and asked to provide detailed individual employment information, on both farm and off-farm work (Book 3). For all the other household members not interviewed in Book 3 and older than 10 years, the head reports whether the individual works, which type of work he/she does, and the total net wages (other than from the farm or non-farm business). Appendix A shows the survey questions used to construct individual labour income in detail.

$$\ln y_i = \alpha_0 + \alpha_1 esper_i + \alpha_2 esper_i^2 + \sum_{j=1}^{J}\sum_{p=1}^{P} \beta_{jp} D_{ji} * P_{pi} \qquad (4.13)$$

where the dependent variable is log monthly income, *esper* and *esper*^{2} is experience and experience square, D_{ji} are educational dummies, and P_{pi} are provincial dummies. In the absence of information on actual experience, *esper* is replaced by "potential experience", measured as $age_i - schooling\ years_i - 7$, assuming people start school at the age of 7.

Equation (4.13) is first estimated without the province interaction terms and including all the dummies to identify w_1, w_2, w_3, x_2, x_3 and z_3 (the Heckman procedure is used to control for sample selection). The income equation is defined as $\ln y_i = \alpha_0 + \sum_j^J \alpha_j D_{ji} + \alpha_{J+1} esper_i + \alpha_{J+2} esper_i^2$, where D_{ji} are dummies to identify the educational-experience categories. The sample is all household members older than age 10 that report information on labour income. Since women and men may face different labour supply decisions, as well as different returns to schooling, we run the model separately for males and females. Table 4.14 reports the number of observations by gender and education-experience categories.

Table 4.14
Number of observations by gender and education-experience categories

Grade completed		*Junior secondary*		*Senior secondary*		*Rest of the life*	
Elementary	*Male*	w_1	N=36	w_2	N=58	w_3	N=732
	Female		N=38		N=38		N=255
Junior secondary	*Male*			x_2	N=23	x_3	N=210
	Female				N=12		N=69
Senior secondary	*Male*					z_3	N=299
	Female						N=126

Sample: all household members older than age 10 interviewed in the sections "labour income"

Because of the small number of observations for cells w_1, w_2, and x_2, coefficients on these dummies are not precisely estimated in the income regression. For females, we cannot reject the hypothesis that coefficients on dummies for w_1, w_2, w_3, x_2, x_3 are equal ($F(4, 1426) = 1.41$, $Prob>F = 0.23$). We therefore generate a single dummy for females graduated from elementary and junior secondary school, for any experience level. This suggests that it is only senior secondary school attendance that substantially increases female income. For males we cannot reject the hypotheses that returns to junior school are equal for the first three years of work and for the rest of the life (F(1, 2625) = 0.14, Prob>F=0.71), and that returns to elementary school are the same for the sub-groups w_1 and w_3 (F(1, 2625) = 0.67, Prob>F=0.41). The coefficient on the dummy for w_2 is found to be lower than w_1 and w_3 at the 10% but not at the 5% level. This finding may be not very reliable given the small number of observations. We hence restrict the coefficients on the three dummies for elementary education to be equal for the male sample; a single dummy for junior secondary school graduates is included.

Equation (4.13) is then estimated, where returns to education vary across provinces. Table 4B.1 (Appendix 4B) reports the results of the selection model (Heckman procedure). Dummies for different levels of completed schooling are included, with no schooling acting as the omitted category (i.e. there is no constant term in the regression equation). Off 5982 females and 5534 males, we observe the log monthly income for 1437 and 2636 working individuals (females and males respectively). The large proportions of zeros, especially for the self-employed sector, suggests that monthly income may contain an important seasonal component. To control for this, the selection equation includes the dummies for the month of interview. Results of the selection equation are summarized as follows.

Household characteristics. The number of children (in particular those aged 0-5) and young adults in the household is significantly and negatively correlated with the probability of women being in the labour force. The presence of infants in the household does not significantly affect the probability of observing income for males, while the number of children aged 6-14 years has a negative and significant effect.

Individual characteristics. Married woman are significantly less likely to work than non married woman, while being married increases the probability of working for males. Education plays an important role in determining labour supply decisions, for both men and women, but with a higher magnitude for females (marginal effects of education are higher for females than for males). Experience increases the probability of working, and it has a non-linear effect.

Month of interview. In rural areas income may have a strong seasonality component. Reported monthly income is the income earned the month before the interview. Results show that being interviewed in the last months of the year (September-December) significantly increases the probability of reporting monthly income with respect to being interviewed in January.

Once we have defined the selection equation, the income regression permits calculation of the returns to schooling. Log individual monthly income is regressed upon experience, experience squared, and educational dummies interacted with provinces (table 4B.1). The impact of experience is greater for females than for males, but females have also a higher negative coefficient on experience squared. Returns to education are higher for men than for women, and they are highest in West Java (where Jakarta is located) and (for men) in the province of South Kalimantan.

The coefficients on education and experience estimated from the income regression are used to construct predicted earnings by gender, provinces and school level[72].

4.4.3 Measures of risk and shocks

We have set out to consider both the ex post and ex ante effects of risk on education. The former are related to income shocks, the latter to income variability. An income shock at time t is defined as the difference between the actual household income at time t and the expected household income estimated as permanent income[73]. The following income equation is estimated[74]:

$$y_h = \alpha_0 + \alpha_1 X_h^P + \varepsilon_h , \tag{4.14}$$

where y_h is actual household income and X_h^P is a set of variables that determine the permanent income component. X_h^P includes the value of household assets, dummies for whether the household owns a farm or non-farm business, dummies for the occupation type of the household head, the education of the household head, the number of income earners in the household, household size and household size squared. Provincial dummies are also included as regressors. Fitted residuals are our measure of income shock[75,76]:

[72] We do not differentiate between income of children working on the family farm and income from wage employment. In Indonesia, as in many developing countries, most children do not work for a wage, but as unpaid family workers in the family farm or business. Moreover, in many developing countries, labour markets are imperfect. It is therefore difficult to infer children's contributions to household income by using observed market wages and self-reported farm or business profits (Menon and Perali, 2006). A shadow wage approach may be more suitable in evaluating the income contribution of child labour, and we will consider this approach in future extensions of this work.

[73] Permanent income is lifetime expected income, and may be taken as an appropriate measure of the expected household income

[74] The procedure used to construct permanent income is similar to that proposed in chapters two and three

[75] The use of cross-section data instead of panel data does not resolve problems of unobservables included in the residual.

$$\hat{\varepsilon}_h = y_{ht} - \hat{y}_h \,, \tag{4.15}$$

where $\hat{y}_h$ is the estimated permanent income component. Shocks may be positive and negative[77].

The results of equation (4.14) are reported in table 4C.1 (Appendix 4C). Because the household income includes child income, the measure of shocks are computed using household income minus child wages as dependent variable in (4.14). Columns A and B of table 4C.1 report the results of the income regression without considering the wage if children aged 12-14 years and 15-17 years respectively. Results are in line with the standard income equation estimates. Residuals from column A are used in the probit model for junior secondary school, and residuals from column B are used in the model for senior secondary school.

We measure uncertainty through income variances. There are two income variances in the model: the variance of life-time incomes of senior high school graduates relative to their expectations, which relates to the variance of η_3, and the variance of income over the period that the child remains at school, as seen by the household, which relates to the variance of η_2. The former is estimated following a two stage procedure:

1. Individual adult income is regressed on individual characteristics for the sub-sample of adults who have completed at least the senior secondary school:

$$y_i = \beta_0 + \beta_1 X_i + u_i \tag{4.16}$$

2. The square of predicted residuals $\hat{u}_i$ from this regression is regressed on X_i, X_i^2 and cross products. The fitted value from this regression is the measure of the variance of income of senior high school graduates ($\hat{\Psi}_i^2$).

The estimation outcome is discussed in Appendix 4C: the income variance is conditional on senior secondary school graduation and on provincial, gender and race dummies (as a proxy for the race we use a dummy that equals one if the individual speaks Indonesian at home). The coefficients are estimated on the sample of adults graduated from senior secondary school. They are employed to calculate the income uncertainty of children once they become adults (they are applied to the decisions to attend junior and senior school). The

[76] The model identifies income deviation as $\eta_h = \frac{\hat{\varepsilon}_h}{\hat{y}_h} = \frac{y_{ht} - \hat{y}_h}{\hat{y}_h}$. We consider both measures of shock in the empirical strategy.

[77] Differently from chapter two and three shocks.

assumption here is that the distribution of income of current adults conditional on being a senior secondary graduate (and on provincial, gender and race dummies) is the same as that of children when they become adults. This may be a strong assumption.

The second source of risk is estimated as the variance of the household period two income deviations η_2. We assume that households base expectations of their future income distribution on current information. We hence measure future income variability as the variance of 1997 income conditional on 1993 information, included 1993 income. Note that 1997-1993 is four years (the junior secondary school lasts three years), and is therefore the appropriate horizon over which to look. This justifies conditioning on 1993 household information. As before, a two stage procedure is used.

1. We first regress future household income on present income and household characteristics:

$$y_h^{97} = \beta_0 + \beta_1 y_h^{93} + \beta_2 Z_h^{93} + u_h \qquad (4.17)$$

where y_h^{97} and y_h^{93} are the 1997 and 1993 household income and Z_h^{93} is a set of 1993 household characteristics[78]. OLS residuals are predicted ($\hat{u}_h$).

2. The square of predicted residuals $\hat{u}_h^2$ is regressed on a constant and all first moments, second moments and cross products of the original regressors (as in the White test for heteroscedasticity):

$$\hat{u}_h^2 = \delta_0 + \sum_{i=1}^{k} \sum_{j}^{k} \delta_{ij} Z_{ih}^{93} Z_{jh}^{93} + v_h \, ,$$

where in this notation Z_h^{93} includes the 1993 household income y_h^{93}. The predicted value from stage two ($\hat{\sigma}_h^2$) is the measure of the household specific risk, relative to period two income. Results are reported in appendix 4C.

[78] Z_h^{93} includes the value of assets (non business assets, liquid and illiquid), the household size and the number of schooling years completed by the household head.

4.4.4 Schooling decisions

Sections 4.4.1-4.4.3 defined the empirical strategies used to construct the cost of schooling, wage levels by education, and the measures of risk and shocks. These predicted variables are now used as regressors in the probit models for school attendance[79] (recall that ability is observed by the household but not by the econometrician).

Children and young adults are divided in two cohorts: children who have graduated from elementary school, aged less than 15 years, and junior secondary graduates aged less than 18 years. The progression to junior secondary school is estimated on the first sub-sample, while the continuation to senior school refers to the second cohort. In those cases in which a household has children in both cohorts the same income variable will define η_1 for the child in the younger cohort and η_2 for the child in the older cohort.

The equation for senior school may be written as:

$$s_{iS} = \alpha_0 + \alpha_1 X_i + \alpha_2 (q + x_{i2}) \sum_{t=0}^{2} \delta^t + \alpha_3 (a_i - x_{i3}) \sum_{t=3}^{T-1} \delta^t + \alpha_4 \eta_{h2} + \alpha_5 \Psi_i^2 + \nu_i \qquad (4.18)$$

where s_{iS} is the school attendance at senior secondary school, X_i is a set of children and household characteristics; $(q + x_{i2})$ is the total cost of attending senior secondary school, with q the cost of the senior school and x_{i2} the income the child will earn if he/she does not go to school. $(a_i - x_{i3})$ is the difference in expected earnings from attending senior secondary school. Both expected earnings and cost are discounted (δ is the discount factor), and T is the length of the working life. The deviation of the household's period 2 income from its expected level is defined as η_{h2}, while Ψ_i^2 is the variance of life-time incomes of senior high school graduates. All regressors with the exception of those included in X_i are estimates, consequently coefficient standard errors may be biased (Davidson and MacKinnon, 2004).

The decision to continue with the junior secondary school after the completion of elementary level is analyzed by estimation of two set of models. First we estimate a reduced form equation that includes all the regressors suggested by the model but without interactions and restrictions, and then, subsequently, a restricted version of the model. The reduced form can be written as:

[79] At this point we are taking the distribution function of ability to be normal. Therefore we estimate probit models for school decisions.

$$s_{iJ} = \alpha_0 + \alpha_1 X_i + \alpha_2 (p + w_{i1}) \sum_{t=0}^{2} \delta^t + \alpha_3 (q + x_{i2}) \sum_{t=3}^{5} \delta^t + \alpha_4 (x_{i2} - w_{i2}) \sum_{t=3}^{5} \delta^t$$
$$+ \alpha_5 (x_{i3} - w_{i3}) \sum_{t=6}^{T-1} \delta^t + \alpha_6 (a_i - w_{i3}) \sum_{t=6}^{T-1} \delta^t + \alpha_7 \eta_{h1} + \alpha_8 \sigma_h^2 + \alpha_9 \Psi_i^2 + \alpha_{10} (1 - \Phi_{iS}) + v_h \quad (4.19)$$

where s_{iJ} is the school attendance at junior secondary school, X_i is a set of children and household characteristics; $(p + w_{i1})$ is the total cost of attending junior school and it is the sum of the cost of schooling (p) and the opportunity cost given by child's earnings (w_{i1}). Similarly $(q + x_{i2})$ is the total cost of attending senior secondary school. The differences in earnings from attending junior secondary school are given by $(x_{i2} - w_{i2})$ for period two and by $(x_{i3} - w_{i3})$ for period three; the difference from attending both junior and senior secondary school is defined as $(a_i - w_{i3})$. η_{h1} is the deviation of the household's period one income from its expected level, σ_h^2 is parent's future income variability and Ψ_i^2 is the individual earnings variability of senior secondary school graduates. Finally, Φ_{iS} is the probability of non-continuation to senior secondary school after the completion of the junior level. All regressors except X_i are estimates. Φ_{iS} is the predicted probability implied by the estimated equation for continuation to senior school. This predicted probability is estimated for the sample of children graduated from the junior secondary school, but is applied to the data for children graduated from the elementary school, and hence for the junior secondary decisions.

The restricted model for the junior secondary school includes all interactions with the probability of continuation, or non-continuation, to the senior school, as suggested by the theoretical model.

$$s_{iJ} = \alpha_0 + \alpha_1 X_i + \alpha_2 (p + w_{i1}) \sum_{t=0}^{2} \delta^t + \alpha_3 (q + x_{i2})(1 - \Phi_{iS}) \sum_{t=3}^{5} \delta^t + \alpha_4 (x_{i2} - w_{i2}) \Phi_{iS} \sum_{t=3}^{5} \delta^t$$
$$+ \alpha_5 (x_{i3} - w_{i3}) \Phi_{iS} \sum_{t=6}^{T-1} \delta^t + \alpha_6 (a_i - w_{i3})(1 - \Phi_{iS}) \sum_{t=6}^{T-1} \delta^t + \alpha_7 \eta_{h1} + \alpha_8 \sigma_h^2 + \alpha_9 \Psi_i^2 + v_h$$

(4.20)

4.5 Probit model estimates

Table 4.15, 4.16 and 4.17 present the set of probit models that estimate the probability of attending senior and junior school respectively for the sub-sample of children and young adults that completed the previous school level. To better capture the continuation decisions made at the time of the interview, the samples for the junior and senior school are restricted to children aged 14 years or less, and 17 years or less respectively[80]. Equations (4.18), (4.19) and (4.20) include predicted earning levels by education as regressors. In section 4.4.2 we found that returns to elementary and junior school (that determine w and x respectively) are not statistically different for girls, whereas for boys $x_{i2} = x_{i3}$ and $w_{i2} = w_{i3}$. Hence in equation (4.19) and (4.20) we have a single variable $(x_i - w_i)$ discounted from t=3 to the end of the working life. Difference in earnings and the cost of schooling are discounted using a discount factor of $\delta = 0.95$ per annum (Zimmerman and Carter, 2003).

We first comment on the results of the probit model for senior school decisions (table 4.15). The probability of attending senior school decreases with the number of younger siblings aged 13-15 in the house[81], and if the household head and the spouse are illiterate[82]. As suggested by other papers, we find that the ownership of a farm decreases the likelihood that a child is in school, because this increases the child's work participation (Cockburn and Dostie, 2007). Other control variables are the number of good senior schools in the village (a positive and significant effect) and the religion of the household (children living in a Muslim or Christian household are more likely to attend senior school compared to household with other religions[83]).

As regard the variables predicted by the model, the difference in earnings from attending senior secondary school (a-x_3) contributes to a higher continuation probability, as suggested by the theory. The deviation of the household's income from its expected level ($y_t - \hat{y}_t$) has

[80] Children should start junior and senior secondary school at the age of 13 and 16 respectively. As noted in section 4.2, this chapter focuses on the sample of children 7-17 years old.

[81] Traditionally the first child drops out school or is not enrolled in school to provide education opportunities for the younger children. Our results suggest this may be the case. In future work we propose to introduce a birth order variable.

[82] Having an illiterate spouse in the household has a higher negative marginal effect than having an illiterate head. The education of the head and the spouse is considered instead of that of parents because some children do not report to have a mother or a father in the household. In most cases the head and the spouse are the father and the mother of the child.

[83] "Christian" and "Muslim" have similar marginal effects (0.445 and 0.49 respectively).

a positive effect with z=1.45 (not statistically significant at the 10% level)[84]. The total cost of senior school, the variance of adult earnings of senior school graduates (it has a negative effect, as predicted by the model, and z=-1.00), and the household permanent income, are not statistically significant. The probability of continuation to the senior school is predicted from this probit regression (setting shocks to zero), and used in the probit model for junior secondary school (it is included as additional regressor in the reduced form model, and it is interacted with other variables in the restricted form, as suggested by the model).

Looking at the unrestricted probit model for junior secondary school attendance, the variables included in the theoretical model that are statistically significant are the residual from the income regression, the parent's income variance and the household permanent income. All three have a positive effect on the probability of a child attending junior school. Moreover, the higher is the probability of continuation to senior secondary school, the higher is the probability of attending junior secondary school. Among the other control variables, the number of siblings 13-15 years old and 16-18 years old have a positive and significant effect (the former at the 10% but not at the 5% level). It is interesting to note that at lower levels of schooling, school progression of a child is positively associated with the number of siblings of the same age or older (Sawada and Lokshin, 2007), while at higher levels of schooling (see the results of senior secondary school) the probability that the child attends school decreases with the number of younger siblings.

Table 4.17 presents the restricted probit for the decision to continue to the junior secondary school. The restricted model includes the interaction terms with the probability of continuation or non-continuation to the senior school. According to this specification the total cost of junior school $(p + w_{i1})$ is negative and significant at the 10% but not at the 5% level. The deviation of the household income from its expected level (the predicted residual from the income regression) has a positive and significant effect[85], as does the predicted variance of the household income. The higher is predicted permanent income, interacted with the probability of continuation, the higher becomes the probability of the child

---

[84] Following the notation of the model, we re-estimated the probit model using $\eta_h = \frac{\hat{\varepsilon}_h}{\hat{y}_h} = \frac{y_{ht} - \hat{y}_h}{\hat{y}_h}$ instead of $\hat{\varepsilon}_h$. The estimated coefficient is 0.296, and the z statistic becomes 1.20.

[85] Following the notation of the model, we re-estimated the probit using $\eta_h = \frac{\hat{\varepsilon}_h}{\hat{y}_h} = \frac{y_{ht} - \hat{y}_h}{\hat{y}_h}$ instead of $\hat{\varepsilon}_h$. The income deviation is no more significant, with a coefficient of 0.059 and z=0.74. Estimated coefficient is similar for the unrestricted model.

attending junior secondary school. All the other variables included in the theoretical model have the predicted sign, but are not statistically significant[86]. As in the reduced form, the number of older siblings in the household (aged 16-18) has a positive and significant effect.

Table 4.15
Probit model for senior secondary school

Variables	*coeff*	*z*	*Marginal effect*	*z*
# of children in the house age 0-6	-0.310	-1.68	-0.121	-1.68
# of siblings age 7-12	-0.089	-0.57	-0.035	-0.57
# of siblings age 13-15	-0.800	-3.40	-0.312	-3.40
# of siblings age 16-18	-0.151	-0.55	-0.059	-0.55
Spouse of the head is illiterate*	-1.362	-2.76	-0.500	-3.44
head is illiterate*	-0.786	-1.63	-0.305	-1.75
Difference in earnings (a-x)	0.038	3.46	0.015	3.47
Total cost of senior school/1000	0.057	0.18	0.022	0.18
$y_t - \hat{y}_t$ (excluding income of children age 15-17)	0.184	1.45	0.070	1.46
Variance of adult earnings of senior secondary school graduates	-0.354	-1.00	-0.139	-1.00
Household permanent income (excluding income of children age 15-17)	0.112	0.65	0.041	0.65
household owns a farm*	-0.672	-1.77	-0.248	-1.89
household religion: islam*	1.332	2.12	0.491	2.56
household religion: cristian*	2.101	2.18	0.444	6.71
# good senior school in the village	1.707	2.15	0.667	2.18
Intercept	1.451	1.31	-	
Pseudo R-squared		0.45		
Log pseudo-likelihood		-50.974		
N		135		

The table reports the results of equation (4.18). Dependent variable is a dummy that equals one if the child is attending school. The sample is children graduated from the junior school. Provincial dummies are included as additional regressors. The cost of schooling is conditional on household income. Household income is in millions rupiah.
(*) dy/dx is for discrete change of dummy variable from 0 to 1

[86] The variance of adult earnings interacted with the probability of continuation to senior school is not significant, but the estimated coefficient is positive (estimated coefficient=0.197, z=0.65).

Table 4.16
Probit model for junior secondary school – reduced form

Variables	*Coeff.*	*z*	*Marginal effect*
Difference in earnings (x-w)	-0.003	-0.33	-0.001
Difference in earnings (a-w)	0.014	0.54	0.005
Total cost of junior secondary school (p+w)/1000	0.438	0.08	0.172
Total cost of senior secondary school (q+x)/1000	-0.993	-0.34	-0.388
Residual from income regression ($y_t - \hat{y}_t$) (excluding wage of children 12-14)	0.266	2.68	0.104
Sigma2 (parent's income variance)	0.067	2.53	0.026
Variance of adult earnings of senior secondary school graduates	-0.107	-0.37	-0.042
Estimated permanent income $\hat{y}_t$ (excluding wage of children 12-14)	0.321	2.22	0.125
Probability of continuation to the senior secondary school	0.796	2.62	0.311
# siblings age 13-15	0.362	1.77	0.142
# siblings age 16-18	0.255	1.91	0.100
intercept	0.579	0.84	-
Pseudo R-squared		0.18	
Log pseudo-likelihood		-175.383	
N		311	

The table reports the results of equation (4.19). Dependent variable is a dummy that equals one if the child is attending school. Sample: children graduated from the elementary school, aged 12-14 years. Provincial dummies are included as additional regressors. The cost of schooling is conditional on household income. Household income is in millions rupiah. Income variance is divided by one million. Whether the head and the spouse of the head are illiterate, the number of younger siblings in the household, the religion of the household, whether the household owns a farm, dummy if the child is female and the number of schools in the village are found to be non statistically significant (their significance is joint tested).

Table 4.17
Probit model for junior secondary school – restricted model

Variables	*Coeff.*	*z*	*Marginal effect*
Difference in earnings (x-w)* Φ_{iS}	0.004	0.66	0.002
Difference in earnings (a-w)*(1- Φ_{iS})	0.016	0.81	0.006
Total cost of junior secondary school (p+w)/1000	-1.107	-1.71	-0.433
Total cost of senior secondary school ((q+x)/1000)*(1- Φ_{iS})	-0.073	-0.26	-0.029
Residual from income regression ($y_t - \hat{y}_t$) (excluding wage of children 12-14)	0.224	2.32	0.089
Sigma2 (period two income variance)	0.079	2.67	0.031
Variance of adult earnings of senior secondary school graduates	-0.025	-0.10	-0.010
Estimated permanent income $\hat{y}_t$ *(1- Φ_{iS}) (excluding wage of children 12-14)	0.514	2.66	0.201
# siblings age 13-15	0.312	1.57	0.122
# siblings age 16-18	0.253	1.91	0.099
intercept	0.857	1.34	-
Pseudo R-squared		0.18	
Log pseudo-likelihood		-175.583	
N		311	

The table reports the results of the restricted probit model for junior secondary decisions (equation 4.20). Dependent variable is a dummy that equals one if the child is attending school. Sample: children graduated from the elementary school, aged 12-14 years. Provincial dummies are included as additional regressors. Φ_{iS} is the probability of non-continuation to the senior secondary school. The cost of schooling is conditional on household income. Household income is in millions rupiah. Income variance is divided by one million. Whether the head and the spouse of the head are illiterate, the number of younger siblings in the household, the religion of the household, whether the household owns a farm, dummy if the child is female and the number of schools in the village are found to be non statistically significant (their significance is joint tested).

4.6 Conclusions

The use of child time allocation as a response to transitory income shocks is well established. Many papers document the adverse effects of negative shocks on child education and child labour in the face of credit constraints. Few papers have analyzed the impact of ex ante risk on children human capital investments. To our knowledge, this is the first study that combines both issues analyzing the role of ex ante income risk and of income shock occurrence on child education, and, in particular, on school progression from a school level to another. The data used in this chapter are from the Indonesian Family Life Survey (mainly the 1993 round). The decision to focus on school progression (that is to analyze the choices to enroll at school after the completion of the previous school level) is motivated by the finding that in 1993 in Indonesia there were low continuation rates from primary to junior secondary school.

The first difference of this study with respect to previous literature regards the sources of risk which we take into account. Fitzsimons (2007), who analyzes the same dataset we use in this chapter, equates risk with the variability of parental income when children become adults. She measures both the household-specific and the village-level component of variability using information on past incomes. Her assumption is that households base predictions of future volatility on the observed volatility of their past earnings streams. We suppose that children's adult earnings, conditional on being graduated from senior secondary school, are random drawing from the provincial income distribution for senior secondary school graduates.

Our chapter examines two different sources of uncertainty: the variance of life-time incomes of senior high school graduates relative to their expectations, and the variance of household income deviation over the time the child will potentially remain at school. The latter component is estimated as the variance of future household income conditional on current household characteristics, assuming that households base expectations of future variability on current information.

The second contribution of our model is that it takes into account the irreversibility of withdrawal from school: children withdrawn from school cannot re-enroll, and hence they lose the opportunity to continue to further levels of schooling and achieve higher earnings when they become adults. State dependence of school attendance has also been modeled by de Janvry et al. (2006), who include a re-entry cost as an additional cost of schooling. There are two main differences relating to our model: first, we consider a "strong" version of state dependence (irreversibility), while de Janvry et al. model re-entry imposing an additional cost. Second, de Janvry et al. focus on the implications of state dependence for the impact of income shocks on child education and child labour, while we analyze both shock responses ex ante, to the anticipation of shocks, and ex post, to their realizations

We have developed a model of household school progression decisions in the face of risk and shocks under the assumption that there are no savings and borrowing, and that children continue to form part of the household once they become adults. The model predicts that household uncertainty with respect to the earnings of the adult (ex-child) once graduated tends to discourage school attendance, while household uncertainty with respect to its own income over the time the child may remain at school should increase school participation. This second effect arises out of irreversibility: in the face of household income variability parents are more likely to send children to school to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults). Finally, as regards ex post responses to shocks, under the assumption that there is no saving and borrowing, negative (positive) shocks are predicted to reduce (increase) the probability of school progression.

We estimate school progression by estimating probit models for both junior and senior secondary school on different cohorts: the sample for junior attendance is those children graduated from the elementary school, aged less than 15 years; the probit for senior school is estimated on the sample of junior secondary graduates with aged less than 18 years. Results suggest that Indonesian rural household adjust child schooling in the face of risk and shocks. Children are more, or less, likely to attend school in response to positive, or negative, income shocks, at both junior and senior secondary school stages. As predicted by the model, we find that the probability of attending junior secondary school depends positively on the variance of household income over the time the child will remain at school, through the option of continuation to senior secondary school. Household uncertainty with respect to the earnings of the adult (ex-child) once graduated has the predicted negative sign in both probit regression models, but is not statistically significant.

The model also predicts that schooling costs and returns to schooling affect progression decisions. The difference in earnings from attending senior school contributes to a higher continuation probability to senior secondary school, while the corresponding variable is not statistically significant for the junior progression decisions. The total cost of junior school has a negative and significant effect (at the 10% level) on the probability of attending the junior level. Finally, the probability of the child attending junior secondary school depends positively on the predicted household permanent income, and on the probability of continuation to senior secondary school.

Appendix 4A - Cost of schooling

Table 4A.1
Poisson model. Dependent variable: # of children in the household attending elementary school

	Male	*Female*
Household socioeconomic status		
Predicted permanent income (excluding wage of children age 7-12)	-0.032 (0.026)	-0.014 (0.028)
Utilize electricity	0.058* (0.028)	0.001 (0.026)
Water source inside the household	0.039 (0.037)	-0.016 (0.039)
Own toilet	-0.001 (0.028)	-0.014 (0.029)
own farm	0.029 (0.032)	-0.068* (0.031)
own non-farm business	0.025 (0.034)	0.028 (0.036)
Household characteristics		
# household members age 0-6	-0.003 (0.014)	0.023 (0.014)
# females age 7-12	-0.076** (0.023)	0.586*** (0.023)
# males age 7-12	0.530*** (0.024)	-0.039 (0.024)
# household members age 13-15	0.124*** (0.028)	0.098*** (0.022)
# household members age 16-18	-0.040 (0.032)	0.006 (0.026)
Islam	0.011 (0.051)	0.075 (0.050)
Christian	0.066 (0.072)	-0.009 (0.083)
Hindu	0.047 (0.072)	0.062 (0.081)
Head and spouse education		
Head illiterate	-0.105 (0.157)	-0.292* (0.134)
Head attended only elementary school	0.006 (0.147)	-0.251* (0.126)
Head attended only junior secondary school	0.016 (0.137)	-0.187 (0.115)
Head attended senior secondary school	0.089 (0.127)	-0.181 (0.112)

Spouse illiterate	-0.023 (0.075)	-0.085 (0.093)
Spouse attended elementary school	0.033 (0.072)	-0.017 (0.088)
Spouse attended junior secondary school	0.048 (0.078)	-0.070 (0.090)
Spouse in HH	0.040 (0.082)	0.079 (0.095)
Village characteristics		
# of elementary school in the village	-0.002 (0.008)	0.018* (0.008)
Constant	-1.003*** (0.288)	-0.475** (0.172)
Observations (sub-sample of households with at least one child aged 7-12)	1065	1101
Log pseudo-likelihood	-1160.58	-1196.81
Measures of fit		
OLS R2	0.379	0.436
McFadden's R2	0.045	0.052
McFadden's Adj R2	0.022	0.030
Maximum Likelihood R2	0.099	0.114

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

The table reports the results of the Poisson model that estimate the number of children attending elementary school. The sample are households with at least one child aged 7-12. High education is the omitted category for education o the head and of the spouse. Permanent income is in millions rupiah. Dummies for the occupation of the household head are included as additional regressors

Table 4A.2
Zero-inflated model. Dependent variable: # of children in the household attending junior and senior secondary school

	Junior secondary		**Senior secondary**	
	Male	*Female*	*Male*	*Female*
Household socioeconomic status				
Predicted permanent income (excluding wage of children in the reference age group)	0.001 (0.071)	0.116* (0.046)	-0.121 (0.1094)	-0.019 (0.1095)
Utilize electricity	-	-	0.716*** (0.213)	1.495*** (0.334)
Water source inside the household	-	-	0.407* (0.202)	0.487* (0.200)
Own toilet	-	-	0.323+ (0.168)	0.498** (0.184)
own farm	-0.263 (0.240)	-0.160 (0.217)	-0.649+ (0.362)	-0.041 (0.277)
own non-farm business	0.048 (0.104)	0.009 (0.123)	0.136 (0.206)	-0.419+ (0.238)
Household characteristics				
# household members age 0-6	-0.011 (0.056)	-0.183** (0.064)	-0.057 (0.122)	-0.170 (0.137)
# household members age 7-12	0.009 (0.057)	0.064 (0.071)	0.035 (0.105)	0.022 (0.138)
# household members age 13-15	-	-	0.050 (0.106)	-0.036 (0.122)
# females age 13-15	-0.084 (0.130)	0.479** (0.166)	-	-
# males age 13-15	0.489*** (0.056)	-0.058 (0.141)	-	-
# household members age 16-18	0.249*** (0.065)	0.099 (0.081)	-	-
# females age 16-18	-	-	-0.228 (0.218)	0.844*** (0.222)
# males age 16-18	-	-	0.547*** (0.131)	-0.011 (0.221)
Islam	-0.146 (0.159)	-0.260 (0.168)	0.016 (0.278)	0.137 (0.310)
Christian	-0.129 (0.231)	0.149 (0.238)	0.814* (0.374)	0.950** (0.356)
Hindu	-0.248 (0.266)	0.021 (0.262)	0.701+ (0.376)	0.411 (0.459)

Head and spouse education				
Head illiterate	-1.404*** (0.298)	-0.279 (0.292)	-0.774 (0.622)	-0.597 (0.637)
Head attended only elementary school	-0.932*** (0.263)	-0.054 (0.241)	-0.428 (0.509)	-0.296 (0.551)
Head attended only junior secondary school	-0.501* (0.207)	0.388+ (0.212)	-0.036 (0.466)	-0.135 (0.535)
Head attended senior secondary school	-0.478** (0.185)	0.084 (0.211)	0.618 (0.453)	0.181 (0.464)
Spouse illiterate	-0.361 (0.233)	-1.027*** (0.231)	0.074 (0.556)	0.126 (0.588)
Spouse attended elementary school	0.069 (0.191)	-0.403* (0.157)	0.476 (0.499)	0.694 (0.504)
Spouse attended junior secondary school	0.183 (0.206)	-0.501** (0.191)	0.743 (0.558)	0.648 (0.633)
Spouse in HH	0.140 (0.220)	0.522* (0.219)	-0.583 (0.505)	-0.985+ (0.539)
Village characteristics				
# of junior school in the village	0.070 (0.049)	0.102* (0.049)	-	-
# of senior school in the village	-	-	-0.023 (0.122)	0.165 (0.139)
Constant	-0.478 (0.470)	-0.911+ (0.502)	-1.617* (0.645)	-3.821** (1.176)
Observations (sub-sample of households with at least one child aged 7-12)	585	561	433	409
Log pseudo-likelihood	-471.03	-409.35	-230.20	-167.31
Measures of fit				
OLS R2	0.304	0.184	0.3098	0.3104
McFadden's R2	0.103	0.076	0.198	0.278
McFadden's Adj R2	0.048	0.009	0.090	0.144
Maximum Likelihood R2	0.172	0.117	0.235	0.276
Vuong test of zip vs. standard Poisson:	z=0.26 Pr>z=0.396	z=0.28 Pr>z=0.389	z=1.42 Pr>z=0.078	z=1.94 Pr>z=0.026

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

The table reports the results of the zero inflated Poisson models that estimate the number of children attending junior and senior secondary school. The sample are households with at least one child aged 13-15 for junior secondary and 16-17 for senior secondary. High education is the omitted category for education of the head and of the souse. Permanent income is in millions rupiah. Dummies for the occupation of the household head are included as additional regressors.

Table 4A.3
Actual and predicted numbers of children attending school, by school levels

		Mean	*s.d.*	*Min*	*Max*
# girls attending elementary school	Actual values	1.131	0.609	0	3
	Predicted – Poisson	1.132	0.408	0.667	3.757
# boys attending elementary school	Actual values	1.141	0.597	0	4
	Predicted – Poisson	1.142	0.374	0.577	3.706
# girls attending junior sec. school	Actual values	0.423	0.557	0	3
	Predicted – Poisson	0.424	0.249	0.074	2.035
	Predicted – ZIP	0.424	0.247	0.070	1.935
# boys attending junior sec. school	Actual values	0.521	0.685	0	7
	Predicted – Poisson	0.521	0.418	0.115	7.173
	Predicted – ZIP	0.567	0.495	0.097	7.967
# girls attending senior sec. school	Actual values	0.228	0.471	0	2
	Predicted – Poisson	0.228	0.307	0.009	2.523
	Predicted – ZIP	0.228	0.321	0.004	2.382
# boys attending senior sec. school	Actual values	0.287	0.564	0	6
	Predicted – Poisson	0.287	0.370	0.025	5.411
	Predicted – ZIP	0.282	0.306	0.017	2.759

The table reports the actual and predicted values of the number of children attending school. Predicted values are from the Poisson and Zero inflated Poisson models.

Table 4A.4
Educational expenditure regression

Number of children attending school (β_j)	**Coeff.**	**t**
# male attending elementary school	130.58	5.60
# female attending elementary school	135.65	5.92
# male attending junior secondary school	206.71	4.47
# female attending junior secondary school	385.00	4.88
# male attending senior secondary school	435.80	5.69
# female attending senior secondary school	794.78	8.62
Income interaction effect (η)	**Coeff.**	**t**
Household income	0.35	4.20

N =1974
R-squared = 0.11

The table reports the IV estimated coefficients from equation 4.21. Dependent variable is the total household annual educational expenditure. Provincial dummies are included in the regression. Household income is instrumented with head's educational dummies and the value of household non-business assets. The number of children in the household attending school are instrumented with the predicted numbers obtained from Poisson and Zero inflated Poisson models. The sample is households with at least one child attending school.

Table 4A.5
Descriptive statistics of predicted cost of schooling

Variable	*Mean*	*Std. Dev.*	*Min*	*Max*
Conditional on household income				
Male-elementary	105.13	47.90	30.34	355.51
Female-elementary	109.21	49.76	31.52	369.31
Male-junior secondary	166.42	75.83	48.04	562.79
Female- junior secondary	309.94	141.23	89.47	1048.17
Male-senior secondary	350.84	159.87	101.27	1186.48
Female-senior secondary	639.84	291.55	184.69	2163.81
Conditional of average income				
Male-elementary	117.50	29.18	71.30	171.82
Female-elementary	122.06	30.31	74.07	178.48
Male-junior secondary	186.01	46.19	112.87	271.99
Female- junior secondary	346.43	86.04	210.22	506.57
Male-senior secondary	392.14	97.39	237.96	573.41
Female-senior secondary	715.16	177.61	433.97	1045.74
# households=2288				

The table reports the predicted costs of schooling, by level of school and gender. The first box presents household specific estimates that are conditional on household income ($y_h^{\hat{\eta}} e^{\hat{\gamma}' P^h} \hat{\beta}_j$), the second shows the predicted values calculated using the average income ($\overline{y}^{\hat{\eta}} e^{\hat{\gamma}' P^h} \hat{\beta}_j$). These latter estimated varies by province, gender and school level.

Appendix 4B - Income levels by education

Table 4B.1
Individual income equation: Heckman selection model

		Female	Male
Dependent variable: log income (robust se in parenthesis)			
Male: dummy elementary school*provincial dummies Female: dummy elementary or junior secondary school*provincial dummies	North Sumatra	1.913*** (0.437)	4.160*** (0.264)
	West Sumatra	1.731*** (0.401)	4.480*** (0.243)
	South Sumatra	1.846*** (0.522)	3.665*** (0.257)
	Lampung	1.501** (0.469)	3.716*** (0.258)
	West Java	1.960*** (0.392)	4.013*** (0.210)
	Central Java	1.525*** (0.373)	3.803*** (0.242)
	Di Yogyakarta	1.449*** (0.433)	3.862*** (0.294)
	East Java	1.774*** (0.375)	4.192*** (0.230)
	Bali	1.858*** (0.426)	4.193*** (0.255)
	West Nusa Tenggara	1.658*** (0.436)	3.832*** (0.281)
	South Kalimantan	1.785*** (0.446)	4.320*** (0.282)
	South Sulawesi	1.490*** (0.389)	4.194*** (0.330)
Male: dummy junior secondary school*provincial dummies	North Sumatra	-	4.233*** (0.316)
	West Sumatra	-	4.742*** (0.344)
	South Sumatra	-	4.025*** (0.292)
	Lampung	-	4.019*** (0.346)
	West Java	-	4.512*** (0.287)
	Central Java	-	4.612*** (0.292)
	Di Yogyakarta	-	4.075*** (0.351)
	East Java	-	4.668*** (0.246)
	Bali	-	4.569*** (0.494)
	West Nusa Tenggara	-	4.599*** (0.312)
	South Kalimantan	-	4.836*** (0.322)
	South Sulawesi	-	4.596*** (0.492)
Dummy senior secondary school*provincial dummies	North Sumatra	2.930*** (0.417)	4.822*** (0.262)
	West Sumatra	3.211*** (0.322)	4.669*** (0.302)
	South Sumatra	2.454*** (0.408)	4.801*** (0.350)
	Lampung	dropped	4.167*** (0.276)
	West Java	3.912*** (0.408)	4.818*** (0.219)
	Central Java	2.895*** (0.363)	4.672*** (0.245)
	Di Yogyakarta	3.402*** (0.374)	4.571*** (0.334)
	East Java	3.142*** (0.303)	4.684*** (0.217)
	Bali	2.604*** (0.467)	4.749*** (0.233)
	West Nusa Tenggara	3.063*** (0.428)	4.699*** (0.224)
	South Kalimantan	2.991*** (0.383)	5.046*** (0.213)
	South Sulawesi	3.656*** (0.305)	4.888*** (0.219)
Experience		0.096*** (0.013)	0.028* (0.011)
Experience square		-0.001*** (0.000)	-0.0004* (0.000)

	Female	Male
Dependent variable: log income (robust se in parenthesis)		
Selection (robust se in parenthesis)		
Non business assets/1000	-0.002 (0.000)	0.001 (0.000)
married	-0.308*** (0.048)	0.660*** (0.067)
unschooled	-1.607*** (0.271)	-0.977*** (0.177)
incomplete primary	-1.541*** (0.267)	-0.803*** (0.169)
highest grade completed: elem	-1.367*** (0.266)	-0.647*** (0.168)
highest grade completed: jrsec	-1.254*** (0.271)	-0.559** (0.173)
highest grade completed: srsec	-0.421 (0.273)	0.012 (0.175)
exper	0.096*** (0.005)	0.098*** (0.005)
exper2	-0.001*** (0.000)	-0.002*** (0.000)
# household members age 0-5	-0.122*** (0.025)	-0.026 (0.025)
# household members age 6-9	-0.023 (0.030)	-0.083** (0.031)
# household members age 10-14	-0.064* (0.025)	-0.063* (0.025)
# household members age 15-17	-0.093* (0.036)	0.002 (0.034)
Month of interview: September	0.235** (0.086)	0.663*** (0.084)
Month of interview: October	0.237** (0.076)	0.614*** (0.074)
Month of interview: November	0.438*** (0.073)	0.586*** (0.074)
Month of interview: December	0.192** (0.069)	0.227*** (0.068)
Intercept	-0.194 (0.271)	-0.994*** (0.176)
Mills (robust se in parenthesis)		
lambda	0.539** (0.192)	-0.868*** (0.127)
N	*5982*	*5534*
Censored	*4545*	*2898*
Uncensored	*1437*	*2636*

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table reports the results of equation (4.13). Dependent variable is log monthly income. Sample: all individuals aged 10 years or old with information on labour. Dummies for "illiterate" and "high education" are included in the income regression, with no schooling acting as the omitted category. In the selection equation omitted categories are "high education" for educational dummies and "January" for the month of interview.

Appendix 4C – Measures of risk and shocks

1) Measure of shocks

Table 4C.1
Household income equation

	A Household income excluding the wage of children 12-14 years old		*B* Household income excluding the wage of children 15-17 years old	
Variables	*Coeff.*	*t*	*Coeff.*	*t*
1992 farm assets	0.010	1.32	0.010	1.33
1992 business non-farm assets	0.019	2.30	0.019	2.30
1992 non business liquid assets	0.029	1.22	0.029	1.22
1992 non-business illiquid assets	0.023	4.28	0.023	4.30
Household owns a farm	-40.886	-0.42	-24.533	-0.25
Household owns a non-farm business	514.629	7.02	519.468	7.12
Household head does not work	536.514	1.75	602.842	1.97
Household head is employee	838.371	3.00	910.185	3.27
Household head is self-employed	192.166	0.70	247.471	0.90
Head primary education	177.365	3.27	194.101	3.62
Head junior secondary education	934.074	5.96	945.656	6.03
Head senior secondary education	1696.077	7.95	1711.204	8.03
Head high education	2896.955	6.32	2886.849	6.31
Nr. of income earner (other than the head)	143.110	4.05	119.823	3.45
Household size	41.248	0.52	34.949	0.44
Household size^(2)	0.949	0.14	1.501	0.21
Intercept	-195.811	-0.52	-248.610	-0.67

Number of obs= 2239
R-squared= 0.30

The table records the results from equation (4.14) and estimate the predicted household income. The dependent variable is 1993 household income excluding the child wage. Whether the household head works as family worker is the omitted category for the working status of the head. Whether the head has no education is the reference category for head's education. This regression includes provincial dummies. Both income and assets are measured in thousands of rupiah. The standard errors are robust.

2) Income variability of senior secondary school graduates

Table 4C.2
Income regression for senior secondary school graduates

Variables	*Coeff.*	*t*
Dummy male	0.246	2.12
Dummy language (1 if language spoken at home is Indonesian)	0.462	2.69
intercept	4.334	22.07
Adj R-squared	0.016	
N	491	

The table reports the results of equation (4.16). The dependent variable is the log of individual monthly income, the sample is individuals graduated at least from the senior secondary school. Provincial dummies are included.

Table 4C.3
Regression to estimate the adult income variance

Variables	*Coeff.*	*t*
Dummy male	-0.418	-0.44
Dummy language (1 if language spoken at home is Indonesian)	-0.165	-0.08
Prov1	1.576	2.90
Prov2	1.322	3.31
Prov3	1.316	0.94
Prov4	1.037	0.93
Prov5	0.692	1.12
Prov6	1.894	4.07
Prov7	1.060	2.07
Prov8	0.529	1.07
Prov9	1.878	3.00
Prov10	1.711	2.83
Prov11	1.015	1.15
Prov12	0.705	0.98
Adj R-squared	0.28	
F(33,458)	6.79	
N	491	

The dependent variable is the square of fitted residuals from equation (4.16). The sample is individuals graduated at least from the senior secondary school. Cross products are included

3) household income variability over the time the child will remain at school

Table 4C.4
1997 income equation

Variables	*Coeff.*	*t*
1993 household income	0.259	9.37
Non business liquid assets	0.069	4.90
Non business illiquid assets	0.018	4.29
# schooling years completed by the household head	50.210	3.97
Household size	-0.675	-0.03
Intercept	713.26	5.18
Adj R-squared	0.12	
F(5,1866)	53.05	
N	1872	

The table reports the results of equation (4.17). The dependent variable is 1997 household income.

Table 4C.5
Regression to estimate the variance of income deviation

Variables	*Coeff.*	*t*
1993 household income	1539.851	1.68
Non business liquid assets	5911.259	3.17
Non business illiquid assets	134.091	0.71
# schooling years completed by the household head	572175.500	1.51
Household size	-388563.900	-0.46
1993 household income^2	0.044	0.99
Non business liquid assets^2	-0.042	-5.22
Non business illiquid assets^2	0.001	1.16
# schooling years completed by the household head^2	-51297.520	-2.37
Household size^2	33426.4	0.50
Adj R-squared	0.052	
F(20,1851)	6.09	
N	1872	

The table reports the results of the second stage of the procedure to estimate the variance of household period two income deviation. The dependent variable is the square of predicted residuals from table 4C.4. Cross products are included as additional regressors.

Appendix 4D – Overall probability of attending senior secondary school

The overall impact of uncertainty on schooling decisions is complicated. Further this impact comes through a number of distinct channels. In this appendix, we focus on the impact of uncertainty on the decision to attend senior secondary school. Only those children who have graduated from junior secondary school are able to continue to senior school. Conditional on graduation, the continuation decision will depend on the income realization at that time (period 2). The uncertainty associated with income at this date will also impact the decision to attend junior secondary school. In order to evaluate the total impact of period 2 income uncertainty on attendance at senior secondary school, we need to control for both these effects.

The total probability of attending senior school may be estimated as $p_s = p_{s/j} \cdot p_j$, where $p_{s/j}$ is the probability of attending senior school conditional on having completed the junior level, and p_j is the probability of continuation to the junior school. In order to consider how probabilities change if income guarantees were in place over the schooling period (so no period two uncertainty and no period two shocks), we calculate the following probabilities. The first is defined as:

$$_1p_s = {}_1p_{s/j} \cdot {}_1p_j$$

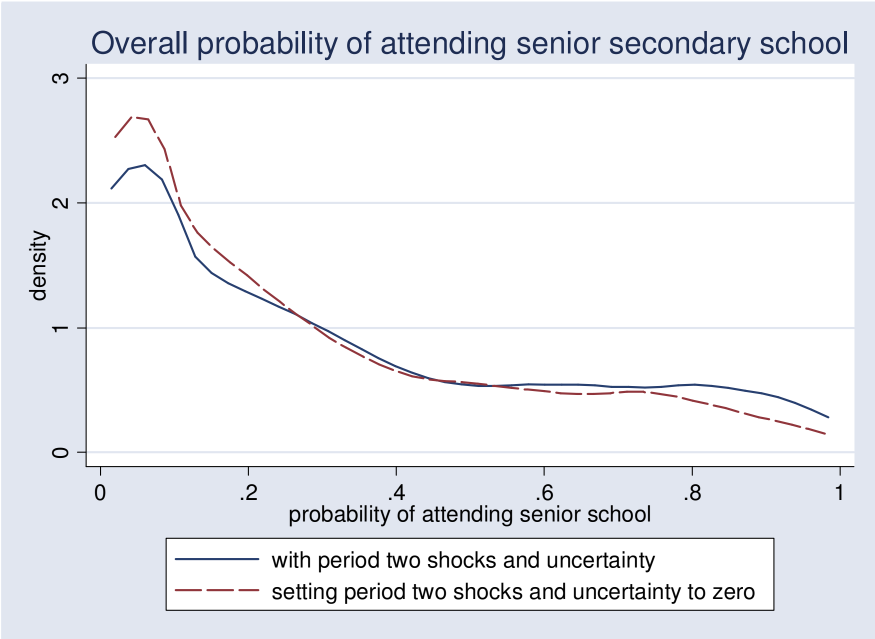
where $_1p_{s/j}$ is the conditional probability predicted by estimating (4.18), and $_1p_j$ the probability predicted from the probit regression model defined in (4.20). $_1p_{s/j}$ depends upon both income uncertainty and shocks. The second measure is calculated setting both period 2 shocks and period 1 uncertainty with respect to period 2 shocks to zero:

$$_2p_s = {}_2p_{s/j} \cdot {}_2p_j$$

where $_2p_{s/j}$ is calculated using the estimates reported in Table 4.15 with $\eta_{h2} = 0$ (i.e. without considering period two shocks), and $_2p_j$ is predicted using the coefficients reported in table 4.17 and setting σ_h^2=0 (where σ_h^2 is the variance of η_{h2}, i.e. period two income uncertainty). $_2p_s$ still depends upon period one shock (η_{h1}) and period three income uncertainty (Ψ_i^2).

The following graph shows the kernel density estimates for the two probabilities: the solid line represents $_1p_s$, the dashed line $_2p_s$. As we can see, the overall probability of attending senior secondary school is lower if we set both period two income risk and shocks to zero. This difference is largely due to the impact of period two income uncertainty (setting only η_{h2} to zero does not significantly change the predicted probability).

Figure 4D.1: Overall probability of attending senior secondary school



Although this result appears paradoxical, it is a direct consequence of our model specification. The option value effect of an increase in the uncertainty with respect to period 2 income on the decision to attend junior high school is always positive, so a reduction in this uncertainty must increase both junior and senior secondary school enrollment. By contrast, the impact of the period 2 income shocks themselves on continuation depends on whether these shocks are positive or negative. Families who experience positive shocks are more likely to allow their children to continue to senior secondary school and the contrary is the case with those families who experience negative shocks. While the impact of these shocks on the conditional continuation probabilities can be large in particular instances, given our linear specification, they are likely to balance out over the entire population. In order to be confident that the effects illustrated by the kernel density functions are in deed valid, we need to look more thoroughly at possible nonlinear shock impacts.

Chapter 5

Conclusions

5.1 Summary

This dissertation investigated the way in which potential and actual shocks affect household economic conditions and choices. We focused on three main issues. First, we explored which strategies are chosen for different types of shocks. Second, we investigated which are the specific mechanisms adopted in the face of the most common shock (crop loss), and the consequences on consumption. Finally, we analyzed the consequences of risk and shocks on a particularly important household decision: how much to invest in children education.

The source of data used in this dissertation is the Indonesian Family Life Survey, mainly the 1993 round. A large amount of information on household economic conditions (income sources, consumption, types of assets), and household and individual characteristics were collected, as well as a detailed section on shocks experienced by the household during the five years prior to the interview. The questionnaire asked households to report a list of adverse events: the death and sickness of a household member, crop loss, unemployment, household or business loss due to a disaster, unanticipated drop in the price of crop products. The coping strategies included in the survey are: taking an extra job, taking a loan, sell assets, family assistance, use of savings, and cutting expenditures.

Chapter one reviewed the theoretical and empirical literature regarding households' behaviour in the face of risk and shocks in developing countries. The theoretical framework that underlines the analysis is an intertemporal model of consumption and asset investment choices, that accounts for the fact that poor households, especially in developing countries, may have no access to perfect credit markets, and they may find it hard to save. Moreover, the vast majority of people living in poor areas do not rely on wage income, rather on farm profits, for which the availability of some productive assets becomes crucial.

Both the literature review and the descriptive analysis of Indonesian data highlighted that household responses to shocks depend on the nature of the shock (whether demographic or economic, idiosyncratic or common,…), and on household characteristics. In particular, we found that informal arrangements (for example family assistance) are important means to overcome demographic and idiosyncratic shocks, while they are marginal as a response to economic hardships. On the other hand, households are more likely to use the labour supply responses to cope with economic and common shocks. In general, relatively few households draw down savings as a coping strategy. A characteristic that has important consequences on the way in which households respond to shocks is clearly household's wealth (and the presence of liquidity constraints that may be related to it). Indonesian data show that poor and non-poor households differ in the choices of coping strategies: the former are more likely to use the labour supply response, while rich households are more likely to sell assets (for demographic shocks) and to use savings. This result is in line with the finding of the theoretical and empirical literature that when liquidity constraints are binding, households need to rely on autarchic savings, both to build a buffer stock of assets and to self-finance profitable investments.

Since households may use more than one strategy to cope with the same shock, the econometric treatment of these choices requires models that allow for non-exclusive and dependent multiple responses. In chapter two we developed two models that take into account these specificities. In the first and simpler specification, the Poisson-multinomial, households first choose the number of responses to a specific shock, and then the specific choices are identified to maximize household utility conditional on the former choice. The second specification, the threshold-multinomial, generalizes the standard multinomial logit model by supposing that agents will choose more than one response if the utility they derive from other choices is "close" to that of the utility-maximizing choice. Both these models are an advance over the standard Marginal Logit Model (MLM) approach in that they see respondents as comparing responses with each other rather than with a common benchmark. A disadvantage of the Poisson multinomial model is that it inherits the IIA property from multinomial logit. Both the MLM and threshold multinomial models, which make the number of responses endogenous to the respondent's decision process, are less vulnerable to this problem. We found the both Poisson and threshold multinomial models outperform the

Marginal Logit Model (MLM). Choice between the two multinomial models is less clear but the data appear marginally more favourable to the threshold model.

Chapter three adds to the literature on self-protection coping strategies investigating whether rural Indonesian households smooth income following a crop loss (i.e. whether they adopt strategies, labour supply adjustments in particular, to recover the income reduction due to the shock), and how much consumption smoothing is achieved through income smoothing strategies (i.e. how much of the increase in income is transferred onto consumption). In particular, we sought to examine two issues. First, which variables influence the adoption of income smoothing strategies and whether these strategies completely or partly recover the income reduction due to the shock. Quantitative measures of income shocks and household's ability to cope with the shock (mainly through labour supply response) were estimated for poor and non-poor households. Using these measures we explored whether the increase in income due to income-smoothing strategies partially offsets, or exceeds, the income loss due to the shock. Second, this chapter investigated whether households smooth consumption and the role played by different coping strategies in mitigating consumption reductions. The analysis was conducted distinguishing between consumption behaviour of poor and non-poor households.

Our results suggest that the income gain given by the extra labour supply response completely compensate the income reduction caused by the crop loss, while the role of non-business assets and "take a loan or sell assets" is in general marginal in recovering the income loss. As regard consumption behaviour, there are two main differences between poor and non-poor households. First, while medium and large farms smooth consumption relative to income, this is not so for small farms: for the latter, the main components of transitory income (crop loss and the extra labour income) have an effect on consumption that is statistically significant and equal to the one associated with permanent income. The second distinction between poor and non-poor households concerns the marginal propensity to consume out of the relevant income measure: the former save about a half of their current income, whereas the latter consume a fraction of their permanent income close to one. This result may confirm the need for poor households, that are more likely to be excluded from financial markets, to rely on autarchic saving strategies.

The final issue examined in this dissertation regards household investments in children education in the face of risk and shocks. We made three contributions to the analysis of these choices. First, the role of both ex ante risk and ex post income shocks on child schooling is considered. Second, we examined two different sources of uncertainty: household uncertainty with respect to the earnings of the adult (ex-child) once graduated, and household uncertainty with respect to parental income over the time the child may remain at school. Third, we proposed a model that accounts for the irreversibility of the decision to withdraw a child from school. The key assumption of the model is that once the child drops out, he/she cannot return to school. In this way, temporary interruptions in schooling in the face of risk or short-term shocks have long term impacts on the child human capital. The model predicts that household uncertainty with respect to the earnings of the adult (ex-child) once graduated tends to discourage school attendance, while household uncertainty with respect to parental income over the time the child may remain at school increases school participation. The latter effect arises out of irreversibility: in the face of household income variability parents are more likely to send children to school to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults). Finally, as regards ex post responses to shocks, under the assumption that there is no saving and borrowing, negative (positive) shocks are predicted to reduce (increase) the probability of school progression.

Results substantially confirm the predictions of the model. Our findings show that Indonesian rural household adjust child schooling in the face of risk and shocks. Children are more, or less, likely to attend school in response to positive, or negative, income shocks, at both junior and senior secondary school stages. The probability of attending junior secondary school depends positively on the variance of household income over the time the child may remain at school, through the option of continuing to senior secondary school. Household uncertainty with respect to the earnings of the adult (ex-child) once graduated has the predicted negative sign in both probit regression models, but is not statistically significant. These results underline that the impact of uncertainty on schooling decisions is more subtle that suggested by much of the existing development literature.

In summary, three sets of main conclusions emerge from this dissertation. First, we have learned that when facing a shock, households choose risk coping strategies by comparing responses with each other rather than with a common benchmark. In such a situation of non-exclusive and dependent multiple responses, the widely used Marginal Logit Model (MLM) suffers from a number of limitations. The two models we developed to take into account these specificities appear to outperform the MLM in describing these type of choices.

The second main conclusion relates to the evidence that the choices between different coping strategies markedly differ between poor and non-poor households. In the face of shocks, the former appear to behave in a very different way. In general, rich households smooth consumption relative to income, whereas the need to accumulate savings to both build a buffer stock of assets and self-finance profitable investments leads poor people to rely more on ex post income smoothing strategies (taking an extra job) and to use part of this extra labour income to preserve their level of assets, even reducing consumption if necessary.

Finally, it is necessary to deepen our knowledge of the long-run consequences of shocks, particularly with respect to the human capital formation of children. We found that the impact of uncertainty on schooling decisions is more subtle than suggested by much of the existing development literature. Taking into account that withdrawal from school is an absorbing state, that is children cannot re-enroll once they stop going to school, temporary interruptions in child schooling have long term impacts on the child human capital. Given irreversibility of withdrawal from school, in the face of household income variability parents are more likely to send children to school to give them the option to continue with higher schooling levels in the future (and hence earn higher earnings when they become adults).

5.2 Further Research

Chapter three constructed quantitative measures of the income reduction caused by the crop loss and of the income gain due to the labour supply response. A further question to be assessed in future work is whether there is persistence of both shock and coping strategy: do crop loss and labour supply responses continue to affect future income? As discussed in chapter one, a variety of theories explain why transitory income shocks may persist. Many of these theories document the link between risk management strategies (both ex ante and ex post) and long-run effects. For example, ex ante strategies may lead poor households to choose safe but low return activities, while households may be forced ex post to sell productive assets or to take children out of school. All these effects may induce poverty persistence. To my knowledge, there is no study that examines possible long-run effects that come from the labour supply responses: the extra income obtained through the shocks-induced labour supply may persist far into the future, and perhaps longer than the shock.

The starting point of this analysis will be the paper of Newhouse (2005). Using the Indonesian Family Life survey, Newhouse (2005) estimates the persistence of transient income shocks (measures by rainfall shocks) to farm households in rural Indonesia. In his model, income shocks persist because current consumption choices affect future level of assets, which in turn affects future income. He finds that, on average, thirty percent of income shocks remain after four years, negative shocks persist no longer than positive ones, and there are no significant differences in the shock persistence between rich and poor.

A second future research question is to analyze the impact of risk and shocks on child labour. Chapter four examined the effects of risk and shocks on child education, but this does not allow me to infer the implications for child labour. As discussed in chapter four, and as suggested by the literature (Ravallion and Wodon, 2000; Probe report, 1999), the trade-off between schooling and working may not be one-to-one, and out of school does not mean necessarily at work. In developing countries children often work part-time, and in many cases as family labourers at home or in the fields. Even a part-time job may be enough to exclude children from the schooling system, but in some cases children may combine work and study (in many developing countries school day is short) (Probe report, 1999). An interesting extension of chapter four will be to include school and labour choices in the analysis and examine the effects of risk and shocks on child time allocation, to account for the fact that children may both attend school and work (de Janvry et al., 2007).

A third and final issue I plan to analyze is the intra-household allocation of resources in the face of shocks. In chapter three I examined consumption smoothing in the face of a crop loss, but without considering how variations in consumption are distributed across household members. As suggested by Dercon (2002), the failure to cope with risk is not only reflected in household consumption, but there are also intra-household consequences. Some household members may be more vulnerable than others in the presence of risk: this is partly addressed in chapter four, where child schooling decisions in the face of risk and shocks are analyzed. This chapter adopted a unitary approach to characterized household choices, and did not examine the intra-household bargaining between parents over the child's time allocation. A possible extension would be to examine child time use in the face of risk using a non-unitary approach. Many papers find that parents may have different preferences regarding children's human capital investments, and the general finding is that the mother cares more than the father about children (Haddad, Hoddinott and Alderman, 1997). Thomas (1990) finds that in Brazil the higher is the mother's bargaining power the healthier are the children. Duflo (2003) shows that income from pensions has a positive impact on girl's nutrition if the income earner is a women, and in particular the mother of the mother. Thomas, Contreras and Frankenberg (2002) show that mother and father's bargaining power have different effects on children's health outcomes, and in particular, sons are healthier than daughters when the mothers are more powerful. An unexpected result comes from Basu and Ray (2002). They analyse the relationship between the allocation of power in the household and child labour, and find that as the woman's power increases, the child's labour will first fall and then rise. The idea is that as the power of one spouse increases, he/she will control household income, included the income earned by the child, and hence will have a greater benefit from the child's labour. The authors' conclusion is that children are better protected in households where the power is more balanced.

References

Abul Naga, R.H. and E. Bolzani (2000), "Poverty and Permanent Income: A Methodology for Cross-Section Data", forthcoming in *Multidimensional Approaches to Poverty Measurement : Theory and Applications*, Deutsch, J. and Silber, J. (eds), Edward Elgar, London

Acedo, C. (Ed.) (2002), *Case Studies in Secondary Education Reform*, Improving educational Quality Project, Washington DC: American Institutes for Research. USAID

Agresti, A. (2003), *Categorical Data Analysis* (2nd edition), New York, Wiley

Agresti, A., and I-M. Liu (1999), "Modeling a categorical variable allowing arbitrarily many category choices", *Biometrics*, 55(3), 936-43

Agresti, A. and I-M. Liu (2001), "Strategies for modeling a categorical variable allowing multiple category choices", *Sociological Methods and Research*, 29(4), 403-34

Alderman, H. (1996), "Saving and Economic Shocks in Rural Pakistan", *Journal of Development Economics*, 51(2), 343-365

Alderman, H., J. Hoddinott and B. Kinsey (2006), "Long Term Consequences of Early Childhood Malnutrition", *Oxford Economic Papers*, 58(3), 450-474

Alderman H. and C. Paxson (1992), "Do the Poor Insurance? A Synthesis of the Literature on Risk and Consumption in Developing Countries", The World Bank, *Policy Research Working Paper Series* No.1008

Asra, A. (1993), "Working Children in Bandung, Indonesia 1993", Jakarta, Central Bureau of Statistics and the International Labour Office

Asra, A. (1996), "Working Children in Medan and Deli Serdang", Jakarta, Central Bureau of Statistics and the International Labour Office

Baland, J. and J. Robinson (2000), "Is child labor inefficient?, *Journal of Political Economy*, 108(4), 663-679

Banerjee, A. (2005), "The Two Poverties", in S. Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER

Bardhan, P. and C. Udry (1999), *Development Microeconomics*, Oxford University Press, Oxford, New York

Barrett, C. and J.G. McPeak (2006), "Poverty Traps and Safety Nets", in A. de Janvry and R. Kambur, eds., *Poverty, Inequality and Development: Essays in Honor of Erik Thorbecke*, Amsterdam: Kluwer

Barrett, C. and M.R. Carter (2005), "Risk and Asset Management in the Presence of Poverty Traps: Implications for Growth and Social Protection", Workshop in Leuven, Belgium

Barrett, C., S. Holden and D.C. Clay, (2002), "Can Food-for-Work Programmes Reduce Vulnerability?", *Wider Discussion Paper* No. 2002/24

Basu, K. and P.H. Van, (1998), "The Economics of Child Labor", *The American Economic Review*, 88(3), 412-427

Basu, K. and R. Ray (2002), "The Collective Model of the Household and An Unexpected Implication for Child Labor: Hypothesis and an Empirical Test", *World Bank Policy Research Working Paper* No. 2813

Beegle, K, R. Dehejia, and R. Gatti (2003), "Do Households Resort to Child Labor to Cope with Income Shocks?", *World Bank Policy Research Working Paper* No. 3075, Washington, D.C.: The World Bank

Beegle, K, R. Dehejia, and R. Gatti (2006), "Child Labour, Crop Shocks, and Credit Constraints", *Journal of Development Economics*, 81(1), 80-96

Browning, M., and A. Lusardi (1996), "Household Saving: Micro Theories and Micro Facts", *Journal of Economic Literature*, 34(4), 1797-1855

Brownstone, D. and K. Train (1999), "Forecasting new product patterns with flexible substitution patterns", *Journal of Econometrics*, 89, 109-129

Cameron, L. and C. Worswick (2003), "The Labour Market as a Smoothing Device: Labour Supply Responses to Crop Loss", *Review of Development Economics*, 7(2), 327-341

Cashdan, E.A. (1985), "Coping with Risk: Reciprocity Among the Basarwa of Northern Botswana", *Man*, 20(3), 454-474

Cochrane, J.H. (1991), "A Simple Test of Consumption Insurance", *Journal of Political Economy*, 99(5), 957-976

Cockburn, J. and B. Dostie (2007), "Child Work and Schooling: The Role of Household Asset Profiles and Poverty in Rural Ethiopia", *Journal of African Economies*, 16(4), 519-563

Davidson, R. and J.B. MacKinnon (1981), "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49(3), 551-65

Davidson, R. and J.B. MacKinnon (2004), *Econometric Theory and Methods*, Oxford University Press, New York

de Janvry A., F. Finan, and E. Sadoulet (2006), "Can Conditional Cash Transfers Serve as Safety Nets to Keep Children at School and Out of the Labor Market?", *Journal of Development Economics*, 79(2), 349-373

De Weerdt, J. (2004), "Risk-sharing and endogenous network formation", in S. Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER

Deaton, A. (1991), "Savings and Liquidity Constraints", *Econometrica*, 59(5), 1221-1248

Deaton, A. (1992), *Understanding Consumption*, Oxford University Press, Oxford, New York

Deaton, A. (1997), *The Analysis of Household Surveys: A Microeconomic Approach*, Johns Hopkins University Press, Baltimore

Dercon, S. (1996), “Risk, Crop Choice and Savings: Evidence from Tanzania”, *Economic Development and Cultural Change*, 44(3), 385-514

Dercon, S. (2002), “Income Risk, Coping Strategies and Safety Nets”, *WIDER Discussion Paper* No. 2002/22

Dercon, S. (2005a), “Risk, Insurance and Poverty: a Review”, in S. Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER

Dercon, S. (2005b), “Risk, Poverty and Public Action”, in S. Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER

Dercon, S. (2005c), “Risk, Vulnerability and Poverty in Africa”, *Journal of African Economies*, 14(4), 483-488

Dercon, S. and J. Hoddinott (2003), “Health, Shocks and Poverty Persistence”, *WIDER Discussion Paper* No. 2003/08

Dercon, S. and P. Krishnan (2000), “In Sickness and in Health: Risk-Sharing within Households in Ethiopia”, *Journal of Political Economy*, 108(4), 688-727

Domencich, T. and D. McFadden (1975), *Urban Travel Demand: A Behavioral Analysis*, Amsterdam, North-Holland

Duflo, E., (2003), “Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa,” *World Bank Economic Review,* 17(1), 1-25

Duryea, S., D. Lam, and D. Levison (2007), “Effects of Economic Shocks on Children’s Employment and Schooling in Brazil”, *Journal of Development Economics*, 84, 188-214

Fafchamps, M. (1999), “Rural Poverty, Risk and Development”, *FAO Economic and Social Development Paper* No.144

Fafchamps, M. and S. Lund (2003), “Risk-sharing Networks in Rural Philippines”, *Journal of Development Economics*, 71(2), 261-287

Fafchamps, M., C. Udry, and K. Czukas (1998), “Drought and Saving in West Africa: are Livestock a Buffer Stock?”, *Journal of Development Economics*, 55(2), 273-305

Fitzsimons, E. (2003), “The Effect of Risk on Education and Child Labor”, *Institute for Fiscal Studies, Working Paper*, vol. 02/07

Fitzsimons, E. (2007), “The Effects of Risk on Education in Indonesia”, *Economic Development and Cultural Change*, 56(1), 1-25

Flavin, M. (1985), “Excess Sensitivity of Consumption to Current Incomes: Liquidity Constraint or Myopia?”, *Canadian Journal of Economics*, 18(1), 117-136

Foster, A.D. (1993), “Household Partition in Rural Bangladesh”, *Population Studies*, 47(1), 97-114

Frankenberg, E. and L. Karoly (1995), “The 1993 Indonesian Family Life Survey: Overview and Field Report” RAND, Santa Monica, CA

Frankenberg E., J.P. Smith, and D. Thomas (2003), "Economic Shocks, Wealth and Welfare", California Center for Population Research, *Working Paper Series* No. 009-03

Frankenberg, E. and D. Thomas (2000), "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2", RAND, Santa Monica, CA. DRU-2238/1-NIA/NICHD

Friedman, M. (1957), *A theory of the consumption function*, Princeton: Princeton U. Press

Gertler, P. and J. Gruber (2002), "Insuring Consumption Against Illness", *American Economic Review*, 92(1), 51-76

Grimard, F. (1997), "Household Consumption Smoothing through Ethnic Ties: Evidence from Cote d'Ivoire", *Journal of Development Economics*, 53(2), 391-422

Guarcello, L., F. Mealli and F.C. Rosati (2003), "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance", *Social Protection Discussion Paper Series* No. 0322

Haddad, L., J. Hoddinott and H. Alderman, eds. (1997), *Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy,* Baltimore: Johns Hopkins University Press

Hoddinott, J. (2006), "Shocks and their consequences across and within households in rural Zimbabwe", *Journal of Development Studies*, 42(2), 301-321

Hoddinott, J. and B. Kinsey (2001), "Child Growth in the Time of Drought", *Oxford Bulletin of Economics and Statistics*, 63(4), 409–436

Jacoby H, and E. Skoufias (1997), "Risk, financial markets, and human capital in a developing country", *Review of Economic Studies*, 64(3), 311–335

Jalan, J. and M. Ravallion (1997), "Are the poor less well insured. Evidence on vulnerability to income risk in rural China", *Policy Research Working Paper* No. 1863, World Bank

Jalan, J. and M. Ravallion (2001), "Behavioral Responses to Risk in Rural China", *Journal of Development Economics*, 66(1), 23-49

Jensen, R. (2000), "Agricultural Volatility and Investments in Children", *American Economic Review*, 90(2), 399–405

Juster, F., J. Smith, and F. Stafford (1999), "The Measurement and Structure of Household Wealth", *Labour Economics*, 6(2), 253-276

Kazianga, H. and C. Udry (2004), "Consumption Smoothing? Livestock, Insurance and Drought in Rural Burkina Faso", *Center Discussion Paper* No. 898, Yale University

Kenjiro, Y (2005), "Why illness causes more serious economic damage than crop failure in rural Cambodia", *Development and Change*, 36(4), 759-783

Kennedy, P. (2003), *A Guide to Econometrics*, Fifth Edition, Cambridge: MIT Press

Kinsey, B., K. Burger and J.W. Gunning (1998), "Coping with Drought in Zimbabwe: Survey Evidence on Responses of Rural Households to Risk", 26(1), 89-110

Kochar, A. (1995), "Explaining Household Vulnerability to Idiosyncratic Income Shocks", *The American Economic Review*, 85(2), 159-164

Kochar, A. (1999), "Smoothing Consumption by Smoothing Income: Hours of Work Responses to Idiosyncratic Agricultural Shocks in Rural India", *The Review of Economics and Statistics*, 81(1), 50-61

Leechor, C. (1996), "Reforming Indonesia's Pension System", Policy Research Working Paper 1677, *The World Bank*

Liu, I. and A. Agresti (2005), "The analysis of ordered categorical data: An overview and survey of recent developments", *Test*, 14(1), 1-73

Long, J.S. and J. Freese (2001), *Regression Models for Categorical Dependent Variables using Stata*, College Station, TX: Stata Press

Loughin, T.M. and P.N. Scherer (1998), "Testing for association in contingency tables with multiple column responses", *Biometrics*, 54, 630-637

Lybbert T.J., C. Barrett, S. Desta and D.L. Coppock (2004), "Stochastic Wealth Dynamics and Risk Management Among a Poor Population", *The Economic Journal*, 114(498), 750-777

Mace, B. J. (1991), "Full Insurance in the Presence of Aggregate Uncertainty", *Journal of Political Economy*, 99(5), 928-956

Maddala, G. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press: Cambridge

Maitra. P. (2001), "Is consumption smooth at the cost of volatile leisure? An investigation of rural India", *Applied Economics*, 33(6), 727-734

Maitra, P., P. Bibhudutta, and S. Sfrangi (2006), "Idle Child: the Household's Buffer", Mimeo, Louisiana State University

Manning, C. (2000), "Labour Market Adjustment to Indonesia's Economic Crisis: Context, Trends and Implications, *Bulletin of Indonesian Economic Studies*, 36(1), 105-136

Mazzocco, M. (2004), "Saving, Risk Sharing and Preferences for Risk", *American Economic Review*, 94(4), 1169-1182

Mazzocco, M. and S. Saini (2006), "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences: Semi-parametric Tests with an Application to Village Economies", Department of Economics, University of Wisconsin, *Working Paper*

McAleer, M. (1987), "Specification tests for separate models: a survey", in M.L. King and D.E.A. Giles eds., *Specification Analysis in the Linear Model*, London, Routledge

McAleer, M. (1995), "The significance of testing empirical non-nested models", *Journal of Econometrics*, 67(1), 149-171

McAleer, M. and M.H. Pesaran (1986), "Statistical inference in non-nested econometric models", *Applied Mathematics and Computation*, 20(34), 271-311

McFadden, D. and K. Train (2000), "Mixed MNL models for discrete response", *Journal of Applied Econometrics*, 15(5), 447-470

McPeak, J. (2004), "Contrasting income shocks with asset shocks: livestock sales in northern Kenya", *Oxford Economic Papers* 56(2), 263-284

Menon, M. and F. Perali (2006), "The Shadow Wage of Child Labour: An Application to Nepal", *UCW Working Paper* No. 11

Milanovic, B. (2000), "Social transfers and social assistance. An empirical analysis using Latvian Household Survey Data", The World Bank, *Policy Research Working Paper Series* No.2328

Mishra, A., H. El-Osta, M. Morehart, J. Johnson, and J. Hopkins (2004), "Income, Wealth, and the Economic Well-Being of Farm Households", Economic Research Service, U.S. Department of Agriculture, *Agricultural Economic Report* No. 812

Morduch, J (1990), "Risk, Production and Saving: Theory and Evidence from Indian Households", Harvard University. Mimeo

Morduch, J. (1994), "Poverty and Vulnerability", *American Economic Review*, 84(2), 221-225

Morduch, J. (1995), "Income Smoothing and Consumption Smoothing", *Journal of Economics Perspectives*, 9(3), 103-14

Morduch, J. (2005), "Consumption Smoothing Across Space: Testing Theories of Risk-Sharing in the ICRISAT Study Region of South India", in S. Dercon (ed.), *Insurance against Poverty*, Oxford: Oxford University Press for WIDER

Newhouse, D.L. (2003), "Household Income Dynamics in Indonesia, 1993-1997", Ph.D. Thesis, Cornell University, Ithaca, NY

Newhouse, D. (2005), "The Persistence of Income Shocks: Evidence from Rural Indonesia", *Review of Development Economics*, 9(3), 415-433

Pan, L. (2007), "Risk Pooling through Transfers in Rural Ethiopia", *Tinbergen Institute Discussion Paper* No. 07-014/2

Paxson, C. (1992), "Using weather variability to estimate the response of savings to transitory income in Thailand", *American Economic Review*, 82(1), 15-33

Ranjan, P. (2001), "Credit constraints and the phenomenon of child labor", *Journal of Development Economics*, 64(1), 81-102

Ravallion, M. and Q. Wodon (2000), "Does Child Labour displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy", *Economic Journal*, 110(462), C158-75

Ravallion, M. and S. Chaudhuri (1997), "Risk and Insurance in Village India: Comment", *Econometrica*, 65(1), 171-184

Rosenzweig, M.R. (1988), "Risk, Implicit Contracts and Family in Rural Areas of Low-Income countries", *The Economic Journal*, 98(393), 1148-1470

Rosenzweig, M.R. and K.I. Wolpin (1993), "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India", *The Journal of Political Economy*, 101(2), 223-224

Sahlins, M. (1972), *Stone Age Economics*, Aldine-Atherton, Inc., Chicago.

Santos, P., and C.B. Barrett (2006), "Informal insurance in the presence of poverty traps: Evidence from southern Ethiopia", Mimeo, Cornell University

Sawada, Y. (2003), "Income Risks, Gender, and Human Capital Investment in a Developing Country", *CIRJE F-Series* No. CIRJE-F-198, CIRJE, Faculty of Economics, University of Tokyo

Sawada, Y. and M. Lokshin (2001), "Household Schooling Decisions in Rural Pakistan", The World Bank, *Policy Research Working Paper Series*

Sawada, Y. and M. Lokshin (2007), "Obstacles to School Progression in Rural Pakistan: An Analysis of Gender and Sibling Rivalry Using Field Survey Data", *CIRJE Discussion Paper* No.CIRJE-F-484

Schultz, P. and A. Tansel (1997), "Wage and Labour Supply Effects of Illness in Cote d'Ivorie and Ghana: Instrumental Variable Estimtes for Days Disabled", *Journal of Development Economics*, 53(2), 251-286

Terza, J.V. and P.W. Wilson (1990), "Analyzing frequencies of several types of events: a mixed multinomial-Poisson approach", *Review of Economics and Statistics*, 72(1), 108-115

The PROBE Team (1999), *Public Report on Basic Education in India*, Oxford University Press, USA

Thomas, D, K. Beegle, E. Frankenberg, B. Sikoki, J. Strauss, and G. Teruel (2004), "Education during a Crisis", *Journal of Development Economics*, 74(1), 53-86

Thomas, D., D. Contreras, and E. Frankenberg (1999), "Child health and the distribution of household resources at marriage", RAND, Santa Monica, Calif., U.S.A.

Townsend, R.M. (1994), "Risk and Insurance in Village India", *Econometrica*, 62(3), 539-591

Townsend, R.M. (1995), "Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies", *Journal of Economic Perspectives*, 9(3), 83-102

Udry, C. (1991), "Credit Markets in Northern Nigeria: Credit as Insurance in a Rural Economy, *World Bank Economic Review*, 4(3), 251-71

Udry, C. (1994), "Risk and Insurance in a Rural Credit Market: An empirical Investigation of Northern Nigeria, *Review of Economic Studies*, 61(3), 495-526

Udry, C. (1995), "Risk and Saving in Northern Nigeria", *The American Economic Review*, 85(5), 1287-1300

Udry, C. and P. Bardhan (1999), "Risk and Insurance in an Agricultural Economy", ch.8 in *Development Microeconomics*, Oxford: Oxford University Press

UNDP (2004), *Indonesia: progress report on the Millenium development goals*, Jakarta: Coordinating Minister for People's Welfare, 2004

UNECE *et al*. (2007), *The Wye Group Handbook. Rural Households' Livelihood and Well-Being*, United Nations Publication

Witoelar, F. (2005), "Inter-household Allocations within Extended Family: Evidence from the Indonesia Family Life Survey", Economic Growth Center, Yale University, *Center Discussion Paper* No. 912

World Bank (2000), *World Development Report 2000/01: Attacking Poverty*, New York: Oxford University Press

Zeldes, S.P. (1989), "Consumption and Liquidity Constraints: An Empirical Investigation", *Journal of Political Economy*, 97(2), 305-346

Zimmerman, F. (1993), *Structural Evolution under Imperfect Markets in Developing Country Agriculture: A Dynamic Programming Simulation*, Department of Economics and Global Studies Research Program, University of Wisconsin, Madison, Mimeo

Zimmerman, F. and M.R. Carter (2003), "Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints", *Journal of Development Economics*, 71, 233-260

Appendix A

Survey questions and selected variable definitions

A.1 Individual and household income, 1993 and 1997

1993 household income

Annual household income has been constructed as the sum of four components:

Table A.1
Components of household income

Component	Level
1) net profit generated by the farm and non-farm business	Household
2) asset income (for non-business assets)	Household
3) individual wages	Individual
4) household non labour income	Household

1) Net profit generated by the farm and non-farm business.

The household head was asked to report the approximate amount of net profit generated by the farm or non-farm business during the 12 months prior to the interview. If respondents did not know the net profit, they were asked to estimate the total revenue received by the household from the farm or non-farm business and the approximate amount spent by the household for the farm business. Net profit can be constructed subtracting the expenses from the gross income.

2) Asset income (for non-business assets).

The household head estimated and reported the total income from the rent/lease/profit-sharing of each type of non-business assets in the 12 months prior to the interview.

3) Individual wages.

Adult household members (the household head and the spouse, if present, are selected, and up to three other members) were randomly selected and asked to provide detailed individual employment information. Private or government workers were asked to report their net salary or wage from both primary and additional job during the 12 months prior to the interview. For all the other household members not interviewed as main respondents and older than 10 years, the head reported the individual net income/wages (other than from the farm or non-farm business). The sum of individual (main and non-main respondents) wages is the total household income from the wage employment sector.

4) Household non labour income.

The household head was asked to estimate the non-labour income received by the household during the 12 months prior to the interview. Sources of income are: pension, scholarship/student loan, insurance claim, winnings (from rotating credit associations, lotteries, etc), gift from family, friend or charities, and other.

1993 individual labour income (used in the analysis conducted in chapter four)

Individual labour income consists of earnings or profits reported by individual respondents or by the head of the household.

Main respondents (adult household members, mainly the household head and the spouse, if present, and up to three other selected members):

self-employed workers were asked to report the net profit/gross income they gained in the 12 months prior to the interview (monthly profits are reported too). The gross income was reported if respondents did not know their net profit. Private or government workers were asked what was the net salary or wage during the 12 months prior to the interview.

Other household members (non-main respondents):

as said above, for each household member over the age of ten and not interviewed in the individual book (as main respondent), the household head reported whether he/she worked to earn income/wage/salary, what was the main work status, and the net income/wage/salary earned from his/her work (other than from the farm or non-farm business) during the 12 months prior to the interview.

1997 household income

Farm and non-farm profit, non labour income and asset income are defined as in 1993 survey. The questionnaire on individual labour income is not available in 1997.

A.2 Household expenditure, 1993 and 1997

1993

This dissertation uses the household expenditure measure created by IFLS users. Total household expenditure is the sum of all expenditures on food and non-food items. Food expenditure includes the value of food purchased, the value of household consumables that were self-produced, and the value of food received from other sources outside the household. Several food items are reported. The questionnaire asked for food expenditures during the week prior to the interview; the original variable is then converted to monthly value.

Non-food items include electricity, water, fuel and the like, personal toiletries, household items, recreation and entertainment, transportation, clothing, medical costs, ritual ceremonies, charities and gifts, taxes, household supplies and furniture and other expenses like the purchase of cars, television sets, beds, livestock and the like. Expenditures of child's education (for children inside and outside the household) are reported. Non-food expenditures are collected as monthly or annual value.

Non-durable expenditure (used in chapter three) is constructed subtracting the value of household supplies, furniture and "other expenditures" (like purchase of car, television, beds, etc) from total household expenditure.

1997

The 1997 survey (IFLS2) asked the same questions on household expenditure as in 1993. The total household expenditure is constructed as the sum of all consumption items.

A.3 Household assets, 1993 and 1992

Total household assets is the sum of 1993 total (market) value of farm, business non-farm and other non-business assets owned by the household, as reported by the head of household. Farm assets are: farm land, hard, stem plants, house or building used for the farm business, livestock/poultry/fish pond, vehicles, tractor, heavy equipments, small tools, other assets.
Business non-farm assets are: land and buildings other that those used for a farm business, vehicles, ships/boat, other vehicles, office equipments, supplies, other assets.
Non business assets are: house occupied by the household, buildings, land and animals not used for farm or non-farm business, vehicles, household appliances, savings/certificate of deposit, stocks, receivables, jewellery, other assets.

For each asset the value of any asset purchased and sold during the 12 months prior to the interview is reported. These variables are used, together with the 1993 value of assets, to generate the value of assets owned by the household the year prior to the interview (1992).

A.4 Household economic shocks, 1993

IFLS1 provides some questions regarding events that have affected the household and caused economic hardship during the five years prior to the interview. 1993 questionnaire is as follow:

Table A.2
Questionnaire for household economic shocks, 1993

Type of economic hardship	Has this household gone through […] within the past 5 years?	When did […] happen?	What measures where taken to overcome the economic hardship of […]?	What was the approximate cost in rupiah to overcome […] during the past 5 years?
Death of a householder	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK
Sickness of a householder that necessitated hospitalization or continuous medical treatment	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK
Crop loss	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK
Household/business loss due to fire or other disasters	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK
Unemployment of a householder	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK
Cut in household income due to falling price of crop	1.yes 3. no	Year ___ Month ___	A B C D E F G H I J K L M	1. value 8. DK

Code for measures used to overcome the shock:

A. A householder's extra job
B. A householder's loan from family or friend/employer
C. A householder's loan with crop as collateral
D. A householder's loan with other assets as collateral
E. Selling next harvest in advance (below market value)
F. Selling food (rice) supply
G. Selling cattle/poultry
H. Selling jewelry

I. Selling other assets
J. Using savings
K. Getting money from family/relatives, friend or employer
L. Getting assistance (in the form of goods) from family/relatives, friend or employer
M. Cutting household expenses

Responses have been aggregated to six indicators: extra job (A), loan (answers B, C, D), asset sale, (answers E, F, G, H, I), family assistance (K, L), use savings (J) and cut expenses (M).

A.5 Years of schooling, 1993

The number of years of schooling has been constructed from respondent's report of highest level of school attended and the highest grade completed within level. Elementary education is typically six years, junior and senior secondary education are three years.

A.6 Working status, 1993

Respondent is define as a worker whether he/she reported "working/trying to work/helping to earn income" as the primary activity during the week prior to the interview. The respondent is defined as a worker even if he/she reported working for at least one hour during the past week, or having a job or business but temporary not working, or working at a family owned business during the past week. Working respondents were then asked to report their work status, for both primary and additional job:

1. self-employed (without help)
2. self-employed (with help of householders/temporary workers)
3. self-employed (with help of regular workers)
4. government worker/employee
5. private worker/employee
6. family worker

A.7 IFLS Provinces

In 1993, about 7200 households were interviewed in 13 provinces representing 83% of the Indonesian population. Indonesia is a country with a high variation in culture, geography and economic conditions. The relevant feature of the IFLS sample is that the 13 provinces included in the survey are able to capture this heterogeneity. The table reports the codes and names of IFLS provinces.

Table A.3
IFLS Provinces, codes and names

Code	Name
12	North Sumatra
13	West Sumatra
16	South Sumatra
18	Lampung
32	West Java
33	Central Java
34	Di Yogyakarta
35	East Java
51	Bali
52	West Nusa Tenggara
63	South Kalimanatan
73	South Sulawesi

Figure A.1: IFLS Provinces

On-line source: http://chd.ucla.edu/IFLS/index.html