Shocks and Coping Strategies in Sub-Saharan Africa

Douglas Scott

(ID 4247928)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

Nottingham, September 2019

Douglas Scott: *Shocks and Coping Strategies in Sub-Saharan Africa*.
Supervised by Simon Appleton, Sarah Bridges and Oliver Morrissey.
September 2019.

Contents

List of Tables

List of Figures

Introduction

Throughout the developing world, the lives of poor individuals are characterised by the constant presence of risk. The potential for misfortune to strike at any time shapes decisions on how to earn income, manage finances and invest for the future. When the worst does happen, the formal financial institutions enjoyed by those in richer countries will often be unavailable or ill-suited to the needs of those who could benefit from them the most. This may leave some groups with no choice but to take actions such as selling productive assets, removing children from school or risking long-term health consequences by reducing vital food consumption. Such actions could potentially trap them or their families in a state of persistent poverty for many years to come. Therefore, in the absence of effective coping strategies, shocks may have enduring consequences.

This thesis presents studies on three Sub-Saharan African countries, where each study considers the broad topic of shocks and coping strategies. Chapter 1 focuses on a recent period of severe drought in Ethiopia and the motivation behind a household's choice of response to the subsequent income shock. This study provides strong evidence that some groups were faced with a choice between protecting short-run consumption and ensuring future wealth and stability. The main contribution of this chapter is to demonstrate how the study of coping strategies can be used as a means of gaining insight into the economic environment in which households operate. In this instance, the observed heterogeneity in shock-responses is shown to be consistent with the presence of a poverty trap threshold, defined by the number of cattle owned prior to the onset of the drought.

The second chapter considers the topic of shocks to early childhood health, resulting from exposure to the 20-year war between the government forces of Uganda and the rebel group known as the Lord's Resistance Army. Alongside the clear physical and psychological damage caused by the fighting, the families of those children caught up in the conflict will often have experienced substantial losses to both income and assets as a result of the war. Such negative shocks will have been exacerbated by the break-down of the systems underpinning traditional coping mechanisms, such as functioning markets to allow the sale of assets or support networks to provide assistance. This study finds evidence of irreversible health deficits amongst children exposed to the conflict for more than six months. However, the findings also suggest that the impact on children's health may be far less severe for shorter periods of exposure.

The final chapter looks at both shocks and coping strategies in the context of two South African townships. The study categorises a wide variety of economic shocks by severity, before analysing both the choice of coping strategy and its effectiveness in protecting vital consumption. Evidence is found that households are capable of effectively responding to even the most severe shocks by utilising existing savings or obtaining assistance. However, when neither of these strategies are put in place, there is evidence of substantial shortfalls in food consumption in the week the event takes place, and further reductions in the week before and after the shock. The results suggest that it is the absence of any potential coping strategy which identifies those most at risk of experiencing the worst outcomes.

Acknowledgements

(In Alphabetical Order)

I am deeply grateful to my supervisors Simon Appleton, Sarah Bridges and Oliver Morrissey for their invaluable support and advice throughout my PhD. I also owe sincere thanks to Gianni De Fraja, Marcus Eberhart, Simon Quinn and Chris Woodruff for comments on the Ethiopia paper and Rumman Khan for his assistance in gathering data for the Ugandan study. In addition, I am grateful for the constructive feedback I received presenting papers at the Royal Economic Society Junior Symposium and European Economic Association Conference, with thanks also going to Abigail Barr, Jake Bradley and Sourafel Girma for their extremely useful comments during seminars here in Nottingham. I would also like to thank the University of Nottingham for the generous financial support provided to me over the course of my PhD.

Finally, I would like to thank my family for their unwavering support over the last few years, without which I would not be where I am today.

Chapter 1

Income Shocks and Poverty Traps: Asset Smoothing in Rural Ethiopia

1.1 Introduction

The dynamics of poverty in developing countries has attracted increased attention in recent years. A small, but growing, literature has sought to identify mechanisms which serve to trap specific groups in a state of persistent deprivation, while others transition to a higher standard of living (Carter and May, 2001; Jalan and Ravallion, 2004; Lokshin and Ravallion, 2004; Lybbert et al., 2004; Adato et al., 2006; Barrett et al., 2006; Carter et al., 2007). The notion of a dynamic system separating those who will move out of poverty, given time, from those who are trapped indefinitely, is of clear interest from a policy perspective. For example, identifying such a process *ex-ante* would allow for tailored insurance instruments or safety nets to be targeted directly towards those most at risk, while also making it possible to identify where a relatively small intervention would be capable of pushing borderline individuals onto a higher trajectory.

This paper employs an innovative approach to detecting such poverty traps, through heterogeneity in the coping strategies used by households in response to an income shock. Applying this methodology to a nationally representative dataset of rural households in Ethiopia, evidence is found of two distinct patterns of response to the onset of a period of severe drought. Households with initial cattle holdings of three or more animals effectively used these assets to insulate consumption expenditure against the drought-induced fall in income. In contrast, those households with smaller herds often elected to preserve their cattle stocks, at the expense of lower levels of consumption. These findings are consistent with the hypothesis of an asset-based poverty trap, defined by the pre-shock level of household cattle holdings.

1.1.1 Asset-based Poverty Traps

The theoretical literature has suggested a wide variety of economic relationships which have the potential to generate some form of poverty trap.[1] However, the specific focus of this study is a bifurcation in the accumulation of assets, whereby a household's convergence to either a high or low equilibrium is determined by initial asset holdings (Lybbert et al., 2004; Carter and Barrett, 2006; Adato et al., 2006; Carter et al., 2007; Quisumbing and Baulch, 2013). This type of separation could potentially occur through a variety of channels, such as a correlation between wealth and attitudes to risk, differential access to credit, or even selective exclusion from informal risk-sharing and lending networks (Banerjee, 2004; Carter, 1988; Santos and Barrett, 2011). Indeed, the theoretical literature has suggested a number of mechanisms which have the potential to generate a separation in the growth of asset wealth, yet the empirical evidence in support of this poverty trap hypothesis is often restricted to specific populations or groups of assets. In addition, the most appropriate strategy for identifying such mechanisms also remains unclear.

[1] Some notable examples include, Eswaran and Kotwal (1986) on credit market imperfections and land efficiency, Dasgupta and Ray (1986) on the links between nutrition, wages and productivity, and Mookherjee et al. (2010) on poverty traps induced by a complementarity between aspirations and educational attainment.

This paper will proceed as follows. Section 1.1.2 presents a review of the current literature on asset-based poverty traps, with a particular focus on those studies relating to Ethiopia. Section 1.2 demonstrates how shock-responses may be influenced by the presence of a poverty trap threshold. An overview of the data used in this analysis can be found in Section 1.3, while Section 1.4 discusses the empirical strategy used to estimate the impact of the drought. The main results of the analysis can be found in Section 1.5, while Section 1.6 tests the robustness of these results to a number of alternative assumptions. The possibility of heterogeneity within groups of pre-shock cattle holdings is considered in Section 1.7, before Section 1.8 concludes.

1.1.2 The Analysis of Poverty Traps

Until recently, studies which have set out to identify asset-based poverty traps have focused on modelling the dynamic relationship between current and lagged asset holdings, over time (Carter and May, 2001; Lybbert et al., 2004; Adato et al., 2006; Barrett et al., 2006). Due to the need to model a potentially complex growth path, containing multiple equilibria, estimating this relationship often relies on non-parametric techniques. Lybbert et al. (2004) employ such an approach in studying a small group of pastoralist households living on the Borana Plateau in southern Ethiopia. The authors find that, where a household's herd size exceeds 15 cattle, in any given period, the future growth path of this asset will tend towards a stable, high equilibrium of 75. In contrast, those households falling below this threshold level (either in the initial period or as a result of a shock to their livestock holdings) will instead converge to a lower equilibrium. Following a similar methodology, Barrett et al. (2006) find comparable results among a group of pastoralist households in northern Kenya, although a lower threshold of 5-6 cattle was found in the Kenyan sample. A study of households in KwaZulu-Natal, South Africa, conducted by Adato et al. (2006), avoids the need to focus attention on one specific asset by constructing a livelihood-weighted asset index, based on the value of assets required to attain a level of consumption equal to the national poverty line. The study concludes that households with sufficient assets to generate a consumption level of approximately twice the poverty line would converge to a higher level of wealth over time, while those below this threshold would instead collapse to a lower, long-run equilibrium.

Although the non-parametric approach used in these studies provides the necessary degree of flexibility to estimate a complex, non-linear relationship, it relies on the assumption that a population's asset dynamics can be accurately represented using only a bivariate relationship (between current and lagged assets). Within such models all heterogeneity in the units of analysis is necessarily disregarded, as are external factors potentially influencing changes in herd size over time. This brings into question the policy value of conclusions based solely on this approach. Nevertheless, to enable comparison with the main results of the paper, the Ethiopian data employed in this study is analysed using a similar non-parametric method to Lybbert et al. (2004). The estimated growth path can be found in the preliminary results section of this paper.

Following the work of Jalan and Ravallion (2004) and Lokshin and Ravallion (2004), recent contributions to the literature have employed more parametric methods to attempt to identify multiple equilibria in the evolution of asset stocks. These studies have the advantage, over the papers mentioned previously, in that they permit the inclusion of covariates within the analysis. In the most commonly applied methodology, the current level of the asset stock is regressed on its lagged value, where the lagged value is expressed as a number of polynomial terms, to allow for both a concave and convex relationship over the range of asset wealth. Quisumbing and Baulch (2013) use this approach, alongside the non-parametric methods previously discussed, in their large-scale study of rural households in Bangladesh. Within this sample, they identify only a single low-level equilibrium in both the parametric and non-parametric estimations. The authors suggest the presence of relatively well-functioning factor markets as the reason for the contrasting results to the African studies. However, using a similar methodology, Giesbert and Schindler (2012) find evidence of only a single low-level equilibrium, using a nationally representative panel of households in rural Mozambique. The presence of only a single equilibrium in these studies mirrors the results obtained by applying the lagged polynomial approach to the data used in this analysis (see Section 1.5.1).

Two important contributions, from the perspective of this paper, come from Van Campenhout and Dercon (2012) and Naschold (2013). Both papers analyse the asset accumulation path in a sample drawn from 15 Ethiopian villages. Unlike the previous studies mentioned, Van Campenhout and Dercon (2012) employ a piecewise linear specification, in an attempt to identify a poverty trap in tropical livestock units. They find evidence of a discontinuity at approximately 7 units (around 7-10 cattle). However, when applying a number of non-parametric and lagged polynomial approaches to a sample from the same dataset, Naschold (2013) finds no evidence of the multiple equilibria growth path suggested in the first paper. Contradictory findings, such as these, highlight the degree to which parametric results may be highly dependent on the functional form assumptions imposed upon the data. In addition, accurately estimating a continuous asset growth path, using a parametric model, requires a sample containing observations at every level of the asset. This is of particular concern, as the dynamic system implied by a poverty trap would see all households converge to one of two extremes. The data must, therefore, contain sufficient exogenous shocks to ensure that the area around the threshold remains populated (Naschold, 2013). This final comment suggests an alternative approach to identifying a level at which asset growth bifurcates.

In an environment where households have limited access to credit, Deaton (1991, 1992) demonstrates that households can effectively protect vital consumption, from all but the worst shocks to income, through the sale and accumulation of asset stocks. His seminal buffer stock model is particularly applicable to low-income, rural communities, who have little or no access to formal credit or savings. There exists a well-established literature testing the implications of the buffer stock theory.[2] However, it is often observed that it is the rich, rather than the poor, who employ such strategies to smooth consumption (Kazianga and Udry, 2006; Carter et al., 2007; Verpoorten, 2009). A potential explanation for such heterogeneity can be found in Hoddinott (2006), whose research on resettled households in rural Zimbabwe found that consumption smoothing through cattle stocks was more likely amongst households with a minimum of three cattle. In an attempt to explain this result, Hoddinott speculates that, in electing to sell some of their herd, households with a higher number of cattle would still be left with sufficient animals to plough their fields in the following season. If this hypothesis were accurate, such a non-linear relationship between herd size and agricultural productivity could plausibly generate the type of separation in asset wealth described in the previous section.

Observable changes in behaviour in response to an income shock opens up the possibility of identifying a poverty trap threshold indirectly, through changes in the shock-responses of households close to this level of assets. This concept is closely related to the work of Carter and Lybbert (2012), who assume the existence of a threshold in asset wealth to identify two separate consumption smoothing regimes, using a sample of households from an earlier study on rural Burkina Faso. While the original analysis found only weak evidence in support of the conventional buffer stock theory (Kazianga and Udry, 2006), Carter and Lybbert concluded that the behaviour of households above the threshold conformed closely to the predictions of the model. Those with smaller herd sizes, however, did not appear to smooth consumption in this manner, in spite of having sufficient animals to do so.[3]

[2]See Rosenzweig and Wolpin (1993), Fafchamps et al. (1998), Kinsey et al. (1998).

[3]The value of livestock sales was sufficient to offset between 80-90% of the shock to income in households above the threshold. Below this threshold, however, only 25-30% of the income shock was smoothed using this strategy (Carter and Lybbert, 2012).

This study makes two key contributions to the literature. Firstly, the discovery of heterogeneity in the evolution of livestock holdings, in different Ethiopian populations and time-periods, provides sufficient motivation for a larger, more representative analysis of rural livestock dynamics. This paper provides such an analysis. The second contribution is methodological. As discussed, the majority of previous attempts to detect multiple equilibria in asset growth, are either limited by an inability to control for the influence of covariates or are overly dependent on the initial assumptions regarding functional form. Carter and Lybbert's finding, that a desire to protect assets may override consumption smoothing motives in the vicinity of a poverty trap threshold, suggests an alternative. This paper seeks to build on this result, by using systematic differences in rainfall shock responses as an indirect approach to detecting the presence of a threshold in the accumulation of asset wealth.

1.2 Theoretical Framework

Angus Deaton's seminal buffer stock model is the most widely applied theory of household consumption smoothing through asset stock adjustment (Deaton, 1991, 1992). The model shares many characteristics with Friedman's Permanent Income Hypothesis (PIH) (Friedman, 1957), while diverging from the standard intertemporal choice model in the underlying assumptions regarding access to credit. Deaton's theory considers risk-averse households, who seek to smooth consumption, given a stochastic stream of income. However, in contrast to the PIH, agents are not presented with an unlimited capacity to borrow and save. Instead, the model considers the case where households are only able to insulate consumption from negative income shocks through the sale and accumulation of a buffer stock of assets. At each point in time, households will attempt to bring their consumption paths as close to the PIH benchmark as the borrowing constraint permits. However, where a household experiences a negative shock to income, with insufficient assets to fully offset the shortfall, all current income and the value of any available buffer assets will be consumed, with nothing carried forward to the next period.

If there exists a threshold level in the accumulation of asset wealth, the use of assets to fully offset the impact of a negative income shock could have potentially serious, long-run implications for households close to this level. As noted in Carter and Zimmerman (2003), if a negative shock is substantial enough that fully smoothing consumption would require a household to reduce its asset stock to a level below the poverty trap threshold, the potential consequences of such an action may instead lead households to consider allowing consumption to fall in the short-run (at least to some degree), in order to protect a level of asset wealth above the threshold. If the capacity exists to withstand a fall in consumption, or households value the long-run accumulation of assets, over the short-run stability of consumption, this strategy of asset smoothing will represent the optimal shock response.[4]

$$\max_{c\ A} \quad E_0\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right] \tag{1.1}$$

Subject to:

$$y_t(A_t\theta_t) = max\ [f^{\ell}(A_t\theta_t), f^h(A_t\theta_t)]$$

$$x_t = y_t(A_t\theta_t) + (1-\delta)A_t$$

$$A_{t+1} = x_t - c_t$$

$$c_t \leq x_t \qquad \forall t$$

$$A_t \geq 0 \qquad \forall t$$

It is possible to simulate this form of heterogeneity in coping strategies using an adaptation of the duel-technology model proposed by Carter and Lybbert (2012). In model (1.1), risk-averse households seek to maximise expected utility, over

[4] It is important to recognise, however, that if consumption levels also lie close to the type of nutritional threshold suggested in Dasgupta and Ray (1986), or reducing consumption leaves vulnerable members at risk of the long-term consequences of malnutrition (Alderman et al., 2006; Dercon and Porter, 2014), a household may simply be side-stepping one poverty trap, only to fall into another.

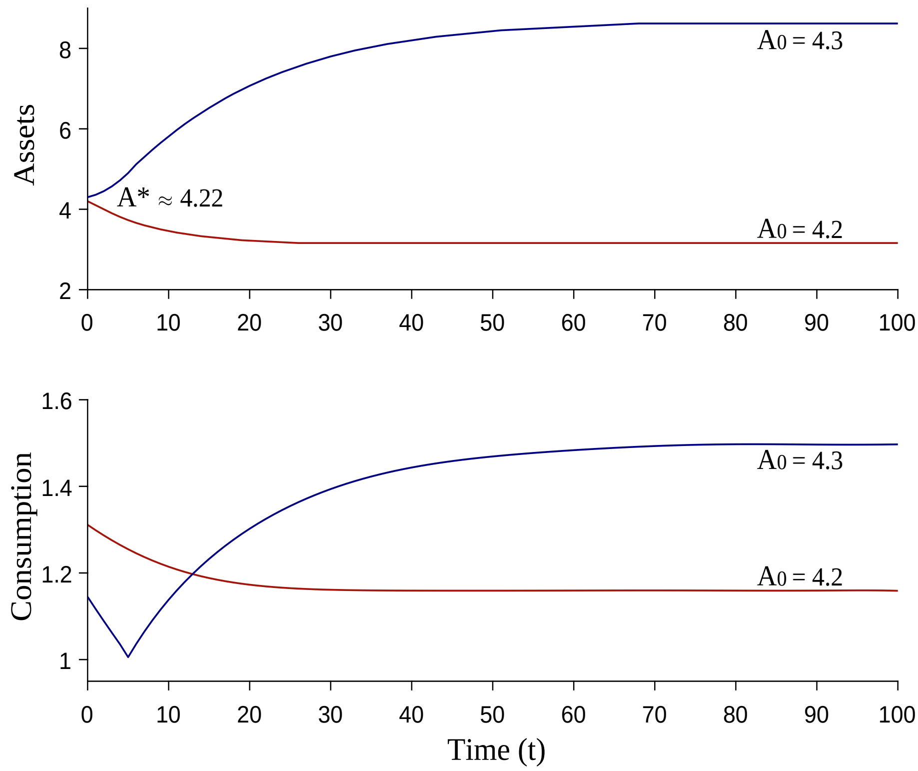
an infinite horizon, through choosing how to allocate current resources between consumption and the accumulation of a productive asset A.

The first constraint defines income as dependent on a household's choice of a low or high technology $f^{\ell}(A_t\theta_t)$ or $f^{h}(A_t\theta_t)$, where the realised income from either choice is a function of the current level of assets A_t and the realisation of a stochastic variable θ. Expected marginal returns are greater under the high technology at all levels of the asset stock, but adoption entails an initial fixed cost. The term x_t, in the second constraint, refers to 'cash-on-hand', a measure of the total resources available to the household in the current period (Deaton, 1991, 1992). This is comprised of current income and the value of any assets held, subject to a depreciation rate δ. The equation for the evolution of the asset stock simply imposes the condition that assets in the following period equate to current period cash-on-hand, less current consumption. The condition $c_t \leq x_t$ indicates that cash-on-hand places an upper limit on consumption, given no access to credit, while the final constraint implies that the value of the asset stock is non-negative in all periods. A set of optimal choices, for different initial asset stocks, can be obtained through expressing model (1.1) in Bellman equation form and solving numerically.[5] The functional form and parameter assumptions used to generate the following growth paths can be found in Appendix 1.1.

Model (1.1) generates the type of bifurcation in asset growth described in Section 1.1.2. The uppermost panel of Figure 1.1 illustrates the movement towards a high and low asset equilibrium, for households with initial asset stocks immediately above and below a separating threshold of $A^* \approx 4.22$. While $A_0 = 4.3$ implies adoption of the high return technology, $A_0 = 4.2$ does not. In the example shown, however, $A_0 = 4.3$ is still not sufficient to allow immediate adoption of the high technology. Instead, it comes at the initial cost of foregone consumption in the early time periods, shown in the lower panel of Figure 1.1.

[5]The Bellman equation relating to model (1.1) is $V(A_t) = \max_{c \leq x}\,[u(c_t) + \beta \int V(x_t - c_t) d\Omega\theta]$, where $\Omega =$ the cdf of $[\theta]$.

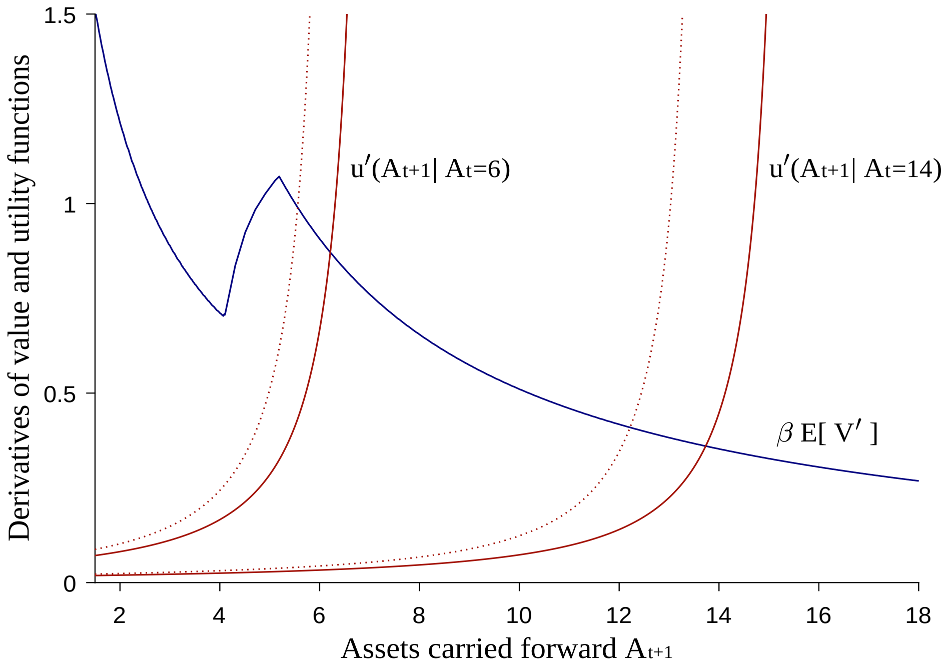
Figure 1.1: Time-Paths of Assets and Consumption



The presence of the bifurcating threshold A^* also generates heterogeneity in income shock responses. The clearest illustration of this comes through analysing optimising behaviour, based on the first-order condition from the Bellman equation. This condition implies that households should consume available resources until the marginal utility of consumption equals the (discounted) expected value of carrying assets forward into the next period, $u'(c_t) = \beta E(V')$. Following Carter and Lybbert (2012), Figure 1.2 provides a graphical representation of this relationship, as a function of A_{t+1}.

The key feature of Figure 1.2 is that the marginal value curve $\beta E(V')$ is not strictly decreasing in assets carried forward. Over the optimal range of choices of A_{t-1}, defined by the threshold $A^* \approx 4.22$ and the level of assets required to adopt the high technology (≈ 5.08) the value of holding additional assets increases rapidly. Over this range, households will be willing to sacrifice consumption to accumulate sufficient assets to optimally choose the high return technology (see Figure 1.1). Having achieved this minimum asset level, however, households faced with having to return to the low technology, in the event of an unexpected shock to income, will also be willing to sacrifice relative more consumption to ensure they remain in this position (compared to those at higher levels of the asset distribution).

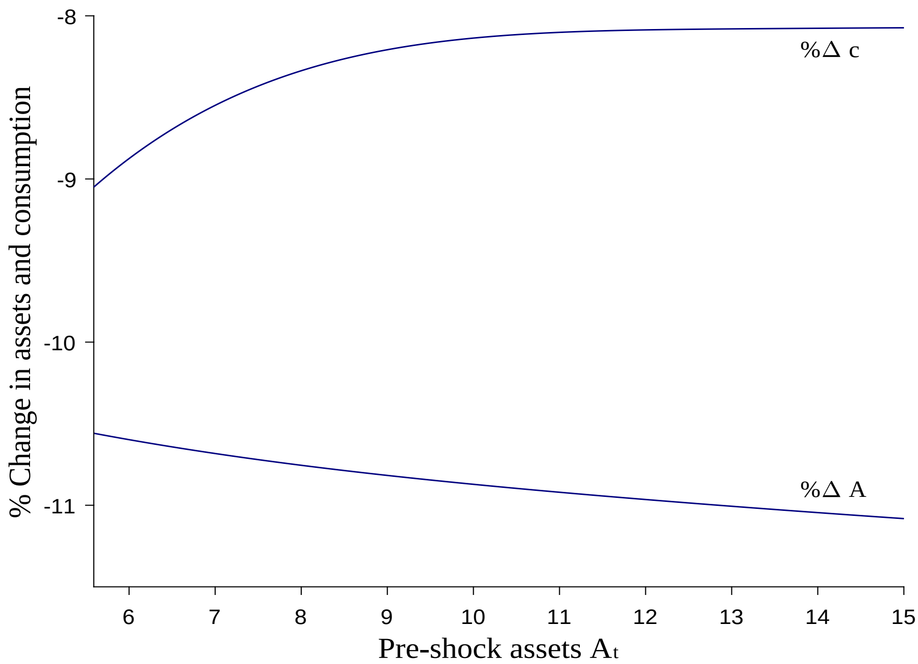
Figure 1.2: The Response of Assets and Consumption to an Income Shock



To illustrate this final point, the marginal utility curves for two representative levels of the initial asset stocks ($A_t = 6$ and $A_t = 14$), are indicated by the solid u' curves in Figure 1.2. The optimal choice of assets carried forward in each case is found at the intersection with the marginal value curve $\beta E(V')$. The dashed lines represent the marginal utility of consumption, for the same initial asset levels, after a severe, unexpected shock to income (introduced via a 0.4 downward shift in the entire distribution of the stochastic term θ in model (1.1)). The two new points of intersection with $\beta E(V')$ imply that a higher proportion of the shock will be absorbed through a reduction in assets by the household with $A_t = 14$, while the household closer to the high technology adoption level will respond with a relatively larger reduction in consumption.[6]

Figure 1.3 illustrates the percentage change in both the asset and consumption variables, induced by the shock to income. The figure clearly illustrates how the model generates heterogeneity in the extent to which households favour asset or consumption smoothing, dependent on pre-shock asset holdings A_t.

Figure 1.3: Heterogeneous Shock Responses



[6] This can be seen in the relative increases in the marginal utility of consumption (the vertical change between the pre-shock and post-shock points of intersection in Figure 1.2).

The model described in this section leads to a number of testable predictions. If an asset threshold A^* does exists, it should be possible to define an initial level of assets, henceforth denoted γ, as the point at which proximity to the threshold induces households to switch from a strategy which prioritises smoothing consumption to a strategy which also values asset protection. The level $\gamma \geq A^*$ would serve to separate households into two distinct shock-response regimes.

Proposition 1: For those households who's pre-shock assets exceed γ, a marginal decrease in transient income should induce a smaller reduction in consumption, relative to those below γ. Instead, asset wealth should be drawn down as a buffer against the shock.

Proposition 2: For those households below γ, an income shock will induce an asset smoothing response and any asset reduction will be of a smaller magnitude than that observed for households in the higher asset regime. As a consequence, consumption in the low asset group will be permitted to fall to a relatively greater degree.

1.3 Data and Descriptive Statistics

The specific years considered in this analysis mark the onset of a period of extreme drought in Ethiopia. Weak or failed rains were reported in some regions as early as mid-2014, and throughout the following year the situation became progressively worse (FAO, 2015). Driven by the unpredictable El Niño weather patterns, the country experienced the initial failure of the early *belg* rains in 2015, followed by low or erratic rainfall in the main *meher* season.[7] The resultant poor

[7]Ethiopia typically experiences two rainy seasons, the relatively smaller *belg* rains, between March and May, before the main *meher* rains, between June and September. 90% of cereal production is derived from the larger *meher* season (Taffesse et al., 2012).

harvests, particularly in the North and East of the country, prompted the need for emergency food assistance to be distributed to an estimated 10.2 million people, in addition to the 7.9 million already protected by the government's Productive Safety Net Programme (HRD, 2015). Although the drought was considered to be one of the worst the country has experienced, improvements in Ethiopia's capacity to respond to such events prevented the type of widespread famine experienced in the 1980s. Nonetheless, the loss of agricultural income suffered by small-scale farmers during the period was extremely high, with the fall in grain output ranging from between 10 and 30%, in the Tigray, Somalie and Gambela region, to between 45 and 66% in Afar and Dire Dawa (CSA, 2016).

The data used in this study comes from two waves of the Ethiopian Socioeconomic Survey (ESS), representing the pre-shock period 2013-14 and the height of the drought in 2015-16. The initial sample was selected to be representative of all households in rural Ethiopia. In accordance with the previous literature, the natural choice for the asset stock variable (over which the sample splitting point γ is defined) is household livestock holdings (Lybbert et al., 2004; Carter et al., 2007; Van Campenhout and Dercon, 2012; Naschold, 2013). Furthermore, as the sale of cattle was found to be the most common coping response in the severe drought periods of 1984-85 and 1990-92 (Webb et al., 1992), this paper focuses specifically on cattle as the potential buffer stock asset.[8]

The two sample waves employed come from an earlier ESS survey of 3466 rural households conducted in 2011.[9] Attrition between this initial data collection and the first panel wave is minimal (approximately 4%), leaving a potential sample of 3323 households in the 2013-14 survey. Although the majority of rural households

[8]Livestock is, arguably, the most important asset in rural Ethiopia (Dercon, 2004) and cattle are found to be present in more than 90% of livestock owning households who live in rural areas (CSA, 2013).

[9]Substantial differences within the agricultural survey instrument prevents the use of the 2011 wave in this analysis.

in Ethiopia are involved in both crop and livestock agriculture, a relatively small proportion are not involved in one or other of these activities. As the population of interest is small-scale agriculturalists, households who owned no cattle in any of the three waves of the ESS survey (including the initial 2011 wave) are omitted, as are households with no cultivated land in any period. In addition, to avoid the inclusion of large-scale, commercial agriculture, households with more than 30 cattle or 5 hectares of cultivated land, in any survey wave, are also omitted from the analysis.[10] The level of attrition between the 2013-14 and 2015-16 wave is, again, minimal (approximately 1.9%). However, it is not possible to generate key variables for all observations in the selected sample. In particular, measurement of land areas and food consumption are commonly reported in local units, for a small number of which there exists no plausible method of conversion.[11] Subsequently, the usable sample of observations falls to 2572 in 2013-14 and 2494 in 2015-16.

1.3.1 Variable Summary

Table 1.1 summarises the key variables used in the following analysis. The section on *Demographics and education* indicates that households are composed of 5-6 members, on average, with a household head who is most likely to be illiterate and aged between 45 and 50 years old. The remarkably poor levels of education in rural Ethiopia are reflected by the proportion of working-age household members within the sample, who have received no formal schooling.

Total farm income reports farm revenue, minus farm expenditure, for the previous 12 months, where the measure of revenue is comprised of income from land rental and the rental of agricultural tools, combined with the (imputed) value of all

[10]The sensitivity of key finding to alternative cut-off levels is considered in Section 1.6.4.

[11]This particular source of attrition appears largely random, however. A comparison of mean household characteristics between observations included in the sample and those with missing values (including attrition between waves) can be found in Appendix 1.2.

crops produced. Deducted from this is a measure of expenditure, comprising of the cost of hired factors of production (land and labour), as well as the imputed value of crops paid to these factors as non-financial reimbursement. The nominal value of farm income is deflated spatially and temporally, following Deaton and Zaidi (2002), and reported in Table 1.1 in (1000) Ethiopian Birr, per adult-equivalent.[12]

Table 1.1: Descriptive Statistics

	Sample		Wave 1 (2013-14)		Wave 2 (2015-16)	
	mean	sd	mean	sd	mean	sd
Demographics and education						
Household size	5.54	2.15	5.53	2.12	5.54	2.18
Age of head	46.62	14.39	45.79	14.30	47.47	14.44
% of members ($15 \leq$ age ≤ 60)						
with no education	65.35	33.98	64.86	33.89	65.86	34.08
Income and consumption (000)†						
Total farm income	4.83	11.80	5.89	12.75	3.75	10.62
Total consumption	4.82	7.18	5.52	7.69	4.11	6.54
Food consumption	3.96	7.07	4.42	7.59	3.48	6.45
Agriculture and rainfall						
Number of cattle	4.09	4.16	4.40	4.23	3.77	4.06
Land cultivated (ha)	0.98	0.92	1.05	0.93	0.92	0.91
Land irrigated (ha)	0.13	0.47	0.14	0.49	0.12	0.45
Land rented (ha)	0.13	0.40	0.18	0.45	0.08	0.32
Plot TWI	12.73	1.58	12.76	1.61	12.70	1.55
Rainfall deviation (cm)	-5.30	35.56	9.90	28.00	-20.96	35.74
Location and social protection						
Distance to bank (kms)	17.50	12.00	18.22	12.78	16.76	11.09
Distance to MFI‡ (kms)	12.97	11.98	12.92	12.00	13.03	11.96
Distance to market (kms)	24.50	5.70	24.06	5.76	24.95	5.61
PSNP‡ in kebele††	0.42	0.49	0.40	0.49	0.43	0.50
Observations	5067		2572		2495	

MFI = Micro-finance Institute, PSNP = Productive Safety Net Programme

† All income and consumption figures reported in 1000 Ethiopian Birr, per adult equivalent

‡ TWI = Topographical Wetness Index. MFI = Micro-finance Institute.
PSNP = Productive Safety Net Programme.

†† A *kebele* is the smallest administrative area in Ethiopia (e.g. a group of villages)

[12] Adult-equivalency scales are reported in Appendix 1.4.

The measure of total consumption in Table 1.1 is the value of household food and non-food consumption over the previous 12 months. Food consumption is comprised of purchased, gifted and own-produced food items, from a list of 26 commonly consumed foodstuffs.[13] Non-food consumption includes everyday items, such as candles, soap, fuel and transportation, but also expenses such as clothing, education and informal funeral insurance (IDDIRs). Expenditures on taxes and levies are excluded, as are ceremonial expenses (such as weddings and funerals), which would represent one-off purchases, rather than the general pattern of household expenditure. Expenditure on durable asset purchases is also omitted. Nominal consumption values are again deflated spatially and temporally, and expressed in 1000 Ethiopian Birr, per adult equivalent.

The average number of cattle reported in Table 1.1 is seen to decline from a mean herd size of 4.4 animals in the 2013-14 wave, to 3.77 at the height of the drought in 2015. Although some of this decline may capture direct livestock mortality (FAO, 2015), it will also reflect the use of cattle sales as a buffer against the subsequent income shock. Figure 1.4 provides a frequency distribution of the herd size in the two survey waves. As would be expected in this sample of small-scale, agricultural households, this distribution is heavily weighted towards cattle holdings of less than 5 animals.[14]

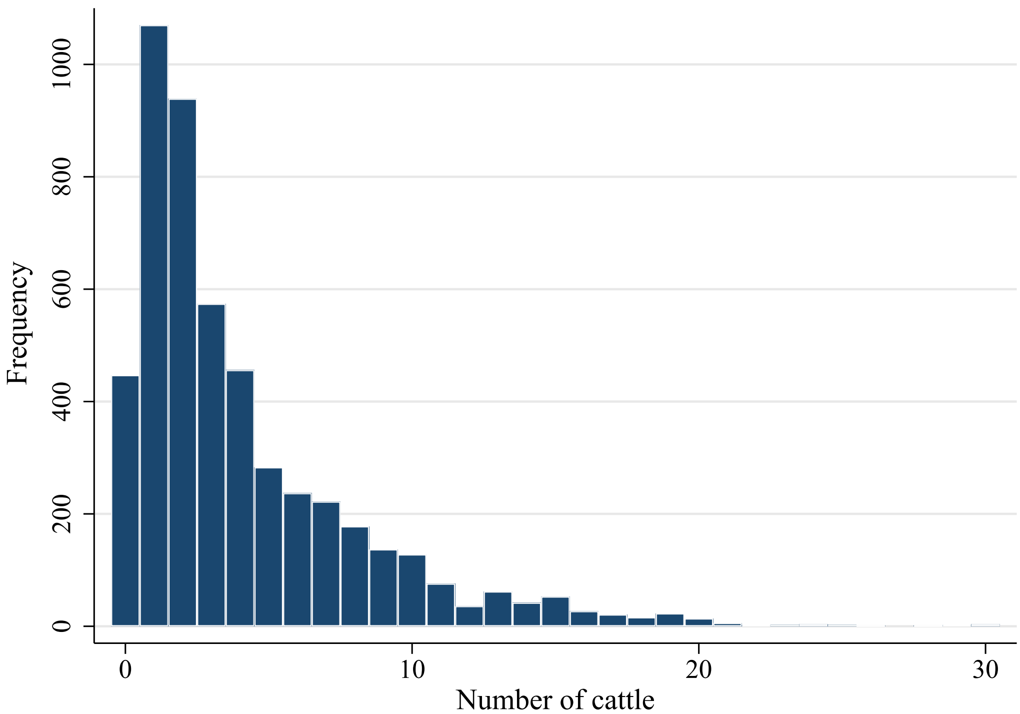
The remainder of the *Agriculture and rainfall* section of Table 1.1 reports the characteristics of cultivated land, such as the area which is irrigated or rented by the sample households. In addition, the topographical wetness index (TWI), provides a measure of the soil saturation and the topographical control of hydrological processes (Sörensen et al., 2006). In the following analysis this variable is

[13] See CSA (2017) for the full list of food and non-food items included in the final calculated value of total consumption.

[14] The distribution does not appear obviously bimodal, which would perhaps have provided early evidence of the presence of a high and low equilibrium in general asset wealth. However, the observed distribution may simply reflect the influence of shocks to cattle holdings disrupting movement towards either extreme.

separated into three groups, representing low, medium and high TWI.[15] The rainfall deviation variable records one-year deviations from the average annual rainfall between 2001 and 2015. The effects of the drought are clearly visible in the second wave of the survey, where the sample households experienced an average rainfall deficit of approximately 21cm, relative to the expected long-run level. Unsurprisingly, the deviation from rainfall is highly correlated with farm income in the two survey waves ($\rho = 0.4239$).

Figure 1.4: Distribution of Herd Size



The final set of variables reported in Table 1.1 indicate the relative location of the households and their access to the Productive Safety Net Programme (PSNP). The distance to a bank and micro-finance organisation (MFI) are intended to

[15]A plot of land classified as *low* TWI will have an index score ≤ 11, while a plot classified as *high* TWI will have an index score ≥ 14. Any area of land with a score between these values is classified as *medium* TWI.

control for variation in access to formal savings and credit facilities, which could provide an alternative to the use of asset sales as a potential coping strategy. Proximity to a large, weekly market is a measure of remoteness, which would be a characteristic of both differential access to asset markets, and a strong indication of household poverty status (World Bank, 2016). As a further control, the PSNP variable indicates whether or not the safety net programme was operating in the local area during the period in which the survey took place.

1.4 Empirical Strategy

1.4.1 The Consumption Response to an Income Shock

The first prediction of the theory presented in Section 1.2 is that the consumption response to the drought-induced, income shock will differ, dependent on the level of pre-shock cattle holdings, relative to γ (see Proposition 1). Model (1.2) uses the deviation from long-run rainfall R as an exogenous measure of the transient shock to agricultural income. The dependent variable is the value of household consumption (per adult equivalent), for household i, at time t. The term H_{it} represents a vector of characteristics, specific to the household, while τ_t and ψ_i represent a common time effect and household fixed effect. μ_{it} denotes the error term.

$$c_{it} = \theta_1 R_{it} + \beta H_{it} + \tau_t + \psi_i + \mu_{it} \tag{1.2}$$

Although the ability to smooth consumption in the sample of households overall is clearly of interest, Section 1.2 puts forward the hypothesis that consumption smoothing behaviour will be more pronounced amongst households with initial cattle stocks above an asset level γ. This level of cattle holdings should mark the point at which proximity to a hypothesised poverty trap induces a switch from consumption smoothing to asset smoothing behaviour. Testing this prediction requires separating the sample into households with a pre-shock herd size $A_{pre} > \gamma$ from those with cattle holdings $A_{pre} \leq \gamma$.[16] This separation is illustrated in equation (1.3), which includes the h and ℓ terms on the rainfall coefficient, to make clear the distinction between the high and low cattle regimes.[17]

$$c_{it} = \begin{cases} \theta_1^h R_{it} + \beta H_{it} + \tau_t + \psi_i + \mu_{it}, & \text{if } A_{it} > \gamma \\ \\ \theta_1^\ell R_{it} + \beta H_{it} + \tau_t + \psi_i + \mu_{it}, & \text{if } A_{it} \leq \gamma \end{cases} \tag{1.3}$$

If the level of cattle holdings γ were known *a priori* estimation of equation (1.3) would be relatively straightforward. As this is not the case however, it is necessary to utilise the sample splitting approach devised by Hansen (1996, 2000). This method relies on minimising the residual sum of squares across a range of candidate values. Therefore, allowing the data itself to determine the most likely value of γ. A detailed description of this approach is provided in Appendix 1.3.

1.4.2 The Cattle Response to an Income Shock

If a poverty trap threshold does exist in asset wealth, Section 1.2 suggests that a household's proximity to this threshold will influence the extent to which asset sales are used to offset a negative shock to income. According to Proposition 2, behaviour above γ should broadly conform with the buffer stock theory, whereby cattle stocks will be drawn upon to achieve a level of consumption which is as

[16] The pre-shock period refers to the 2013/14 survey wave.

[17] To ease notation, no h or ℓ script is applied to the other terms in (1.3), although all coefficients are permitted to vary between the two regimes.

close to smooth as the availability of the buffer stock permits. In contrast, those households below γ may employ a strategy of asset smoothing, when faced with a similar shock to income.

$$A_{it} = \begin{cases} \alpha_1^h R_{it} + \beta H_{it} + \tau_t + \psi_i + \upsilon_{it}, & \text{if } A_{pre} > \gamma \\ \\ \alpha_1^\ell R_i + \beta H_{it} + \tau_t + \psi_i + \upsilon_{it}, & \text{if } A_{pre} \leq \gamma \end{cases} \tag{1.4}$$

The estimation strategy for determining the effect of the rainfall shocks on cattle holdings is similar to the approach used to assess the extent of consumption smoothing in Section 1.4.1. In model (1.4), the dependent variable is replaced with the number of cattle A_{it}, while α_1^h and α_1^ℓ measure the impact of the rainfall shock in the high and low cattle regimes, respectively. Again, the hypothesis of separate response regimes implies that all coefficients in the model may assume different values above and below γ. This separation is modelled in (1.4), where all independent variables have the same interpretation as previously. As in the consumption model (1.3), the value of the sample splitting level of cattle γ will be replaced with an estimated value $\hat{\gamma}$, representing the herd size which minimises the residual sum of squares from a single-equation estimation of (1.3).

Versions of models (1.3) and (1.4) are also estimated using ordinary least squares, where the model includes a regional fixed effect (in place of ψ_i) and the variables indicating distance (kms) to i) a market ii) a bank iii) a micro-finance organisation, and the presence of the PSNP in the local area, are added to the list of controls. The OLS estimates also include variables intended to capture heterogeneity in the agricultural land available to households in the 2013/14 survey. This includes the area of cultivated land (ha) classified by irrigation and rental status, as well as high, medium and low TWI (see Section 1.3).

1.5 Results

1.5.1 Preliminary Results

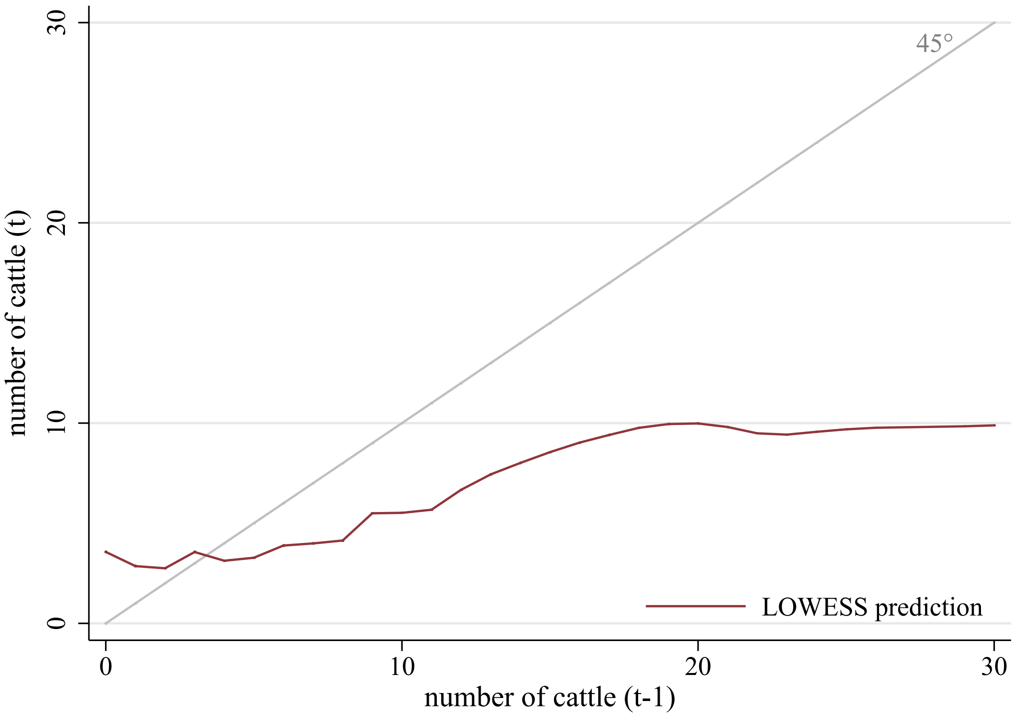
This section provides an initial investigation into the possibility of multiple equilibria in the growth-path of cattle stocks, using more commonly applied methods (see Section 1.1.2). Results obtained using a similar non-parametric approach to Lybbert et al. (2004) and Adato et al. (2006) are reported below, while results from the lagged polynomial method of Jalan and Ravallion (2004), Barrett et al. (2006) and Giesbert and Schindler (2012) are reported in the following sub-section.

Before comparing these preliminary results with the main findings of the paper, it is important to recognise that the methods used below do not necessarily provide answers to precisely the same question. Directly modeling the evolution of asset stocks has the potential to identify a poverty trap that generates measurable changes in asset stocks. However, if agents engage in avoidance behaviour, through asset smoothing, multiple equilibria in the underlying growth path may not be detected. In contrast, it is the detection of precisely this type of avoidance behaviour which motivates the main results of this analysis.

Non-Parametric Estimation of the Asset Growth-Path

Figure 1.5 maps the relationship between current and lagged cattle holdings using the Locally Weighted Scatter-plot Smoothing (LOWESS) method of Lybbert et al. (2004) and Adato et al. (2006), whereby local regressions are estimated for each of the n data-points within the sample. In generating Figure 1.5, each regression uses only those observations within a 8% bandwidth of the data-point in question. Unlike more parametric methods, this approach makes no assumptions regarding the functional form underlying the relationship $A_t = f(A_{t-1})$.

Figure 1.5: Non-Parametric Estimation of Herd Size Dynamics



It is clear from Figure 1.5 that estimating the growth-path using this approach implies a relationship which does not generate multiple equilibria in the accumulation of cattle stocks. Instead, the figure indicates the presence of only a single (low-level) equilibrium. This can be found at approximately 3 cattle, and represents a point of convergence, as opposed to bifurcation.

Parametric Estimation of the Asset Growth-Path

Turning now to the lagged polynomial approach of Jalan and Ravallion (2004), Barrett et al. (2006) and Giesbert and Schindler (2012). Figure 1.6 graphs the predicted values of current cattle holdings against lagged herd size, with the predicted value of A_t obtained from an estimation of model (1.5). This equation is similar to that employed by Barrett et al. (2006), whereby polynomials of the lagged herd size (up to the fourth-order) are included as regressors, allowing the relationship $A_t = f(A_{t-1})$ to generate the hypothesised non-linearities.

$$A_{it} = \sum_{p=0}^{4} \beta_p A_{it-1}^p + \beta G_i + \beta T_i + \tau_t + \xi_i + \varepsilon_{it} \tag{1.5}$$

The first term on the right-hand side of (1.5) represents the polynomials of lagged herd size (and a constant β_0). G is a selection of control variables, measured in the 2013/14 wave, including demographic and education characteristics, as well as the general location characteristics of the household. T_i represents the variables that measure the characteristics of cultivated land in the initial wave, while ξ_t and τ_t indicate a region fixed effect and common time effect. The full regression results are reported in Appendix 1.5.

Figure 1.6: Parametric Estimation of Herd Size Dynamics

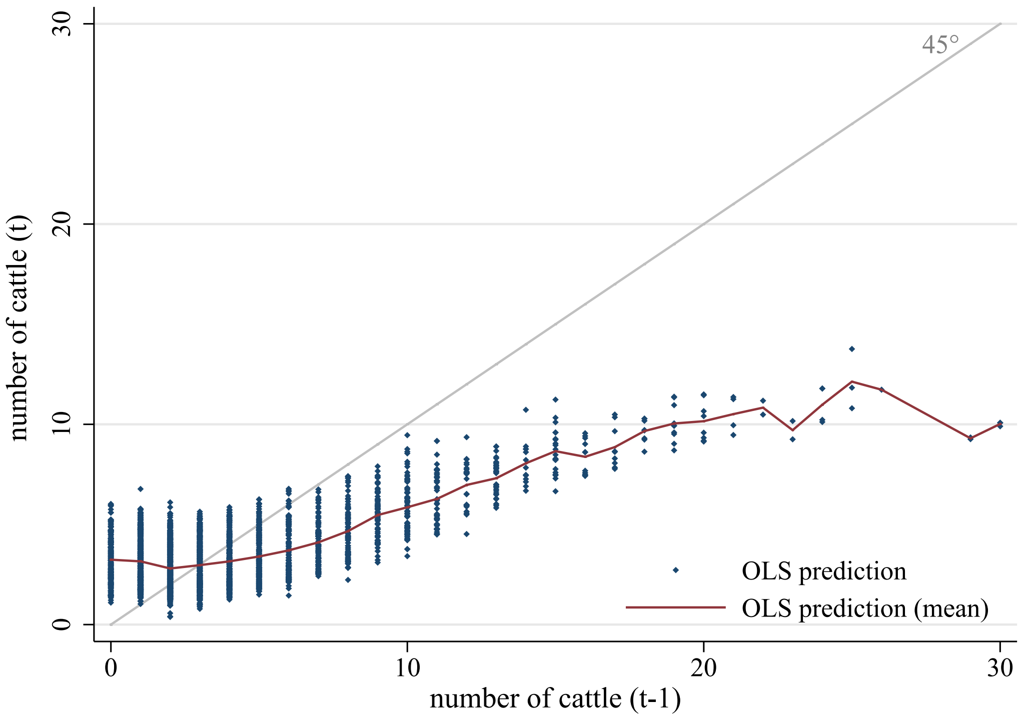


Figure 1.6 summarises the relationship between current and lagged herd size, estimated using the lagged polynomial approach described above. As the lagged herd size variable is discrete, each value on the x axis corresponds to a number of predictions at any given level of A_{t-1}. Figure 1.6 reports these predicted values, but also fits a line through the mean prediction at each level of lagged cattle holdings.[18] Although few observations make for an imprecise estimation of the time-path at larger herd sizes, the results of the parametric approach appear to mirror those of the non-parametric estimation. Again, the 45-degree line is intercepted only once, and from above, indicating a pattern of converge rather than bifurcation.

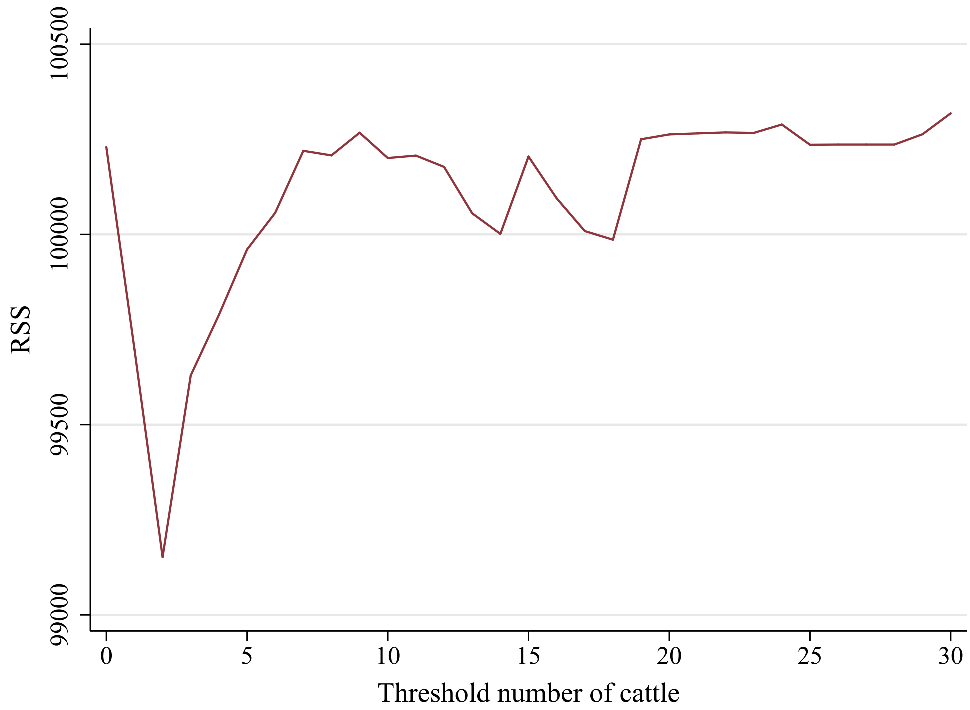
1.5.2 Estimation of γ

The main findings of this paper begin by reporting the results of the estimation used to determine the sample splitting value γ. Having found the herd size most likely to represent this value, Sections 1.5.3 and 1.5.4 report the effect of the drought-induced income shock on consumption and cattle holdings for households above or below (equal to) this level.

Using the methodology described in Appendix 1.3, the most likely candidate value for γ is found through comparison of the residual sum of squares (RSS), from a single-equation estimation of model (1.3), across the range of pre-shock (2013/14) cattle holdings A_{pre}. If two distinct consumption smoothing regimes exist, this implies that households can be separated into either regime, dependent on whether their pre-shock cattle holdings lie above or below the predicted level. Figure 1.7 graphs the residual sum of squares obtained from splitting the sample at all potential values of γ, between $1 \geq A_{pre} \geq 30$. The graph is minimised at a value of 2, indicating that this herd size is the level of cattle holdings, around which the coefficients in the model are most likely to assume different values.

[18] If instead, the median predicted value at each lagged herd size is used, this does little to alter the relationship in Figure 1.6.

Figure 1.7: Residual Sum of Squares at Potential Levels of γ



Having estimated the most likely value, at which to split the sample, it is also necessary to determine whether being above or below $\hat{\gamma} = 2$ alters the impact of the rainfall shock in a statistically significant manner. Based on Hansen's approach (Hansen, 1999), the p-value of 0.008 in Table 1.2 provides strong evidence against the null hypothesis of no difference between coefficients in the two regimes. Therefore, as Section 1.2 suggests, splitting the sample at $\hat{\gamma} = 2$, is required to accurately model household's behaviour, in response to the negative shock to income.

Table 1.2: Significance of the Difference between Regime Coefficients

Threshold estimate	RSS	MSE	Bootstrap repetitions	p-value
$\hat{\gamma} = 2$	99175	19.77	1000	0.008***

* p<0.1, ** p<0.05, *** p<0.01

Table 1.3 reports the mean values and standard deviations for key variables above and below (equal to) $\hat{\gamma} = 2$. F-test 1 reports a standard test of the significance of inclusion in either regime, in predicting the values of the listed variables, while F-test 2 conducts the same test, with the inclusion of regional fixed effects. This second test is necessary when assessing the similarity of rainfall deviations experienced by the two regimes, as both cattle holdings and the impact of the drought are likely to vary between different regions. Although the impact of the drought may not be equivalent at the national level, at the regional level, households should have experienced similar rainfall deficits.

Table 1.3: Characteristics of Households in High and Low Cattle Regimes

	$A_{pre} > 2$		$A_{pre} \leq 2$		F-test 1	F-test 2
	mean	sd	mean	sd	p-value	p-value
Household size	5.78	2.10	5.20	2.18	0.000***	0.000***
Age of head	46.82	14.03	46.35	14.87	0.256	0.200
Head is female	0.17	0.38	0.19	0.40	0.070*	0.051*
Total farm income	4.90	8.93	4.75	14.81	0.655	0.364
Total consumption	4.75	6.85	4.93	7.60	0.374	0.354
Food consumption	3.93	6.84	4.00	7.38	0.725	0.720
Land cultivated (ha)	1.00	0.93	0.96	0.91	0.129	0.084*
Rainfall deviation (cm)	-4.98	37.10	-5.65	33.40	0.504	0.122
Distance to bank (kms)	18.11	12.08	16.66	11.84	0.000***	0.000***
Distance to MFI (kms)	13.19	11.99	12.71	11.97	0.159	0.050*
Distance to market (kms)	23.79	5.58	25.44	5.71	0.000***	0.000***
PSNP in kebele	0.40	0.49	0.43	0.50	0.060*	0.162
Observations	2899		2161			

* p<0.1, ** p<0.05, *** p<0.01
All income and consumption figures reported in 1000 Ethiopian Birr, per adult equivalent
F test 1: Significance of high or low regime in determining the value of the listed variable
F test 2: Includes the addition of region-specific fixed-effects

The table indicates significant differences in demographics between the two regimes. Those households in the small herd group have fewer members, on average, and appear more likely to have a female household head. The rainfall deviations experienced during the survey do not appear to be more negative in the small herd group, even before partialling out the regional fixed effect. Therefore, there is no evidence that either regime was significantly worse affected by the drought.

Proximity to a formal bank, weekly market or micro-finance organisation also varies between the groups, with households in the low cattle regime more likely to be found further from these facilities. In summary, Table 1.3 indicates that variables generally considered to be good predictors of overall poverty status increase the probability of inclusion in the lower cattle regime.[19]

1.5.3 Smoothing Consumption

Table 1.4 present the results from the consumption smoothing model shown in equation (1.3), with OLS and household-level fixed effects estimations reported for each group. The dependent variable in Table 1.4 is the total value of household consumption, measured in 1000 Ethiopian Birr, per adult equivalent. The reported standard errors are clustered at the level of the household and the common time effect (τ_t) is included, but not reported.

Focusing on the coefficients on the rainfall deviation in Table 1.4, it is not possible to reject the hypothesis of full consumption smoothing, amongst the large herd group, in either the household fixed effects or OLS estimations (at any common level of significance). For those households with smaller numbers of cattle, however, the coefficients on rainfall, shown in columns 2 and 4, indicate that the shock has a highly significant impact on consumption expenditure in this group. A test of the equality of the rainfall coefficients between the two regimes (test $\theta_1^\ell = \theta_1^h$ in Table 1.4) indicates the two effects are statistically different from each other, in both sets of estimations, lending early support to the poverty trap hypothesis discussed in Section 1.2.

[19] The possibility that γ may actually be detecting heterogeneity, due to more general patterns of household wealth, is considered in Section 1.6.

Table 1.4: The Effect of Rainfall Shocks on Household Consumption

	(1)	(2)	(3)	(4)
	$A_{pre} > 2$ OLS	$A_{pre} \leq 2$ OLS	$A_{pre} > 2$ FE	$A_{pre} \leq 2$ FE
Rainfall deviation (cm)	0.008	0.059***	0.001	0.042***
	(0.005)	(0.007)	(0.005)	(0.015)
Household size	−0.589***	−0.647***	−0.343***	−0.542***
	(0.072)	(0.072)	(0.106)	(0.146)
Age of head	0.106*	0.064	0.194	0.081
	(0.059)	(0.071)	(0.127)	(0.164)
Age of head squared	−0.001*	−0.001	−0.002	−0.001
	(0.001)	(0.001)	(0.001)	(0.002)
% No education	−0.012***	−0.020***	−0.010	−0.033*
	(0.004)	(0.006)	(0.007)	(0.017)
Distance to market (kms)	−0.036*	−0.018		
	(0.021)	(0.026)		
Distance to bank (kms)	−0.018	−0.007		
	(0.012)	(0.012)		
Distance to MFI (kms)	0.013	−0.004		
	(0.013)	(0.015)		
PSNP in kebele	−0.129	−0.072		
	(0.364)	(0.320)		
Land low TWI (ha)	0.241	−0.362		
	(0.334)	(0.334)		
Land medium TWI (ha)	0.808***	0.497		
	(0.243)	(0.311)		
Land high TWI (ha)	1.053***	0.893**		
	(0.283)	(0.374)		
Land irrigated (ha)	0.092	0.306		
	(0.296)	(0.338)		
Land rented (ha)	−0.716**	−0.295		
	(0.294)	(0.399)		
Household fixed effects	No	No	Yes	Yes
Region fixed effects	Yes	Yes	No	No
Observations	2899	2166	2899	2166
R^2	0.085	0.149	0.034	0.111
R^2 *within*			0.025	0.090
R^2 *between*			0.041	0.123
Test: $\theta_1 = 0$ p-value	0.120	0.000***	0.813	0.005***
	OLS models (1) and (2)		FE models (3) and (4)	
Test: $\theta_1^\ell = \theta_1^h$ p-value	0.000***		0.009***	

Clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Cultivated land characteristics are based on the 2013/14 wave only

1.5.4 Smoothing Assets

Establishing that households with a higher level of cattle holdings are more capable of smoothing consumption is not sufficient evidence to conclude the existence of an asset-based poverty trap. Similar results could indicate a variety of different underlying mechanisms (see Section 1.1.2). It is, therefore, necessary to determine whether the extent to which cattle are used as a buffer against the shock also depends on a household's level of pre-shock cattle holdings. Table 1.5 reports the results of an estimation of the asset model (1.4), where the dependent variable in all estimations is the herd size A_{it}.

Again, the coefficient on the rainfall shock indicates a large degree of heterogeneity between the two regimes. For those households with a pre-shock herd size of three or more animals, the positive coefficients in columns 1 and 3 imply that a fall in rainfall is associated with a decline in cattle holdings. In contrast, column 2 and 4 indicates that the shock has no significant impact on the herd size of households below (equal to) $\gamma = 2$. Testing whether the effect of the rainfall shock differs between the estimations in models (1) and (2), and models (3) and (4) (see test $\alpha_1^\ell = \alpha_1^h$), again, leads to a strong rejection of the null of homogeneity in shock responses within the two regimes.

To summarise the key findings of the previous sections, those observations coming from households with pre-shock cattle holdings of three or more animals experience a reduction in herd size, as a result of the drought. For this group, it is also impossible to reject the hypothesis that consumption is fully insulated from the shock. In the case of households with two or fewer cattle, however, there is no evidence of any adjustments to herd size, yet the rainfall shock is associated with a significant decline in household consumption. This heterogeneity in coping strategies is consistent with the existence of the type of asset-based poverty trap described in Section 1.2.

Table 1.5: The Effect of Rainfall Shocks on Household Cattle Holdings

	(1)	(2)	(3)	(4)
	$A_{pre} > 2$ OLS	$A_{pre} \leq 2$ OLS	$A_{pre} > 2$ FE	$A_{pre} \leq 2$ FE
Rainfall deviation (cm)	0.038***	0.003	0.041***	0.002
	(0.003)	(0.002)	(0.003)	(0.003)
Household size	0.321***	0.023	0.043	0.001
	(0.043)	(0.031)	(0.073)	(0.050)
Age of head	0.075**	−0.005	0.054	0.051
	(0.034)	(0.024)	(0.071)	(0.051)
Age of head squared	−0.001**	0.000	0.000	−0.001
	(0.000)	(0.000)	(0.001)	(0.001)
% No education	−0.005*	−0.004***	0.003	−0.009***
	(0.003)	(0.001)	(0.003)	(0.003)
Distance to market (kms)	−0.029*	−0.007		
	(0.016)	(0.009)		
Distance to bank (kms)	0.005	−0.002		
	(0.007)	(0.005)		
Distance to MFI (kms)	0.001	−0.010**		
	(0.007)	(0.004)		
PSNP in kebele	−0.771***	0.053		
	(0.171)	(0.098)		
Land low TWI (ha)	0.212	−0.048		
	(0.194)	(0.117)		
Land medium TWI (ha)	0.203	0.050		
	(0.162)	(0.095)		
Land high TWI (ha)	0.262	0.178		
	(0.209)	(0.147)		
Land irrigated (ha)	−0.091	0.083		
	(0.220)	(0.124)		
Land rented (ha)	−0.117	0.205		
	(0.216)	(0.169)		
Household fixed effects	No	No	Yes	Yes
Region fixed effects	Yes	Yes	No	No
Observations	2899	2166	2899	2166
R^2	0.200	0.160	0.145	0.118
R^2 *within*			0.297	0.212
R^2 *between*			0.062	0.009
Test: $\alpha_1 = 0$ p-value	0.000***	0.218	0.000***	0.528
	OLS models (1) and (2)		FE models (3) and (4)	
Test: $\alpha_1^\ell = \alpha_1^h$ p-value	0.000***		0.000***	

Clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Cultivated land characteristics are based on the 2013/14 wave only

1.6 Robustness

1.6.1 Choice of Sample Splitting Variable

In using cattle herd size as the variable defining γ, it is possible that the estimation procedure may actually be detecting heterogeneity defined by more general patterns of household wealth. For example, Table 1.3 suggests there exists a possible correlation between cattle holdings and measures such as area of land cultivated and remoteness. To test this hypothesis, the sample splitting procedure is repeated, using the variables measuring land ownership and distance to a weekly market.[20]

Table 1.6: Test for a Threshold Effect in Other Variables

	$\hat{\gamma}$	RSS	MSE	Bootstrap repetitions	p-value
Owned land (ha)	0.36	99083	19.92	1000	0.232
Distance to market (kms)	16.00	98947	19.89	1000	0.149

* p<0.1, ** p<0.05, *** p<0.01

The table indicates that the most likely candidate for γ in cultivated land is at approximately 3600m^2. However, splitting the sample around this value does not generate significant differences in the high and low regime coefficients (p-value = 0.232). Similarly, splitting the sample, based on the distance of a household from a weekly market also fails to generate a significant difference in the coefficients either side of the respective candidate γ value (16kms). These findings suggest that it is household cattle holdings, specifically, which generate the heterogeneity in shock responses.

[20] A household's distance from a weekly market should be a strong indicator of poverty status. In rural Ethiopia, the proportion of those falling below the national poverty line is found to increase by 7% for every 10kms travelled from a market town (World Bank, 2016).

1.6.2 Availability of Cattle

Although rural households with no involvement in cattle rearing are omitted from the sample (see sample description in Section 1.3), the dataset still contains observations where a household owns no cattle in one of the two survey waves. The estimations in Table 1.7 are intended to test whether the non-significant effect, in the small herd group, is a result of some households entering the period of the drought with no cattle available to sell (whether they would otherwise, choose to or not). Table 1.7 reports that it is still not possible to reject the hypothesis of no effect in this adjusted sample, at any of the significance levels reported.

Table 1.7: Omitting Observations with No Cattle

	Rainfall coefficients and standard errors reported			
	(1)	(2)	(3)	(4)
Dependent variable:	Total consumption		Change in herd size	
	$A_{pre} > 2$	$0 < A_{pre} \leq 2$	$A_{pre} > 2$	$0 < A_{pre} \leq 2$
	FE	FE	FE	FE
	0.003	0.044**	0.040***	0.001
	(0.005)	(0.020)	(0.003)	(0.003)

Clustered standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

1.6.3 Uniqueness of γ=2

When assessing the residual sum of squares (RSS) generated by splitting the sample at possible values of γ, a value of $\hat{\gamma} = 3$ provided the second-lowest RSS, above $\hat{\gamma} = 2$ (see Figure 1.7). It is, therefore, worth considering the implications of splitting the sample at the alternative herd size of 3.

Table 1.8 reports the coefficients on the rainfall shock for households with pre-shock cattle holdings in a high and low asset regime, separated by $\hat{\gamma} = 3$. Columns

1 and 2 indicate that the consumption response follows the same pattern as Table 1.4. Again, only the coefficient in the high cattle regime is significantly different from zero.

The coefficients representing the effect of the rainfall shock on herd size (columns 3 and 4) indicate a significant response in cattle holdings, for households in the lower cattle regime, where none was present when splitting the sample at $\hat{\gamma} = 2$ (see Table 1.5). This is evidence that at least some households with a pre-shock herd size of 3 animals were willing to use cattle as a means of protecting consumption. Testing the null hypothesis that the effect of the drought shock is equivalent in both regimes, however, generates a strong rejection (test $H0 : \alpha_1^h = \alpha_1^\ell$ yields p-value $= 0.000$).

Table 1.8: Splitting the Sample at $\gamma = 3$

	Rainfall coefficients and standard errors reported			
	(1)	(2)	(3)	(4)
Dependent variable:	Total consumption		Change in herd size	
	$A_{pre} > 3$	$A_{pre} \leq 3$	$A_{pre} > 3$	$A_{pre} \leq 3$
	FE	FE	FE	FE
	0.002	0.029***	0.036***	0.012***
	(0.006)	(0.011)	(0.003)	(0.003)

Clustered standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

In light of both sets of results, therefore, it is possible to conclude that splitting the sample at a pre-shock herd size of $\hat{\gamma} = 3$ would also generate significant heterogeneity in coping strategies between the high and low cattle groups. Although, the difference in shock-responses appears more pronounced under the condition that $\hat{\gamma} = 2$.

1.6.4 Choice of Inclusion within the Sample

To avoid the inclusion of large-scale, commercial farms in the analysis, it was necessary to set upper-bounds on cattle holdings and cultivated land (see Section 1.3). The initial cut-off points were at 5 hectares of land or 30 head of cattle. The uppermost section of Table 1.9 tests the sensitivity of key findings to alternative cut-off levels. As previously, only the effect of the rainfall variable on total consumption (columns 1 and 2) and the change in cattle holdings (columns 3 and 4), is reported in each case.

Table 1.9: Robustness of Results to Alternative Choices of Sample

	Rainfall coefficients and standard errors reported			
	(1)	(2)	(3)	(4)
Dependent variable:	Total consumption		Change in herd size	
	$A_{pre} > 2$	$A_{pre} \leq 2$	$A_{pre} > 2$	$A_{pre} \leq 2$
	FE	FE	FE	FE
Alternative cut-off for sample inclusion				
Land<4 ha and cattle<25	0.004	0.055***	0.040***	0.003
	(0.005)	(0.009)	(0.003)	(0.003)
Land<6 ha and cattle<35	0.001	0.041***	0.041***	0.002
	(0.005)	(0.015)	(0.003)	(0.003)
Omission of regions				
Tigray omitted	0.004	0.037**	0.044***	0.003
	(0.005)	(0.016)	(0.003)	(0.003)
Amhara omitted	−0.002	0.041**	0.032***	0.000
	(0.006)	(0.018)	(0.003)	(0.003)
Oromia omitted	0.001	0.047***	0.036***	0.004
	(0.005)	(0.008)	(0.003)	(0.003)
SNNP omitted	0.001	0.038**	0.039***	0.001
	(0.006)	(0.019)	(0.003)	(0.004)
Other regions omitted†	0.003	0.045**	0.052***	0.001
	(0.006)	(0.020)	(0.003)	(0.004)

Clustered standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

† Afar, Somalie, Benishagul Gumuz, Gambela, Harari, Dire Dawa

Omitting households who possess more than 4 hectares of land or 25 cattle, in any wave, reduces the sample from 5067 to 4875. However, Table 1.9 indicates this does little to alter the interpretation of the results. Similarly, increasing the cut-off for sample inclusion to 6 hectares of land or 35 cattle, increases the sample to 5076 households, but again, does not alter the overall pattern of results.

As the type of farming practised in Ethiopia varies substantially between different agro-climatic zones, it is important to determine whether any specific region is responsible for driving the main findings of the analysis. The lower section of Table 1.9 tests the robustness of key results to the omission of observations from the four most populous rural regions.[21] Table 1.9 also reports tests from a sample which omits all observations from the remaining six regions (listed below Table 1.9). Any changes in the coefficients are not sufficient to warrant any new interpretation of the results (although, the significance of the coefficients does vary to a small degree in some cases). It is, therefore, possible to conclude that no specific, single region is driving the main results of the analysis.

1.7 Heterogeneity within Regimes

Significant differences in smoothing behaviour, amongst households with two or fewer cattle, mirrors the findings of Hoddinott (2006), in a study conducted in rural Zimbabwe. The key reason suggested for this effect was the need to provide sufficient animal traction to ensure fields could be ploughed in the following season. The wide variety of crops and farming practices found in Ethiopia provides an opportunity to test this hypothesis.

In the Ethiopian case, the traditional plough (*maresha*) requires oxen to operate. If the household does not possess these animals they must either obtain them by other means (short-term hire, for example) or cultivate the land without

[21] Tigray, Amhara, Oromia and SNNP, representing approximately 10.2%, 21.6%, 20.4% and 27.3% of the sample, respectively.

animal traction. Either option is likely to have a negative impact on the following season's net farm profits. In turn, lower profits in future harvests would be expected to hinder a household's efforts to bring its herd size back to the minimum number of animals required. It is, therefore, highly plausible that such a non-linear relationship between cattle stocks and productivity could generate the bifurcating growth path illustrated in Figure 1.1.

Formulating a suitable test of this hypothesis requires utilising the wide variety of crops and farming methods employed in Ethiopia. Within the country, approximately three-quarters of the total cultivated land is dedicated to the production of cereal crops, such as teff, wheat and maize (Taffesse et al., 2012). In almost all areas, this form of agriculture is highly dependent on the traditional ploughing methods described above. In contrast, farmers who specialise in the cultivation of non-cereal crops are far less reliant on cattle for the purpose of ploughing and preparing land. For example, in the southwest of the country, households rely heavily on the enset palm as a principal source of food, a crop which is cultivated using a wooden hoe, and requires no animal traction (Westphal and Stevels, 1975). Similarly, many farmers dedicate significant proportions of their land to tree crops or vegetables (Taffesse et al., 2012). As a result, if the need for animal traction determines γ, the requirement to retain a minimum number of cattle should be more evident amongst those households growing predominantly cereal crops.

$$A_{it} = \begin{cases} \alpha_1^h R_{it} + \alpha_2^h (R_{it} * C_{pre}) + \beta H_{it} + \tau_t + \psi_i + \upsilon_{it}, & \text{if } A_{pre} > \gamma \\ \alpha_1^\ell R_{it} + \alpha_2^\ell (R_{it} * C_{pre}) + \beta H_{it} + \tau_t + \psi_i + \upsilon_{it}, & \text{if } A_{pre} \leq \gamma \end{cases} \tag{1.6}$$

Model (1.6) modifies the original asset equation, by interacting the rainfall shock R_{it} with a continuous variable, indicating the share of a household's cultivated land dedicated to cereal crops in the pre-shock period C_{pre} (taking values between 0 and 1). Based on this model, a significant interaction effect in the coefficient α_2 would indicate that the response of cattle holdings to the shock is dependent on the share of cultivated land dedicated to cereal crops.[22]

Table 1.10: Heterogeneous Effects of Rainfall Shocks on Cattle Holdings

	(1)	(2)	(3)	(4)
	$A_{pre} > 2$ OLS	$A_{pre} \leq 2$ OLS	$A_{pre} > 2$ FE	$A_{pre} \leq 2$ FE
Rainfall deviation (cm)	0.037***	0.009***	0.042***	0.013***
	(0.004)	(0.003)	(0.004)	(0.005)
% Cereal crop	0.063	0.051		
	(0.236)	(0.126)		
Rainfall dev * % Cereal	0.001	−0.010***	−0.004	−0.016***
	(0.005)	(0.004)	(0.005)	(0.005)
Household fixed effects	No	No	Yes	Yes
Region fixed effects	Yes	Yes	No	No
Observations	2899	2166	2899	2166
R^2	0.200	0.168	0.144	0.126
R^2 *within*			0.300	0.227
R^2 *between*			0.060	0.014
Test: $\alpha_1 = 0$ p-value	0.000***	0.000***	0.000***	0.000***
Test: $\alpha_2 = 0$ p-value	0.900	0.000***	0.424	0.000***

Clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
All variables in table 1.5 included, but not reported
Cereal crop share is based on the 2013/14 wave only

Table 1.10 reports the results of an estimation of model (1.6), where all control variables in Table 5 are included (but not reported). The results of both the OLS and household fixed effects estimations indicate the interaction term is only significant in the low cattle regime. The coefficient on the main effect of the shock now also represents a significant effect in this group, where none was present in the initial estimations (see Table 1.5), implying that cattle holdings will respond to a rainfall shock in the small herd sub-sample, but only when the share of cereal crops approaches zero. The coefficient on the interaction term indicates

[22] The cereal crops considered are teff, wheat, maize, barley, millet, sorghum and oats.

that the strength of this correlation decreases, the greater the share of land these households dedicate to cereal crops. Overall, these findings lend strong support to the hypothesis that heterogeneity in the degree to which cattle are used to smooth consumption exists due to the need to protect a minimum number of animals for efficient land preparation in the following season.

1.8 Conclusion

Using a nationally representative dataset of agricultural households in rural Ethiopia, this paper finds evidence of two distinct methods of response to the income shock caused by the onset of a period of severe drought. A threshold estimation approach is used to separate the sample into two groups, according to each household's pre-shock level of cattle holdings. Analysis of these sub-samples concludes that households with an initial herd size of 3 or more animals effectively used these assets as a means of protecting consumption from the impact of the shock. In contrast, those households with 2 or less cattle did not reduce their holdings in response to the drought, instead choosing to protect their current herd size, at the cost of lower consumption. These results are consistent with the hypothesis of a poverty trap threshold, defined by household cattle ownership.

The key findings of this analysis contradict results obtained from direct estimations of the asset growth path (see Section 1.5.1). Modelling the relationship between current and lagged herd size, directly, produces no evidence of the type of multiple equilibria, which would characterise a poverty trap. The likely explanation for this disparity lies in this paper's use of behavioural changes to indirectly identify the existence of an asset threshold. With this approach, it is possible to identify such a threshold, even where no household's asset stock ever falls below it. Modelling the growth path directly however, will only identify a threshold which is observed to have measurable consequences for a household's asset wealth.

Extending the main results uncovers further heterogeneity within the small herd group. The degree to which cattle are used as a buffer stock within this group is shown to depend on the proportion of land these households dedicate to the cultivation of cereal crops. As cattle are an essential source of animal traction in cereal cultivation, this result suggests that a change in smoothing behaviour, between a herd size of 3 and a herd size of 2, is a response to the need to retain sufficient animals to allow for the preparation of land in the following season.

This analysis is not without its limitations, however. In particular, the treatment of consumption shortfalls as uniform to all household members. While convenient, this approach does not attempt to ascertain how the negative cost of protecting cattle stocks impacts on vulnerable household members. If such costs are borne disproportionately by younger generations within the households, short term drops in consumption may have very long term consequences (Dercon and Porter, 2014). Determining the extent to which asset smoothing strategies influence the outcomes of future generations will be a topic considered in future research.

The findings of this study highlight some of the complexities of assessing vulnerability to drought in rural populations. Even with the availability of social protection and safety nets, it is inevitable that some households will still be faced with a choice of either reducing consumption now or undermining productivity in the future. Such stark choices draw attention to the need for improvements in the coverage and design of agricultural insurance programmes, particularly in more drought-prone rural areas. Encouraging, the implementation and uptake of insurance products, where financial literacy and monitoring capacity are low, may require some degree of public sector involvement. Yet, if the country is able to effectively manage the substantial weather risk to small-scale agricultural, the removal of such a binding constraint may see Ethiopia turn recent development gains into long-term economic growth.

References

Adato, M., Carter, M. R., and May, J. (2006). Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data. *The Journal of Development Studies*, 42(2):226–247.

Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford economic papers*, 58(34):450–474.

Banerjee, A. (2004). The two poverties. *Insurance against poverty*, pages 59–75.

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J. C., Rasambainarivo, J., and Wangila, J. (2006). Welfare dynamics in rural Kenya and Madagascar. *The Journal of Development Studies*, 42(2):248–277.

Carter, M. R. (1988). Equilibrium credit rationing of small farm agriculture. *Journal of Development Economics*, 28(1):83–103.

Carter, M. R. and Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies*, 42(2):178–199.

Carter, M. R., Little, P. D., Mogues, T., and Negatu, W. (2007). Poverty traps and natural disasters in Ethiopia and Honduras. *World Development*, 35(5):835–856.

Carter, M. R. and Lybbert, T. J. (2012). Consumption versus asset smoothing: testing the implications of poverty trap theory in Burkina Faso. *Journal of Development Economics*, 99(2):255–264.

Carter, M. R. and May, J. (2001). One kind of freedom: Poverty dynamics in post-apartheid South Africa. *World development*, 29(12):1987–2006.

Carter, M. R. and Zimmerman, F. J. (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraint. *Journal of Development Economics*, 71(2):233–360.

CSA (2013). *Ethiopian Rural Socioeconomic Survey Report*. Central Statistics Agency, Addis Ababa.

CSA (2016). *Agricultural sample survey 2015/2016 (2008 E.C.) - Volume I. Report on area and production of major crops (Private peasant holdings, meher season).* Central Statistics Agency, Addis Ababa.

CSA (2017). *LSMS—Integrated surveys on agriculture Ethiopia socioeconomic survey (ESS).* Central Statistical Agency of Ethiopia.

Dasgupta, P. and Ray, D. (1986). Inequality as a determinant of malnutrition and unemployment: Theory. *The Economic Journal*, 96(384):1011–1034.

Deaton, A. (1991).Saving and liquidity constraints.*Econometrica*, 59(5):1221–1248.

Deaton, A. (1992). *Understanding Consumption.* Oxford University Press, Oxford.

Deaton, A. and Zaidi, S. (2002). *Guidelines for constructing consumption aggregates for welfare analysis.* World Bank Publications.

Dercon, S. (2004). Growth and shocks: evidence from rural ethiopia. *Journal of Development Economics*, 74:309–329.

Dercon, S. and Porter, C. (2014). Live aid revisited: long-term impacts of the 1984 ethiopian famine on children. *Journal of the European Economic Association*, 12(4):927–948.

Eswaran, M. and Kotwal, A. (1986). Access to capital and agrarian production organisation. *The Economic Journal*, 96(382):482–498.

Fafchamps, M., Udry, C., and Czukas, K. (1998). Drought and saving in West Africa: Are livestock a buffer stock? *Journal of Development economics*, 55(2):273–305.

FAO (2015). *El Niño response plan 2016.* Food and Agriculture Organization, Addis Ababa.

Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function*, pages 20–37. Princeton University Press.

Giesbert, L. and Schindler, K. (2012). Assets, shocks, and poverty traps in rural Mozambique. *World Development*, 40(8):1594–1609.

Hansen, B. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 1:413–430.

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics*, 93(2):345–368.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, 68(3):575–603.

Hoddinott, J. (2006). Shocks and their consequences across and within households in rural Zimbabwe. *The Journal of Development Studies*, 42(2):301–321.

HRD (2015). *Joint Government and Humanitarian Partners' Humanitarian requirement document: Ethiopia 2016*. Federal Democratic Republic of Ethiopia, Addis Ababa.

Jalan, J. and Ravallion, M. (2004). Household income dynamics in rural China. In Dercon, S., editor, *Insurance against poverty*, pages 105–124. "Oxford University Press".

Kazianga, H. and Udry, C. (2006). Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, 79(2):413–446.

Kinsey, B., Burger, K., and Gunning, J. W. (1998). Coping with drought in zimbabwe: Survey evidence on responses of rural households to risk. *World Development*, 26(1):89–110.

Lokshin, M. and Ravallion, M. (2004). Household income dynamics in two transition economies. *Studies in Nonlinear Dynamics and Econometrics*, 8(3).

Lybbert, T. J., Barret, C. B., Desta, S., and Coppock, D. L. (2004). Stochastic wealth dynamics and risk management among a poor population. *The Economic Journal*, 114(498):750–777.

Mookherjee, D., Ray, D., and Napel, S. (2010)Aspirations, segregation, and occupational choice*Journal of the European Economic Association*,8(1):139–168.

Naschold, F. (2013). Welfare dynamics in Pakistan and Ethiopia–Does the estimation method matter? *The Journal of Development Studies*, 49(7):936–954.

Quisumbing, A. and Baulch, B. (2013). Assets and poverty traps in rural bangladesh. *The Journal of Development Studies*, 49(7):898–916.

Rosenzweig, M. R. and Wolpin, K. (1993). Credit market constraints, consumption smoothing and the accumulation of durable production assets in low-income countries: Investments in bullocks in india. *Journal of Political Economy*, 101(2):223–244.

Santos, P. and Barrett, C. B. (2011). Persistent poverty and informal credit. *Journal of Development Economics*, 96(2):337–347.

Sörensen, R., Zinko, U., and Seibert, J. (2006). On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences Discussions*, 10(1):101–112.

Taffesse, A. S., Dorosh, P., and Gemessa, S. A. (2012). Crop production in Ethiopia. In Dorosh, P. and Rashid, S., editors, *Food and Agriculture in Ethiopia*. "University of Pennsylvania Press", Philadelphia.

Van Campenhout, B. and Dercon, S. (2012). Nonlinear dynamics of livestock assets: evidence from Ethiopia. IFPRI Discussion Paper 01215.

Verpoorten, M. (2009). Household coping in war-and peacetime: Cattle sales in Rwanda, 1991–2001. *Journal of Development Economics*, 88(1):67–86.

Webb, P., Von Braun, J., and Yohannes, Y. (1992). *Famine in Ethiopia: policy implications of coping failure at national and household levels.* International Food Policy Research Institute.

Westphal, E. and Stevels, J. M. C. (1975). *Agricultural systems in Ethiopia*, volume 826. Centre for Agricultural Publishing and Documentation Wageningen.

World Bank (2016). *Ethiopia : Priorities for ending extreme poverty and promoting shared prosperity.* World Bank, Washington, DC.

Appendices

Appendix 1.1

Low technology: $f^{\ell}(A_t\theta_t) = \theta_t A_t^{0.3}$

High technology: $f^h(A_t\theta_t) = \theta_t A_t^{0.45} - 0.45$

Depreciation rate: $\delta = 0.08$

Discount factor: $\beta = 0.95$

$u(c_t) = \frac{c_t^{1-\sigma}-1}{1-\sigma}, \quad \text{with: } \sigma = 1.5$

Probability distribution of θ approximated as:

$$\theta = \begin{cases} 0.05 \;\; if \;\; \theta_t = 0.8 \\ 0.10 \;\; if \;\; \theta_t = 0.9 \\ 0.70 \;\; if \;\; \theta_t = 1 \\ 0.10 \;\; if \;\; \theta_t = 1.1 \\ 0.05 \;\; if \;\; \theta_t = 1.2 \end{cases}$$

Probability distribution of θ (post-shock) approximated as:

$$\theta = \begin{cases} 0.05 \;\; if \;\; \theta_t = 0.4 \\ 0.10 \;\; if \;\; \theta_t = 0.5 \\ 0.70 \;\; if \;\; \theta_t = 0.6 \\ 0.10 \;\; if \;\; \theta_t = 0.7 \\ 0.05 \;\; if \;\; \theta_t = 0.8 \end{cases}$$

A_t and A_{t+1} approximated using a grid of discrete values $A_t = \{1, \; 2 \; ..... \; 45\}$

Appendix 1.2

The upper panel of Table 1.A1 indicates the mean and standard deviation of important household characteristics. These are reported for observations included in the sample and those which were omitted due to missing values (see Section 1.3). The F-test in the right-hand column of the table indicates that omission from the sample is not a significant predictor of any of the variables listed.

Table 1.A1: Tests for Non-Random Attrition Due to Missing Values

	Sample		Omitted		F-test
	mean	sd	mean	sd	p-value
Household size	5.54	2.15	5.42	2.28	0.255
Age of head	46.62	14.39	46.67	14.80	0.933
Head is literate	0.40	0.49	0.39	0.49	0.874
Number of cattle	4.09	4.16	4.39	3.88	0.111
Land cultivated (ha)	0.98	0.92	1.01	0.93	0.574
Total consumption	4.82	7.18	5.14	4.22	0.150
Food consumption	3.96	7.07	4.03	3.23	0.705
Total farm income	4.83	11.80	5.77	22.33	0.477
Rainfall deviation (cm)	-5.30	35.56	-5.68	27.52	0.758
Regions					
Tigray	0.10	0.29	0.09	0.29	0.858
Afar	0.01	0.10	0.01	0.12	0.536
Amhara	0.22	0.41	0.20	0.40	0.399
Oromia	0.20	0.40	0.21	0.41	0.737
Somalia	0.07	0.26	0.09	0.28	0.273
Benishagul-Gumuaz	0.03	0.16	0.02	0.15	0.636
SNNP	0.27	0.44	0.27	0.44	0.899
Gambela	0.03	0.17	0.04	0.19	0.553
Harari	0.04	0.19	0.03	0.17	0.305
Dire Dawa	0.04	0.19	0.04	0.20	0.716
Observations	5067		591		

* p<0.1, ** p<0.05, *** p<0.01
All income and consumption figures measured in 1000 Ethiopian Birr, per adult equivalent

Appendix 1.3

The following describes the procedure used to estimate γ. It is useful to follow Hansen (1999) and represent the two parts of (1.3) in a single equation. Collecting the right-hand side variable $\{R_{it}, H_{it}, \tau_t\}$ in a vector X_{it}, and representing all associated coefficients for the low and high asset regimes as β_ℓ and β_h, respectively, the empirical model (1.3) can be written as follows.

$$c_{it} = \beta_\ell' X_{it} I(A_{it} \leq \gamma) + \beta_h' X_{it} I(A_{it} > \gamma) + \psi_i + e_{it} \tag{1.a1}$$

In equation (1.a1) the bracketed terms indicates the position of the household, relative to γ, where $I(.)$ is an indicator function. A further simplification combines the coefficient vectors for the high and low regimes, and represents the separation of the data as follows.

$$c_{it} = \beta' X_{it}(\gamma) + \psi_i + e_{it} \tag{1.a2}$$

where

$$\beta = (\beta_\ell', \beta_h')'$$

$$X_{it}(\gamma) = \left\{ \begin{array}{c} X_{it} I(A_{it} \leq \gamma) \\ X_{it} I(A_{it} > \gamma) \end{array} \right\}$$

Subtracting the mean, over the time-index t, for each household i, yield the transformed, fixed effects model, where * denotes the transformed variables, and the fixed effect ψ_i is eliminated.

$$c_{it}^* = \beta' X_{it}^*(\gamma) + e_{it}^* \tag{1.a3}$$

For a given γ, equation (1.a3) can be estimated via ordinary least squares (OLS), yielding both the vector of regression residuals $\hat{e}^*_{it}(\gamma)$ and the residual sum of squares $S(\gamma)$.

$$\hat{e}^*_{it}(\gamma) = c^*_{it} - \hat{\beta}'(\gamma)X^*_{it}(\gamma) \tag{1.a4}$$

$$S(\gamma) = \hat{e}^*_{it}(\gamma)'\hat{e}^*_{it}(\gamma) \tag{1.a5}$$

The most likely candidate for a possible splitting point is found through a process of searching over the range of values of the pre-shock herd size variable A_{pre}. The candidate value $\hat{\gamma}$ is the herd size which minimises the residual sum of squares from an OLS estimation of (1.a3) (Hansen, 1999).

$$\hat{\gamma} = \min_{\gamma} S(\gamma) \tag{1.a6}$$

Testing for the significance of differences in the high and low regime coefficients is a test of the null hypothesis, $H0 : \beta_h = \beta_\ell$ in equation (1.a1), with γ replaced by $\hat{\gamma}$. However, under the null hypothesis, the indicator function $I(.)$ has no place in the estimation and the pooled model (1.2) is the appropriate choice. Similarly to equation (1.a3), under the null hypothesis, model (1.2) can be written as follows, where the subscript p indicates the vector of coefficients relate to the pooled model.

$$c^*_{it} = \beta'_p X^*_{it} + \eta^*_{it} \tag{1.a7}$$

Estimation of (1.a7) again generates the predicted standard errors $\hat{\eta}$ and the residual sum of squares S_p.

$$\hat{\eta}^*_{it} = c^*_{it} - \hat{\beta}'X^*_{it} \tag{1.a8}$$

$$S_p = \hat{\eta}^{*\prime}_{it}\hat{\eta}^*_{it} \tag{1.a9}$$

The test statistic for determining the significance of differences in regime coefficients takes the form in (1.a10). However, as the sample splitting point is not identified under $H0$, the asymptotic distribution of F is non-standard and strictly dominates χ_k^2 (Hansen, 1999).

$$F = \frac{S_p - S(\hat{\gamma})}{\hat{\sigma}^2} \tag{1.a10}$$

Hansen (1996, 1999) describe a bootstrap procedure that can be used to generate an asymptotically valid p-value for the test of the significance of differences in the regime coefficients. The procedure involves cluster re-sampling (with replacement) of the predicted errors $\hat{e}^*$ from an estimation of (1.a3). These errors are then used to generate a bootstrapped dependent variable in equation (1.a7), by holding the values of X^* and the variable A fixed (in all bootstrap samples), while allowing the coefficients β_p' to assume any arbitrary value.[23] The generated bootstrap sample is then used to estimate the residual sum of squares, under the null and alternative hypotheses, and compute a bootstrapped version of the F statistic (1.a10). In repeated samples, the proportion of draws, in which this bootstrapped statistic exceeds F, is the bootstrapped p-value for the null hypothesis of no threshold effect. Further details of this procedure can be found in Hansen (1999), pages 350-351.

[23] This is permitted, as F is not dependent on β_p', under the null hypothesis (Hansen, 1999).

Appendix 1.4

All income and consumption variables used in the analysis are expressed per adult equivalent. The weighting used to construct these variables is shown in Table 1.A2. These weights are based on the suggested construction of consumption aggregates, provided with the ESS data (2011-12).

Table 1.A2: Adult Equivalency Scales

Age range	Male	Female
age $\leq$ 1 year	0.33	0.33
1 year $<$ age $\leq$ 2 years	0.46	0.46
2 year $<$ age $\leq$ 3 years	0.54	0.54
3 year $<$ age $\leq$ 5 years	0.62	0.62
5 year $<$ age $\leq$ 7 years	0.74	0.70
7 year $<$ age $\leq$ 10 years	0.84	0.72
10 year $<$ age $\leq$ 12 years	0.88	0.78
12 year $<$ age $\leq$ 14 years	0.96	0.84
14 year $<$ age $\leq$ 16 years	1.06	0.86
16 year $<$ age $\leq$ 18 years	1.14	0.86
18 year $<$ age $\leq$ 30 years	1.04	0.8
30 year $<$ age $\leq$ 60 years	1	0.82
60 year $<$ age	0.84	0.74

Appendix 1.5

Table 1.A3 reports the results of an estimation of equation (1.5). The fitted values from this estimation are used to generate the herd size growth path, shown in Figure 1.6. The dependent variable in Table 1.A3 is the current number of cattle owned by each household A_{it} and the model is estimated using OLS. All control variables come from the first survey wave only.

Table 1.A3: Lagged Polynomial Estimation of Herd Size Dynamics

	(1)		
$A(t-1)$	−0.249 (0.169)	Distance to bank (kms)	−0.004 (0.007)
$A^2(t-1)$	0.074** (0.037)	Distance to MFI (kms)	−0.011 (0.007)
$A^3(t-1)$	−0.003 (0.003)	PSNP	−0.519*** (0.151)
$A^4(t-1)$	0.000 (0.000)	Land low TWI (ha)	0.052 (0.176)
Household size	0.073* (0.040)	Land medium TWI (ha)	0.428*** (0.132)
Age of head	0.005 (0.033)	Land high TWI (ha)	0.420* (0.218)
Age of head squared	0.000 (0.000)	Irrigate land (ha)	0.1560 (0.196)
% No education	0.001 (0.002)	Rented land (ha)	−0.126 (0.194)
Distance to market (kms)	0.021 (0.013)		
Region fixed effects			Yes
Observations			2488
R^2			0.230

Clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
All control variables are measured from 2013/14 wave

Chapter 2

Early Childhood Health During Conflict: The Legacy of the Lord's Resistance Army in Northern Uganda

2.1 Introduction

Few conflicts on the continent of Africa have gained as much notoriety as the twenty-year war fought between the government forces of Uganda and the Lord's Resistance Army (LRA). Under the leadership of the self-proclaimed spiritual medium Joseph Kony, the shockingly brutal tactics employed by the LRA against the civilian population drew international condemnation. Although attempts have been made to analyse the post-war recovery of those caught up in the conflict (Blattman, 2009; Blattman and Annan, 2010; Bozzoli and Brück, 2010; Adelman and Peterman, 2014; Fiala, 2015), there exists no direct study of the impact of the war on the physical health of the civilian population. This paper exploits spatial and temporal variation in the spread of the conflict to measure the causal effect of the war on the health outcomes of children living in the affected districts. Evidence is found of irreversible health deficits for children exposed to the conflict for a period of more than 6 months, with children in this group experiencing an average shortfall in height-for-age z-scores of 0.65 standard deviations.[1] In contrast, however, there is no evidence of significant height deficits amongst those exposed to the fighting for shorter periods of time. These results are found to be robust to alternative samples, aimed at addressing potential sources of bias, and alternative definitions of conflict exposure, with a similar pattern of results also observed in other anthropometric measures of health status. Given potential links between early childhood health and outcomes in later life (Strauss and Thomas, 2007), the deficits experienced by these children are likely to impact the economic prospects of war-affected regions for many years to come.[2]

The remainder of this section presents a review of the relevant literature to which this study contributes. Section 2.2 provides a brief history of the LRA conflict which took place in northern Uganda, while a description of the data used, and the characteristics of the sample, can be found in Section 2.3. The empirical strategy employed to estimate the causal effect of the conflict on health outcomes is outlined in Section 2.4, with the main results of this analysis reported in Section 2.5. A comprehensive study of robustness and potential sources of heterogeneity can be found in Section 2.6 and 2.7, before a summary of the key findings and concluding remarks are presented in Section 2.8.

2.1.1 Previous Studies

This paper makes a contribution to three specific strands of the literature. Firstly, the analysis adds to the growing number of studies which endeavour to quantify the impact of conflict on early childhood health. For example, Bundervoet et al. (2009) provide evidence of a cumulative effect of exposure to the 1993-2005 Burundian

[1] Height-for-age z-scores measure the number of standard deviations from the median height of a child in a healthy reference population and are an indication of long-term nutritional status.

[2] See Alderman et al. (2006) on malnutrition in Zimbabwe and completed grades of education, Maccini and Yang (2009) on the impact of rainfall shocks on health, education and asset wealth in Indonesia, and Dercon and Porter (2014) on the long-run health and income effects experienced by survivors of the 1984 Ethiopian famine.

civil war on a sample of children aged between 6 months and 5 years. Each additional month of exposure to the war is found to decreases children's height-for-age z-scores by 0.047 standard deviations, with an average differential between exposed and non-exposed children of -0.348 to -0.525. Akresh et al. (2012b) use a similar methodology to estimate the effect of the 1998 Eritrean-Ethiopian war on children's health, yielding comparable results. Their findings indicate that children alive during the war, and living in a war-affected region, would be between 0.447 and 0.454 standard deviations shorter than those who were not exposed to the conflict. More recently, Minoiu and Shemyakina (2014) show similar results for children exposed to the 2002-2007 civil war in the Côte d'Ivoire.[3]

Secondly, these results contribute to the literature looking specifically at the effects of the LRA insurgency in northern Uganda. Amongst these studies, Blattman and Annan (2010) analyse the labour market outcomes of children who were previously abducted by the rebels to serve as child-soldiers.[4] The authors estimate the loss of human capital from time spent away from education and employment, coupled with the psychological distress of increased exposure to violence, led to 33% lower wages for former abductees. Much attention has also been given to the economic consequences of the widespread displacement caused by the conflict. For example, Fiala (2015) finds evidence of a negative impact on both consumption and asset wealth amongst previously displaced households, with only wealthier households showing later signs of recovery. In another analysis of conflict-driven displacement, Adelman and Peterman (2014) find that resettled households experienced significant losses in access to agricultural land upon return to their former

[3]Other notable contributions on the impact of conflict on children's health can be found in Akresh et al. (2011), which covers the civil war preceding the 1994 Rwandan genocide, Shemyakina (2017), in relation to politically motivated violence in Zimbabwe, and Akresh et al. (2012a) who focus on the height deficits of adults born during the Nigerian civil war of 1967-1970.

[4]It is estimated that between 60,000 and 80,000 children and young adults were abducted during the conflict, mostly from the northern, Acholi districts, which were formerly Kitgum and Gulu (Blattman and Annan, 2010).

locations. Evidence also exists of less tangible impacts of the LRA conflict. For example, Rohner et al. (2013) find an increase in ethnic identity following the fighting, hampering economic recovery in more fractionalised communities, while Bozzoli et al. (2011) show that exposure to the conflict reduces individual's expectations of their future life situation and economic prosperity. Links have also been found between exposure to violence and increased political participation amongst ex-combatants (Blattman, 2009), although this may only take place locally, as a response to the immediate needs placed upon communities, rather than as a result of an increased concern over politics at the national level (De Luca and Verpoorten, 2015).

Finally, this study contributes to the literature on gender bias in childhood health outcomes and the potential for heterogeneity in the impact of negative shocks. For example, Baird et al. (2011) analyse a sample of 59 low-income countries and find relatively higher infant mortality amongst female children in response to fluctuations in per-capita GDP, while Rose (1999) finds that positive rainfall shocks increase the probability of survival for girls in rural India (relative to male children). There is, however, little evidence of a gender bias in the small number of papers that specifically address childhood health outcomes in response to conflict shocks. For example, Minoiu and Shemyakina (2014) uncover no evidence of significant heterogeneity in height-for-age z-scores of male and female children who were exposed to conflict in Côte d'Ivoire. Similarly, Akresh et al. (2011) do not find evidence of a gender bias as a result of the Rwandan civil war, in spite of a clear bias towards male children in response to crop failures in other parts of the country. A more recent study by Dagnelie et al. (2018) does find lower survival rates among female children during the 1997-2003 civil war in the Democratic Republic of Congo. However, the authors attribute this to an adverse selection effect, driven by the lower probability of survival for male children in utero, rather than a gender bias influencing their subsequent survival prospects.

2.2 The War in Northern Uganda

As with many other violent disputes in East Africa, the origins of Uganda's LRA conflict can be traced to the long-standing political and ethnic divisions within the country. Historically, Uganda was divided between the predominantly Bantu South, and the Nilotic (Nilo-Hamitic) and Central Sudanic people of the North and Northwest (Rohner et al., 2013). Following independence from British rule in 1962, the North provided the majority of Uganda's military power, paving the way for a series of brutal Northern dictatorships, which would serve to concentrate political power within the hands of those loyal to the current head of state. In 1986 this dominance ended with the defeat of the de facto government forces (largely comprised of the Northern Acholi and Langi ethnic groups) by the National Resistance Army (NRA) led by Yoweri Museveni, a Southerner and veteran of the campaign which ousted the notorious dictator Idi Amin (Allen, 2013). Following the NRA victory, defeated Northern soldiers retreated back to their homelands, with many crossing the border into southern Sudan. From the remnants of these forces a number of armed resistance groups initially formed in opposition to the new government. By 1988 most of these groups had either signed peace agreements with Museveni's government or been defeated by forces loyal to the new regime. The decision to stop fighting was not unanimous, however, and a small number of soldiers, who were unwilling to accept the outcome of peace negotiations, gathered under the leadership of Joseph Kony, in what would become the Lord's Resistance Army (Allen, 2013).

Kony belonged to the Acholi ethnic group and it was the districts of Gulu and Kitgum (known as Acholi-land) which would initially become the centre of the LRA insurgency.[5] There existed little local support for the poorly financed

[5]These two districts were later subdivided into seven districts (Amuru, Nwoya, Gulu, Agago, Lamwo, Pader and Kitgum). However, for clarity and consistency with the empirical analysis which follows, district names are recorded as they stood in 1991, which corresponds to the earliest birth year recorded for children in the studied sample (see Section 2.3.1).

rebel group prior to 1995, making them almost entirely reliant on the abduction and forced conscription of civilians (often children) to maintain their numbers (Blattman, 2009). Although the LRA's stated aim was to overthrow the Ugandan government, increasing Kony began to target civilian populations, proclaiming that Acholi society must be 'purified' to overcome its oppressors (Doom and Vlassenroot, 1999; Allen, 2013). Following an initially slow campaign, the number of violent attacks started to escalate in 1995, when the LRA received support from the Sudanese government, in the form of arms, provisions and land to establish bases in southern Sudan (Dolan, 2009). Attacks on civilians increased notably during this period, occurring seemingly at random, and with overwhelming brutality (Doom and Vlassenroot, 1999; Blattman and Annan, 2010). Many of the rural inhabitants of Acholi-land and neighbouring districts sought protection closer to local towns or military outposts and from 1996 the government began forcibly relocating the Acholi population to internally displaced people (IDP) camps situated in these locations. The camps were intended to shield civilians from the LRA attacks, but in reality, they offered little protection. Instead, they were characterised by overcrowding, poor sanitation and an abundance of disease (Bozzoli and Brück, 2010).

During the late 1990s, the plight of northern Uganda was gaining international attention, placing pressure on the Sudanese government to cease support for the LRA. In 2002 the Ugandan military (with support from the US) obtained permission from Khartoum to launch operation 'Iron Fist', a military incursion against the LRA bases in southern Sudan. However, Kony and almost all senior LRA members survived the raid and the rebels outflanked the government forces, attacking new territories in Lira, Apac and Soroti (Allen, 2013). The failure of this operation marked a rapid increase in the number of attacks and fatalities attributed to the LRA, with the period between 2002 and 2005 constituting the height of the violence experienced during the campaign (Rohner et al., 2013; De Luca and Verpoorten, 2015). By 2004 approximately 1.5 million people were estimated to be internally displaced as a result of the fighting (International Crisis Group, 2004) and increased global awareness of this humanitarian crisis, along with the numerous atrocities committed by the LRA, led the International Criminal Court to issue an arrest warrant against Kony and four of his senior commanders (Dolan, 2009). With global attention fixed on the conflict, pressure was placed on both sides to reach a peaceful solution. In mid-2006 the government of Uganda agreed to engage in peace talks with the rebels, resulting in the signing of a cessation of hostilities agreement in Juba (Sudan) on 26th August 2006. Although Kony never signed the final peace agreement, these talks began the effective end of the LRA war in northern Uganda (Dolan, 2009).

2.3 Data

2.3.1 Health and Demographic Data

The individual-level health data used in this analysis comes from three waves of the Uganda Demographic and Health Survey (DHS), collected in 1995, 2000 and 2006.[6] Data is provided on a number of key childhood health indicators, including height-for-age, weight-for-age and weight-for-height. The DHS surveys also contain demographic, health and education measures relating to the child's mother and the characteristics of the household. In order to accurately link this data to specific locations and events, the sample considered only includes observations for children whose mothers were present in the same DHS sample cluster since the child's birth.[7] This initial sample consists of a cross-section of 9496 children, born between 1991 and 2006, and aged less than 5 years old at the time of the survey.[8] The geographical coverage of the three surveys varies to some degree, as does

[6] At the time of writing, all Uganda DHS data used in this study is available on request from dhsprogram.com/data.

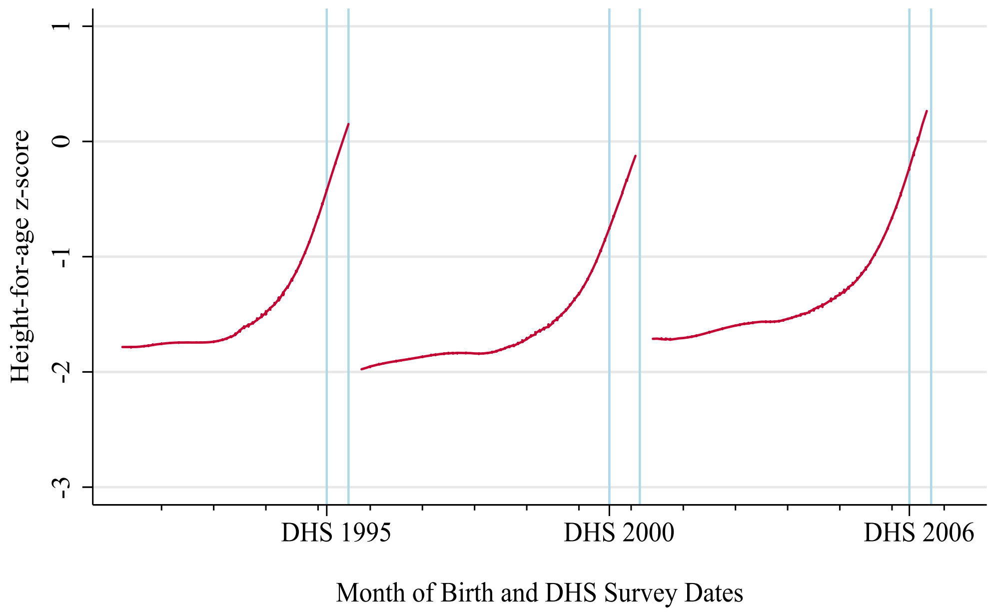
[7] The likelihood of bias in the estimated impact of conflict on health, due to a correlation between household relocation and health status, is discussed in Section 2.6.1.

[8] However, the main analysis is conducted using only a sub-sample of 9202 children who were born either before or during the fighting (see Section 2.3.3).

the ages of the children selected for height and weight measurements (children were only measured before 48 months in 1995 survey). However, the identification strategy employed, coupled with evidence from estimations on alternative samples (Section 2.6.2), suggests that the main findings of this analysis are not unduly influenced by this.

The primary outcome of interest is the height-for-age z-score of children within the sample, which measures the number of standard deviations from the (age and gender-specific) median height of a child in a healthy reference population.[9] This is a long-run measure of exposure to poor health and nutrition, implying that an accumulation of past health deficits should still be visible in the data, even amongst older children.[10]

Figure 2.1: Height-for-Age z-scores by Month of Birth



[9]The impact of conflict exposure on alternative anthropometric measures is considered in Section 2.6.6.

[10]This measure is also indicative of an individual's future health and economic status. An extensive literature has found strong links between early childhood height and physical and cognitive development, morbidity, mortality, schooling and economic productivity in later life. See Strauss and Thomas (2007) for a review.

The curves in Figure 2.1 illustrate the relationship between height-for-age z-scores and a child's month of birth, generated using locally weighted, scatter-plot smoothing (LOWESS). Each separate curve represents data points from a different DHS survey, with the start and finish dates of the three data collections shown by the vertical lines in Figure 2.1. In each case, the far left of the curve represents the oldest children at the time of survey, while the z-scores of younger children will be represented by the area where the curve is closest to (or within) the time period where the survey took place. Figure 2.1 clearly shows that each curve slopes upward towards the dates of the respective survey, indicating that as children in the sample grow older their height lags increasingly behind that of the reference population.[11] As only around 1 in 10 children in the sample were directly exposed to the fighting, this suggests that, even in the absence of conflict, children born in Uganda during this period are likely to have faced substantial health challenges.

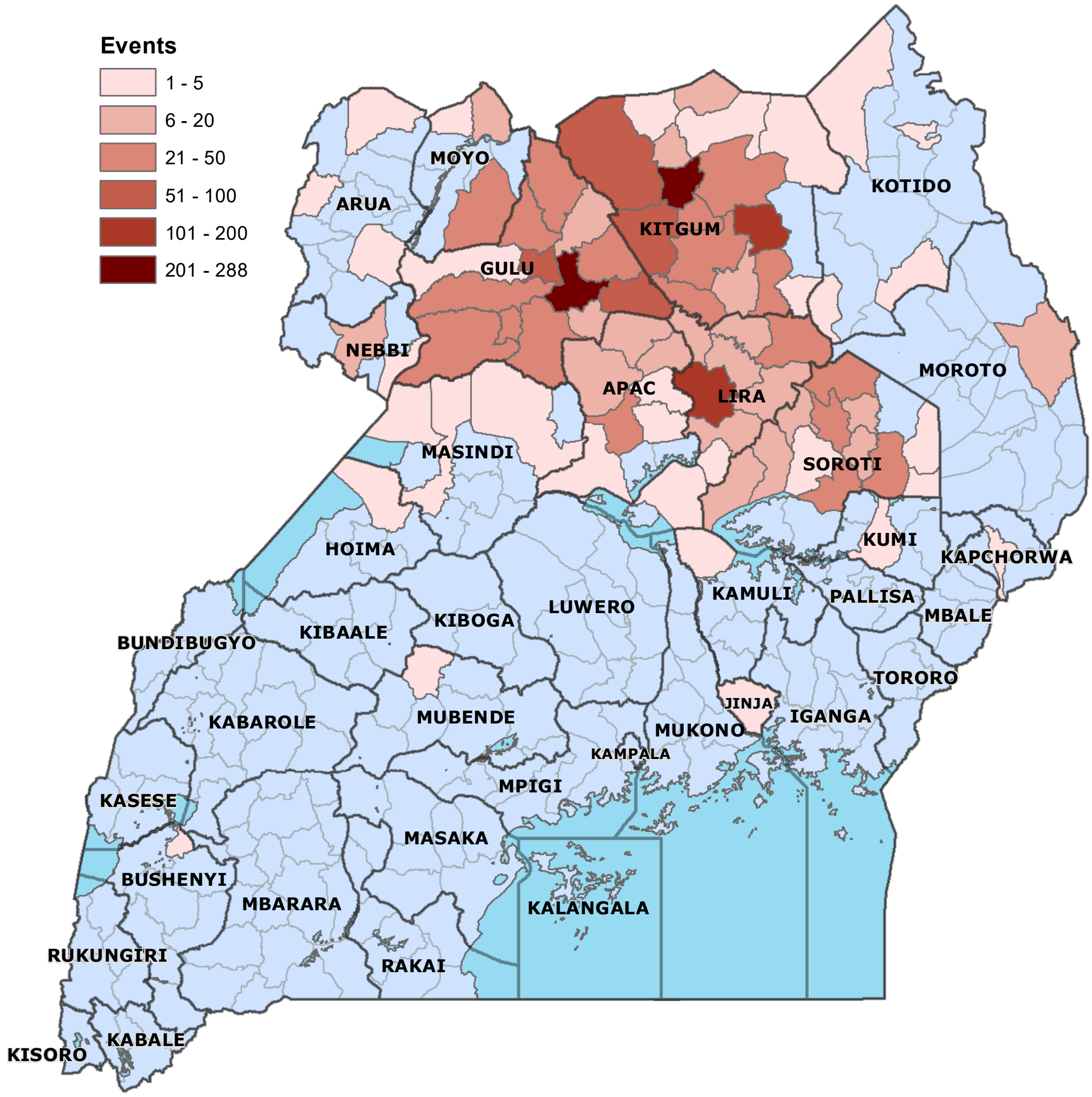
2.3.2 Conflict Data and Localities

Using the sub-county locations of the DHS clusters and the children's dates of birth, health outcomes are linked to information on LRA conflict events obtained from the Armed Conflict Location Events Database (ACLED).[12] For the purposes of this analysis, conflict events are defined as any recorded battle or act of one-sided violence, where the LRA is listed as one of the actors. The ACLED dataset records 1947 such events, occurring between January 1987 and March 2007.

[11] A cumulative deficit in height-for-age, especially before 3 years of age (see Fig 2.1), conforms with a long-recognised pattern in low-income countries (Martorell and Habicht, 1986).

[12] Events occurring from the 1st of January 1997 onwards are taken from the current version of ACLED, accessed from https://www.acleddata.com/data/ on 17th June 2019 (Raleigh et al., 2010). Events taking place before this date are obtained from an earlier version of the same data, compiled by the Peace Research Institute Oslo (Raleigh and Hegre, 2005).

Figure 2.2: Number of Recorded LRA Conflict Events, by Locality



Source: Based on ACLED dataset (Raleigh and Hegre, 2005; Raleigh et al., 2010)

The twenty-year duration of the LRA insurgency presents clear challenges for identification. For example, no children in the DHS data were measured prior to the first event taking place.[13] A second concern relates to the geographical proximity of the children's households to the recorded locations of where events

[13] One earlier DHS survey was conducted in Uganda in 1988/89. However, other than the West Nile region (Arua, Moyo and Nebbi districts in Fig 2.2), the survey was only conducted in the south and southwest of the county.

actually took place.[14] Therefore, to mitigate these concerns, while fully exploiting the spatial and temporal variation of the fighting, Uganda is sub-divided into 309 *localities*, along county and subcounty administrative boundaries. These localities are constructed by sub-dividing any county, where the distance between two border locations exceeds 50kms, along subcounty administrative boundaries (if possible). For counties/municipalities where this distance never exceeds 20kms, the area is combined with an adjoining county.[15] A description of the areas covered by the localities can be found in Appendix 2.1.

Figure 2.2 shows the extent of the LRA conflict, which took place between 1987 and 2007, along with the sub-division of Uganda's 38 districts (as of 1991) into the 309 localities. Darker shaded areas represent localities that experienced a greater number of events, based on the ACLED dataset (Raleigh and Hegre, 2005; Raleigh et al., 2010). The highest intensity of fighting can clearly be observed in the Northern Acholi and Langi sub-regions (Kitgum, Gulu, Apac and Lira), as well as the district of Soroti, further to the southeast.

2.3.3 Exposure to Conflict

The initial approach to assigning whether or not a child is considered as being exposed to the fighting is based on the methodology used by Bundervoet et al. (2009) and requires first defining a *conflict window* for each locality where at least one event took place. This conflict window represents the time period between (and including) the calendar month where the first and last recorded events occurred in the locality. The simplest definition of exposure assigns any child who was alive during the conflict window as exposed to the conflict, while omitting from the sample all children who were born after the conflict window. Initial estimations

[14] Studies using a similar methodology have classified children living more than 100kms from any conflict event as exposed to the fighting. For example, see Akresh et al. (2012b) and Minoiu and Shemyakina (2014).

[15] One exception to this approach occurs in the case of the geographically small area of Gulu municipality, which occupies a location directly between Aswa and Omoro counties. In this instance, the three counties/municipalities are combined, and then subdivided along subcounty lines thereafter.

will, therefore, only utilise observations from children who were either exposed to the conflict or measured prior to any fighting taking place. In the analysis which follows, a number of alternative definitions of conflict exposure are considered, both in the initial estimations and by way of robustness checks in Section 2.6. The table provided in Appendix 2.2 provides a more detailed description of the 49 localities in the sample which experienced at least one event, henceforth, referred to as *conflict localities*.

Columns 1 to 3, in Table 2.1, compare the characteristics of children in conflict and non-conflict localities, whereas columns 4 to 6 report the same variables for exposed and non-exposed children, within conflict localities. One of the clearest disparities lies in the difference between the average heights of children and their mothers, across the two sets of results. Surprisingly, Table 2.1 suggests that, on average, mothers and children are actually taller in the areas where the fighting took place (relative to areas which were unaffected). Furthermore, children who were exposed to the fighting, within these areas, appear significantly taller than those who were not exposed. This seemingly counterintuitive finding can be understood by recognising that the majority of the fighting took place in the districts of Gulu, Kitgum, Lira, Apac and Soroti (see *Household Characteristics* in Table 2.1), where the Nilotic ethnic groups, which dominate these areas, have far closer historical links to the famously tall Dinka people of southern Sudan or the Maasai of Kenya and northern Tanzania than those in the south and southwest of Uganda (Shoup, 2011). On closer examination of Table 2.1, it is also clear that, within the observations from conflict localities, children exposed to the fighting have a higher probability of coming from one of these five districts (see columns 4 to 6). Therefore, the over-representation of Nilotic children in the exposed group (within the conflict areas) would also potentially explain the significant height (and possibly weight) disparity in column 6.[16]

[16]For example, repeating the tests reported in Table 2.1, column 6, with the addition of district fixed effects, fully negates any significant difference in height or weight, between exposed and non-exposed children (height-for-age p-value = 0.730, weight-for-age p-value = 0.606).

Table 2.1: Descriptive Statistics

Variable	(1) Conflict	(2) Non-Conflict	(3) Diff. (1)-(2)	(4) Exposed	(5) Not Exposed	(6) Diff. (4)-(5)
	Locality			*Conflict Exposure*		
	Mean	Mean	*p*-value	Mean	Mean	*p*-value
Child's Characteristics						
Height-for-age z-score	−1.34	−1.55	0.065*	−1.19	−1.57	0.009***
Weight-for-age z-score	−0.99	−1.11	0.134	−0.87	−1.18	0.010**
Height-for-weight z-score	−0.19	−0.18	0.818	−0.16	−0.24	0.228
Child is female	0.52	0.51	0.347	0.51	0.54	0.202
Child age (years)	2.14	2.04	0.030**	2.24	1.98	0.012**
Birth order 1-3	0.46	0.46	0.853	0.44	0.49	0.155
Birth order 4-6	0.32	0.33	0.608	0.35	0.29	0.042**
Birth order 7+	0.22	0.20	0.506	0.21	0.22	0.761
Mother's Characteristics						
Mother's age (years)	28.54	28.17	0.190	28.99	27.85	0.036**
Mother's height (cms)	160.85	157.87	0.000***	161.18	160.36	0.379
Mother is married/cohabiting	0.91	0.91	0.909	0.91	0.91	0.902
Mother has no education	0.31	0.26	0.207	0.26	0.39	0.036**
Mother has primary education	0.58	0.61	0.235	0.61	0.54	0.150
Mother has secondary education	0.11	0.13	0.596	0.13	0.08	0.232
Mother is head	0.90	0.92	0.118	0.92	0.87	0.135
Mother is daughter of head	0.07	0.05	0.046**	0.06	0.08	0.329
Mother not head/daughter	0.03	0.03	0.639	0.02	0.05	0.139
Mother's time in location	13.04	11.79	0.155	11.96	14.66	0.053*
Household Characteristics						
Urban area	0.20	0.18	0.844	0.30	0.04	0.011**
Gulu, Kitgum, Lira, Apac, Soroti	0.61	0.02	0.000***	0.66	0.53	0.487
Household size (time of survey)	6.78	6.54	0.196	6.62	7.03	0.337
Household head is male	0.82	0.83	0.485	0.80	0.85	0.015**
Household head's age (years)	36.20	35.64	0.194	35.96	36.55	0.485
Observations	1447	7755		869	578	

Test standard errors clustered at the locality * p<0.1, ** p<0.05, *** p<0.01

To formally test the likelihood of underlying anthropometric differences between ethnic groups, Appendix 2.3 reports an analysis of the relationship between ethnicity and height, conducted on a sample of women who would have achieved adult stature before the LRA insurgency began in 1987. This simple empirical analysis confirms that adult women from the groups which dominate the five most affected districts have significant height advantages over those in other conflict localities and, indeed, elsewhere in Uganda. A naïve analysis which does not recognise this would (at best) understate the health impacts of the LRA insurgency on those who were most affected by the conflict.

The measures in Table 2.1 also suggest significant variation in the mean ages of children between groups, with evidence of relatively older children present in conflict localities and also in the group who were exposed to the fighting. In the latter case, a higher probability of being exposed to the war, amongst relatively older children, would almost certainly be a consequence of having a longer period of time in which an event can occur, whereas the fact that this group are older, on average, would account for a higher birth order and also mothers who are relatively older when the survey is conducted. The reason behind the difference in the mean ages of children in conflict and non-conflict localities is not initially so clear, however. One explanation is that this may be related to the differences in the sampling methodology in the three DHS survey waves. In the 1995 wave, only children who were younger than 4 provided height and weight measurements (as opposed to children younger than 5 in the 2000 and 2006 surveys). The 1995 survey also did not cover the district of Kitgum, where all but one of the 17 localities were affected by the insurgency. Therefore, in 1995, both older children and conflict localities are relatively under-sampled.

A significant difference in the mean ages of children between the groups in Table 2.1 clearly presents a possible threat to identification.[17] However, it is important to recognise that the identification strategy outlined in the following section addresses this by only estimating the effects of the LRA conflict using variation within birth cohorts. Similarly, all variables which show significant heterogeneity between the

[17]Assuming relatively older children will have accumulated larger height deficits (see Figure 2.1), the measured effect of conflict exposure would be downward biased if this disparity in ages was not acknowledged in the identification strategy outlined in Section 2.4.

groups in Table 2.1 (such as the gender of the household head or the education level of the mother) are added as additional controls in a separate set of estimations, to assess the effect of their inclusion on the main coefficients of interest.[18]

2.4 Empirical Strategy

To attempt to measure the causal effect of the war on the height-for-age z-scores of affected children, three alternative definitions of conflict exposure are considered. The baseline model (2.1) follows Bundervoet et al. (2009) in measuring the impact of the LRA insurgency on children who were both living in affected localities and exposed to the conflict. Given a higher relative mean age of exposed children, coupled with possible differences in the underlying anthropometrics of those in conflict and non-conflict localities (see Appendix 2.3), the following baseline specification identifies the effect of exposure to the war, using only variation within localities and within birth cohorts.

$$haz_{ijt} = \alpha_j + \delta_t + \beta_1(Conflict\ Locality_j * Exposed_i) + \phi_1 female_i + \varepsilon_{ijt} \quad (2.1)$$

In model (2.1), haz_{ijt} represents the height-for-age z-score of child i, living in locality j, who was born in year t, while the terms α_j and δ_t represent locality and birth cohort (year of birth) fixed effects. $Conflict\ Locality$ takes value 1 if the child comes from an area which experienced at least one conflict event (0 otherwise) and the $Exposed$ term in this baseline model is simply a binary

[18]Differences in the educational attainment of the children's mothers may reflect a relatively higher number of children in urban areas who were exposed to the fighting (see 2.1). Repeating the tests reported in Table 2.1, controlling for rural/urban status, produces p-values in excess of 0.25 for all education groups.

indicator for whether or not a child was alive during the conflict window. In the first set of estimations, only a control for the child's gender is included. The final term ε in model (2.1) denotes a random, idiosyncratic error.

One potential drawback of model (2.1) is that only a relatively small number of conflict localities contain both exposed and non-exposed children (see Appendix 2.2). Given the inclusion of the locality fixed effect α_j, the coefficient of interest β_1 will only be estimated using variation within this small subset of localities. Furthermore, due to the longer duration of the fighting in the worst affected districts (resulting in fewer pre-war observation) the areas where variation in the binary measure of exposure exist will commonly be found on the peripheries of the most intense fighting.[19] In light of this, model (2.2) replaces *Exposed* with the variable *Ex Duration*, measuring the number of months a child was alive during the conflict window. In (2.2) the coefficient β_2 measures the effect of an additional month of exposure, under the implicit assumption of a linear relationship between months of exposure and childhood health. All other terms in (2.2) have the same interpretation as before.

$$haz_{ijt} = \alpha_j + \delta_t + \beta_2(Conflict\ Locality_j * Ex\ Duration_i) + \phi_1 female_i + \varepsilon_{ijt} \quad (2.2)$$

A third definition of exposure is shown in model (2.3), which relaxes the assumption that the marginal health impact of an additional month in the conflict window is uniform across all possible exposure durations. In (2.3), children in affected localities who were exposed to the violence are assigned to one of two

[19]The mean duration of the conflict windows in localities from the worst affected districts (Gulu, Kitgum, Lira, Apac and Soroti) is approximately 110 months, compared to only around 35 months outside of these districts. Furthermore, a relatively lower percentage of the sampled children in these districts were not exposed to the fighting (only 16.4%, as opposed to 39.7% elsewhere, based on the table in Appendix 2.2).

groups, dependant on the time they were alive during the conflict window. The coefficients β_3 and β_4 measure the effect of exposure to the fighting in a conflict locality (for the given durations), relative to those children who were not directly exposed to the war. The initial estimation of (2.3) separates the exposed sample at a conflict duration of 6 months, while in the results in Section 2.5 also defined the two groups around a split point of 12 months.

$$\begin{aligned} haz_{ijt} = \alpha_j + \delta_t + \beta_3(Conflict\ Locality_j * Ex\ 1 - 6\ months_i) \\ +\beta_4(Conflict\ Locality_j * Ex > 6\ months_i) + \phi_1 female_i + \varepsilon_{ijt} \end{aligned} \tag{2.3}$$

In each of the previous models, identification relies on the assumption that the underlying time trend in children's heights would be the same in both conflict and non-conflict localities, had the insurgency not taken place. This may be an overly strong requirement, given the focus of the fighting in the Northern districts and the large economic, cultural and ethnic divide between the North and South of the country. Following the literature, therefore, this condition can be modified by the assumption that children in conflict and non-conflict localities, *within districts*, would display similar growth pattern in the absence of exposure. This requires the inclusion of a district-specific time trend in models (2.1) to (2.3), which is indicated in the case of the baseline model below by the term γ_{jt}.

$$haz_{ijt} = \alpha_j + \delta_t + \gamma_{jt} + \beta_1(Conflict\ Locality_j * Exposed_i) + \phi_1 female_i + \varepsilon_{ijt} \tag{2.4}$$

The main results of this analysis also report estimations for each of the three definitions of exposure including a number of additional controls, relating to the child, their mother and the household in which they live. Following a discussion of the robustness of findings from the previous specifications, Section 2.7 also extends the main results to assess heterogeneity in the effect of conflict between different sub-sample groups. This section focuses on three potential sources of heterogeneity, the intensity of the fighting experienced by the child and the gender of either the head of the household or the child themselves.

2.5 Results

Table 2.2 reports results from estimations based on the models described in the previous section. The first three columns show measures of the β_1 interaction coefficient from the baseline model (2.1) (all terms in (2.1) are included in the estimations, but not reported in Table 2.2). Under the assumption of a common time trend in height-for-age z-scores, column 1 reports little impact of the war on children residing in conflict localities and exposed to the fighting (relative to those not directly affected). With the inclusion of the district-specific time trends in column 2, however, there is evidence of a 0.37 standard deviation height deficit amongst those who were exposed. Table 2.1 suggests several underlying characteristics may differ systematically between children in conflict and non-conflict localities (as well as between exposed and non-exposed children). After the inclusion of variables capturing the characteristics of the child, their mother, and the household,[20] column 3 indicates the estimate of β_1 is only minimally affected, suggesting only that the causal effect of exposure to the war may be closer to -0.39.

[20]An alternative version of Table 2.2, including the estimated coefficients on all additional controls, is provided in Appendix 2.4.

Table 2.2: The Effects of Conflict Exposure on Height-for-Age

Dep. Variable: height-for-age z-score	(1)	(2)	(3)	(4)	(5)	(6)
Conflict Locality * Exposed	-0.239	-0.369**	-0.388**			
	(0.187)	(0.186)	(0.195)			
Conflict Locality * Ex Duration				0.001	-0.011**	-0.011**
				(0.003)	(0.005)	(0.005)
Observations	9202	9202	9202	9202	9202	9202
R^2	0.079	0.147	0.205	0.079	0.147	0.205
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	No	Yes	Yes	No	Yes	Yes
Controls†	No	No	Yes	No	No	Yes

	(7)	(8)	(9)	(10)	(11)	(12)
Conflict Locality * Ex 1-6m	-0.191	-0.200	-0.234			
	(0.213)	(0.215)	(0.218)			
Conflict Locality * Ex > 6m	-0.315*	-0.659***	-0.651***			
	(0.165)	(0.194)	(0.222)			
Conflict Locality * Ex 1-12m				-0.231	-0.298	-0.317
				(0.197)	(0.201)	(0.204)
Conflict Locality * Ex > 12m				-0.258	-0.566***	-0.585**
				(0.171)	(0.200)	(0.227)
Observations	9202	9202	9202	9202	9202	9202
R^2	0.079	0.148	0.206	0.079	0.147	0.205
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	No	Yes	Yes	No	Yes	Yes
Controls†	No	No	Yes	No	No	Yes
H0: Ex 1-6(12)m = Ex > 6(12)m						
p-value	0.293	0.005***	0.020**	0.710	0.083*	0.096*

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

In (7) to (9): Not exposed ($n = 578$), 1-6m exposure ($n = 218$), exposure > 6m ($n = 651$)
In (10) to (12): Not exposed ($n = 578$), 1-12m exposure ($n = 344$), exposure > 12m (n=525)
†Child controls include: female child and birth order (grouped)
Mother's controls include: height, age, marital and education status, and relationship to head.
Household controls include: urban DHS survey cluster, age and gender of the head
The full table of results is shown in Appendix 2.4

The magnitude of the coefficients in columns 2 to 3 could be considered somewhat small, relative to comparable studies.[21] As previously noted, however, the localities which contain both exposed and non-exposed children tend to be found on the edges of the worst fighting.[22] As an alternative to the simple (yes/no) exposure variable, columns 4 to 6 report the effect of an additional month of exposure to the conflict, measured by β_2 in model (2.2). Again, significant coefficients are found in the estimations including district-specific time trends, with both column 5 and 6 reporting a 0.011 standard deviation deficit in height for each additional month a child spends within the conflict window.

Columns 7 to 9 report results from estimations based on model (3). This approach splits the sample of children into groups, based on whether a child's duration of exposure is less than (equal to) or more than 6 months. The results in columns 8 and 9 show clear evidence of a greater deficit for those children who were exposed to the fighting for more than 6 months, with the results suggesting this may be in excess of -0.65 standard deviations (relative to those not affected). In contrast, it is not possible to reject the hypothesis of no impact for those exposed over shorter periods of time (at any of the levels reported). In columns 8 and 9, a Wald test of the equality of the impact of different exposure durations also confirms heterogeneity in the negative effects experienced by the two groups (p-values reported in Table 2.2).

The final set of results in columns 10 to 12 splits the sample of exposed children at 12 months, as opposed to 6. The results appear broadly similar to those in columns 7 to 9, although the estimated effect of inclusion in the longer exposure group appears less negative. There also exists weaker evidence of a significant difference between the shorter and longer periods of exposure when separating

[21] Bundervoet et al. (2009) found a -0.348 to -0.525 standard deviation height deficit in their study of the 1993-2005 civil war in Burundi, Akresh et al. (2012b) reported coefficients of -0.447 to -0.454 for the border war between Eritrea and Ethiopia, while Minoiu and Shemyakina (2014) found an effect of exposure to conflict between -0.250 and -0.414 (for the Côte d'Ivoire).

[22] For example, only a single locality in the Acholi-land districts of Gulu and Kitgum provides variation in the initial definition of exposure, while these two districts account for 1384 of the recorded events (71.1% of the total 1947 events).

observations at 12 months duration (see p-values in Table 2.2). This pattern of results suggests an acceleration in the negative health effects of exposure to the LRA conflict between 6 and 12 months, and minimal evidence of further detrimental effects thereafter.

With no variance in the binary exposure measure in the majority of localities from the worst affected districts, the models which exploit variation in the length of exposure are preferable. Having split the sample of affected children at 6 months exposure, Table 2.2 also suggests little additional insight can be gained from also splitting the sample at 12 months. Therefore, the remainder of the analysis will focus attention on the estimations used to generate the results in column 6 and column 9 of Table 2.2 (with both additional controls and district-specific time trends).

2.6 Robustness

This section considers the robustness of the results reported in Table 2.2 (columns 6 and 9) to the use of samples intended to address potential sources of bias, alternative definitions of conflict localities (and the conflict window), and the use of alternative anthropometric health measures.

2.6.1 Endogenous Selection out of Conflict Exposure

Arguably, the clearest source of potential bias comes from an endogenous selection process governing which children are exposed to conflict. For example, if relatively wealthier households were more easily able to relocate to safer areas, and in the likelihood that there exists a positive correlation between the health status of children and the economic status of households, the results in Table 2.2 will be biased (downwards). Unfortunately, there exists no information on household migration in the Ugandan DHS surveys. However, this information is recorded in the Ugandan National Household Survey 2005/06 (UNHS), which was collected just prior to the 2006 DHS wave. The UNHS records whether household members moved to their current location since 2001 and provides information on the district of origin and their main reason for relocation.

Of the 2749 households in the UNHS, who either resided in one of the 16 conflict-affected district or moved from one of these areas, 283 had relocated with the stated reason of avoiding insecurity (see Appendix 2.5). The vast majority of these households (89.4%) had settled within the same district, however, implying that the bulk of conflict-driven relocation would have occurred at a relatively local level.[23] If households with healthier children were able to leave conflict areas, either before the fighting started or at a relatively earlier stage, there should exist a negative correlation between time in the current location and the child's height, within localities yet to be affected by the fighting. Testing for this correlation indicates that the relationship between time in location and height-for-age z-scores is found to be significantly negative ($p < 0.05$) only in three districts, Gulu, Hoima and Kotido.[24] To establish the extent of any possible bias in the results reported in Table 2.2 (columns 6 and 9), these estimations are repeated on a sample which omits observations from the three districts listed above. The results can be found in columns 1 and 2 of Table 2.3.

Comparing the estimated effects in Table 2.3 to the initial results suggest that the linear exposure duration effect (in column 1) may not be wholly robust to the omission of the three districts above, although the magnitude of the effect is only reduced slightly. However, in the case of the grouped results in column 2, any downward bias in the measured effects appears to be relatively small and certainly does not warrant any re-interpretation of the findings.[25]

[23] A summary of the migration pattern of households from the affected districts, taken from the Uganda National Household Survey 2005/06, is provided in Appendix 2.5, where the migration history of the household is assumed to follow the relocation history of the household head.

[24] A description of the estimation used to test for a negative correlation between time in location and height-for-age z-scores can be found in Appendix 2.6.

[25] The reason that selection effects appear negligible may be linked to the seemingly random nature of the LRA attacks (Blattman and Annan, 2010; Adelman and Peterman, 2014). Furthermore, in the case of relocation to IDP camps, this was often involuntary, with little or no notice, making endogenous selection out of conflict exposure even less likely (Fiala, 2015).

Table 2.3: Alternative Samples and Assumptions on Exposure

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable: height-for-age z-score	*Gulu, Hoima and Kotido Omitted*		*Districts and Ages Consistent all DHS*		*Re-classify Conflict Localities in South*	
Conflict Locality * Ex Duration	-0.009		-0.015**		-0.011**	
	(0.005)		(0.007)		(0.005)	
Conflict Locality * Ex 1-6m		-0.220		-0.135		-0.369
		(0.224)		(0.246)		(0.303)
Conflict Locality * Ex > 6m		-0.614***		-0.598***		-0.742***
		(0.234)		(0.223)		(0.269)
Observations	8889	8889	7973	7973	9202	9202
R^2	0.200	0.201	0.211	0.212	0.205	0.206
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
H0: Ex 1-6m = Ex > 6m						
p-value		0.040**		0.011**		0.055*

	(7)	(8)	(9)	(10)	(11)	(12)
	Children in IDP Camps Omitted		*Extending Conflict Window + 6m*		*Extending Conflict Window + 12m*	
Conflict Locality * Ex Duration	-0.009*		-0.012**		-0.014***	
	(0.005)		(0.005)		(0.005)	
Conflict Locality * Ex 1-6m		-0.213		-0.079		0.165
		(0.228)		(0.183)		(0.163)
Conflict Locality * Ex > 6m		-0.638***		-0.617***		-0.611***
		(0.232)		(0.186)		(0.171)
Observations	9091	9091	9244	9244	9291	9291
R^2	0.205	0.206	0.204	0.205	0.204	0.205
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
H0: Ex 1-6m = Ex > 6m						
p-value		0.036**		0.002***		0.000***

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

2.6.2 Variation in Survey Coverage between Waves

As the population of interest is Ugandan children aged less than 60 months, two further potential sources of bias can be found in the districts selected for the DHS surveys and the ages of children measured in these surveys. The 1995 wave only measured children who were aged less than 4 years old and did not survey the Northern district of Kitgum. In contrast, the survey conducted in 2000 omitted both Kitgum and Gulu, along with the districts of Kasese and Bundibugyo, while increasing the age range to 5 years of age.[26] Similarly, the final 2006 survey measured children up to 5 years old, but in this final wave, all districts of Uganda were surveyed.

The districts which were not surveyed in either 1995 or 2000 were omitted due to fears of insecurity, implying an obvious correlation between the probability of inclusion in the original DHS sample and conflict exposure. This suggests the estimated impact of the LRA conflict in Table 2.2 may contain a positive bias (Children in the worst affected areas were not surveyed). The absence of children aged 4 or over in the 1995 sample implies a further potential selection effect, although in this case, the sign of the bias is less clear.[27] To address these concerns, the key models are re-estimated on a sample of children which is consistent (both in age and geographical coverage) across all three DHS survey waves.

The results in columns (3) and (4) of Table 2.3 show the impact of the war on children aged less than 4 years, and not living in the districts of Gulu, Kitgum, Kasese and Bundibugyo. This alternative sample should be fully representative of

[26]The latter two Western districts were experiencing insecurity during this period, due to another rebel group ADF-NALU operating along the border of Uganda and the Democratic Republic of Congo.

[27]For example, the relative under-sampling of older children in 1995, and the likelihood of cumulative health deficits from longer periods of exposure, would be expected to generate a positive bias. However, if the effects of a given length of exposure are more detrimental to children at a younger age, the bias would be in the opposite direction.

a population with these characteristics. Unsurprisingly, given that this sub-sample omits over 1200 observations and considers children over a different age range, the magnitude of the coefficients varies to some degree (relative to those in column 6 and 9 of Table 2.2). However, the fact that the broad pattern of results remains unchanged is further reassurance that any sample selection bias is not sufficient to warrant a new interpretation of the previous results.

2.6.3 Isolated Conflict Localities

Three conflict localities in Fig 2.2 stand out as being much farther South than the districts where the vast majority of the fighting took place. The localities in Bushenji, Mubende and Jinja, experienced only five events in total during the entirety of the LRA war. Arguably, it may not be appropriate to treat children exposed to these more isolated pockets of violence as experiencing a continuation of the main conflict in the North.[28] Therefore, as an alternative definition of exposure to conflict, the results reported in columns 5 and 6 of Table 2.3 are derived from re-classifying all children in these areas as not exposed to the conflict.

Table 2.3 indicates that the effect of the continuous measure of exposure (column 5) remains unchanged. However, the more negative coefficients in column 6 suggest that the inclusion of the children in these Southern localities (in the original estimations) may have served to moderate the estimated impact of the war. With the limited number of events in these areas, exposed and unexposed children (the base category in column 6) are likely to be more similar, relative to those in areas more central of the conflict, decreasing the measured effect. Although this clearly leads to a change in the magnitude of the coefficients, again, the overall interpretation of the results remains the same.

[28]In the case of Jinja, in particular, where the two events which took place occurred more than 15 years apart, many of those children categorised as exposed to the conflict will have had relatively little direct experience of the fighting.

2.6.4 The Health Impacts of IDP Camps

Given the dire conditions experienced by those in the government-sanctioned IDP camps (WHO, 2005; Bozzoli and Brück, 2010), it initially seems plausible that the negative health impacts attributed to the war are, in reality, capturing the consequences of children being subjected to these conditions during displacement. Were this the case, the same detrimental health effects could potentially be observed anywhere where poor sanitation and disease were prevalent (even in the absence of armed conflict).

Within the final sample, only 111 children were measured while residing in an IDP camp, making it unlikely that these observations are responsible for determining the results in Section 2.5. However, this small number of observations does highlight that a relatively larger number of IDP children were not included in the sample due to the mother arriving in the DHS cluster location after the child was born.[29] If the impact of conflict exposure would have been worse for these children, the results in Table 2.2 will underestimate the true effect of the war. Therefore, to establish the magnitude of any bias, while ruling out the possibility that the remaining IDP observations exert undue influence on the results, columns 7 and 8 of Table 2.3 report estimates obtained from a sample which omits IDP children.

When considering the grouped lengths of exposure in column 8 of Table 2.3, the estimated health effects of the conflict, on children who had never lived in IDP camps, are similar to those obtained from the full sample. Where the impact of exposure duration on children's heights is assumed to be linear, column 7 indicates a slightly smaller effect, relative to Table 2.2. Overall, the results suggest that any positive bias, due to a higher probability of IDP children being dropped from the sample, is minimal. Furthermore, it is clear that the remaining observations from the IDP camps are not solely responsible for driving the results in Section 2.5.

[29]From the original 205 observations of IDP children, 63 were omitted due to the mother relocating since the child's birth (30.7%), while this reason accounted for only around 19% of non-IDP children dropped from the original DHS data.

2.6.5 Extending the Conflict Window

In the empirical strategy defined in Section 2.4, the negative effects of being exposed to the war are assumed to be limited to the period between the calendar months in which the first and last events take place. Although this simplifying assumption avoids the need to place an arbitrary limit on how long after the conflict window children's health may still be affected, it risks understating the full impact of the fighting. To determine the sensitivity of results to alternative assumptions regarding the duration of exposure, the conflict window is extended by 6 months after the last event in columns 9 and 10 of Table 2.3, and by 12 months in columns 11 and 12.[30]

The linear effects reported in columns 9 and 11 appear to show an increasingly negative impact of an additional month of fighting, when extending the limit for exposure beyond the original conflict window. These results would suggest that the initial measure of 0.011 (column 6 in Table 2.2) may represent a conservative estimate of the cumulative effect of the conflict. The results in columns 10 and 12, instead, show evidence of a reduction in the estimated impacts for both grouped durations of exposure. This will come as a result of shifting the distribution of exposure duration upwards, such that some children born after the conflict (and dropped in the initial estimation) will now fall within the 1 to 6 month category, while some of those in this group will now be considered as exposed for a longer duration of time. Although the estimated effects of more than 6 months exposure are smaller in both columns 9 and 11, the results still imply a height deficit in excess of 0.6 standard deviations, with no significant impact for shorter periods of exposure.

[30] As children who were born after the last event were dropped from the original sample, extending the conflict window also increases the sample size used in these estimations.

2.6.6 Alternative Anthropometric Health Measures

Height-for-age z-scores are intended to detect evidence of long-term poor nutrition and exposure to disease. However, they represent only one potential anthropometric measure of health status available in the DHS data. If the height deficits uncovered in Section 2.5 are, indeed, measuring the causal effect of conflict exposure on childhood health, evidence should exist in more short-term measures of health status, in particular amongst children who were surveyed while the fighting was still taking place. Table 2.4 presents an alternative set of estimations of the impact of the war on children's health, where the results shown in columns 3 to 6 replace the dependent variable with measures of short-run health status, based on children's weight (as opposed to their height). As these measures are intended to capture short term health deficits, observations coming from children who were measured after the conflict window are dropped from the sample for these estimations. This should serve to limit the possibility of any unrelated health shocks impacting a child's weight during the post-conflict period.

As the results in Table 2.4 are estimated using a subset of the original sample, the first two columns represent identical estimations to columns 6 and 9 of Table 2.2 (including the original height-for-age measure of health status as the dependent variable). The pattern of results is broadly similar to that in Table 2.2, although the estimated coefficients on the grouped exposure terms are less negative (even positive in the shorter exposure group). This is likely due to a shift in the age distribution resulting from dropping those children measured after the conflict window. This approach will favour the inclusion of younger children in the two exposure groups, who should have experienced shorter durations of conflict, on average.[31]

[31] The mean age of children exposed to the war in the original sample was 26.9 months, yet this falls to 23.7 months with the omission of those measured after the conflict window.

Table 2.4: Alternative Measures of Health Status

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable:	*Height-for-Age z-score*		*Weight-for-Age z-score*		*Weight-for-Height z-score*	
Conflict Locality * Ex Duration	-0.013**		-0.016***		-0.007**	
	(0.006)		(0.004)		(0.003)	
Conflict Locality * Ex 1-6m		0.219		0.065		-0.251
		(0.236)		(0.249)		(0.211)
Conflict Locality * Ex > 6m		-0.580***		-0.947***		-0.672***
		(0.215)		(0.193)		(0.180)
Observations	8978	8978	8978	8978	8978	8978
R^2	0.206	0.207	0.192	0.195	0.125	0.126
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
H0: Ex 1-6m = Ex > 6m						
p-value		0.000***		0.000***		0.022***

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

The coefficients in columns 3 and 4 report estimates of the effect of conflict exposure on weight-for-age z-scores, which measure the deviation of a child's weight from that of the median child (of the same age and gender) in the reference population. Under the linear assumption in column 3, the results indicate that each additional month of exposure reduces a child's weight by 0.016 standard deviations, while the grouped duration coefficients in column 4 still show little evidence of health impacts for an exposure duration of 6 months or less. For children exposed to the conflict for more than 6 months, however, the estimated weight deficit is remarkably large at 0.947 standard deviations, indicating even greater heterogeneity in health impacts between the two groups, when using weight-for-age as a measure of health status.

The final set of estimations employ weight-for-height as the dependent variable, which records the deviation of a child's weight from the median child, of the same height, in the reference population. The results in columns 5 and 6 follow the same pattern as those in the previous estimations. Again, there is evidence of a significant deficit in weight for each additional month of exposure in column 5, while a strong negative impact is found in column 6, but only for those exposed for more than 6 months. In light of these results and those in columns 3 and 4, the similarity in the interpretation of effects obtained using weight, as opposed to height, as an indication of health status, provides further evidence in favour of the findings in Section 2.5.

2.7 Heterogeneity

Having established to what extent the main results in columns 6 and 9 of Table 2.2 are robust to alternative assumptions, this section considers heterogeneity in the estimated impacts of the war on childhood health status. Three potential sources of variation in the effect of the conflict are considered, heterogeneity between male and female children, heterogeneity between children belonging to male and female-headed households and heterogeneity over the intensity of the fighting experienced.

2.7.1 Heterogeneity by Child's Gender

To test for evidence of a difference in the effect of the conflict between male and female children, this study follows Minoiu and Shemyakina (2014) by interacting the conflict exposure terms in models (2.2) and (2.3) with a female indicator variable. As it is most likely that any gender bias would favour male children (Baird et al., 2011; Dagnelie et al., 2018), a negative coefficient on this interaction term, when included in model (2.2), would be evidence of an additional month of exposure to the war exerting a relatively higher toll on the health of female children, than males. Similarly, a negative coefficient on the interaction of the female indicator variable, with either of the grouped exposure terms in model (2.3), would signal inclusion in these groups implied relatively worse health outcomes for female children.

Table 2.5: Heterogeneous Effects of Conflict Exposure

Dep. Variable: height-for-age z-score	(1)	(2)	(3)	(4)	(5)	(6)
Conf Loc * Ex Duration	-0.010*		-0.011**		-0.011*	
	(0.006)		(0.005)		(0.005)	
Conflict Loc * Ex 1-6m		-0.238		-0.177		-0.204
		(0.278)		(0.210)		(0.298)
Conflict Loc * Ex > 6m		-0.630***		-0.668***		-0.612**
		(0.236)		(0.220)		(0.244)
Conflict Loc * Ex Duration * Female Child	-0.002					
	(0.003)					
Conflict Loc * Ex 1-6m * Female Child		0.010				
		(0.222)				
Conflict Loc * Ex > 6m * Female Child		-0.036				
		(0.121)				
Conflict Loc * Ex Duration * Female Head			0.001			
			(0.004)			
Conflict Loc * Ex 1-6m * Female Head				-0.328		
				(0.259)		
Conflict Loc * Ex > 6m * Female Head				0.111		
				(0.121)		
Conflict Loc * Ex Duration * Intensity					-0.003	
					(0.006)	
Conflict Loc * Ex 1-6m * Intensity						-0.054
						(0.239)
Conflict Loc * Ex > 6 months * Intensity						-0.193
						(0.217)
Observations	9202	9202	9202	9202	9202	9202
R^2	0.205	0.206	0.202	0.203	0.202	0.203
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Table 2.5 reports results for the two sets of estimations, based on the models used to generate the results in columns 6 and 9 of Table 2.2. These results are shown in the first two columns of the table. Although the interaction terms are negatively signed in column 1 and the longer duration group in column 2, the results show no evidence of any additional effects of exposure to conflict amongst female children.[32]

2.7.2 Heterogeneity by the Gender of the Household Head

It is possible that gender bias in the impact of the war may operate, not through the gender of the child, but through the gender of the head of the household. For example, if limited resources available during conflict were disproportionately under the control of males, or women experience increased vulnerability to crime as a result of the breakdown of law and order (Byrne, 1996), children within female-headed households may suffer additional health consequences not shared by those in households where the head was male.

Columns 3 and 4 of Table 2.5 report the interaction of measures of conflict exposure with the gender of the household head. In both sets of estimations, it is not possible to discern any significant health disadvantages amongst children in female-headed households. Taken in conjunction with the results in columns 1 and 2, these findings suggest that the magnitude of effects were not defined by the gender of those affected.

2.7.3 Heterogeneity by the Intensity of Fighting

The final results in Table 2.5 report estimates intended to determine whether the negative health impacts of the war are weighted towards children in areas where the intensity of the conflict was greatest. The measure used to represent the intensity

[32]The absence of any evidence of a gender bias in conflict-related health outcomes mirrors the findings of Akresh et al. (2011) and Minoiu and Shemyakina (2014).

of the fighting experienced is the average number of days where a conflict event took place (event days), across all months where the child was alive during the conflict window.[33] Interacting this intensity measure with the conflict exposure term in column 5 generates a coefficient which measures how the health impact of an additional month of exposure changes, as a result of a child experiencing one more event day per month exposed. In the grouped duration model (column 6), the coefficients on the interaction terms measure how an additional event day per month modifies the effects of inclusion in either of the two groups.

In a similar manner to the gender interactions in the previous sections, there is no evidence of significant heterogeneity, associated with the intensity of the fighting experienced (at the levels shown in Table 2.5). Instead, these results suggest negative health effects are predominantly determined by length of exposure, as opposed to the concentration of events within this time period.

2.8 Conclusion

This study exploits spatial and temporal variation in the spread of the LRA insurgency in northern Uganda to estimate the causal effect of exposure to conflict on the health status of children. Linking data on 9202 children, aged between 0 to 5 years old, to the locations of 1497 LRA conflict events, evidence is found of irreversible health deficits amongst those exposed to the fighting for more than 6 months. The main results of the analysis imply that, on average, children within this group display height-for-age z-scores which are 0.65 standard deviations lower than those who were not directly impacted by the war. Furthermore, amongst children in this group (aged 0 to 5) a deficit in excess of 0.6 standard deviations is

[33]Based on this approach, the mean value of conflict intensity is 0.3 event days per month exposed, although the median is only 0.03 event days. Skewing of the distribution towards lower conflict intensity occurs due to 47% of those children born within the conflict window experienced a conflict intensity of 0, implying that all events occurred outside of the period between the month they were born and the month they were surveyed.

found using a number of alternative samples and definitions of exposure.[34] When considering the height deficits amongst the sample of children impacted by the war for shorter periods of time, the negative effects are found not to be significantly worse than those experienced by children who were unaffected by the fighting.[35]

The heterogeneity in health impacts between affected children in the two duration of exposure groups can also be observed in more short-run anthropometric measures of health status. Estimates employing both weight-for-age and weight-for-height measures follow an identical pattern amongst children measured in areas still exposed to the fighting. The negative effects of more than 6 months exposure on weight-for-age z-scores are particularly alarming, suggesting an average deficit approaching 1 standard deviation, relative to those who were not directly exposed to the conflict. Heterogeneous effects within the different groups of exposure duration appear largely absent, however, with no evidence in the sample of any gender bias in the negative health impacts of the war, either between male and female children or those living in male of female-headed households, and no significant heterogeneity observed amongst children who experienced different intensities of fighting. While the absence of such heterogeneity implies the negative health effects of the conflict were not disproportionately weighted towards one specific group, it also suggests that no group was capable of protecting themselves from these effects better than any other.

The extensive literature on the long-run negative impacts of early childhood health shocks suggests the indirect effect of the LRA insurgency may be felt in the loss of health, education and economic wellbeing of those affected by the conflict at an early age. Although *ex-post* interventions aimed at rebuilding the war-affected

[34]Interestingly, one such alternative sample found that results were robust to the omission of observation from IDP camps, suggesting camp conditions did not generate additional long-run health deficits, over and above those observed for children who were exposed to the fighting elsewhere.

[35]The main results also suggest a cumulative deficit in height, of 0.11 standard deviations, for each additional month of exposure, although the linear effect is found not to be robust to the omission of districts where endogenous selection out of conflict exposure is more likely.

communities show promising results (Blattman et al., 2013, 2016), the findings of this study, instead, draw attention to the need for policies aimed at ensuring a swifter resolution to such long-running conflicts. Indeed, it appears likely that the irreversible health deficits of such events can be largely avoided if civilian populations can be returned to a state of stability within a relatively short period of time. In the absence of a timely resolution, however, communities affected by armed violence may be forced to carry the burden of war, long after the last shots have been fired.

References

Adelman, S. and Peterman, A. (2014). Resettlement and gender dimensions of land rights in post-conflict northern Uganda. *World Development*, 64:583–596.

Akresh, R., Bhalotra, S., Leone, M., and Osili, U. O. (2012a). War and stature: growing up during the Nigerian Civil War. *American Economic Review*, 102(3):273–77.

Akresh, R., Lucchetti, L., and Thirumurthy, H. (2012b). Wars and child health: Evidence from the Eritrean–Ethiopian conflict. *Journal of development economics*, 99(2):330–340.

Akresh, R., Verwimp, P., and Bundervoet, T. (2011). Civil war, crop failure, and child stunting in Rwanda. *Economic Development and Cultural Change*, 59(4):777–810.

Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford economic papers*, 58(3):450–474.

Allen, T. (2013). *Trial justice: The international criminal court and the Lord's Resistance Army.* Zed Books Ltd.

Baird, S., Friedman, J., and Schady, N. (2011). Aggregate income shocks and infant mortality in the developing world. *Review of Economics and statistics*, 93(3):847–856.

Blattman, C. (2009). From violence to voting: War and political participation in Uganda. *American political Science review*, 103(2):231–247.

Blattman, C. and Annan, J. (2010). The consequences of child soldiering. *The review of economics and statistics*, 92(4):882–898.

Blattman, C., Fiala, N., and Martinez, S. (2013). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *The Quarterly Journal of Economics*, 129(2):697–752.

Blattman, C., Green, E. P., Jamison, J., Lehmann, M. C., and Annan, J. (2016). Returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda. *American Economic Journal: Applied Economics*, 8(2):35–64.

Bozzoli, C. and Brück, T. (2010). Child morbidity and camp decongestion in post-war Uganda.

Bozzoli, C., Brueck, T., and Muhumuza, T. (2011). Does war influence individual expectations? *Economics letters*, 113(3):288–291.

Bundervoet, T., Verwimp, P., and Akresh, R. (2009). Health and civil war in rural Burundi. *Journal of human Resources*, 44(2):536–563.

Byrne, B. (1996). Towards a gendered understanding of conflict. *IDS bulletin*, 27(3):31–40.

Dagnelie, O., De Luca, G. D., and Maystadt, J.-F. (2018). Violence, selection and infant mortality in Congo. *Journal of health economics*, 59:153–177.

De Luca, G. and Verpoorten, M. (2015). Civil war and political participation: Evidence from Uganda. *Economic Development and Cultural Change*, 64(1):113–141.

Dercon, S. and Porter, C. (2014). Live aid revisited: long-term impacts of the 1984 Ethiopian famine on children. *Journal of the European Economic Association*, 12(4):927–948.

Dolan, C. (2009). *Social torture: the case of northern Uganda, 1986-2006*, volume 4. Berghahn Books.

Doom, R. and Vlassenroot, K. (1999). Kony's message: a new koine? the Lord's Resistance Army in northern Uganda. *African affairs*, 98(390):5–36.

Fiala, N. (2015). Economic consequences of forced displacement. *The Journal of Development Studies*, 51(10):1275–1293.

International Crisis Group (2004). Northern Uganda: Understanding and solving the conflict.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–26.

Martorell, R. and Habicht, J.-P. (1986). Growth in early childhood in developing countries.

Minoiu, C. and Shemyakina, O. N. (2014). Armed conflict, household victimization, and child health in Côte d'Ivoire. *Journal of Development Economics*, 108:237–255.

Raleigh, C. and Hegre, H. (2005). Introducing ACLED: An armed conflict location and event dataset (selection of ACLED maps), paper presented at the conference on disaggregating the study of civil war and transnational violence, University of California Institute of Global Conflict and Cooperation, San Diego, CA.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: an armed conflict location and event dataset: special data feature. *Journal of peace research*, 47(5):651–660.

Rohner, D., Thoenig, M., and Zilibotti, F. (2013). Seeds of distrust: Conflict in Uganda. *Journal of Economic Growth*, 18(3):217–252.

Rose, E. (1999). Consumption smoothing and excess female mortality in rural India. *Review of Economics and statistics*, 81(1):41–49.

Shemyakina, O. N. (2017). Political violence and child health: Results from Zimbabwe. *World Development*.

Shoup, J. A. (2011). *Ethnic groups of Africa and the Middle East: An encyclopedia*. ABC-CLIO.

Strauss, J. and Thomas, D. (2007). Health over the life course. *Handbook of development economics*, 4:3375–3474.

WHO (2005). Health and mortality survey among internally displaced persons in Gulu, Kitgum, and Pader districts, northern Uganda.

Appendices

Appendix 2.1

Figure 2.A1: Overview of Localities

Table 2.A1: Overview of Localities

	District	*Locality*		*District*	*Locality*		*District*	*Locality*
1	Apac	Chawente	78		Tingey County	155		Rugaaga
2		Kole North	79	Kasese	Bukonjo North	156		Ruhaama County
3		Kole South	80		Bukonjo South	157		Rwampara County
4		Kwania North	81		Busongora North	158		Sanga
5		Maruzi North	82		Lake Katwi	159	Moroto	Checkwii East
6		Oyam North	83		Muhokya	160		Iriri
7		Oyam South	84	Kibaale	Bugangaizi West	161		Karita
8	Arua	Ajia	85		Buyaga East	162		Lorengedwat
9		Aringa South	86		Buyaga West	163		Lotome
10		Ayivu County and Arua Mn	87		Buyanga West	164		Matheniko South/Moroto Mn
11		Beleafe	88	Kiboga	Kiboga Central	165		Nabilatuk
12		Kei	89		Kiboga East	166		Namalu
13		Koboko County	90	Kisoro	Kisoro District	167		Rupa
14		Madi-Okollo South	91	Kitgum	Aruu Central	168	Moyo	Adropi
15		Maracha County	92		Awere	169		Obongi South
16		Odupi	93		Chua Central	170		West Moyo County East
17		Rigbo	94		Chua West	171		West Moyo County West
18		Terego West	95		Lira - Palwo	172	Mpigi	Busiro North
19		Vurra South	96		Lokung	173		Busiro South and Entebbe Mn
20	Bundibugyo	Bwamba North	97		Padibe	174		Butambala East
21		Harugali	98		Paimol	175		Butambala West
22	Bushenyi	Buhweju County	99		Parabongo	176		Gomba East
23		Bunyaruguru Central	100		Patongo	177		Kabulasoke
24		Igara County	101	Kotido	Abim	178		Kyadondo North
25		Kajara County	102		Alerek	179		Maddu
26		Katerera	103		Jie North	180		Mawokota North
27		Ruhinda County	104		Kaabong	181		Mawokota South
28		Rushenyi County	105		Kapedo	182	Mubende	Bukuya
29		Sheema County	106		Kathile	183		Busujju East
30	Gulu	Amuru	107		Lolelia	184		Buwekula Central
31		Anaka (Payira)	108		Nakaperimoru	185		Buwekula North
32		Aswa/Omoro East	109		Napore (Karenga)	186		Kasambya
33		Aswa/Omoro South	110		Sidok (Kopoth)	187		Kassanda South
34		Aswa/Omoro West/Gulu Mn	111	Kumi	Bukedea North	188		Kitenga
35		Bobi	112		Bukedea South	189		Maanyi
36		Lamogi	113		Ngora North	190		Mityana County
37		Pabbo	114		Ngora South	191	Mukono	Bbaale South
38	Hoima	Bugahya East	115	Lira	Dokolo North	192		Buikwe North
39		Bugahya North	116		Erute South and Lira Mn	193		Buikwe South
40		Buhanguzi East	117		Kioga East	194		Kome Islands
41		Buhanguzi West	118		Moroto County East	195		Mukono North
42	Iganga	Bugweri County	119		Otuke East	196		Mukono South
43		Bukooli North	120	Luwero	Bamunanika South	197		Nakifuma Central
44		Bukooli West	121		Baruli Central	198		Ntenjeru County
45		Bunya North	122		Kakooge	199		Seeta
46		Bunya South	123		Kalungi	200	Nebbi	Jonam South
47		Busiki County	124		Katikamu North	201		Okoro County
48		Buyinja	125		Katikamu South	202		Padyere Central
49		Kigulu County	126		Kikyusa(Kamira)	203	Pallisa	Budaka County
50		Luuka North	127		Nakaseke Central	204		Butebo County
51		Luuka South	128		Nakaseke South	205		Kibuku County
52	Jinja	Jinja District	129	Masaka	Bukomansimbi County	206		Pallisa County Central
53	Kabale	Ndorwa County/Kabale Mn	130		Bukoto East and Masaka Mn	207	Rakai	Kabula North
54		Rubanda East	131		Bukoto West	208		Kabula South
55		Rubanda West	132		Kalungu County	209		Kooki North
56		Rukiga County	133		Lwemiyaga County	210		Kooki South
57	Kabarole	Bunyangabu County	134		Mawogola South	211		Kyebe
58		Burahya Central/Fort Portal	135		Mijwala	212		Kyotera County
59		Kahunge	136	Masindi	Budongo	213	Rukungiri	Bujumbura South
60		Kamwenge	137		Buruuli Central	214		Bwambara
61		Kitagwenda County	138		Kibanda South	215		Kinkiizi North
62		Kyaka North	139		Kiryandongo	216		Kinkiizi South
63		Mwenge North	140		Mutunda	217		Rubabo County
64		Mwenge South	141	Mbale	Bubulo County	218	Soroti	Kaberamaido County
65		Mwenge West	142		Budadiri County	219		Kalaki County
66		Nkoma	143		Bulambuli South	220		Kasilo County
67	Kalangala	Kyamuswa East	144		Bungokho County/Mbale Mn	221		Orungo
68		Mazinga	145		Manjiya County	222		Soroti County North
69	Kampala	Kampala District/Makindye	146	Mbarara	Buremba	223		Soroti County South/Soroti Mn
70	Kamuli	Budiope East	147		Ibanda Central	224		Usuk
71		Budiope West	148		Ibanda North	225		Usuk County West
72		Bugabula East	149		Isingiro North	226	Tororo	Bunyole County
73		Bugabula West	150		Isingiro South	227		Kisolo (West Budama) County
74		Bulamogi County	151		Kashari East/Mbarara Mn	228		Samya - Bugwe County
75		Buzaaya County	152		Kashari West	229		Tororo County and Tororo Mn
76	Kapchorwa	Binyini	153		Kashongi			
77		Kongasis County	154		Kazo Central			

Localities in Figure 2.A1 but not listed above, were not surveyed in the DHS waves used in this study

Appendix 2.2

Table 2.A2: A Description of the Conflict Localities

			Conflict Window			Exposed		Months Exposed	
District	Locality Name	Obs.	Start	End	Total Months	Yes	No	Mean	s.d.
Apac	Kole North	48	2002m7	2006m6	48	5	43	2.54	9.31
Apac	Kwania North	44	2005m1	2005m1	1	6	38	0.14	0.35
Apac	Maruzi North	33	1994m4	2006m4	145	33	0	24.00	16.08
Apac	Oyam North	28	1987m7	2005m3	213	28	0	25.75	17.83
Apac	Oyam South	104	2002m3	2005m3	37	10	94	2.09	6.99
Gulu	Amuru	1	2005m2	2006m8	19	1	0	17.00	.
Gulu	Anaka (Payira)	9	1998m11	2006m8	94	9	0	31.11	18.33
Gulu	Aswa/Omoro East	9	1998m6	2005m7	86	5	4	11.78	12.72
Gulu	Aswa/Omoro South	7	1998m11	2006m2	88	7	0	20.29	20.96
Gulu	Aswa/Omoro West/Gulu Mn	52	1988m1	2006m8	224	52	0	20.12	12.29
Gulu	Bobi	3	2000m5	2005m10	66	3	0	12.67	3.51
Gulu	Lamogi	5	1998m9	2006m7	95	5	0	29.00	16.00
Gulu	Pabbo	16	1987m4	2005m2	215	16	0	18.13	14.14
Kitgum	Aruu Central	11	2001m4	2006m7	64	11	0	24.91	13.67
Kitgum	Awere	5	1998m4	2004m12	81	5	0	15.00	15.28
Kitgum	Chua Central	9	1996m4	2006m8	125	9	0	22.00	11.94
Kitgum	Chua West	7	1990m5	2007m3	203	7	0	15.14	9.96
Kitgum	Lira - Palwo	5	1998m4	2005m12	93	5	0	24.40	16.53
Kitgum	Lokung	7	2004m4	2005m7	16	7	0	12.43	3.46
Kitgum	Padibe	3	1991m5	2005m5	169	3	0	25.00	11.53
Kitgum	Parabongo	7	1997m2	2006m8	115	7	0	30.43	23.35
Kitgum	Patongo	3	1988m6	2005m12	211	3	0	14.33	11.02
Lira	Dokolo North	42	2003m8	2004m2	7	2	40	0.33	1.51
Lira	Erute South and Lira Mn	118	1987m3	2005m4	218	118	0	23.34	14.87
Lira	Kioga East	42	2004m1	2004m2	2	5	37	0.21	0.61
Lira	Moroto County East	78	1988m6	2004m1	188	78	0	25.91	18.29
Lira	Otuke East	25	1998m5	2005m10	90	13	12	12.04	12.93
Soroti	Kaberamaido County	39	1987m8	2003m11	196	39	0	23.64	15.22
Soroti	Kalaki County	1	1987m8	2003m11	196	1	0	16.00	.
Soroti	Orungo	4	2003m3	2005m9	31	4	0	15.75	14.08
Soroti	Soroti County North	2	2003m6	2003m12	7	2	0	7.00	0.00
Soroti	Soroti County South/Soroti	64	1987m3	2005m9	223	64	0	24.05	14.20
Soroti	Usuk County West	4	1998m4	2004m11	80	4	0	26.50	14.84
Arua	Kei	41	1998m7	1998m7	1	9	32	0.22	0.42
Moyo	West Moyo County East	2	2001m5	2005m4	48	2	0	19.00	19.80
Moyo	West Moyo County West	18	1996m9	2001m3	55	10	8	17.67	19.83
Nebbi	Jonam South	19	1997m10	1997m10	1	8	11	0.42	0.51
Nebbi	Padyere Central	38	1996m9	2006m8	120	26	12	19.32	19.81
Kotido	Abim	3	2003m5	2003m5	1	3	0	1.00	0.00
Kotido	Kaabong	5	2004m6	2004m6	1	5	0	1.00	0.00
Moroto	Matheniko South/Moroto Mn	21	1998m12	2004m2	63	9	12	5.48	8.65
Jinja	Jinja District	154	1987m10	2003m6	189	154	0	24.65	15.60
Kamuli	Budiope West	65	2004m2	2004m2	1	5	60	0.08	0.27
Kapchorwa	Binyini	2	2003m3	2003m4	2	2	0	2.00	0.00
Hoima	Bugahya North	24	1995m10	1996m7	10	1	23	0.21	1.02
Masindi	Budongo	26	2001m3	2005m12	58	3	23	4.77	13.80
Masindi	Kiryandongo	4	2005m11	2005m11	1	4	0	1.00	0.00
Mubende	Buwekula North	46	2000m9	2000m9	1	28	18	0.61	0.49
Bushenyi	Katerera	34	2000m12	2000m12	1	17	17	0.50	0.51

Appendix 2.3

As discussed in Section 2.3.3, the majority of LRA conflict events occurred in districts dominated by Nilotic ethnic groups, with the former districts of Gulu, Kitgum, Lira, Apac and Soroti accounting for 1842 of the 1947 recorded battles or acts of one-sided violence listed in the ACLED dataset (94.6%). The ethnic Acholi (Gulu and Kitgum), Langi (Lira and Apac) and Iteso (Soroti), within these areas, share anthropological links to some of the tallest people in Africa, notably the Dinka of southern Sudan, and the Maasai (Samburu) and Turkana, of neighbouring Kenya (Shoup, 2011).

To establish whether significant height differences exist between the adult women of these Nilotic groups and the predominantly Bantu African groups further South, a sample of 1720 adult women is taken from the 1995 DHS survey. This sample only includes women who were at least 18 years old (to 49 years old) in 1986 and, therefore, should have achieved full adult height before the LRA conflict began.[36]

The first model in Table 2.A3 reports results from a simple bivariate regression of women's ethnicity on height, where the dependent variable is a binary indicator of belonging to one of the three ethnic groups (Acholi, Langi or Iteso). The second estimation reports similar results, disaggregating the groups to establish whether height disparities are apparent in all three ethnicities (relative to those not in these groups). Based on this simple analysis, the average height advantage of these women appears to be upward of 4.7 centimetres, with significant results for all groups individually. Furthermore, controlling for variations in the ages of

[36]Unfortunately, the 1995 survey is the only relevant DHS wave where ethnicity is recorded. However, the dominance of the Nilotic groups, in the districts listed, is supported by the self-reported ethnic affiliation of the sample of women from this survey. In Gulu, 94.87% of the women identified as being a member of the Acholi group, in Lira and Apac, 89.5% and 93.6% identified as Langi, while the Iteso ethnic group accounted for 73.3% in Soroti district.

women, in models 3 and 4, does nothing to change this interpretation. The results in models 5 to 8 repeat the estimations for the sub-sample of women coming only from conflict localities. The similarities in the magnitude of the coefficients strongly suggests that the disparity in height between these ethnic groups and others is also present within conflict localities.

Table 2.A3: Ethnicity as a Determinant of Height - Women 18+ in 1986

Dep. Variable: height (cms)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Acholi/Langi/Iteso	4.757***		4.817***		4.282***		4.538***	
	(0.415)		(0.418)		(0.617)		(0.653)	
Acholi		4.012***		4.173***		3.460***		3.895***
		(0.539)		(0.575)		(0.725)		(0.888)
Langi		5.478***		5.466***		4.722***		5.005***
		(0.518)		(0.545)		(0.696)		(0.716)
Iteso		4.357***		4.450***		3.894***		4.051***
		(0.554)		(0.563)		(0.812)		(0.845)
Observations	1720	1720	1720	1720	295	295	295	295
R^2	0.070	0.071	0.083	0.084	0.101	0.104	0.173	0.176
Age Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes

Standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Appendix 2.4

Table 2.A4: Table 2.2 Results - Control Coefficients Reported

Dep. Variable: height-for-age z-score	(1)	(2)	(3)	(4)	(5)	(6)
Conflict Locality * Exposed	-0.239	-0.369**	-0.388**			
	(0.187)	(0.186)	(0.195)			
Conflict Locality * Ex Duration				0.001	-0.011**	-0.011**
				(0.003)	(0.005)	(0.005)
Female child	0.146***	0.149***	0.152***	0.147***	0.148***	0.151***
	(0.031)	(0.032)	(0.031)	(0.031)	(0.032)	(0.031)
Birth order 4-6			-0.101**			-0.102**
			(0.046)			(0.046)
Birth order 7+			-0.296***			-0.296***
			(0.065)			(0.066)
Mother's height (cms)			0.049***			0.049***
			(0.003)			(0.003)
Mother married/cohabiting			0.036			0.037
			(0.067)			(0.067)
Mother's age (years)			0.017***			0.017***
			(0.004)			(0.004)
Mother primary educated			0.053			0.053
			(0.046)			(0.046)
Mother secondary educated			0.251***			0.252***
			(0.060)			(0.060)
Mother is daughter of head			-0.041			-0.040
			(0.101)			(0.101)
Mother is not head/daughter			-0.240**			-0.238**
			(0.104)			(0.106)
Mother's time in location			-0.001			-0.001
			(0.002)			(0.002)
Urban DHS cluster			0.424***			0.421***
			(0.068)			(0.068)
Head's age (years)			0.005**			0.005**
			(0.002)			(0.002)
Head is male			0.057			0.055
			(0.051)			(0.051)
Observations	9202	9202	9202	9202	9202	9202
R^2	0.079	0.147	0.205	0.079	0.147	0.205
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	No	Yes	Yes	No	Yes	Yes

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Table 2.A4 Continued

Dep. Variable: height-for-age z-score	(7)	(8)	(9)	(10)	(11)	(12)
Conflict Locality * Ex 1-6m	-0.191	-0.200	-0.234			
	(0.213)	(0.215)	(0.218)			
Conflict Locality * Ex > 6m	-0.315*	-0.659***	-0.651***			
	(0.165)	(0.194)	(0.222)			
Conflict Locality * Ex 1-12m				-0.231	-0.298	-0.317
				(0.197)	(0.201)	(0.204)
Conflict Locality * Ex > 12m				-0.258	-0.566***	-0.585**
				(0.171)	(0.200)	(0.227)
Female child	0.146***	0.149***	0.153***	0.146***	0.148***	0.152***
	(0.031)	(0.032)	(0.031)	(0.031)	(0.032)	(0.031)
Birth order 4-6			-0.101**			-0.102**
			(0.046)			(0.046)
Birth order 7+			-0.295***			-0.295***
			(0.066)			(0.066)
Mother's height (cms)			0.049***			0.049***
			(0.003)			(0.003)
Mother married/cohabiting			0.037			0.036
			(0.067)			(0.067)
Mother's age (years)			0.017***			0.017***
			(0.004)			(0.004)
Mother primary educated			0.053			0.054
			(0.046)			(0.046)
Mother secondary educated			0.250***			0.253***
			(0.060)			(0.060)
Mother is daughter of head			-0.040			-0.040
			(0.101)			(0.101)
Mother is not head/daughter			-0.239**			-0.241**
			(0.105)			(0.105)
Mother's time in location			-0.001			-0.001
			(0.002)			(0.002)
Urban DHS cluster			0.423***			0.422***
			(0.069)			(0.068)
Head's age (years)			0.005**			0.005**
			(0.002)			(0.002)
Head is male			0.057			0.056
			(0.051)			(0.051)
Observations	9202	9202	9202	9202	9202	9202
R^2	0.079	0.148	0.206	0.079	0.147	0.205
Locality Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Birth Cohort Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Trend	No	Yes	Yes	No	Yes	Yes

Locality clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Table 2.A4 reports identical results to those found in Table 2.2, with the inclusion of the coefficients on the control variables (listed beneath Table 2.2). Focusing on controls at the level of the child, significant differences in height-for-age z-scores are notable between children's gender groups, although, this is likely to only represent the extent to which Ugandan children (of either gender) fit the growth patterns of the reference population. The variables representing birth order clearly influences the child's height, with those in the first three positions having significant height advantages over sibling born later. The characteristics of the child's mother also clearly exert influence on the results, while both the age of the head and the mother appear positively correlated with childhood height.[37] Interestingly, the relationship between these two household members shows evidence of a relative height deficit for children who are neither, the child or grandchild of the head (However, this does not necessarily imply a causal effect). Unsurprisingly, children in urban locations (as defined by the original DHS clusters) are taller than their rural counterparts, on average, yet there is no evidence of any overall effect on children's height associated with the gender of the household head, the marital status of the child's mother or the time she has spent in the current DHS cluster location.

[37]The head of the household and the mother of the child are different household members for 86.4% of the observations within the sample used in Table 2.2.

Appendix 2.5

Table 2.A5 summarises the migration data taken from the Ugandan National Household Survey 2005/06.[38] Of those 669 households who migrated from one of the affected districts (listed in Table 2.A5), only 147 left the original district (21.4%). Furthermore, within the group of 283 households who relocated specifically to avoid insecurity, only 30 (10.6%) moved to a new district.

Table 2.A5: Uganda National Household Survey 2005/06 Migration History

Migrated Since 2001:		*For Any Reason*				*To Avoid Insecurity*			
District From	Sample Total	Total	Same District	Same Region	New Region	Total	Same District	Same Region	New Region
Mubende	289	44	39	1	4	0	0	0	0
Jinja	164	46	20	10	16	0	0	0	0
Kamuli	320	37	23	12	2	0	0	0	0
Kapchorwa	39	4	3	1	0	0	0	0	0
Soroti	141	39	27	4	8	15	10	1	4
Kotido	68	3	1	0	2	1	0	0	1
Moroto	57	5	2	0	3	3	1	0	2
Apac	316	61	38	22	1	11	5	6	0
Gulu	213	65	58	3	4	46	43	2	1
Kitgum	166	129	121	3	5	118	115	1	2
Lira	240	135	122	11	2	82	72	10	0
Arua	367	34	31	1	2	0	0	0	0
Moyo	57	16	15	1	0	7	7	0	0
Nebbi	116	17	12	0	5	0	0	0	0
Hoima	86	11	5	4	2	0	0	0	0
Masindi	110	23	9	0	14	0	0	0	0
Total	2749	669	526	73	70	283	253	20	10

[38] Uganda Bureau of Statistics (2008), Uganda National Household Survey 2005/2006, Version 1.0 of the public use dataset, provided by the National Data Archive. www.ubos.org

Appendix 2.6

Table 2.A6: Mother's Time in Location on Height-for-Age, by District

District	Obs.	dy/dx	se	t stat	p-value
Apac	337	−0.001	0.014	−0.094	0.926
Arua	536	−0.004	0.004	−0.896	0.373
Bushenyi	378	−0.002	0.005	−0.442	0.660
Gulu	102	−0.215	0.020	−10.578	0.000***
Hoima	141	−0.033	0.017	−1.999	0.049**
Kamuli	651	0.000	0.010	0.010	0.992
Kapchorwa	85	−0.001	0.003	−0.248	0.805
Kitgum	63	−0.005	0.008	−0.664	0.509
Kotido	111	−0.031	0.011	−2.816	0.006***
Lira	305	0.000	0.003	−0.048	0.962
Masindi	117	−0.006	0.005	−1.054	0.295
Moroto	115	−0.007	0.011	−0.623	0.535
Moyo	119	0.000	0.005	0.082	0.935
Mubende	398	0.001	0.008	0.191	0.849
Nebbi	139	0.013	0.007	1.963	0.053*
Soroti	224	0.004	0.002	1.618	0.110
Total	3975				

Locality clustered standard errors reported * p<0.1, ** p<0.05, *** p<0.01

The results shown in Table 2.A6 report the marginal effects and standard errors of an additional year in the current location (based on the mother's time in location) on height-for-age z-score, by district. These marginal effects are based on an estimation of the model shown below.

$$\begin{aligned} haz_{idt} &= \delta_t + District_d + \theta_1(Time\ in\ Location_i) \\ &\quad + \theta_1(District_d * Time\ in\ Location_i) + \theta X_{idt} + \upsilon_{idt} \end{aligned} \tag{2.a1}$$

In model (2.a1) haz_{idt} measure the height-for-age z-score of a child who resides in district d. The variable $District$ represents a district fixed effect, while $Time\ in\ Location$ measures the number of years the child's mother has lived in the current DHS cluster. The term X represents the same control variables used in generating the results in Table 2.2. As the majority of relocation appears to have occurred within districts (see Appendix 2.5), only observations taken from children in the conflict-affected districts, who lived in localities which had not yet been affected by the insurgency, are used in the estimations.

The results indicate the presence of a negative correlation between time in location and height in observations from pre-conflict localities in Gulu, Hoima and Kotido districts. This negative correlation could represent a selection effect, driven by a relocation away from more severely affected localities, amongst relatively healthier children (see discussion in Section 2.6.1).

Chapter 3

Consumption Shocks and Coping Strategies in Two South African Townships

3.1 Introduction

The vast majority of poor and vulnerable households throughout the world face an economic environment which is both volatile and unforgiving. When households lack the capacity to effectively insure against risk, shocks to income patterns or consumption needs may leave only very limited options available to those affected. In many cases, households may be forced to take actions such as selling land or productive assets, removing children from education or risking irreversible malnutrition by cutting back on food consumption. The formal instruments used to smooth income over time or different states of the world, such as credit, savings or insurance, are often unavailable or unsuited to the poorest households. Instead, these communities commonly rely on informal mechanisms to avoid the need for measures such as those outlined above. In particular, households often fall back on social and family networks to provide them with monetary and non-monetary loans and gifts should the worst befall them. It is also common that the accumulation of assets in favourable times acts as both a store of wealth and insurance for households wishing to protect themselves from income volatility. The consensus is that these informal measures are, at best, only partially able to insulate households from exposure to risk.[1] As a result, many low-income households face the possibility that events as commonplace as the sickness of a family member could signal the first step towards chronic poverty or destitution.

With such high stakes, it is perhaps unsurprising that many households are willing to forego potentially more productive activities, in favour of more predictable returns. A preference for diversification (rather than specialisation) of income sources or holding unproductive liquid assets would both be common examples of strategies used by households to lower their real exposure to risk. These examples of the trade-off between security and productivity illustrate the pervasive link between risk-management and poverty. Where public safety nets are weak or missing, formal instruments are unavailable (or ill-suited to household's needs) and informal systems are unable to effectively shield consumption from shocks, households may be forced to adopt behaviours which could potentially trap them in a state of persistent poverty.

This study utilises high-frequency panel data, collected as part of the Southern Africa Labour and Development Research Unit's (SALDRU) Financial Diaries Project, with the aim of quantifying the effects of shocks on household welfare.[2] The following analysis seeks to evaluate how the impact of these shocks varies depending on the type of coping strategy the household employs.[3] The following paper focuses specifically on densely populated urban settlements, where the lives

[1] For example, the full-insurance model is rejected in Townsend (1994); Udry (1994) and Jalan and Ravallion (1999), amongst others, while complete consumption smoothing through assets or savings is also commonly rejected in low-income communities (Paxson, 1992; Alderman, 1996; Kazianga and Udry, 2006).

[2] For the purposes of this study, a shock is defined as an event which has a negative influence on a household's cash flow or asset position. The full list of shocks is provided Appendix 3.1.

[3] Following Yilma et al. (2014), coping strategies are defined as actions undertaken by a household to accommodate the effect of a shock. This definition, therefore, excludes any risk-management behaviour intended to reduce exposure to risk *ex-ante*, such as income smoothing through diversification or seasonal migration. See Alderman and Paxson (1994) for a discussion of income smoothing as a risk-management strategy.

of the inhabitants are arguably more intertwined and the sources of risk more idiosyncratic. The analysis finds that both assistance/insurance and savings-based strategies effectively smooth consumption in the case of moderate or severe shocks. In contrast, where no specific coping strategy is used, these shocks result in a 47% reduction in food consumption (on average) in the week the shock occurs. Shortfalls in consumption are also present in both the week following the event and the week prior to the shock taking place, a result which suggests that some shocks can be pre-empted (at least within this time-horizon). The choice of coping strategy is found to follow a general pattern, whereby the least costly shocks result in no coping strategy, moderate shocks elicit a savings-based response and severe shocks lead to either the use of an assistance/insurance-based or savings-based response.

The remainder of the paper will be presented as follows. Section 3.2 will discuss the literature which looks specifically at household choices of coping strategy in response to a variety of shocks. Section 3.3 will present an overview of the data used in this analysis, including a description of the events experienced by the household within the sample and the coping responses employed. Section 3.4 provides a description of the empirical identification strategy, while Section 3.5 reports the main results of this study. A discussion of the robustness of these results can be found in Section 3.6, with a summary of the findings and concluding remarks presented in Section 3.7.

3.2 Previous Literature

The need for more effective policy to protect low-income households from risk should be self-evident, yet our understanding of exposure to different types of risk, and the strategies employed to mitigate the effects of shocks, is still relatively incomplete (Toye, 2007). Much of the prominent literature relating to shocks, risk and insurance, seeks either to establish the validity of the theory that households seek to smooth consumption over time, through adjusting stocks of savings or liquid assets (Paxson, 1992; Rosenzweig and Wolpin, 1993), or smooth consumption across different states of the world, through risk-sharing arrangements and informal insurance (Mace, 1991; Townsend, 1994; Udry, 1994; Jalan and Ravallion, 1999).[4]

The predominant theory relating to the use of savings to smooth consumption relies on the hypothesis that (absent borrowing or saving constraints) individuals seek to equate the marginal utility of consumption over any two consecutive periods of time (dependent on the real interest rate, and the individual/household's time discount rate). A number of influential studies have attempted to test this hypothesis, by disaggregating income into permanent and transitory components. Such tests have found that consumption smoothing, although substantial, falls short of complete self-insurance (Paxson, 1992; Deaton, 1992; Udry, 1995).

A separate strand of the literature relates to the use of risk-sharing arrangements to smooth consumption over different states of the world. The most commonly applied theory describes a model of *perfect* risk-sharing, within a network of households, where the network achieves a Pareto optimal allocation of consumption by maximising a social welfare function. Of the many studies which have followed this type of approach, the majority have found some evidence of a positive relationship between individual and network consumption, but have been unable to reject a significant correlation between household income and household consumption (Morduch, 1991; Townsend, 1994; Jalan and Ravallion, 1999; Kazianga and Udry, 2006). The consensus within the literature is, therefore, that risk-sharing does take place, but is unable to fully insulate consumption against idiosyncratic shocks.

[4] In considering these two approaches, it is important to recognise that inference drawn when testing one method of consumption smoothing often relies on the absence of the other.

In contrast to the studies mentioned above, this paper considers a wide spectrum of shocks and a wide variety of responses. In this respect, the analysis presented here is more closely related to a far smaller body of literature. For example, Heltberg and Lund (2009) measure the impact of a variety of shocks in their survey of rural households in Pakistan. The authors deliberately focus on the worst potential consequences of these shocks, such as school drop-out (of children) or bonded-labour. They consider a number of potential coping mechanisms, including selling assets, using savings or obtaining credit, and how well these strategies protect households from more serious outcomes. They find that most households rely heavily on informal credit or self-insurance, rather than weak public or private safety net programmes, suggesting potential benefits from more effective policy aimed at mitigating risk. Similarly, in a multi-shock study of households in Laos, Wagstaff and Lindelow (2010) also report that informal credit, assistance and savings (in particular) are commonly employed method of response to incidences of crop loss, pest infestation, livestock disease or prolonged periods of illness.

In Sub-Saharan Africa, Yilma et al. (2014) use a retrospective panel survey in rural Ethiopia to study the link between different types of shocks and choices of coping strategy. They find that almost all shocks resulted in some reduction in savings, but covariate shocks such as drought and crop-damage also triggered reductions in food consumption. In addition, they concluded that more idiosyncratic shocks are likely to result in borrowing on the part of households. Further insight also comes from the work of Heltberg et al. (2014), who summarise household survey data from sixteen developing countries (including five from Sub-Sahara Africa). The authors find that dissaving is a common strategy in the majority of countries considered, with credit, assistance and reductions in food consumption also likely in most cases. In contrast, reductions in health and education spending were relatively rare, overall.

This study contributes to this literature, while focusing attention on the short-term effects of shocks, on a week-by-week basis. The results allow for a comparison of the relative effectiveness of coping strategies, in terms of short-run recovery, as well as immediate impacts. The use of high-frequency, panel data to mitigate bias, due to unobserved household heterogeneity and the potential endogeneity of both shocks and coping strategies, should yield an accurate evaluation of the relative effectiveness of these strategies in poor, urban (and peri-urban) communities.

3.3 Data

3.3.1 Overview and Sampled Locations

Although the majority of poor South Africans still reside in rural areas, the share of these households is falling relative to that of their urban counterparts, indicating that poverty in South Africa is increasingly becoming an urban problem (Bhorat and Kanbur, 2006). In light of this shift towards urban poverty, this study moves the focus away from rural communities, and instead focuses on two of South Africa's townships, where the poorest households are increasingly likely to be found. This analysis is based on the data collected by the Southern Africa Labour and Development Unit (SALDRU), as part of the Financial Diaries Project conducted in 2003/04.[5] The Financial Diaries survey was designed to improve understanding of the financial needs of the poor and inform the government and financial sector on how best to improve the quality (and availability) of the financial services offered to low-income communities. This data provides 4161 weekly observations, relating to 85 households in two study locations.

[5]The initial survey also covered the rural area of Lugangengi, in South Africa's Eastern Cape. However, only data from the urban and peri-urban Langa and Diepsloot townships is used in the following analysis. Observations recorded prior to January 2004 are also not used, due to missing expenditure information for some sampled households (although income measure for this period are employed in measuring shock severity – see Section 3.3.3). Subsequently, the period under consideration covers January the 1st 2004 to December the 16th 2004.

The first of these locations is the peri-urban township of Diepsloot, which lies approximately 20kms North of central Johannesburg and was established in 1995 as a relocation camp for families from Alexandra and other informal settlements. In spite of only gaining formal recognition as a permanent settlement in 1999, the densely populated township is estimated to be home to over 85,000 inhabitants (2001 census data, Statistics South Africa). The township suffers from inadequate infrastructure and housing to provide for its rapidly growing population, with the majority of the households living in single-room shacks (Collins et al., 2010). Due to its proximity and direct transport links to central Johannesburg (and other economic centres, such as Sandton and Randburg), further pressures on infrastructure and housing come as a result of economic migrants from other parts of the country seeking work, along with foreign nationals from many other parts of Africa (most notably Zimbabwe and Mozambique). High levels of crime, unemployment and the presence of large numbers of foreigners are regularly cited as the underlying causes of the many incidents of xenophobic violence in the township's recent past (Pillay et al., 2008).

The second location is the urban township of Langa, which is located approximately 11 kms from the centre of Cape Town and was established as a residence for black African migrant workers in 1927. In contrast to Diepsloot, Langa is one of the oldest and most established townships in South Africa, with many of the current residents directly descended from the original inhabitants. The population prior to the survey was estimated at approximately 50,000 and, although the township still contains the original permanent housing, many of these residents live in informal dwellings or constructed shacks. Access to infrastructure is poor and unevenly distributed, with only 35% of households having access to running water within their homes (2001 census data, Statistics South Africa). Relative to the inhabitants of Diepsloot, those in Langa are more culturally homogeneous, generally identifying themselves as South African and members of the Xhosa tribe.

Although most of the residents have long-standing connections to the township, the sample also includes a small number of households who migrated from rural areas, seeking work in the urban centres around Cape Town (Collins et al., 2010).

Ten representative districts were randomly selected within the two locations, and from these selected districts, senior members of the community were asked to rank all households in the district, according to their perceived wealth. These rankings were then used to assign each household in the district a participatory wealth ranking (PWR) of poor, upper-poor or non-poor. The final sample selected from each district was stratified to contain an equal number of households from each of these three PWR rankings.[6] For the purposes of this study, the original categories of upper-poor and non-poor are combined into one classification.

3.3.2 Household Characteristics

A comparison of the general characteristics of the households, across the two townships and wealth categories, can be found in Table 3.1. As expected, given the sampling methodology, the proportions of poor and upper-poor households are broadly equivalent in both locations. The household income and consumption categories indicate mean values, per adult equivalent, across all weeks of the survey.[7] Average, weekly income and consumption appear slightly higher in the Diepsloot sample, although not significantly so (at the given levels). Similarly, the size, employment and gender composition of households appears statistically similar across the two samples, although the mean household size is larger in Langa (again, not significantly so).[8] There is some evidence of both a higher dependency

[6]A description of the participatory wealth ranking methodology, along with the potential benefits of this approach to measuring wealth, can be found in Simanowitz and Nkuna (1998).

[7]Where possible, the construction of these variables follows closely the guidelines described in Deaton and Zaidi (2002) and McKay and Round (1990).

[8]Within the survey the a household member was defined as someone who lived under the same roof for at least 15 days in the past year; and when together, shared food, and contributed to the same resource pool (Financial diaries survey 2003-2004: First initial questionnaire).

ratio and also a higher likelihood of a household containing a chronically ill or disabled member in Langa. Arguably the clearest disparity in household characteristics, however, comes from the higher probability of those in Diepsloot being resettled to their current location from elsewhere.[9] Heads of households from the Langa sample are more likely to be both female and older than those in Diepsloot, a result which may contribute to the higher relative incidence of disability or chronic illness in this location. Household heads (or their spouses) are also more likely to belong to an informal funeral society in Langa, relative to those in the other township.[10]

Poor households earn total incomes which are, on average, 54% less than upper-poor households, while consuming 45% less (in monetary terms) than their relatively wealthier counterparts.[11] The expenditure value of food consumption is only 33% lower, however, suggesting that the additional incomes of less poor households are allocated relatively more towards non-food spending. The average percentage of household members in regular work is also far lower in the poor wealth category. As this measure is based on household members receiving a regular payment from an employer, this disparity may be capturing both differences in employment and differences in the formality of employment between the two groups. The household heads in the poor category are more likely to be female and unemployed than those in the relatively wealthier group, while the probability of belonging to an informal funeral society is also lower for these individuals and their spouses. As a measure of a household's ability and willingness to use formal financial instruments, the

[9] As previously mentioned, Diepsloot functions primarily as a centre for relocation of households from other overcrowded settlements. This is clearly represented in the data, with around 70% of the households in the sample having been relocated to their current location, relative to less than 15% in Langa.

[10] One of the most costly, life-cycle events faced by South African households is the funeral of one of its members. In South African culture, the various stages of the ceremony take place over many weeks or months, with the household of the deceased often expected to host relatives during the ceremonies, in addition to paying all costs involved (Collins and Leibbrandt, 2007).

[11] Both wealth categories indicate mean total incomes which exceed the value of mean consumption. However, it was not possible to include an (imputed) value of rental expenditure within the consumption measure, due to a lack of information on rental costs within the sample.

number of formal bank accounts held (per adult household member, 16 or over) is also reported in Table 3.1. The table indicates that being an adult member in the poorer group of households corresponds to having one less formal bank account, on average, relative to those in the upper-poor wealth category.

Table 3.1: Descriptive Statistics

Variable	(1) All mean	(2) Diepsloot mean	(3) Langa mean	F-test (2)-(3) p-value	(4) Poor mean	(5) Upper poor mean	F-test (4)-(5) p-value
Wealth							
% Poor	36.50	36.40	36.60	0.983			
% Upper-poor	63.50	63.60	63.40	0.983			
Household income and expenditure (Rand per week)							
Regular income	174.31	178.85	169.44	0.783	90.99	222.15	0.000***
Total income	179.53	189.17	169.18	0.532	102.84	223.55	0.000***
Food consumption	56.94	59.79	53.88	0.354	43.47	64.67	0.000***
Total consumption	144.62	154.82	133.67	0.289	94.68	173.29	0.000***
Other household characteristics							
Household size	3.01	2.74	3.31	0.193	2.99	3.03	0.938
% Male	41.60	43.00	40.10	0.676	36.70	44.40	0.276
% Regular job	43.00	43.60	42.40	0.864	27.10	52.20	0.000***
Dependency ratio†	0.23	0.19	0.28	0.093*	0.25	0.22	0.570
Disabled/chronic ill	0.35	0.25	0.46	0.041**	0.39	0.33	0.626
Resettled household	0.44	0.71	0.15	0.000***	0.36	0.48	0.257
Household head characteristics							
Head female	0.44	0.30	0.59	0.007***	0.65	0.32	0.003***
Head age	41.98	38.07	46.17	0.002***	43.68	41.00	0.344
Head employed	0.73	0.77	0.68	0.359	0.45	0.89	0.000***
No junior school	0.54	0.48	0.61	0.225	0.68	0.46	0.053*
Junior school	0.46	0.52	0.39	0.225	0.32	0.54	0.053*
Financial instruments							
In burial society	0.67	0.57	0.78	0.037**	0.52	0.76	0.027**
Bank accounts	1.04	1.16	0.90	0.209	0.42	1.39	0.000***
Households	85	44	41		31	54	
Observations	4161	2154	2007		1510	2651	

* p<0.1, ** p<0.05, *** p<0.01
All weekly income and expenditure figures measured in South African Rand, per adult equivalent
†Dependency ratio: (members $<$ 16y + members $\geq$ 65y) / household size

3.3.3 Shocks

In addition to providing regular information on financial activity, the households were also required to record any unusual events which influenced cash flow, specifying which event had taken place from a list of sixteen possible shocks. For each event, the respondents were also asked to estimate the cost, in terms of assets lost and additional expenses incurred. These self-reported costs are used to classify the shocks as mild, moderate or severe, by calculating a measure of severity which expresses total costs as a percentage of the household's regular monthly income prior to January 2004.[12]

$$\text{Shock Severity} = \frac{\text{Value of assets lost} + \text{Additional expenses incurred}}{\text{Household monthly income}}$$

Where the cost of the shock was reported as less than 10% of regular income (severity < 0.1) the event was classified as *mild*, where this cost was 10% to 50% of regular income ($0.1 \leq$ severity ≤ 0.5) the shock was recorded as *moderate*, and where costs accounted for more than 50% of monthly income (severity > 0.5), the shock was considered *severe*. In this way, the magnitude of the event experienced was considered in relation to the financial position of the household and its ability to absorb the additional costs imposed.

During the 50 weeks, between January the 1$^{\text{st}}$ and December the 16$^{\text{th}}$ of 2004, the 85 households within the sample experienced 107 shocks in total. 38 of these were mild, 41 were moderate and 28 of the shocks were severe. As previously noted,

[12] The sample used in the main analysis does not include observations before January the 1$^{\text{st}}$ 2004, due to missing expenditure information for some households. However, complete data on incomes for an additional two months before this date are used in the calculation of shock severity.

the original survey considered 16 types of shocks. However, the overwhelming dominance of health shocks within the sample, and mortality shocks in particular (see Appendix 3.1), suggested the distinction be made between health-related shocks occurring to members of the household, and health shocks occurring to family members living elsewhere. The category of household health shocks in Table 3.2, therefore, refers to events occurring within the household unit, such as the death of a member of the household or unveiling of a household member's tombstone.[13] The second category of shock in Table 3.2 refers to health events which occurred to family members outside of the household, all of which involve the death of a family member living elsewhere.

Table 3.2: Frequencies of Shocks, by Magnitude

	Total	Mild	Moderate	Severe
Household health shocks	32	14	13	5
Indirect health shocks	41	14	14	13
Economic shocks	7	3	4	0
Crime shocks	13	4	5	4
Family shocks	14	3	5	6
Total	107	38	41	28

Mild = Cost $<$ 0.1*Income, Moderate = 0.1*Income $\leq$ Cost $\leq$ 0.5*Income, Severe = 0.5*Income $<$ Cost

In contrast to the two health shock categories, the other classifications contain relatively few occurrences. Economic shocks, including loss of employment and failure of a household business, account for only 7 of the 107 shocks, none of which were severe. Crime shocks, which combine theft of property and violent crime,

[13]The unveiling of the tombstone is a ceremony of great importance within South African culture. It usually takes place a few weeks after the funeral itself and can also entail high costs for the family of the deceased. Aside from the immediate cost of the ceremony and stone itself, proceedings will usually be followed by the provision of food for friends, family and neighbours (Collins et al., 2010).

occurred 13 times overall, with severe consequences on 4 occasions. The category of family shocks, such as divorce or loss of household property, also included weddings, which accounted for 5 of these shocks within the sample.[14]

Table 3.3: Shocks Experienced, by Township and Wealth Ranking

Variable	(1) All mean	(2) Diepsloot mean	(3) Langa mean	F-test (2)-(3) p-value	(4) Poor mean	(5) Upper poor mean	F-test (4)-(5) p-value
Number of mild shocks	0.45	0.52	0.37	0.370	0.32	0.52	0.222
Number of moderate shocks	0.48	0.61	0.34	0.112	0.45	0.50	0.794
Number of severe shocks	0.33	0.30	0.37	0.611	0.42	0.28	0.332
Total number of shocks	1.26	1.43	1.07	0.181	1.19	1.30	0.704
Reported cost of shocks	837.68	856.16	813.04	0.888	713.37	901.51	0.490

* p<0.1, ** p<0.05, *** p<0.01
Mild = Cost $<$ 0.1*Income, Moderate = 0.1*Income $\leq$ Cost $\leq$ 0.5*Income,
Severe = 0.5*Income $<$ Cost

Table 3.3 shows there is no significant difference between the number of shocks experienced by the households in the two different locations. This is the case for the overall number of shocks and for the three magnitudes of shocks considered. Furthermore, there also appears to be no significant difference in the average cost of these events between the two locations (the figure used to categorise the shocks by magnitude). Perhaps surprisingly, there also appears to be little difference in the number of shocks experienced between the two wealth categories. Although, the abundance of mortality shocks outside of the household will place less weight on the relative wealth of the households themselves, as a determinant of exposure to risk.

[14] It is important to note that weddings, initiations and births (listed as household health shocks) are unlikely to constitute unexpected shocks to the household. This point is taken up further in Section 3.6.1.

3.3.4 Coping Strategies

In the event of a shock, enumerators asked respondents to list all the ways in which the total amount of money was found to deal with the event. The questionnaire allowed for up to three responses. However, as only a single strategy was used in over 85% of the shocks recorded, only one coping strategy is associated with each shock in the following analysis.[15] Table 3.4 reports the frequency with which the different strategies were used, in response to the three magnitudes of shock severity. The specific coping strategies (within the three grouped categories shown below) are listed in more detail in Appendix 3.2.[16]

Table 3.4: Frequencies of Coping Strategies, by Shock Magnitude

Shock magnitude:	Any	Mild	Moderate	Severe
Savings-based strategy	28	5	15	8
Assistance/Insurance-based strategy	35	8	11	16
No coping strategy	44	25	15	4
Total	107	38	41	28

Mild = Cost $<$ 0.1*Income, Moderate = 0.1*Income $\leq$ Cost $\leq$ 0.5*Income, Severe = 0.5*Income $<$ Cost

Savings-based strategies contain only the incidences where the household used savings to respond to the shock. This was the most common single response, used in just over one-quarter of the shocks in the survey (see Appendix 3.2). The second

[15] Where more than one strategy was employed, the strategy used was deemed to be that which contributed the greatest proportion of the funds raised to offset the negative effects of the shock.

[16] The table in Appendix 3.2 also indicates that households were provided with a choice of selecting an asset-based strategy (selling or disposing of assets), as the coping strategy which most accurately reflected their response. However, as no household reported using this type of response for any of the 107 shocks, the categories of responses are restricted to the three groups listed in Table 3.4.

group of responses refers to the (external) strategies of assistance and insurance, which include the receipt of both borrowing and gifts, and represent both money and non-monetary goods.[17] This category also contains insurance payments, which were mostly received as the result of the death of a family member. The third group contains all shock responses which imply no coping strategy, in the context of this analysis. As the extent to which households are capable of insulating consumption from additional costs or income shortfalls will be used as the measure of how effectively they respond to shocks, this category contains instances where households report a consumption reduction as the coping strategy used (going without meat or school uniforms, for example). This category of response also contains instances where the household claimed no specific coping strategy was required, or income was sufficient to cover the additional expenses.

According to Table 3.4, assistance/insurance strategies appear to be a more common response than the use of savings, overall (although the use of neither strategy is more common still). As the cost of mild shocks constitutes a relatively small proportion of regular income, it is unsurprising that the majority of these shocks elicit no coping strategy (other than the use of income or a reduction in consumption). Moderate shocks appear to lead to a mixed pattern of responses, although using savings accounts for a higher proportion than assistance/insurance responses. The most severe shocks, however, are far more likely to result in the household receiving assistance or an insurance pay-out, with households relying on this type of coping strategy in 16 of the 28 severe shocks occurring within the survey period.

[17] Within this category, gifts and loans are not treated separately, as gifts received may, in fact, represent reciprocal quasi-credit arrangements (Fafchamps, 1999).

3.4 Empirical Strategy

Determining the effect of shocks on household consumption requires addressing certain problems of endogeneity. The first of which refers to heterogeneity in the probability of experiencing a given shock. There is no reason to expect that shocks would be uniformly distributed across households, and any analysis which proceeds as though this were the case would be likely to generate biased estimates of their impact. In a similar vein, where the household's choice of response is expected to influence the subsequent effect of a shock, treating the choice of coping strategy as though it were exogenously determined would also be inappropriate. Subsequently, this study takes the more general position that shocks do not occur entirely at random but are instead caused by the underlying characteristics of the household and economic environment, which will often be correlated with outcomes (and may or may not be observed within the data). In addition, the identification strategy also assumes that the choice of response is not determined fully by the nature of the shock, but is instead a function of numerous underlying factors (available savings, social capital and engagement with formal financial instruments, for example), many of which will be strongly correlated with the household's consumption decisions.

As suggested by the previous discussion, identifying the effect of shocks on welfare will proceed in three broad stages. The first stage will analyse the co-variates of shocks, with the aim of isolating which characteristics of the sample households are correlated with the probability of experiencing any of the events considered. The second stage will focus on the strategies used by the households who experienced these adverse events and will attempt to determine which characteristics are correlated with specific types of coping strategies (controlling for factors related to the probability of a shock taking place). The final part of the analysis will combine the insights from the first two sections to identify the impact of the shocks, conditional on specific types of response.

3.4.1 Which Characteristics Are Correlated with Shocks?

Model (3.1) estimates the probability of experiencing a shock S as a function of household characteristics and monthly time effects (common to all households). Under the assumption that the zero-mean error term ε follows a standard normal distribution, the link function F represents the standard normal CDF and (3.1) is estimated as a probit model.

$$\text{Prob.}(S_{it} = 1) = F(\alpha_0 + \beta X_i + \beta K_{it-2} + \phi Langa_i + \delta_t + \varepsilon_{it}) \tag{3.1}$$

The left-hand side of equation (3.1) represents the probability of experiencing a shock S, for household i, in week t. On the right-hand side, X represents a selection of time-invariant characteristics measured during the baseline interviews. The term K represents potentially endogenous (time-dependent) variables, lagged by two weeks to reduce the possibility of reverse causality, due to households pre-empting events taking place (weddings, births or initiations, for example).[18]

3.4.2 Which Characteristics Determine Coping Strategies?

To determine the relative probability of households selecting certain coping strategies in the event of a shock, the possible responses a household could choose are used as the dependent variables in a multinomial logit model. This choice of coping strategy is best motivated with a discussion of latent variables.

Consider a household with a pre-shock level of utility U_i. If the household experiences a specific magnitude (m) of negative shock S_{im}, it selects a coping strategy CS_c, from K available strategies ($CS_1, CS_2 \ldots CS_K$). In the absence of

[18]Evidence of the ability to adjust consumption in the week prior to the shock taking place is found in Section 3.5. However, these results (and the results in Section 3.6.2) also suggest no pre-shock adjustments are made before this period.

any (effective) strategy, the negative shock would lead to a reduction in utility $\Delta\underline{U}_{im} < 0$. The selection of a coping strategy therefore depends on the extent to which this reduction in utility can be limited. Denoting this adjusted reduction in utility as $\Delta\overline{U}_{im} \leq 0$, it is possible to define a variable $\tilde{w}_{imc} = \Delta\overline{U}_{im} - \Delta\underline{U}_{im}$ which would be positive for any strategy that effectively reduces the impact of the shock. For any given shock, the relationship between the latent variable and the observed choice of coping strategy can be expressed as follows.

$$\tilde{w}_{imc}\,(C_{ic}\,|\,S_{im}) = \beta_0 + \beta H_i + \mu_{imc}$$

$$C_{ic}\,|\,S_{im} = \begin{cases} CS_1 & \text{if} \quad \tilde{w}_{im1} > \tilde{w}_{imc} \qquad \forall\, c \neq 1 \\ & \qquad \vdots \\ CS_K & \text{if} \quad \tilde{w}_{imK} > \tilde{w}_{imc} \qquad \forall\, c \neq K \end{cases}$$

As a specific response is only observed when an event occurs, the following econometric model is estimated using only the sub-sample of observations containing weeks where household experienced a shock.

$$\text{Prob.}(C_{ic}\,|\,S_{im}) = \frac{\exp\left(\gamma_c S_{imt} + \beta_c H_{it}\right)}{\sum_{k=0}^{K} \exp\left(\gamma_k S_{imt} + \beta_k H_{it}\right)} \tag{3.2}$$

$$\text{Where } H_{it} = X_i + K_{it-2} + Langa_i + Quarter_t, \quad \text{and } c = 0, \ldots, K$$

The left-hand side of equation (3.2) indicates the probability of household i responding to a shock by using a coping strategy of type c (in week t). As a response occurs only as a result of a shock, this probability is conditional on a shock occurring to household i in that week. On the right-hand side of equation (3.2), S_{imt} represents a set of indicator variables, taking the value 1 if household i experienced a shock of magnitude m in week t. The variables contained in

H include the fixed and time-variant characteristics, again denoted X_i and K_i, and will also include any variables correlated with the probability of a household experiencing a shock (determined by estimations of equation 3.1). The model controls for location (through the dummy variable *Langa*), as well as flexible time trend expressed through a set of quarterly dummy variables. The coefficient estimates will represent the change in the relative log-odds of selecting a particular coping strategy, relative to the base category (no coping strategy), as a result of a change in the associated dependent variable.

3.4.3 Which Strategies Effectively Smooth Consumption?

To estimate the welfare effects of shocks on the sample households, the measure used as an indicator of household welfare is the value of food consumption per adult equivalent. The use of a consumption measure as a money-metric of welfare is standard within the literature on household shocks, although the use of food consumption specifically is preferable in this analysis for two reasons. Firstly, even though a reduction in (the value of) food consumption is still a fairly common strategy amongst poor households (Heltberg et al., 2014), food consumption is often found to be more effectively protected from income fluctuations than non-food consumption (Skoufias and Quisumbing, 2005). Therefore, a reduction in food consumption would indicate a last resort strategy on the part of households (Anand and Harris, 1994). Secondly, many shocks could imply an increase in total consumption, even though the household may have been forced to make drastic reductions in essential day-to-day spending to offset additional expenses. For example, where a household is required to pay for medicine or additional travel expenses, in the event of a health shock.

In the likelihood that the negative effects of shocks can be spread beyond (or anticipated before) the period in which they occur,[19] the lagged (and lead) effects of shocks must be incorporated into the proposed model to estimate their full effect on consumption. This approach is also desirable to estimate the impacts of the shock beyond the week the event actually takes place. All variables which exert a significant influence over the probability of either experiencing a shock or employing a particular response in equations (3.1) and (3.2) will be incorporated in the estimation of (3.3), as will a household-level fixed effect term to control for any unobserved (time-invariant) heterogeneity between households.

$$
\begin{aligned}
\log C_{it} = \delta_t + \alpha_i + \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_s S_{is} + \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{ms}(S_{is} * ModerateSevere_{mis}) \\
+ (\delta_t * Langa) + \gamma_1 \log Y_{it-2} + \gamma_2 \log Y_{it-3} + \gamma G_{it-2} + \upsilon_{it}
\end{aligned}
\tag{3.3}
$$

The dependent variable in the proposed baseline model (3.3) is the log of food consumption (per adult equivalent) for household i, in week t. The term δ_t represents a set of monthly time effects, while α_i represents a household-level fixed effect. δ_t is also interacted with the variable indicating a household comes from the Langa sample, to allow for a township-specific time trend. In this initial baseline model, only the impact of the different magnitudes of shock are estimated (coping strategies are not considered). This is achieved through the inclusion of the variable S_{it} indicating whether or not a household experienced any shock at time t, alongside the interaction of this variable with an indicator of whether the shock was either moderate or severe.[20] As well as the effect of shocks being

[19]For example, if existing bulk food purchases were able to provide sufficient food for the household during the week of the shock itself, the true impact of the shock may only become apparent when these stocks become depleted and new purchases must be made (presumably with lower funds available).

[20]Shocks which are either moderate or severe are considered as a single magnitude, to limit the disaggregation of shocks into as few groups as possible, when interacting shock magnitudes with coping strategies in model (3.4).

considered contemporaneously, these terms are included with weekly lags, to week $t - \underline{\tau}$, and weekly leads, to week $t + \overline{\tau}$, where the number of weeks (from $t - \underline{\tau}$ to $t + \overline{\tau}$) should represent the full effect of the shock, from the earliest week the shock could be pre-empted (if *ex-ante* consumption adjustments are made) to the week where food consumption returns to pre-shock levels. The vector of controls G_i will incorporate all time-variant characteristics, which were found to influence either the probability of experiencing a shock or the probability of employing a specific coping strategy, where, again, these are lagged by 2 weeks (see Section 3.4.1). The log of household regular income, per adult-equivalent, Y_{it} is included on the right-hand side of (3.3) and lagged by both 2 and 3 weeks.[21]

To assess the relative degree to which specific coping strategies protect households from shocks, a further layer of interactions is included in model (3.4).

$$\begin{aligned}
\log C_{it} = \delta_t + \alpha_i + \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_s S_{is} + \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{sav\ s}(S_{is} * CS_{sav\ is}) \\
+ \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{ass\ s}(S_{is} * CS_{ass\ is}) + \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{ms}(S_{is} * ModerateSevere_{mis}) \\
+ \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{sav\ ms}(S_{is} * ModerateSevere_{mis} * CS_{sav\ is}) \\
+ \sum_{s=t-\underline{\tau}}^{t+\overline{\tau}} \beta_{ass\ ms}(S_{is} * ModerateSevere_{mis} * CS_{ass\ is}) \\
+ (\delta_t * Langa) + \gamma_1 \log Y_{it-2} + \gamma_2 \log Y_{it-3} + \gamma G_{it-2} + \upsilon_{it}
\end{aligned} \tag{3.4}$$

The variables CS_{sav} and CS_{ass} indicate whether a household employed a savings-based or assistance/insurance-based strategy in responding to the shock, with the base category being no response. Given this approach, it is possible to measure the estimated change in food consumption, given a specific magnitude of shock and a specific response, as a linear combination of the coefficients.

[21] The addition of the 3 week lag of income was found to be a significant determinant of food consumption in early estimations of model (3.3).

3.5 Results

Before reporting the impact of the various shocks on household food consumption, the results from estimating the binary dependent variable model, aimed at determining the probability of experiencing any shock are reported in Section 3.5.1, while Section 3.5.2 presents results from the multinomial logit model of coping strategy choice. Sections 3.5.3 and 3.5.4 then report the effects of shocks on food consumption, based on models (3.3) and (3.4), where these effects are first estimated, unconditional on the coping strategy employed, and then conditional on household's method of response. To conclude the main results of this analysis, Section 3.5.5 considers the variation between the estimated impact of shocks, under different coping strategies.

3.5.1 The Probability of Experiencing a Shock

Table 3.5 shows the results of estimating equation (3.1), to determine which factors influence the probability of households experiencing a shock. The household-specific variables, considered as potentially correlated with the probability of a shock occurring, consist of those which should intuitively influence exposure to shocks and those suggested by the literature discussed in Section 3.2. Specifically, the included covariates are the term indicating whether or not the household is considered relatively poor (based on the PWR ranking) and the gender, age and education level of the household head. A variable measuring the number of shocks a household experienced in the year prior to the survey is also included, along with whether or not the household contained a chronically ill or disabled member. Lagged values of the size of the household and dependency ratio are also added to the estimations. All results in Table 3.5 indicate marginal effects, with the standard errors clustered at the household level.

Table 3.5: The Probability of Experiencing a Shock

Dependent variable: Shock =1	(1)	(2)	(3)	(4)
Shock magnitude:	Any	Mild	Moderate	Severe
	dy/dx	dy/dx	dy/dx	dy/dx
Poor household	−0.0014	−0.0034	−0.0004	0.0021
	(0.006)	(0.0036)	(0.0042)	(0.0025)
Household size $(t-2)$	0.0010	0.0015*	−0.0002	−0.0017*
	(0.0017)	(0.0008)	(0.0011)	(0.0009)
Dependency ratio $(t-2)$	0.0038	0.0070	−0.0051	0.0071
	(0.0135)	(0.0088)	(0.0079)	(0.0070)
Head female	−0.0015	−0.0029	0.0036	−0.0022
	(0.0062)	(0.0041)	(0.0044)	(0.0029)
Head age	−0.0001	−0.0001	0.0000	0.0001
	(0.0003)	(0.0001)	(0.0002)	(0.0001)
Head junior school	0.0033	−0.0017	0.0050	0.0005
	(0.0060)	(0.0034)	(0.0038)	(0.0029)
Disabled/chronic ill	0.0035	0.0050	−0.0007	−0.0015
	(0.0064)	(0.0044)	(0.0038)	(0.0030)
Shocks in previous year	0.0045	0.0004	0.0033**	0.0009
	(0.0029)	(0.0015)	(0.0015)	(0.0013)
Langa township	−0.0133**	−0.0057	−0.0079*	−0.0017
	(0.0064)	(0.0042)	(0.0043)	(0.0027)
Month dummies	Yes	Yes	Yes	Yes
Observations	3821	3821	3821	3821
Households	85	85	85	85
Log-likelihood	-445.11	-193.84	-199.64	-136.94
McFadden's R^2	0.031	0.070	0.063	0.060

Household clustered standard errors in parentheses   * p<0.1, ** p<0.05, *** p<0.01

The first column of Table 3.5 indicates the marginal effect of selected variables on the probability of experiencing any shock (of any magnitude). In-line with Table 3.3, the coefficient on the Langa variable indicates that a household in this township faces a 1.3% lower risk (on average) of experiencing a shock than a household in Diepsloot. However, other than this result, no covariates have a significant effect on the probability of a sample household experiencing a non-specific magnitude of shock (at any common significance level). Again, in-line with the summary of shocks presented in Table 3.3, it appears that (even controlling for additional covariates) inclusion within the lower wealth category does not necessarily predict a higher risk of experiencing a shock. As previously noted, however, many of the shocks within the sample do not occur directly to the households themselves.

Columns (2) to (4) of Table 3.5 show the marginal effects of the same covariates on the probability of experiencing a particular magnitude of shock. Again, few of the independent variables appear to exert any strong influence on these probabilities, perhaps with the exception of the variable measuring the number of shocks a household experienced in the year prior to the survey, which appears to be positively correlated with the probability of experiencing a shock of moderate magnitude. It is important to recognise, however, that this lack of any strong correlation could be seen as a desirable result at this stage of the analysis, as this simply implies that general exposure to the type of shocks considered is largely random. In addition to this, finding little correlation between exposure to risk and commonly used household and economic characteristics does not imply that no such characteristics exist. Instead, it is likely that the probability of experiencing a shock is at least partially determined by other unobserved factors. Provided these characteristics are time-invariant, the estimations of the impact of shocks on food consumption should still be consistent. Those variables which do show some evidence of being correlated with some magnitudes of shock (household size, location and past shocks) will be included in the following stages of the analysis, to limit any remaining bias relating to which households experience these events.

3.5.2 The Probability of a Specific Coping Strategy

The results of estimating model (3.2), on the sub-sample of weeks where households experienced a shock, are shown in Table 3.6.[22] As the correlates of the probabilities of choosing a particular response are estimated within the same model, all columns of Table 3.6 represent a single estimation of model (3.2). The variables of household size (lagged 2 periods) and shocks in the previous year are included in the estimation, due to the finding that these variables were (albeit minimally) correlated with the probability of the sample households experiencing some magnitudes of shock (see Table 3.5). Although not correlated with the probability of experiencing a shock, the variable indicating if a household is relatively poor, along with the gender (and age) of the household head, could potentially influence the likelihood of certain responses being favoured over others.[23] These variables are, therefore, included in the estimation of model (3.2). In a similar manner, the fact that a household member is disabled or chronically ill may prompt households to choose (or avoid) a particular type of coping strategy. The estimation also includes two additional explanatory variables, the number of bank accounts per adult member (as a proxy for access to formal financial services) and whether or not the household head or spouse is a member of a funeral society.

The first column of Table 3.6 shows the predicted probability of responding to a shock using a savings-based strategy. The results indicate that using this type of strategy is more likely where the household has experienced a moderate or severe shock (relative to a mild shock), a result that mirrors the summary figures in Table 3.4. Column (1) also indicates that an additional household member reduces the probability of utilising a savings-based strategy by approximately 6%,

[22]Of the 107 events which occurred during the survey, the inclusion of lagged regressors reduces the number of shocks considered to 100.

[23]In their study of data on coping strategies from 16 developing countries, Heltberg et al. (2014) find that (in general) poorer households are more likely to employ 'bad' coping strategies, such as reducing food consumption or selling productive assets, whereas the gender of the household head may influence coping choice through channels such as differential access to credit (Okurut, 2006).

whereas experiencing an additional shock, in the year before the survey, reduces the probability of using this type of response by around 11%. The latter result could be evidence of households depleting their available savings when responding to past events, leaving this coping strategy unavailable. The negative coefficient on both household size and the age of the household head may represent life-cycle effects. For example, if younger household heads are more likely to be working, and have fewer dependents, this could arguably allow them to accumulate savings more easily. The final row of Table 3.6 also indicates an increased likelihood of a savings-based strategy in the sample from Langa, relative to the households in Diepsloot. This is the case in spite of no significant difference between the incomes or wealth ranking in the two samples (If anything, the average household in Langa has lower income). One possible explanation for this is that, where keeping money generally takes place outside of formal financial institutions, Langa may be a more conducive environment for accumulating savings.[24]

Column (2) of Table 3.6, indicates that the only variables exerting a significant influence on the probability of using an assistance/insurance based strategy are the dummy variable for the most severe shocks and (to a lesser extent) the size of the household. The dominance of this type of response to the most severe events can clearly be explained as a result of the worst shocks overwhelming the ability of a household to mobilise sufficient funds internally.[25] Although the household head or spouse's membership of a burial society would be expected to be positively correlated with assistance/insurance-based strategies, the results indicate that membership of this type of organisation is more likely to prompt a savings-based approach (albeit at low significance). If this measure is viewed as a general indicator of household attitudes towards risk, however, it could also be seen as a potential signal of precautionary saving behaviour.

[24] Diepsloot has a reputation for extremely high levels of crime, including burglary (Mahajan, 2014). Therefore, there may be greater security concerns over savings being held within the household.

[25] A reliance on friends or family to provide assistance, in the event of the most severe shocks, is a common finding in the literature on coping strategies. For example, in their study of rural Ethiopian households, Yilma et al. (2014) find that the proportion of households borrowing more than doubles, when a shock's severity goes from the lowest to the highest category.

Table 3.6: The Probability of Using a Specific Coping Strategy

Multinomial Logit	(1)		
Outcome: Coping strategy =	Savings-based	Assistance/Insurance	No strategy
	dy/dx	dy/dx	dy/dx
Moderate shock	0.215**	0.0359	−0.251***
	(0.108)	(0.124)	(0.081)
Severe shock	0.206**	0.304***	−0.510***
	(0.095)	(0.107)	(0.083)
Poor household	0.120	−0.023	−0.097
	(0.095)	(0.123)	(0.128)
Household size $(t-2)$	−0.066**	0.049*	0.017
	(0.030)	(0.030)	(0.024)
Head female	−0.063	0.098	−0.035
	(0.081)	(0.141)	(0.135)
Head age	−0.009**	0.008	0.000
	(0.004)	(0.009)	(0.009)
Head junior school	−0.058	0.163	−0.105
	(0.070)	(0.130)	(0.118)
Disabled/chronic ill	0.170*	−0.032	−0.139
	(0.089)	(0.097)	(0.100)
Shocks in previous year	−0.109***	0.014	0.095***
	(0.033)	(0.041)	(0.036)
Bank accounts	0.080*	0.027	−0.107*
	(0.046)	(0.054)	(0.056)
In burial society	0.225*	−0.090	−0.135
	(0.118)	(0.109)	(0.116)
Langa township	0.257***	−0.017	−0.240***
	(0.077)	(0.093)	(0.080)
Quarter dummies			Yes
Observations (number of shocks)			100
Log-likelihood			-63.523
McFadden's R^2			0.368

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

The final column in Table 3.6 shows the determinants of the probability of not using either of the two types of coping strategies considered. The frequencies in Table 3.4 indicate that using no coping strategy is far more common for the lowest magnitude of shocks, and the multivariate analysis in Table 3.6 supports this finding. Table 3.6 indicates that experiencing an additional shock in the previous year implies the probability of a household using no coping strategy increases by approximately 9.5%. Again, this result is indicative of the depletion of a household (or a risk-sharing network's) stock of resources, as a result of previous shocks. The coefficient on the Langa township dummy variable also shows that the probability of enacting no specific coping response is significantly higher in the Diepsloot sample.[26]

3.5.3 The Unconditional Effects of Shocks

Before estimating the impact of shocks, conditional on the different coping strategies, Table 3.7 presents the unconditional effect of shocks on food consumption. In the following estimations, household size is the only additional control included, via the term G in model (3.3), other characteristics determining exposure to shocks or choice of response are time-invariant and, therefore, incorporated into the household fixed effect term. All other variables follow directly from the description of the unconditional model in Section 3.4.3. The full results of estimating (3.3) can be found in column (1) of Appendix 3.4. However, as the contemporaneous, delayed and pre-empted effects of a particular magnitude of shock are more clearly expressed as a linear combination of the estimated coefficients, Table 3.7 reports only these marginal effects.

The first column of Table 3.7 indicates that food consumption is, in general, well insulated from the effects of mild shocks. Neither the immediate effect (in week t), or adjustments prior to the shock, appear significantly different from zero. There is, however, some evidence of a drop in consumption at $t-3$ of approximately

[26]Results from an alternative estimation approach can be found in Appendix 3.3, where the probability of each choice of strategy is modelled separately in a series of binary outcome (probit) regressions. This approach changes little in the general interpretation of the results, although the increased probability of a household employing a savings-based strategy, in response to a severe shock, is no longer a significant result (p-value = 0.176).

9.3%, on average (p-value = 0.070).[27] Overall, the results suggest that either these shocks are generally not severe enough to warrant large reductions in food consumption, or the coping strategies employed are sufficient to avoid this course of action. In the case of the moderate/severe shocks, however, there is stronger evidence that these events do result in significant shortfalls in food consumption, with a decrease in food consumption of approximately 11.1%, in the week the shock occurs, and a smaller shortfall of around 8.8%, in the following week.

Table 3.7: The Unconditional Impact of Shocks on Food Consumption

Dependent variable: Log food consumption	(1) Mild shock	(2) Moderate or Severe shock
S $(t+2)$	−0.086	−0.029
	(0.076)	(0.062)
S $(t+1)$	−0.054	−0.055
	(0.050)	(0.046)
S (t)	−0.008	−0.111**
	(0.051)	(0.052)
S $(t-1)$	0.011	−0.088**
	(0.061)	(0.042)
S $(t-2)$	−0.017	−0.046
	(0.054)	(0.048)
S $(t-3)$	−0.093*	−0.067
	(0.051)	(0.053)
S $(t-4)$	−0.044	−0.005
	(0.051)	(0.053)
Controls†	Yes	Yes
HH Fixed Effects	Yes	Yes
Township trends	Yes	Yes
Observations	3566	3566
Households	85	85
R^2	0.114	0.114

Household clustered standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

[27] The results in the following section suggest this is related to shocks which prompt an assistance/insurance based response. Therefore, this may be capturing repayments on loans or reciprocal arrangements obtained during the event itself.

3.5.4 The Effects of Shocks, Conditional on Response

Having estimated the unconditional impact of the shocks on food consumption, the next stage in the analysis is to include the full list of interactions described in model (3.4). The results of estimating this model can be found in column (2) of Appendix 3.4, whereas the marginal effects (specific to each shock magnitude and coping strategy pair) are shown in Table 3.8 below. These effects are also presented graphically in Figure 3.1.

Table 3.8: The Impact of Shocks on Food Consumption, by Coping Strategy

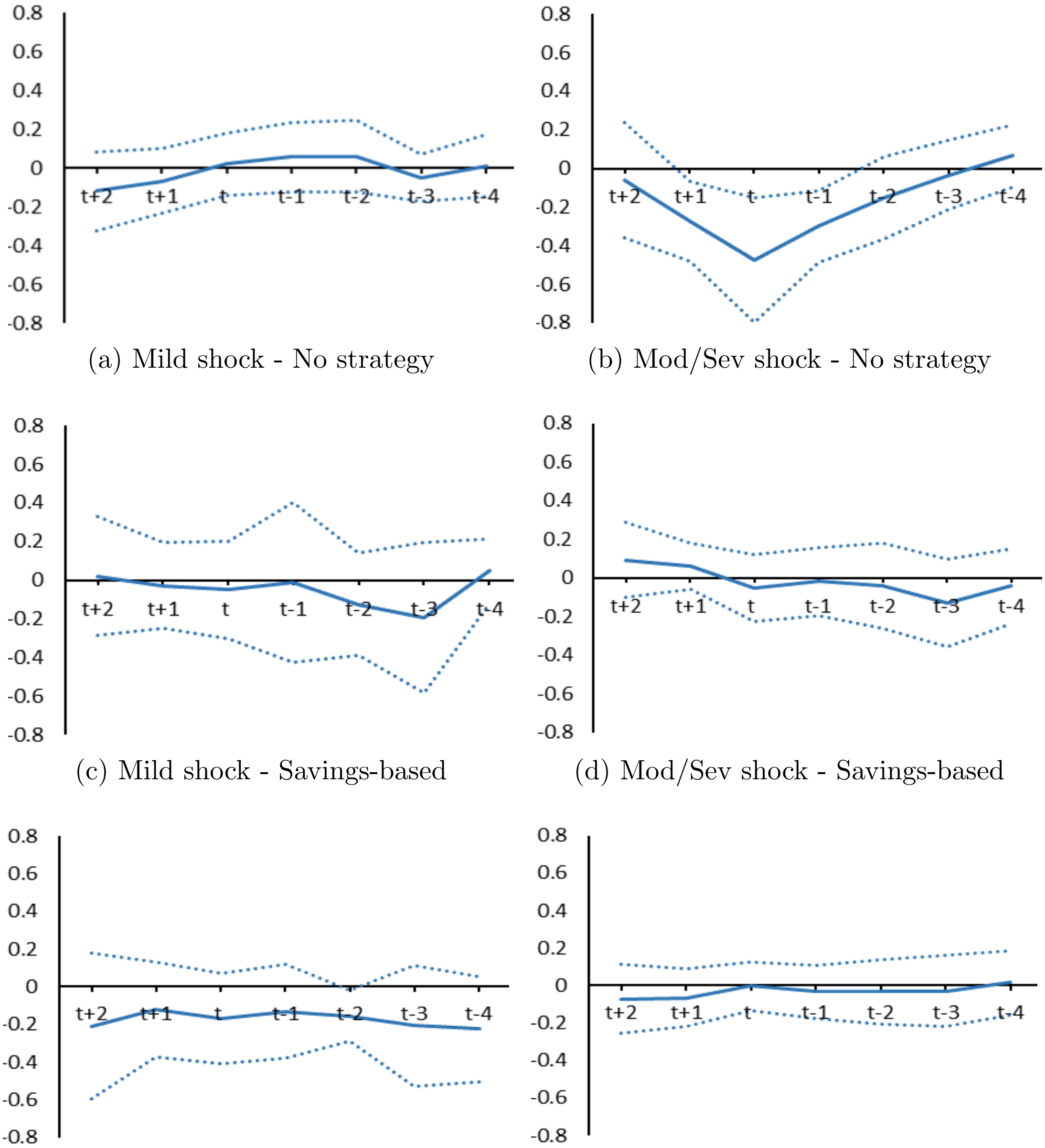
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Log food consumption	Mild shock			Moderate or Severe shock		
Coping strategy:	No Strategy	Savings-based	Assistance/ Insurance	No Strategy	Savings-based	Assistance/ Insurance
S $(t+2)$	−0.119	0.021	−0.210	−0.062	0.093	−0.071
	(0.102)	(0.154)	(0.195)	(0.150)	(0.097)	(0.094)
S $(t+1)$	−0.065	−0.027	−0.121	−0.275***	0.061	−0.065
	(0.084)	(0.111)	(0.126)	(0.103)	(0.061)	(0.078)
S (t)	0.022	−0.050	−0.168	−0.475***	−0.053	−0.002
	(0.081)	(0.128)	(0.121)	(0.162)	(0.088)	(0.066)
S $(t-1)$	0.058	−0.011	−0.130	−0.300***	−0.018	−0.034
	(0.089)	(0.209)	(0.125)	(0.094)	(0.088)	(0.072)
S $(t-2)$	0.063	−0.124	−0.156**	−0.152	−0.039	−0.034
	(0.092)	(0.133)	(0.068)	(0.107)	(0.111)	(0.085)
S $(t-3)$	−0.050	−0.193	−0.207	−0.036	−0.132	−0.028
	(0.062)	(0.195)	(0.162)	(0.090)	(0.115)	(0.095)
S $(t-4)$	0.013	0.049	−0.224	0.065	−0.042	0.015
	(0.081)	(0.084)	(0.140)	(0.081)	(0.097)	(0.087)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3566	3566	3566	3566	3566	3566
Households	85	85	85	85	85	85
R^2	0.125	0.125	0.125	0.125	0.125	0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

The broad interpretation of the results in columns (1) to (3) of Table 3.8 is that food consumption appears relatively well insulated from mild shocks. The exception to this is a 15.6% reduction in food consumption two weeks after a shock occurs, as a result of households employing an assistance/insurance-based coping strategy. This could represent evidence of households reducing consumption to make repayments on loans (or gifts), used to smooth consumption during the week the shock takes place. It is also likely that households relying on assistance for such mild shocks, will do so, not out of choice, but rather due to an inability to mobilise funds internally (from either savings or regular income). A household in such circumstances may be limited in their choices of assistance and may, therefore, be subjected to higher repayment rates, resulting in a stronger *ex-post* reduction in consumption to meet these expenses (Collins et al., 2010).

Considering the results for the combined category of moderate/severe shocks, in columns (4) to (6), it is apparent that the use of either of the two coping strategies (savings-based or assistance/insurance-based) results in the effective insulation of food consumption from the negative impact of shocks. In contrast, however, column (4) indicates that the choice to respond to a moderate/severe shock with no coping strategy results in a substantial reduction in food consumption, not only in the week the event takes place, but in both the following week and the week prior to the shock. The immediate effect (at time t) of a moderate/severe shock, which is met with no coping strategy, is a 47.5% reduction in food consumption. This reduction is even more striking when considering that (on average) households experience an additional shortfall of around 27.5%, in the week before a shock occurs, while the week following the event also sees food consumption reduced from pre-shock levels by approximately 30%. It is only at $t-2$ that the impact of the shock on food consumption is no longer significantly different from zero (at the levels given).

Figure 3.1: The Impact of Shocks on Food Consumption



(a) Mild shock - No strategy

(b) Mod/Sev shock - No strategy

(c) Mild shock - Savings-based

(d) Mod/Sev shock - Savings-based

(e) Mild shock - Assistance/Insurance

(f) Mod/Sev shock - Assistance/Insurance

Dotted lines represent a 95% confidence interval around the estimates in Table 3.8

The results described in Table 3.8 are also presented in Figure 3.1, where the estimated effects are graphed against the lags and leads of the shocks. Each panel represents a combination of shock magnitude and response, and includes a 95%

confidence interval (bounded by the dotted lines). The drastic reduction in food consumption, as a result of a moderate/severe shock, under no coping strategy, is clearly evident in the panel (b) of Figure 3.1. Aside from this result, however, only the $t-2$ period in panel (e), representing the use of an assistance/insurance-based strategy (to respond to a mild shock) shows evidence of a negative impact which is statistically different from zero at the 5% level.

3.5.5 Difference in the Effectiveness of Responses

The results of the previous section indicate whether or not the impact of a shock is significantly different from zero, given a specific choice of coping strategy. However, perhaps a more relevant question is whether the impact of a shock differs between responses. Table 3.9 reports the difference between the predicted impacts of the shocks, for all pairings of coping strategy. Columns (1) to (3) report how the impact of mild shocks differs between responses, whereas columns (4) to (6) report these differences for moderate/severe shocks. At any particular lag (or lead) of the specified magnitude of shock, a positive change in the predicted effect between two strategies $\Delta_{ab} = (CS_a - CS_b)$ implies the first strategy CS_a is relatively more effective at mitigating the impact of the shock (the decline in consumption is relatively less severe). Whereas, a negative change implies the second strategy CS_b is relatively more effective.

The first three columns of Table 3.9 indicate that the ability to mitigate the negative impact of a mild shock differs little between the three possible responses. The only exception can be found in column (2), where the use of an assistance/insurance-based strategy (relative to no coping strategy) is associated with a further negative impact of around 22% two weeks after a shock takes place (p-value = 0.071). It is possible that, if income is sufficient to cover the costs of these smaller shocks, and seeking assistance entails an obligation of repayment, the use of regular income (defined as no coping strategy in Section 3.3.4) may be a more effective choice of response.[28]

Table 3.9: The Difference in the Impact of Shocks between Coping Strategies

	(1)	(2)	(3)	(4)	(5)	(6)
	Mild shock			Moderate or Severe shock		
Diff. in effect:	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance
Table 3.8 column:	(2)-(1)	(3)-(1)	(2)-(3)	(5)-(4)	(6)-(4)	(5)-(6)
S $(t+2)$	0.139	−0.091	0.231	0.155	−0.009	0.164
	(0.787)	(0.179)	(0.592)	(0.577)	(0.005)	(0.206)
S $(t+1)$	0.039	−0.055	0.094	0.336***	0.210	0.126
	(0.047)	(0.079)	(0.700)	(0.131)	(0.176)	(0.147)
S (t)	−0.071	−0.190	0.119	0.421**	0.472**	−0.051
	(0.173)	(0.275)	(0.342)	(0.213)	(0.207)	(0.128)
S $(t-1)$	−0.069	−0.188	0.119	0.282**	0.266**	0.016
	(0.097)	(0.322)	(0.392)	(0.153)	(0.154)	(0.013)
S $(t-2)$	−0.187	−0.219*	0.032	0.113	0.119	−0.006
	(0.378)	(0.147)	(0.032)	(0.229)	(0.490)	(0.003)
S $(t-3)$	−0.143	−0.157	0.014	−0.096	0.008	−0.104
	(0.760)	(0.468)	(0.008)	(0.125)	(0.005)	(0.177)
S $(t-4)$	0.036	−0.237	0.273	−0.107	−0.050	−0.057
	(0.048)	(0.262)	(0.281)	(0.454)	(0.113)	(0.138)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3566	3566	3566	3566	3566	3566
Households	85	85	85	85	85	85
R^2	0.125	0.125	0.125	0.125	0.125	0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

[28]It is important to note that in re-estimating the results in Table 3.9, with the inclusion of additional leads and lags of the shocks, this coefficient is no longer found to be significantly different from zero (see Section 3.6.2).

Columns (4) to (6) report the heterogeneity in the effectiveness of coping strategies in protecting food consumption from moderate/severe shocks. In the case of both the savings-based and assistance/insurance-based strategies, the reduction in the impact of these shocks, relative to the use of no strategy, is significant, both in the week during and following the shock. In the period of the shock (week t), the use of no coping strategy is associated with an additional negative reduction in food consumption in excess of 40%, relative to either of the alternatives. Furthermore, in the week immediately following the shock, the use of no coping strategy is associated with an additional reduction in excess of 25%. The predicted change in impact between a savings-based strategy (relative to no strategy) in column (4), also shows evidence of a 33.6% relative reduction, in the week prior to the shock taking place (p-value = 0.006). The final column in Table 3.9 reports the difference in the impact of a moderate/severe shock, between households using a savings-based or an assistance/insurance-based strategy. The difference in the impact of a moderate/severe shock between these two strategies is never statistically different from zero, either in the week of the shock itself, or in the *ex-ante*/*ex-post* period surrounding the event. Therefore, neither strategy appears to be a more (or less) effective response to these more serious shocks.

3.6 Robustness

This part of the analysis aims to test the robustness of the main findings presented in Sections 3.5.4 and 3.5.5. The first robustness test below addresses the extent to which shocks can be pre-empted, by re-estimating model (3.4), having first removed the three specific types of shocks which households should (reasonably) be able to predict. In Section 3.6.2, the time period under consideration is extended by two additional weeks, both before and after the shock occurs, in order to determine if results are sensitive to the number of lags or leads incorporated in the estimations.

3.6.1 The Omission of More Predictable Shocks

Where households can foresee events that will place additional strain on income or expenditure, it is reasonable to assume that they will seek to take action to smooth essential consumption.[29] It is likely, therefore, that the extent to which shocks can be predicted may influence both the choice of coping strategy and the impact of these shocks on consumption. For example, households may elect to make adjustments to saving behaviour, or perhaps attempt to earn additional income by seeking secondary employment, either of which actions would increase the probability that the household elects to employ a savings-based strategy in response to the shock (assuming additional labour income is saved). To assess the likelihood that more predictable shocks unduly influence the results in Tables 3.8 and 3.9, the estimations of model (3.4) are repeated, with the omission of weddings, births and initiation ceremonies. These should, arguably, constitute the more predictable shocks included in the list of events in Appendix 3.1. Table 3.10 presents results for the impact of the remaining sub-set of shocks on food consumption, conditional on coping strategies, while Table 3.11 reports the results of the difference in effectiveness between the possible responses.

The original estimations of model (3.4) found a 15.6% reduction in food consumption two weeks after a mild shock occurred, as a result of employing an assistance/insurance-based coping strategy (see Table 3.8). For the same period, Table 3.9 also indicated that this response is significantly less effective than the use of no coping strategy (p-value = 0.071). Both results are robust to the omission of the more predictable shocks in Tables 3.10 and 3.11, although the magnitude of the estimated coefficients varies to some degree.

[29] The 27.5% decline in food consumption in the week before the shock, in Column (4) of Table 3.8, provides evidence that some shocks can be pre-empted by household, at least one week before they take place.

Table 3.10: The Impact of Shocks - Omission of More Predictable Events

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Log food consumption	Mild shock			Moderate or Severe shock		
Coping strategy:	No Strategy	Savings-based	Assistance/ Insurance	No Strategy	Savings-based	Assistance/ Insurance
S $(t+2)$	-0.122	0.019	-0.251	-0.076	0.081	-0.140
	(0.105)	(0.156)	(0.200)	(0.180)	(0.101)	(0.092)
S $(t+1)$	-0.062	-0.027	-0.081	-0.247*	0.044	-0.092
	(0.087)	(0.114)	(0.150)	(0.129)	(0.063)	(0.090)
S (t)	0.014	-0.051	-0.161	-0.433**	-0.064	-0.032
	(0.084)	(0.131)	(0.133)	(0.202)	(0.085)	(0.076)
S $(t-1)$	0.058	-0.011	-0.159	-0.330***	-0.036	-0.080
	(0.094)	(0.213)	(0.145)	(0.110)	(0.087)	(0.070)
S $(t-2)$	0.063	-0.128	-0.179**	-0.179*	-0.058	-0.090
	(0.099)	(0.137)	(0.074)	(0.107)	(0.110)	(0.076)
S $(t-3)$	-0.060	-0.196	-0.062	-0.066	-0.157	-0.089
	(0.064)	(0.199)	(0.123)	(0.089)	(0.116)	(0.087)
S $(t-4)$	-0.007	0.046	-0.183	0.026	-0.056	-0.024
	(0.081)	(0.080)	(0.160)	(0.081)	(0.105)	(0.087)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3566	3566	3566	3566	3566	3566
Households	85	85	85	85	85	85
R^2	0.125	0.125	0.125	0.125	0.125	0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

In relation to the moderate/severe shock category, the results in Table 3.8 indicated that the use of either a savings-based or an assistance/insurance-based coping strategy was an effective means of smoothing food consumption. The results of Table 3.9 also found that the difference in the impact of a moderate/severe shock between these two strategies was never statistically different from zero. Again, both results are robust to the omission of the more predictable shocks. The main results in Table 3.8 also showed a substantial decline in food consumption as a result of employing no coping strategy, both in the period of the shock and in the weeks before and after the shock occurred. In general, this result is not substantially altered by the omission of weddings, births and initiations. However, the negative effect of these shocks appears pushed more towards the *ex-post* period, both in terms of significance and the magnitude of the coefficients, suggesting that the more predictable shocks were indeed responsible for much of the *ex-ante* consumption adjustment.

Table 3.11: The Difference in Impacts - Omission of More Predictable Events

	(1)	(2)	(3)	(4)	(5)	(6)
	Mild shock			Moderate or Severe shock		
Diff. in effect:	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance
Table 3.10 column:	(2)-(1)	(3)-(1)	(2)-(3)	(5)-(4)	(6)-(4)	(5)-(6)
S $(t+2)$	0.140	-0.129	0.270	0.158	-0.064	0.221
	(0.845)	(0.602)	(0.464)	(0.140)	(0.099)	(0.177)
S $(t+1)$	0.035	-0.020	0.055	0.291*	0.154	0.136
	(0.038)	(0.014)	(0.077)	(0.177)	(0.365)	(0.165)
S (t)	-0.066	-0.175	0.109	0.369	0.401*	-0.032
	(0.129)	(0.372)	(0.840)	(0.292)	(0.269)	(0.041)
S $(t-1)$	-0.069	-0.217	0.148	0.293**	0.250*	0.044
	(0.094)	(0.340)	(0.104)	(0.171)	(0.161)	(0.084)
S $(t-2)$	-0.191	-0.241*	0.051	0.121	0.090	0.032
	(0.420)	(0.154)	(0.076)	(0.106)	(0.367)	(0.034)
S $(t-3)$	-0.136	-0.002	-0.134	-0.091	-0.023	-0.069
	(0.140)	(0.001)	(0.763)	(0.694)	(0.021)	(0.209)
S $(t-4)$	0.053	-0.176	0.229	-0.082	-0.051	-0.031
	(0.127)	(0.472)	(0.256)	(0.746)	(0.114)	(0.034)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3566	3566	3566	3566	3566	3566
Households	85	85	85	85	85	85
R^2	0.125	0.125	0.125	0.125	0.125	0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

The results in Table 3.11 also show less clear evidence of a strong advantage for either of the two active responses, over no coping strategy (relative to Table 3.9). One potential explanation for this is that either a savings or an assistance/insurance-based approach is most effective where financial shocks can be foreseen to some degree. Although, this may also represent larger standard errors driven by a reduction in the number of shocks from which heterogeneity in impacts can be assessed.

3.6.2 The Inclusion of Additional Lags and Leads

Another important test of the robustness of the results estimated in Section 3.5 relates to the number of lags and leads included in models (3.3) and (3.4). The accurate estimation of the impact of shocks is dependent on the assumption that 4 lags and 2 leads ($t - \underline{\tau} = 4$ and $t + \overline{\tau} = 2$) are sufficient to capture the full impact of the events. To test this assumption, the following estimations extend the number of both lags and leads by two weeks to $t - \underline{\tau} = 6$ and $t + \overline{\tau} = 4$.

The results of re-estimating model (3.4), with the inclusion of the additional lags and leads are reported in Tables 3.12 and 3.13. These findings show that the additional time periods do little to alter the overall impact of the moderate/severe shocks, perhaps with the exception of a reduction in the contemporaneous difference in the impact between a savings-based strategy and no strategy (column 4 of Table 3.13) relative to the same coefficient in Table 3.9. All originally significant coefficients remain so in the new specification, while the coefficients on the additional lags and leads are wholly insignificant, at the levels shown.

In the case of the mild shocks, the inclusion of additional lags suggests a further 16.8% negative effect on consumption at $t - 6$, given an assistance/insurance-based strategy (see Table 3.12 column 3). Although, this result is only marginally significant (at 10%). In addition to this, the original negative impact at $t - 2$ is found to be slightly stronger (more negative) at 17.2%. In spite of this, however, the overall interpretation of the mild shock results does not change. Furthermore, the result that no coping strategy implies a less negative impact at $t - 2$, relative to an assistance/insurance strategy (see Table 3.9, column 2), is shown not to be robust to the inclusion of the additional time periods. Indeed, Table 3.13 shows that, in the case of mild shocks, no coping strategy performed significantly better than any other (in any period before, after or during the event).

Table 3.12: The Impact of Shocks - Additional Lags and Leads

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Log food consumption	Mild shock			Moderate or Severe shock		
Coping strategy:	No Strategy	Savings-based	Assistance/ Insurance	No Strategy	Savings-based	Assistance/ Insurance
S $(t+4)$	−0.118	−0.006	−0.102	−0.113	−0.061	−0.138
	(0.154)	(0.144)	(0.137)	(0.115)	(0.075)	(0.106)
S $(t+3)$	−0.030	0.011	−0.056	−0.180	0.038	−0.098
	(0.097)	(0.154)	(0.145)	(0.162)	(0.080)	(0.108)
S $(t+2)$	−0.069	−0.009	−0.211	−0.085	0.094	−0.097
	(0.097)	(0.192)	(0.222)	(0.160)	(0.109)	(0.104)
S $(t+1)$	−0.056	−0.025	−0.183	−0.306***	0.043	−0.069
	(0.086)	(0.123)	(0.135)	(0.115)	(0.073)	(0.080)
S (t)	0.015	−0.135	−0.175	−0.454**	−0.115	0.003
	(0.090)	(0.119)	(0.130)	(0.173)	(0.102)	(0.075)
S $(t-1)$	0.034	−0.110	−0.140	−0.330***	−0.035	−0.032
	(0.094)	(0.223)	(0.131)	(0.111)	(0.093)	(0.085)
S $(t-2)$	0.041	−0.201	−0.172**	−0.145	−0.056	−0.060
	(0.090)	(0.146)	(0.086)	(0.124)	(0.130)	(0.093)
S $(t-3)$	−0.073	−0.273	−0.222	−0.148	−0.142	−0.035
	(0.064)	(0.243)	(0.186)	(0.100)	(0.134)	(0.104)
S $(t-4)$	−0.012	0.039	−0.239	−0.032	−0.044	0.034
	(0.083)	(0.083)	(0.176)	(0.096)	(0.121)	(0.098)
S $(t-5)$	0.028	−0.050	−0.023	0.037	−0.022	−0.042
	(0.072)	(0.108)	(0.092)	(0.124)	(0.101)	(0.075)
S $(t-6)$	−0.004	−0.039	−0.168*	0.157	−0.004	0.009
	(0.068)	(0.057)	(0.087)	(0.127)	(0.114)	(0.091)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3226	3226	3226	3226	3226	3226
Households	85	85	85	85	85	85
R^2	0.124	0.124	0.124	0.124	0.124	0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

Table 3.13: The Difference in Impacts - Additional Lags and Leads

	(1)	(2)	(3)	(4)	(5)	(6)
		Mild shock			Moderate or Severe shock	
Diff. in effect:	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance	Savings-based - No Strategy	Assistance/ Insurance - No Strategy	Savings-based - Assistance/ Insurance
Table 3.12 column:	(2)-(1)	(3)-(1)	(2)-(3)	(5)-(4)	(6)-(4)	(5)-(6)
S $(t+4)$	0.112	0.016	0.096	0.174	0.251	0.077
	(0.431)	(0.011)	(0.280)	(0.258)	(0.235)	(0.653)
S $(t+3)$	0.041	-0.026	0.067	0.218	0.082	0.137
	(0.041)	(0.021)	(0.095)	(0.308)	(0.183)	(0.301)
S $(t+2)$	0.060	-0.142	0.202	0.179	-0.012	0.191
	(0.080)	(0.718)	(0.715)	(0.531)	(0.008)	(0.236)
S $(t+1)$	0.031	-0.127	0.158	0.349**	0.237	0.113
	(0.031)	(0.091)	(0.482)	(0.152)	(0.186)	(0.221)
S (t)	-0.149	-0.190	0.040	0.339*	0.457**	-0.118
	(0.413)	(0.347)	(0.044)	(0.258)	(0.227)	(0.370)
S $(t-1)$	-0.144	-0.174	0.030	0.295**	0.298**	-0.003
	(0.116)	(0.447)	(0.023)	(0.172)	(0.178)	(0.001)
S $(t-2)$	-0.242	-0.213	-0.030	0.089	0.084	0.005
	(0.310)	(0.174)	(0.027)	(0.286)	(0.293)	(0.002)
S $(t-3)$	-0.200	-0.149	-0.051	0.006	0.114	-0.107
	(0.139)	(0.117)	(0.045)	(0.003)	(0.797)	(0.841)
S $(t-4)$	0.051	-0.227	0.277	-0.012	0.066	-0.078
	(0.101)	(0.394)	(0.262)	(0.008)	(0.215)	(0.240)
S $(t-5)$	-0.078	-0.051	-0.027	-0.058	-0.079	0.021
	(0.414)	(0.131)	(0.026)	(0.096)	(0.386)	(0.018)
S $(t-6)$	0.043	-0.164	0.207	-0.161	-0.147	-0.014
	(0.152)	(0.157)	(0.136)	(0.432)	(0.396)	(0.009)
Controls†	Yes	Yes	Yes	Yes	Yes	Yes
HH Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3226	3226	3226	3226	3226	3226
Households	85	85	85	85	85	85
R^2	0.124	0.124	0.124	0.124	0.124	0.124

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01
†Household size $(t-2)$, log household regular income $(t-2)$ and $(t-3)$

3.7 Discussion and Concluding Remarks

This paper presents a three-stage approach to estimating the welfare impact of shocks, conditional on a household's choice of coping strategy. Initially, significant predictors of shocks were identified, so as to limit selection bias in the analysis of both the choice of strategy and the conditional impact of the shocks considered. Correlates of the choice of coping strategy were then assessed, having first controlled for factors affecting the probability of any shock taking place. Finally, the analysis employed food consumption as a money-metric measure of welfare, in order to estimate the impact of different magnitudes of shock, conditional on the household's choice of response.

When attempting to determine which variables increased or decreased a household's probability of experiencing a shock, few covariates provided significant insight. In contrast to other related studies, the gender and education level of the household head, the household size and the wealth of the household did not appear to affect exposure to shocks (Heltberg and Lund, 2009; Yilma et al., 2014; Heltberg et al., 2014). Instead, events appeared to occur largely at random, with the only determinant of increased susceptibility being the inclusion of households in the sample from the Langa township. One key reason for this result is likely to be the fact that 38.3% of the shocks recorded related to family members living outside of the household, making the characteristics of the household itself less influential as a predictor of these events taking place.

The second part of the analysis focused on determining which variables were correlated with the choice of coping strategy. These results indicated that the relationship between the magnitude of the shock and the choice of coping strategy appeared to followed a general pattern, whereby mild shocks would result in no coping strategy, moderate shocks would elicit a savings-based response and severe shocks would lead to either the use of an assistance/insurance or savings-based response. The choice of a savings-based strategy was found to be approximately 26% more likely in the sample from Langa, relative to the households in Diepsloot, while experiencing an additional shock in the year prior to the survey reduced the probability of using this approach by approximately 11%. There was also evidence that households who were relatively larger, or where the household head was relatively older, were less likely to rely on savings. In addition, the results showed that an increase in the number of events experienced in the year before the survey decreased the probability of a household using either of the two forms of coping strategy considered.

Turning finally to the results on the impact of shocks. In the case of mild shocks, using no strategy (a category which includes the use of regular income) or using a savings-based strategy effectively protects households from the negative impact of such shocks. Both responses result in no significant reduction in food consumption. Furthermore, the impact of a mild shock, under a savings-based strategy is not statistically different from the impact of such a shock under no coping strategy. There is, however, evidence that using an assistance/insurance-based strategy to respond to a mild shock, results in a significant decrease in food consumption in at least one of the periods after the shock has occurred. This may be explained by a need to repay loans (or reciprocate gifts), used to offset the original impact of the shock.

Where households are exposed to moderate or severe shocks, the use of no coping strategy results in a reduction of 47.5% in food consumption (per adult-equivalent) during the week the shock occurs. In addition to this, household food consumption is 30% lower in the week after the event, and 27.5% lower in the week prior to the shock taking place. In contrast, the use of either a savings-based or assistance/insurance-based strategy effectively protects households from the negative impacts of moderate/severe shocks. Furthermore, the difference in the impact of such a shock, between these two strategies, is never statistically different from zero.

This pattern of results is broadly robust to the inclusion of additional lags and leads within the estimations, and also the omission of more predictable shocks (weddings, births and initiations). However, the benefits of either of the two responses, relative to no coping strategy, is less clear in the latter case, suggesting that much of the gain from these strategies comes as a result of being able to formulate a response to foreseeable events *ex-ante*.

Although the random sample selected should be a suitable representation of the households in the two townships, there is no reason to believe that the 107 shocks experienced during the course of the survey are a perfect reflection of the average experiences of households such as these. An important limiting factor of this analysis, therefore, is the small number of instances from which these effects can be measured (shocks, by definition, are rare events). For example, estimating the impact of a mild shock, given a savings-based strategy, requires considering the outcomes of only five events fitting this description (see Appendix 3.2). Therefore, in spite of the representative nature of the households in the sample, the paucity of information on specific magnitude/coping-strategy combinations would suggest that the results presented in this paper may not necessarily be representative of all possible samples drawn from the two townships. In acknowledgement of this, these results should be viewed as an analysis of both the choice of coping strategy, and the subsequent variation in the impact of the shocks, for the specific shocks experienced by the sample. The extent to which these results represent general findings for the two townships is entirely dependent on how representative these shocks are to the events experienced by this population as a whole.

The results of this study give a good indication that those sample households who are able to formulate a specific response to income or expenditure shocks are sufficiently capable of mitigating the short-run, negative effects of even the most costly shocks. However, this analysis also highlights the substantial negative impacts, where no specific coping strategy can be put into effect. Based on the sample of households and shocks considered, the households which are at the highest risk of the potential negative consequences of reducing food consumption are those who both lack the capacity to accumulate sufficient savings to protect consumption *and* are unable to utilise external assistance or insurance mechanisms. This paper has presented some indications of which factors may contribute to households finding themselves in this position. However, further research, on a larger sample of households (and shocks) is required to make definitive claims regarding which groups are most at risk.

References

Alderman, H. (1996). Saving and economic shocks in rural Pakistan. *Journal of Development economics*, 51(2):343–365.

Alderman, H. and Paxson, C. H. (1994). Do the poor insure? a synthesis of the literature on risk and consumption in developing countries. In *Economics in a changing world*, pages 48–78. Springer.

Anand, S. and Harris, C. J. (1994). Choosing a welfare indicator. *The American Economic Review*, 84(2):226–231.

Bhorat, H. and Kanbur, S. R. (2006). *Poverty and policy in post-apartheid South Africa*. HSRC press.

Collins, D., Morduch, J., Rutherford, S., and Ruthven, O. (2010). *Portfolios of the poor: how the world's poor live on $2 a day*. Princeton University Press.

Collins, D. L. and Leibbrandt, M. (2007). The financial impact of HIV/AIDS on poor households in South Africa.

Deaton, A. and Zaidi, S. (2002). *Guidelines for constructing consumption aggregates for welfare analysis*, volume 135. World Bank Publications.

Deaton, A. S. (1992). Saving and income smoothing in Côte d'Ivoire. *Journal of African Economies*, 1(1):1–24.

Fafchamps, M. (1999). Risk sharing and quasi-credit. *Journal of International Trade & Economic Development*, 8(3):257–278.

Heltberg, R. and Lund, N. (2009). Shocks, coping, and outcomes for Pakistan's poor: Health risks predominate. *The Journal of Development Studies*, 45(6):889–910.

Heltberg, R., Oviedo, A. M., and Talukdar, F. (2014). What are the sources of risk and how do people cope? insights from household surveys in 16 countries. *Background paper for the World Bank*.

Jalan, J. and Ravallion, M. (1999). Are the poor less well insured? evidence on vulnerability to income risk in rural China. *Journal of development economics*, 58(1):61–81.

Kazianga, H. and Udry, C. (2006). Consumption smoothing? livestock, insurance and drought in rural Burkina Faso. *Journal of Development economics*, 79(2):413–446.

Mace, B. J. (1991). Full insurance in the presence of aggregate uncertainty. *Journal of Political Economy*, 99(5):928–956.

Mahajan, S. (2014). *Economics of South African townships: special focus on Diepsloot*. The World Bank.

McKay, A. D. and Round, J. I. (1990). *Income and expenditure in a system of household accounts. Concepts and estimation.* World Bank.

Morduch, J. (1991). *Consumption Smoothing Across Space: Tests for Village-Level Responses to Risk.* Harvard University.

Okurut, F. N. (2006). Access to credit by the poor in South Africa: evidence from household survey data 1995 and 2000. *Stellenbosch: University of Stellenbosch*.

Paxson, C. H. (1992). Using weather variability to estimate the response of savings to transitory income in Thailand. *The American Economic Review*, pages 15–33.

Pillay, S., Barolsky, V., Naidoo, V., Mohlakoana, N., and Hadland, A. (2008). Citizenship, violence and xenophobia in South Africa: Perceptions from South African communities.

Rosenzweig, M. R. and Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India. *Journal of political economy*, 101(2):223–244.

Simanowitz, A. and Nkuna, B. (1998). Participatory wealth ranking operational manual. *Small Enterprise Foundation*.

Skoufias, E. and Quisumbing, A. R. (2005). Consumption insurance and vulnerability to poverty: A synthesis of the evidence from Bangladesh, Ethiopia, Mali, Mexico and Russia. *The European journal of development research*, 17(1):24–58.

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, 62:539–539.

Toye, J. (2007). Poverty reduction. *Development in practice*, 17(4-5):505–510.

Udry, C. (1994). Risk and insurance in a rural credit market:An empirical investigation in northern Nigeria. *The 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.

Wagstaff, A. and Lindelow, M. (2010). *Are health shocks different? Evidence from a multi-shock survey in Laos*. The World Bank.

Yilma, Z., Mebratie, A., Sparrow, R., Abebaw, D., Dekker, M., Alemu, G., and Bedi, A. S. (2014). Coping with shocks in rural Ethiopia. *Journal of Development Studies*, 50(7):1009–1024.

Appendices

Appendix 3.1

On two separate occasions during the survey, a household experienced more than one shock on the same day. For example, Birth *and* Death of a household member or other family member/Unveiling of tombstone. Rather than treat these as separate shocks, for the purpose of this study, they are treated as one shock, with the combined cost used to calculate the shock magnitude.

Table 3.A1: Frequencies of Specific Shocks, by Magnitude

Shock	Total	Mild	Moderate	Severe
Household health shocks				
Death of household member/Unveiling of tombstone	12	3	6	3
Injury/illness keeping household member from normal activities	15	8	5	2
Birth	4	2	2	0
Birth *and* Death of household member/Unveiling of tombstone	1	1	0	0
Indirect health shocks				
Death of family living in another place/Unveiling of tombstone	41	14	14	13
Economic shocks				
Loss of regular job of a household member	3	2	1	0
Cut-off or decrease to government grants (not due to death)	3	1	2	0
Loss of crop or livestock	0	0	0	0
Failure of business or bankruptcy of business	1	0	1	0
Did not receive regular remittance from someone outside home	0	0	0	0
Crime shocks				
Theft of household property	6	1	3	2
Victim of violent crime	7	3	2	2
Family shocks				
Abandonment or divorce	2	2	0	0
Fire or destruction of household property	3	0	1	2
Wedding	5	0	3	2
Initiation	2	0	1	1
Repossession of home or physical assets	1	1	0	0
Fire or destruction of property *and* Theft of property	1	0	0	1
Total	107	38	41	28

Mild = Cost $<$ 0.1*Income, Moderate = 0.1*Income $\leq$ Cost $\leq$ 0.5*Income,
Severe = 0.5*Income $<$ Cost

Appendix 3.2

The strategies were initially separated into six groups. Although households were provided with the option of listing asset sales or disposal as a coping strategy, no household recorded using this approach.

Table 3.A2: Frequencies of Specific Coping Strategies, by Shock Magnitude

		Magnitude of shock		
Copping strategy	Total	Mild	Moderate	Severe
Savings-based strategies				
Use savings	28	5	15	8
Asset-based strategies				
Sell assets	0	0	0	0
Dispose of assets (for example, slaughter livestock)	0	0	0	0
Assistance/Insurance-based strategies				
Borrow money	13	3	4	6
Receive gift/donation of money	12	4	3	5
Receive other help such as time/labour	1	0	0	1
Borrow goods other than money	2	0	0	2
Receive gift/donation of goods other than money	1	0	1	0
Insurance payment (cash)	6	1	3	2
Insurance payment (other goods)	0	0	0	0
No coping strategy				
Go without meat	2	1	1	0
Go without other food	6	4	1	1
Go without school uniforms	0	0	0	0
Take children out of school	0	0	0	0
Didn't affect them (no strategy used)	15	12	2	1
Take from income (enough left over to meet expenses)	21	8	11	2
Total	107	38	41	28

Mild = Cost $<$ 0.1*Income, Moderate = 0.1*Income $\leq$ Cost $\leq$ 0.5*Income,
Severe = 0.5*Income $<$ Cost

Appendix 3.3

The following results were obtained by estimating separate binary outcome (probit) models, as an alternative approach to assessing the likelihood of different shock responses.

Table 3.A3: Alternative Estimations of Coping Strategy Choice

Probit	(1)	(2)	(3)
	dy/dx	dy/dx	dy/dx
Outcome: Coping strategy =	Savings-based	Assistance/Insurance	No strategy
Moderate shock	0.201**	0.074	−0.243***
	(0.090)	(0.119)	(0.078)
Severe shock	0.142	0.318***	−0.496***
	(0.097)	(0.108)	(0.077)
Poor household	0.097	−0.072	−0.056
	(0.091)	(0.128)	(0.125)
Household size ($t-2$)	−0.066**	0.035	0.009
	(0.026)	(0.023)	(0.021)
Head female	−0.065	0.119	−0.064
	(0.083)	(0.139)	(0.125)
Head age	−0.007*	0.005	0.002
	(0.004)	(0.009)	(0.007)
Head junior school	−0.051	0.070	−0.058
	(0.077)	(0.115)	(0.111)
Disabled/chronic ill	0.126	−0.036	−0.156*
	(0.083)	(0.110)	(0.091)
Shocks in the previous year	−0.089***	0.014	0.080**
	(0.028)	(0.045)	(0.034)
Bank accounts	0.064*	0.032	−0.103**
	(0.037)	(0.051)	(0.049)
In burial society	0.201**	−0.132	−0.107
	(0.096)	(0.099)	(0.106)
Langa township	0.240***	0.040	−0.255***
	(0.071)	(0.102)	(0.077)
Quarter dummies	Yes	Yes	Yes
Observations (number of shocks)	100	100	100
Log-likelihood	-34.464	-50.414	-37.516
McFadden's R^2	0.371	0.152	0.381

Household clustered standard errors in parentheses   * p<0.1, ** p<0.05, *** p<0.01

Appendix 3.4

Table 3.A4: The Impact of Shocks on Food Consumption - Full Results

Dependent variable: Log food consumption	(1)	(2)		(2) cont.
S($t+2$)	−0.086	−0.119	S * ModerateSevere * CS (sav) ($t+2$)	0.016
	(0.076)	(0.102)		(0.255)
S($t+1$)	−0.054	−0.065	S * ModerateSevere * CS (sav) ($t+1$)	0.297
	(0.050)	(0.084)		(0.188)
S(t)	−0.008	0.022	S * ModerateSevere * CS (sav) (t)	0.493**
	(0.051)	(0.081)		(0.247)
S($t-1$)	0.011	0.058	S * ModerateSevere * CS (sav) ($t-1$)	0.351
	(0.061)	(0.089)		(0.273)
S($t-2$)	−0.017	0.063	S * ModerateSevere * CS (sav) ($t-2$)	0.300
	(0.054)	(0.092)		(0.247)
S($t-3$)	−0.093*	−0.050	S * ModerateSevere * CS (sav) ($t-3$)	0.047
	(0.051)	(0.062)		(0.271)
S($t-4$)	−0.044	0.013	S * ModerateSevere * CS (sav) ($t-4$)	−0.143
	(0.051)	(0.081)		(0.183)
S * ModerateSevere ($t+2$)	0.057	0.057	S * CS (ass) ($t+2$)	−0.091
	(0.093)	(0.184)		(0.232)
S * ModerateSevere ($t+1$)	−0.001	−0.209	S * CS (ass) ($t+1$)	−0.055
	(0.067)	(0.128)		(0.178)
S * ModerateSevere (t)	−0.104	−0.496***	S * CS (ass) (t)	−0.190
	(0.072)	(0.181)		(0.162)
S * ModerateSevere ($t-1$)	−0.099	−0.358**	S * CS (ass) ($t-1$)	−0.188
	(0.076)	(0.142)		(0.173)
S * ModerateSevere ($t-2$)	−0.029	−0.215	S * CS (ass) ($t-2$)	−0.219*
	(0.072)	(0.138)		(0.120)
S * ModerateSevere ($t-3$)	0.026	0.014	S * CS (ass) ($t-3$)	−0.157
	(0.076)	(0.115)		(0.174)
S * ModerateSevere ($t-4$)	0.039	0.052	S * CS (ass) ($t-4$)	−0.237
	(0.080)	(0.118)		(0.177)
S * CS (sav) ($t+2$)		0.139	S * ModerateSevere * CS (ass) ($t+2$)	0.082
		(0.176)		(0.295)
S * CS (sav) ($t+1$)		0.039	S * ModerateSevere * CS (ass) ($t+1$)	0.265
		(0.148)		(0.219)
S * CS (sav) (t)		−0.071	S * ModerateSevere * CS (ass) (t)	0.662***
		(0.161)		(0.243)
S * CS (sav) ($t-1$)		−0.069	S * ModerateSevere * CS (ass) ($t-1$)	0.454**
		(0.225)		(0.227)
S * CS (sav) ($t-2$)		−0.187	S * ModerateSevere * CS (ass) ($t-2$)	0.337*
		(0.183)		(0.184)
S * CS (sav) ($t-3$)		−0.143	S * ModerateSevere * CS (ass) ($t-3$)	0.165
		(0.215)		(0.227)
S * CS (sav) ($t-4$)		0.036	S * ModerateSevere * CS (ass) ($t-4$)	0.187
		(0.124)		(0.228)
Household size ($t-2$)	−0.013			−0.009
	(0.069)			(0.068)
log Regular income ($t-2$)	0.031**			0.032**
	(0.014)			(0.013)
log Regular income ($t-3$)	0.013			0.012
	(0.011)			(0.011)
HH Fixed Effects	Yes			Yes
Township trends	Yes			Yes
Observations	3566			3566
Households	85			85
R^2	0.114			0.125

Household clustered standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

Conclusion

The three chapters presented in this thesis have approached the topic of shocks and coping strategies in very different ways. In the study of Ethiopia, the analysis focused on the potential short-run and long-run consequences of the choices made to limit the impacts of an income shock. Using a panel of 2572 rural households, evidence was found of two distinct patterns of response to the onset of a period of severe drought. The results indicated that households with pre-shock cattle holdings of three or more animals effectively used these assets as a buffer against the drought-induced fall in income. In contrast, households with smaller herds elected to preserve their current herd size, at the expense of reduced consumption. In an extension to the main findings, further heterogeneity was uncovered within the small herd group, suggesting that a need to retain a minimum stock of two cattle was motivated by a requirement to use these animals in the ploughing and preparation of land for the following season. These results were shown to be consistent with the theoretical predictions of shock-response behaviour in the vicinity of a poverty trap threshold.

Chapter 2 considered the impact of a prolonged conflict shock on the health status of children. The study found evidence of irreversible health deficits amongst young children who were exposed to the LRA insurgency in northern Uganda (1987-2007). The causal effect of the conflict was found to be a 0.65 standard deviation fall in height-for-age z-scores amongst those exposed for a period of more than six months. In contrast, the health impacts of shorter periods of exposure were found to be relatively minimal. This suggests that, even in situations where many of the conventional coping strategies used by households are unavailable, the most detrimental impacts of shocks can be avoided, if the duration of exposure can be minimised.

The final chapter estimated the short-run effects of income and expenditure shocks on household food consumption, per adult equivalent, conditional on the method of response employed. Utilising high-frequency panel data on households in two urban and peri-urban South African townships, the analysis found that responding to the most costly shocks by either dis-saving or seeking external assistance was an effective means of smoothing vital consumption. In contrast, where neither strategy was used, these shocks resulted in a 47% reduction in food consumption (on average) in the week the shock occurred and additional shortfalls in the weeks immediately following and preceding the shock.

In summary, this thesis has served to highlight the stark choices made by some groups when faced with the need to respond to unexpected events. In the case of the rural Ethiopian households, this represented the choice to either reduce consumption today or potentially undermine productivity in the future. For those in the two South African townships, it was a lack of any effective coping strategy which identified those most at risk of experiencing the worst outcomes. This study has also drawn attention to the need for a swift response to large-scale, covariate shocks (such as armed conflict) if these events are to be prevented from impacting negatively on the most vulnerable groups in society.

In much of the world outside of Sub-Sahara Africa, we are able to rely on a wide array of financial risk-management instruments to limit the negative consequences of shocks, while benefiting from a relatively high level of social protection. In developing countries, however, not only are individuals more exposed to risk, but the potential coping strategies available to them are far fewer and, in the worst cases, could potentially trap them in a state of persistent poverty. Where groups of individuals or households are unable to effectively respond to widespread shocks (such as drought or conflict), this could result in the economic growth of entire regions being hampered for many years to come. From a policy perspective, therefore, the provision of more effective coping strategies for vulnerable groups should be a priority for any government seeking to promote economic growth at the national level. This may come in the form of a drive for improved insurance instruments and formal savings provision, or better targeting and availability of social protection and safety nets. In this respect, much has already been done in the three countries considered here. However, as the findings presented in this thesis demonstrate, there still remains much work left to do.