The development of vulnerability functions relating household poverty outcomes to crop failures in Ethiopia with the prospect of developing a probabilistic catastrophe risk model

Suggested short title: CAT risk models and poverty analysis

ABSTRACT (150 WORDS)

We analyse the potential to combine catastrophe (CAT) risk modelling with economic analysis of vulnerability to poverty using the example of drought hazard impacts on the welfare of rural households in Ethiopia. The aim is to determine the potential for applying a derived set of damage (vulnerability) functions based on realized shocks and household expenditure/consumption outcomes, onto a forward-looking view of drought risk. We outline the CAT risk modelling framework and the role of the vulnerability module. We present results of a regression model estimating ex-post drought impacts on consumption for heterogeneous household types. We assess the generalizability of the derived functions to infer applicability to a CAT risk modelling framework. We stress-test the model using statistical models of resampling to establish external validity: whether the relationships established in the dataset can be used for forecasting. We conclude with caution that a full CAT risk model could be applied.

Keywords: Catastrophe risk models, disaster risk financing, poverty, vulnerability, drought, Ethiopia

1. Introduction

Probabilistic catastrophe risk models (CAT risk models), used extensively in the international insurance and reinsurance markets, develop a view of risk beyond the historical occurrence of catastrophes. CAT risk models generate thousands of synthetic stochastic events whose characteristics, evolution and pathways are calibrated based on historical event occurrence and a physical knowledge of the potential of the system that generates them. This framework is powerful as it allows for changes in exposed population and assets over time, and considers an extensive range of possible event scenarios well beyond the historical record. This is of particular value when evaluating low recurrence frequency catastrophe events, which by nature have a sparse historical record.

Whilst by now a large body of evidence using historical data has been amassed to show the actual and potential effects of natural disasters on poverty (Hallegatte et al, 2017), CAT risk models have yet to be applied to forecasting poverty outcomes due to stochastic events at the household level.[1] If CAT risk models could provide perspectives on the potential future relationship between the local severity of hazards arising from natural catastrophes and indicators of welfare, practitioners would have tools to assess the impacts of shocks under a forward-looking view of potential catastrophe occurrence, with a view to providing early assistance or insurance.[2] The principal challenge is the development of

[1] The potential to use probabilistic catastrophe risk models outside of the context of insurance, has been recognized in recent years. This has resulted in developments such as the Pacific Risk Information System[1] and the CAPRA[1] Program which both apply the probabilistic catastrophe risk modelling framework developed by the insurance markets for disaster risk management. A similar framework has also been applied to estimate food security needs through the Africa RiskView platform. See http://pcrafi.sopac.org/about/;
http://www.ecapra.org/;http://www.rockefellerfoundation.org/uploads/files/fa08d48b-08ef-4fc7-8991-4872f6e929b0-africa.pdf

[2] Muir-Wood, 2014; Anttila -Hughes and Sharma, 2014-hereafter AHS, 2014.

general form relationships between hazard occurrence and indicators of welfare.[3] Such relationships would comprise the 'vulnerability module' in a CAT risk modelling framework.

Vulnerability to poverty in the economics literature is most often conceptualized as a high probability that at some future point, household consumption of essential goods will fall below the poverty threshold.[4] The main empirical constraint to forward looking economic analysis of vulnerability has been an imperfect ability to model future states of the world. The few studies that are nationally representative rely on cross-sectional distribution of wellbeing and shocks (and their correlation) to model the probability of any given household falling below the poverty line, though a number of microeconomic studies have begun to link to specific stochastic events such as hurricanes, droughts or floods *ex-post*.[5]

The vulnerability module within a catastrophe risk model contains damage functions[6] that represent a mean response to a given hazard intensity, with response typically given as a damage ratio (the level of damage expressed as a percentage of total potential damage[7]). The relationships are continuous functions, increasing monotonically with hazard intensity – their shape denoting the form of response of the impacted asset to the shock. Some models account for the uncertainty in the damage ratio, by describing a probability distribution around the mean damage ratio for a given level of hazard intensity.

---

[3] AHS, 2014.
[4] Hill and Porter, 2014. Poverty, for the purpose of policy analysis, is defined as consumption (or expenditure) that falls below some pre-specified level, considered the minimum for an acceptable standard of living.
[5] AHS, 2014.
[6] Also referred to as fragility or vulnerability curves or functions
[7] For buildings, the damage ratio is the cost of repair as a proportion of the total replacement value. Jain, V, 2010.

In considering a microeconomic model as a potential application for the vulnerability module we build on a now large body of economic studies that estimate the ex-post impact of shocks on welfare (Hallegatte et al, 2017). We ask whether it is possible to reduce these complex relationships into functions that can be used to model poverty outcomes under a forward-looking view of catastrophe occurrence.

In order to evaluate the feasibility of this approach to the vulnerability module, we develop:

1) **A regression model to derive quantitative relationships between a selected drought hazard measure and household poverty outcome** for rural households in Ethiopia. This is a survey-weighted regression model combining historical household data with historical data on drought hazard, which effectively constitutes our 'vulnerability module';
2) **Tests of the derived 'vulnerability module' to evaluate its robustness**, and therefore the validity of its future application onto a forward-looking probabilistic view of drought occurrence generated from a catastrophe risk modelling framework. This evaluation is conducted through the application of Statistical Learning Methods.

The regression uses a "reduced form" approach. Essentially, this evaluates impacts of the shock on household consumption, after households have used every strategy available to them to mitigate its effect (diversification, asset sales). This can be contrasted with "structural approaches" (Chetty, 2008) that could be an alternatively developed. Household welfare outcomes after a catastrophe are influenced by many complex, often interacting, factors beyond the direct damage to household physical assets. Such factors include different adaptive behaviours that lead to different outcomes for a given level of physical asset damage. If such strategies include opting for low-risk, low-return activities , then risk also carries an ex-ante cost, which cannot be measured with our approach, and is likely quite high (Elbers et al, 2007).

Our evaluation of validity of the resulting damage functions will be centred on concepts of internal and external validity.[8] The framework for damage function development described above, allows for the application of the derived functions *out of the context in which they were derived*. This is a fundamental feature of catastrophe risk models, as their purpose is to provide a forward-looking view of risk beyond the historical events used in their development, and is the key difference with economic approaches. This out-of-context application is more challenging when considering the impact of hazards on poverty, rather than physical damage, outcomes.

2. Methodology for deriving vulnerability relationships for the impact of drought on poverty in Ethiopia

Within catastrophe risk modelling there are two principal methods used to derive damage functions; empirical and analytical derivation. We follow a purely empirical approach, developing a survey-weighted regression model that combines household survey data with historical data on drought hazard. The analogy for drought damage to crops, for

[8] AHS, 2014.

example, would be the use of mechanistic agro-meteorological models based on a process approach versus statistical methods using historical crop loss and drought hazard data.

The model also attempts to separate out observed relationships for different vulnerability-determining characteristics, analogous to how catastrophe risk models incorporate distinct damage functions for different classes of exposed asset; for example, by structural 'class' for buildings or by crop type. Damage functions developed for buildings in one region can be modified for use in another, if sufficient detail around differences in construction types and quality is available.

The measure of drought chosen for this study is an index of crop yield shortfall, taken from the World Food Programme's LEAP (Livelihoods Early Assessment and Protection) software, as this is used in Ethiopian policy analysis.[9] The yield shortfall calculation uses time-variable meteorological recordings combined with data tables on soil and crop characteristics to calculate yield reductions relative to the expected production under non-limiting water conditions. It is available at the Woreda[10] level through creation of a composite index for the relevant crop basket. The methodology for the calculation has been developed over several years (Doorenbos and Kassam, 1979; Frère and Popov, 1979; Frère and Popov, G, 1986; Hoefsloot and Calmanti, 2012; Abraha, M, 2013).

LEAP calculates yield shortfall by combining a model for water balance (the FAO's Water Requirement Satisfaction Index (WRSI)) with a model describing crop yield response to water stress: 100-((1-(1-A/B)*Ky)*100)

[9] Drechsler and Soer, 2016.
[10] *Woreda* is an administrative unit in Ethiopia, equivalent to a county in the United States.

Where: *A* is the Actual Evapotranspiration; *B* is the Total Water Requirement without water stress; and *Ky* is a crop specific factor – “Yield Response Factor” - for growing seasons or stages of growing seasons, derived empirically from actual measured crop yield responses to water under good growing conditions . The WRSI gives the ratio of Actual Evapotranspiration to Total Water Requirement (i.e. A/B in the above) for a season. The Actual Evapotranspiration represents the actual amount of water withdrawn from the soil water reservoir and is calculated indirectly using rainfall data within a model of soil water balance. It requires inputs of time-variable rainfall data and soil and crop-specific data. The Total Water Requirement is calculated as (Potential Evapotranspiration x Crop Coefficient). It requires time-variable climate data inputs on solar radiation (sunshine), air temperature, humidity and wind speed.

In the context of a probabilistic catastrophe risk modelling framework, the key ‘hazard’ input into the process is the rainfall data used in the calculation of Actual Evapotranspiration. High levels of rainfall variability determine the occurrence or otherwise of meteorological drought (translated into crop yield loss), and this is where the large covariate shocks arise that CAT risk models seek to capture.

Using a rainfall-based index confers the advantage of objectivity in the measure, removing challenges such as reporting bias that are present when working with reported crop yield statistics. It is also more plausibly exogenous than a self-reported measure. However, we acknowledge the limitations of working with a modelled estimate rather than a direct measure of yield impact. Also, the LEAP protocol for assessing crop yield loss estimates only meteorological drought arising from rainfall variability. This has been criticized as limiting the extent to which the LEAP crop loss figures capture experience on the ground, as they do not capture the many other factors (such as pests) impacting yields.[11] We also acknowledge that extreme drought is not the only cause of poverty and hunger, which can have its root causes in the failure of entitlements.[12] People can also be pushed into hunger even when rainfall fails only moderately, but if asset prices fall considerably, or food prices increase rapidly.

The limited number of survey years available for the regression poses a number of challenges. One such challenge is the impossibility of capturing impacts arising from more severe seasonal drought on a national or regional scale. For example, in the aftermath of a large drought, impacts on food prices (and other commodities for which demand has increased), can be expected. This can impact household consumption. Changes in the labour market should also be considered, as a large drought event may push a large number of households to turn to alternative employment as a coping mechanism even if they were not directly affected (Noy and Patel, 2011). A large drought may also prompt a political response that influences household consumption in a way not seen in the historical dataset used to derive damage functions. These non-local factors are complex and would be difficult to capture, even with a more comprehensive dataset. However, the fact that the impact on a household is influenced by the scale of a disaster, as well as by its local intensity, is indisputable, and not accounted for in the methodology presented in this paper.

[11] As demonstrated by a 2007 ground-truthing exercise described in: FSCB/WFP Workshop Livelihood Early Assessment and Protection (LEAP): its potential application, benefits and limits, NOTE FOR THE RECORD. 21st January 2008.

[12] Sen, 1981.

It is worth reiterating here that the interest of this paper is not the causal relationship between household poverty and crop yield losses. Rather, the purpose is to establish the validity of relationships derived between a selected poverty indicator and a proxy measure of natural hazard that could be modelled within established probabilistic catastrophe risk modelling frameworks. Other measures of meteorological drought may therefore serve our objectives.

The key questions to be answered in terms of the household regression model are:

- To what extent can impacts statistically associated with a drought hazard measure be attributed to the drought (internal validity/robustness)?
- To what extent can the regression results be applied to predict outcomes from hazard in other contexts (external validity/robustness)?

The regression model specified is: *y=D(h,s)* where *y* our outcome of interest is the log of consumption per adult equivalent, *h* is the community level annual crop loss as predicted by LEAP (defined above), and *s* are other household and community characteristics (including other shocks experienced by the household). We note also that the relationship between *h* and *y* is attenuated by certain household and community characteristics *(s)*, and the regression model therefore seeks to draw out these attenuating impacts separately to increase the external validity/robustness of the model.

The regression model calculates a relationship between consumption and crop loss. Households are separated into categories according to the characteristics known to attenuate the impact of drought on consumption. The relationship in all cases is defined by a selected functional form and coefficient output from the regression model. The base specification is based on initial work[13] that derived a general model of consumption for Ethiopian households using all areas, rural and urban, and focused on the impact of drought, food prices, and other idiosyncratic shocks on ln consumption per adult. We make several modifications due to our focus, mainly concerned with achieving precision on the relationship between drought and consumption for differing household characteristics.

The **dependent variable** is the natural logarithm of consumption per adult equivalent at household level, as used by Hill and Porter (2014). The drought-crop-loss variable or **hazard** is the LEAP estimated crop losses at woreda level (see above for definitions). Household (HH) characteristics included in the model are: HH head gender, age and education, and a dummy for household head not in agriculture; HH assets including cattle, sheep, chickens, land, good roof, toilet; Idiosyncratic shocks including crop-loss, animal illness or death, HH member illness or death, food price shocks; other characteristics including financial capital and household composition. Community characteristics include: agro-climactic zone, region, distance to town, market access.

As noted above, the model also seeks to capture what in risk-modelling are termed attenuating factors, in econometrics as heterogeneous impacts, through interaction terms included in the model (e.g. LEAP*varname). Table 2.1 shows these variables. Interacting variables proxy for characteristics of the household/community that may affect household ability to cope with shock, these include: ability of head to access coping strategies (head

[13] Hill and Porter, 2014.

education, head not in agriculture, HH doesn't own cattle, dependency ratio); other shocks that may already be stretching household ability to cope (crop-damage, livestock shock, illness); access to institutional coping strategies (distance to town/market, access to financial capital, public safety net). Finally, agro-climactic zone is discussed in a separate section below. Statistical learning methods are used to perform model selection for the appropriate specification of the (potentially non-linear) and heterogeneous relationship between drought-crop-loss and consumption.

Internal validity

In the absence of a randomized experiment assigning some households to better community-level crop yields than others, we must assume implicitly that (in some specific sense) the crop loss in a given year is exogenous to (uncorrelated with) *unobservable* household and community characteristics that affect household consumption. If households more susceptible to the impacts of shocks are locating themselves in hazard-prone areas this also limits the extent to which internal robustness/validity can be examined using cross-sectional data (Winsemius et al, 2015), despite our rich controls and county fixed effects. The purpose of the research sidesteps this issue somewhat, as the question of interest is not primarily about causality per se but rather, an externally replicable association.

The question of internal validity may be more weakly stated as being a "stable" observable relationship between drought-crop-loss and consumption over several specifications (and external validity over several datasets), and answerable by looking at whether the point estimates and associated p-values established through the regression work, and statistical learning through the bootstrapping method described later. By giving a level of significance around the relationship between consumption and crop loss (defined by the coefficients output from the regression, and their associated functional form), the derived p-values give us information on the strength of fit of the modelled relationship. The bootstrapping method generates a mean estimate of the relationship and a related standard error (see below for more detailed methodology) which allows us to examine the stability of the coefficients that have been derived.

External validity

External validity is concerned with whether results of any individual piece of analysis are generalizable to other contexts. It is helpful to define three specific sub-topics within overall external validity that are relevant in our context. These are referred to as EV1) out-of-sample shock estimates EV2) time-stability of estimated relationships and EV3) concerns around over-fitting to the training data. We define this in more detail below.[14]

The main threat to external validity (EV1) in the context of specifying the damage function D(.) as outlined above, lies in mis-specifying the function for values of the hazard (shock) that were unavailable for the analysis- in particular, extreme values that occur only very infrequently (e.g. a once-in-a-generation severe drought, typhoon, tsunami etc.).[15]

[14] EV1) is discussed as overlapping within the scope of the internal validity checking.
[15] AHS, 2014.

The initial model estimation combined the years 2004-5 and 2010-11 in order to allow the broadest (artificially) cross-sectional range of values for our shock.[16] In some pre-analysis, we conducted a careful analysis of the range of data (see LEAP document). The issue in the context of our dataset is the limited number of observations available at some of the high levels of crop loss (e.g. greater than 60 percent crop loss).

Figure 2.1 shows the distribution of the crop loss data for 2011 and 2005 combined, and it can be seen that the distribution is left-skewed. Where we identify data paucity that challenges the strength of fit of the modelled relationship at this extreme end of the crop loss spectrum, we need to highlight a caveat regarding the sparse data above 60% crop loss. 2005 is somewhat worse year than 2011, with a higher mean crop loss, and 300 observations above 50%. However in 2011, just under 100 observations lie between 50 and 60%, and none at all are above 60%. The distribution is somewhat different between years, and we therefore examine the difference between the pooled dataset that has a broader distribution, with the use of one round of data as the training data, and another round as the testing data, in order to test the stability of the drought-consumption relationship over different drought distributions.

The second issue is over what length of time relationships established can be considered as valid (EV2). For the purpose of CAT risk modelling, it is helpful to consider relationships to be stable over a five-year horizon. To establish this, we attempt to use a dataset from a different survey year (2012, as opposed to 2005 and 2010-11) as validation data. We do not anticipate the external validity of the derived relationships to be applicable

[16] Hill and Porter, 2014

over very broad timeframes (i.e. 20 years or more), due to structural changes in the economy that lead to changes in lifestyle/behaviour and potential exposure to shocks. We discuss the testing dataset for 2012 in a later section of the paper.

Also related to temporal aspects of shock exposure is the concern of recurrence times and macroeconomic (or second order) effects.[17] An example of recurrence times is that, in one crisis households with livestock may sell something in order to protect their consumption, but if another crisis hits, the impact of the second shock is likely to be higher. We are not likely able to address this latter point with the data available, and so must simply take the ex-post distribution as being standard.

The concept of *ecological validity* is useful in the context of EV3: "a study has ecological validity to the extent that the context in which subjects cast decisions is similar to the context of interest."[18] We restrict ourselves to a fairly narrower tractable question whether the specified model valid across (non-pastoralist) Ethiopia. In one specification, the regression model will also explicitly treat agro-climatic region as an attenuating factor, testing whether the relationship between crop loss and consumption will hold across regions. This is consistent with the idea that the efficacy of adaptive behaviour to weather shocks varies with average climate conditions.[19]

---

[17] AHS, 2014.

[18] Roe and Just, 2009, p1267.

[19] Hsiang and Narita, 2012.

A final note of caution around the estimation results is that there is most likely measurement error in our variables that could potentially lead to attenuation bias in the estimates.

3. Results: Vulnerability relationships

This section discusses the results of the estimated vulnerability relationships. We first present the descriptive statistics, and then the tables of results. The next section discusses the statistical learning methods to validate the regression results.

We use the data collected in the 2005 and 2011 rounds of the nationally representative Household Income and Consumption Expenditure and Welfare Monitoring Surveys (HICES/WMS). The key information recorded in the HICES used to calculate vulnerability is expenditure on food and other items. The WMS records household assets and characteristics as well as a fairly detailed module on self-reported adverse events (referred to as shocks throughout). In both years they were administered by Ethiopia's Central Statistics Agency (CSA).

The advantage of using the HICES-WMS for vulnerability analysis is that they are relatively large, nationally representative, and comparable across years and allow measures of vulnerability to be estimated at the household level. This allows us to look at the relative importance of geographic and household factors in determining vulnerability, and it also allows us to examine how vulnerability varies across certain groups of households. The descriptive statistics for the two datasets are shown in table 3.1.

Our four specifications include the baseline (no interactions). **Model 1** is a parsimonious model including only female-headed household, head schooling and access to the social safety net (PSNP). **Model 2** includes further interaction terms of distance to market and the dependency ratio. **Model 3** includes a full set of interaction terms, including the financial access indicator, and the other shocks that may compromise household ability to cope with drought. Note that the latter set of variables are self-reported, and therefore potentially endogenous, if having experienced a drought in any way affects the response to the questions asked.

Table 3.2 presents the results. We show the coefficient on the LEAP variable divided by 10, to ease interpretation. As the dependent variable is in log form, we can read the coefficients as "a 10% worsening of the LEAP crop loss leads to an X% reduction in consumption per adult". For example, the baseline result shows that a 10% worsening of LEAP is associated with consumption falling on average by 1.5%. We computed bootstrapped standard errors across all of the statistical learning components (see next section) and therefore also present a 95% confidence interval for the point estimate based on bootstrapping results. The LEAP drought variable is significant in all specifications. In the following columns, interaction terms are included, so the point estimate is interpreted as the drought impact for households that are not defined by any of the characteristics included. In column 2, we include female headed, PSNP, and schooling. The coefficient of 2.0% represents the impact of a 10% increase in the LEAP drought on consumption of male headed households with no schooling and no access to the PSNP. Schooling and female headed do not show differential impacts. However PSNP access mitigates the drought impact by 0.5%: therefore households with PSNP access would experience a 1.5% decrease in consumption (rather than 2% with no access). We find similar results for households with cattle and access to finance. However, households that suffered other crop damage experienced a heavier impact of the drought. Note above the potential concern that financial access and shocks such as crop damage are self-reported. Some households have networks and receive transfers, but other coping mechanisms that are used to protect consumption such as selling assets, borrowing, may be detrimental in the longer run (Hill and Porter, 2014). We note also that the targeting of PSNP to the poorest communities may mean that these results are not unbiased.

To explore whether the impact of drought differs across regions we created dummy variables for the four non-pastoral agro-ecological zones of Ethiopia ((EDRI 2009: drought prone highlands, moisture-reliable cereals areas, moisture-reliable enset areas, humid moisture-reliable lowlands), and included these in the final regression model, shown in the final column of table 3.2. The results show that the relationship between drought and crop loss is different across agro-ecological zones.

We estimated the same model and performed statistical testing on a quadratic and cubic form to test for nonlinearity of the relationship between drought and consumption, which is plausible. The full set of results is available on request. Adding these higher powers did not change the coefficients on the interaction terms, so we present in figure 3.1 a graph showing the simulated shape of the curve using the squared and cubic models for ease of interpretation. The cubic model appears to have a second turning point around 70% crop loss- which is around the point at which we lose support for the data in 2011, so we may not have enough values of the data to create a plausible estimate for any further nonlinearity than a squared term.

4. Validation Methodology

The proposed methodology to testing the predictive power of the vulnerability relationships is to use Statistical Learning Methods (re-sampling and cross-validation). Statistical learning is used for Assessing model accuracy; Checking the performance of different functional forms of D(.); Choosing the model with the lowest “test” MSE (the best prediction) instead of just the lowest “training” MSE (the best fit on the currently used historic data*)*.[20]

This validation methodology has not been widely used in economics; one example uses “holdout data” to choose the empirical model that performs best in terms of RMSE.[21] Recently, the World Bank has applied the method of training and testing datasets in the validation of poverty scorecard methodology.[22]

The resampling method works by examining the fit of the model when we apply the results from the regression using a training dataset (where the model is initially fitted) to a testing dataset (a separate dataset or subset of the dataset that has never been used to fit the model).[23] K-fold cross validation randomly allocates all n observations of the data into k-parts (folds or groups) of approximately equal sizes. The first fold is treated as the testing

[20] *James et al 2013*
[21] Todd and Wolpin, 2007.
[22] Diamond et al, 2015
[23] *James et al 2013*

dataset and withheld while the model is fitted to the remaining k-1 folds. The observations in the first fold are then fitted to the estimated model and the mean squared error (MSE) MSE calculated (MSE1). The procedure is repeated, each time using a different fold as the validation dataset. The testing MSE is defined as the average of MSE1, MSE2… MSEk.[24]

$$CV_k = \frac{1}{k}\sum_{i=1}^{k} MSE_i \qquad (5.1)$$

This methodology gives us information on the strength of fit of the models with respect to each other, and in absolute terms, using the testing MSE. In particular we examine the relative improvement in the MSE relative to the complexity of the model. We also use the Akaike Information Criteria (AIC), which penalizes a model for including "unnecessary" variables or complexity, to complement the k-fold cross validation method of model selection.

For the bootstrapping approach, M distinct data sets with n observations are drawn (with replacement), by repeatedly sampling observations from the original data set, the same observation can occur more than once in the bootstrap data set. The model is fitted to each of these M datasets. The coefficient of interest is computed M times to compute the standard error. We then compare the bootstrap estimates of the drought parameter across the candidate models. The model with the least bootstrap variance would be selected.

Note here that either method (bootstrapping or k-fold) can be used to validate the parameters of an estimated model.

---

[24] *James et al 2013*

Outputs from the analysis

Recall that D(.) comprises the hazard ***I*** and household characteristics ***X***. The shortlisted functions of D(.) are candidate models 1-4 outlined above. We proposed that the control variables should be selected on the basis of economic theory/empirical findings/common sense. The interactions should also be selected on that basis, however the number of interactions, and the functional form is decided using the statistical analysis. The next step is to test the performance of the candidate models in predicting consumption out of sample using different vectors of X.

The first round of testing discussed above involved random division into k parts. We also split groups according to characteristics that vary across the dataset, but that we have not explicitly identified as determinant of vulnerability, to test *external validity, or generalizability across ecological contexts*. We use non-random sampling for the k-Fold Cross Validation method, separating the k groups according to geographical area (e.g. Woreda) but controlling for differences expected from changes in agro-climatic conditions. This will tell us if, when we separate out households according to what we have determined to be the key vulnerability characteristics, the relationships hold across (non-pastoral) geographies.

5. Results from the validation analysis

We began with the combined 2005/11 dataset, and conducted q-fold (10 rep) cross-validation on the shortlisted 4 models with a linear specification for the drought variable. To be specific, this meant that for the pooled dataset, we fit the model onto 9/10 of the data (randomly selected), and assess the fit on the one remaining tenth. The exercise is repeated 10 times. As discussed in section 4, the models are assessed based on the average mean squared error (MSE) over all ten repetitions as defined in equation (5.1). The results are shown in table 5.1. There is little difference between the models, with model 4 having the lowest average MSE by a small margin.

We then repeat the exercise, but considering the whole dataset for 2005 (2011) as the training (testing) dataset respectively (table 5.2). In this way one might imagine going back in time to 2005, predicting consumption with that dataset for 2011, and assessing the results. When 2005 (2011) is the training dataset, Model 2 (1) is the best predictor (lowest MSE). For all the models, training on the 2011 dataset predicts the 2005 data slightly more accurately than the reverse, suggesting that the relationship between drought and consumption is fairly homogenous, and stable. The results using the squared models appear very similar to those of the linear model. The cross-validation results are identical to three decimal places.

For the regional validation we non-randomly hold out one region at a time to use as a testing dataset, shown in table 5.3.[25] Overall, the full model with all interactions and a quadratic crop loss function is the preferred specification. We use this model to illustrate

[25] Todd and Wolpin, 2007.

below for a simulation of CAT risk modelling outcomes, using the pooled 2005-2011 model as the training data.

6. Application of the derived vulnerability relationships within a probabilistic catastrophe risk modelling framework

The purpose of this section is to demonstrate how the derived functions would be applied in practice, and to propose future work on the basis of the findings of the paper.

In the absence of a probabilistic hazard model for rainfall variability in Ethiopia at the resolution required, and compatible with the LEAP protocol, we have produced illustrative examples based on the results from the regression model.

Figure 6.1 shows the simulated crop loss with examples of heterogeneous impacts: The baseline (in dark blue) shows crop losses up to 100% and subsequent impact on consumption (e.g. at 80% crop loss, consumption would fall by just under 30%). The most effective mitigation is cattle ownership.

For the simulated impacts on consumption overall, net impact of the drought should be calculated given each household’s characteristics (e.g. households with PSNP and cattle will be impacted even less than those just PSNP, but another idiosyncratic shock would exacerbate impact).

Finally, if the policy interest is ***poverty impacts of drought***, then this should be incorporated into the module. E.g. a 20% average drop in consumption will push households already below the poverty line into deeper poverty, plus other households will fall below the line.

The components for the vulnerability module that would be needed are as follows:

- A model of geo-referenced exposure, as assets or population at risk (exposure module);

- A model of the frequency, severity and location of possible hazard occurrence (hazard module);

- A model of the relationship between the modelled hazard occurrence and the impact on the exposure (vulnerability module).

An illustration is shown in figure 6.2 – which shows simulated headcount poverty for each level of crop loss as predicted by the model. The full vulnerability module would thus combine these impacts with a more fully developed risk model and the final result would allow policymakers to understand the likely poverty burden in future time periods.

Conclusion and extensions

We have explored the possibility of combining a regression-based model of shock impact on consumption with a catastrophe risk model for the purposes of producing a forward-looking instrument for policy. The relationship between consumption and a drought measure comprised of crop losses based on water adequacy has been calculated. This relationship has then been stress tested using statistical validation techniques.

The results show that the impact of drought is significant across all models examined; with the baseline result showing that for every 10% worsening of the LEAP drought variable, consumption falls on average by 1.5%. The results also show an apparent mitigating impact on this relationship from certain community and household characteristics; for example, access to a safety net (PSNP) mitigates the drought impact on consumption by 0.5%. However, whilst the results show consistency, we advise caution in interpretation of this relationship, given that the two years of data available do not show the worst rainfall experienced in Ethiopia.

As an extension a full probabilistic catastrophe risk model could be developed. This would entail development and application of a stochastic model of rainfall that could be applied into the LEAP framework to produce values for actual evapotranspiration in the index calculation. Sensitivity analyses could be applied within the hazard modelling to consider potential outcomes in the longer term under climate change scenarios. For example, increases in the rates of occurrence of extreme rainfall variability could be used to look beyond the near term view. Similarly, projections of population increase and composition change could be applied to the exposure dataset.

References

Abraha, M, (2013). Assessment of Drought Early Warning in Ethiopia, a comparison of WRSI by surface energy balance and water balance. PhD Thesis. University of Twente, The Netherlands.

Anttila-Hughes and Hsiang (2011) Destruction, disinvestment, and death: economic and human losses following environmental disaster. Working Paper.

Anttila -Hughes and Sharma (2014) Linking Risk Models to Microeconomic Indicators, World Bank publications, 2014.

Chetty, R (2008). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. No. w14399. National Bureau of Economic Research,

D’Ayala D., Meslem A., Vamvastikos D., Porter, K., Rossetto T., Crowley H., Silva V. (2014) Guidelines for Analytical Vulnerability Assessment of low-mid-rise Buildings – Methodology, Vulnerability Global Component project.

Dercon, S. "Risk, crop choice, and savings: Evidence from Tanzania." Economic development and cultural change 44.3 (1996): 485-513.

Diamond, P, et al (2015) Estimating Poverty Rates in Target Populations: An Assessment of the Simple Poverty Scorecard and Alternative Approaches. World Bank Group.

Doorenbos, J. & Kassam, A, (1979). Yield response to water. FAO Irrigation and Drainage Paper No. 33. Rome, FAO

Drechsler, M. B. S., & Soer, W. (2016). Early warning, early action: the use of predictive tools in drought response through Ethiopia's productive safety net programme. World Bank Policy Research Working Paper, (7716).

Elbers, C., J. W. Gunning, and B. Kinsey. 2007. “Growth and Risk: Methodology and Micro Evidence.” World Bank Economic Review 21: 1–20.

FEMA (2002). Rapid visual screening of buildings for potential seismic hazards. FEMA, 2002. Retrieved from: http://www.fema.gov/media-library-data/20130726-1646-20490-8071/fema_154.pdf.

Fermont, A, and Benson, T, (2011). Estimating Yield of Food Crops Grown by Smallholder Farmers: A Review in the Uganda Context. International Food Policy Research Institute, IFPRI Discussion Paper 01097, June 2011

Hallegatte, S., A. Vogt-Schilb, M. Bangalore, and J. Rozenberg. 2017. Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters. Washington, DC: World Bank.

Hill, Ruth and C. Porter (2014). The Changing Nature of Vulnerability in Ethiopia. In Ethiopia Poverty Assessment 2014. World Bank Group. 2015. Washington, DC. https://openknowledge.worldbank.org/handle/10986/21323

Jain, V (2010). Anatomy of a Damage Function, AIR Worldwide online publications. Retrieved from: www.air-worldwide.com, February 2015.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning. New York: Springer, 2013.

Leong, Airriess, Li, Chen, and Keith (2007) Resilient History and the Rebuilding of a Community: The Vietnamese American Community in New Orleans East, Journal of American History, 2007.

Muir-Wood (2014) The use of probabilistic catastrophe risk models to assess the impact of disasters on poverty, World Bank publications, 2014.

Noy, I., and P. Patel. 2014. "Floods and Spillovers: Households after the 2011 Great Flood in Thailand." Working Paper Series No. 3609, School of Economics and Finance, Victoria University of Wellington.

Rossetto T., Ioannou I., Grant D.N., Maqsood T. (2014) Guidelines for empirical vulnerability assessment, GEM Technical Report 2014-X, GEM Foundation, Pavia

Smith M, and Steduto, P, Yield response to water: the original FAO water production function Food and Agricultural Organisation Irrigation and Drainage Paper No. 66. Rome, FAO

Todd, P E., and K. I. Wolpin. "The production of cognitive achievement in children: Home, school, and racial test score gaps." *Journal of Human capital* 1.1 (2007): 91-136.

Verma, V, Marchant, T, and Scott, C, (1988). Evaluation of Crop-Cut Methods and Farmer Reports for Estimating Crop Production: Results of a Methodological Study in Five African Countries. London, Longacre Agricultural Development Centre Ltd

Winsemius, H., B. Jongman, T. Veldkamp, S. Hallegatte, M. Bangalore, and P. J. Ward. (2015). "Disaster Risk, Climate Change, and Poverty: Assessing the Global Exposure of Poor People to Floods and Droughts." Policy Research Working Paper 7480, World Bank, Washington, DC.

World Bank (2014) Ethiopia Poverty Assessment, World Bank Working Paper.

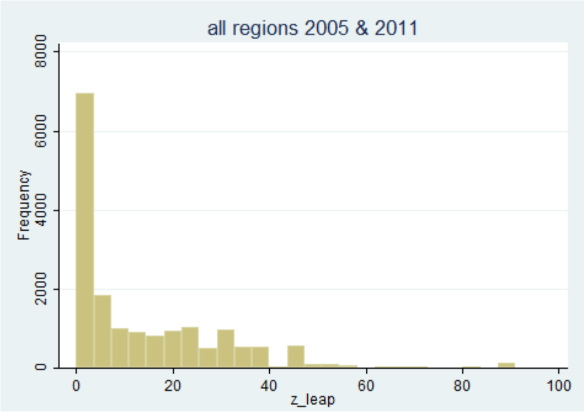
World Food Programme (2008). FSCB/WFP Workshop Livelihood Early Assessment and Protection (LEAP): its potential application, benefits and limits, NOTE FOR THE RECORD. 21st January 2008.

TABLES AND FIGURES FOR TEXT

Table 2.1: Household characteristics and proxy variables

Household characteristics	Interacting variables
Ability of head to access coping strategies	Head education level, sector of occupation, gender
Household composition that allows labour response	Dependency ratio, ratio of able-bodied
Household assets that mitigate shock	Cattle, other livestock, jewellery
Other shocks that compromise ability to mitigate shock	Illness, livestock disease, crop damage from pests
Access to institutional coping strategies	Distance to market, access to financial products (insurance, credit), public safety net access.

Figure 2.1 Frequency and 5% bin of Drought-Crop-Loss Data (LEAP)[26]



[26] NB: Distributions use 2004 LEAP (Meher and Belg) for 2005 survey,
2010 LEAP (Meher and Belg) for 2011 survey (interviews Jan - June 2011)
2009 Meher and 2010 Belg LEAP for 2011 survey (interviews July-Dec 2010) .

Table 3.1: Descriptive statistics

Variable	2005	2011
Ln adult monthly	7.27	7.28
	(0.50	(0.50
LEAP crop loss	16.3	11.5
	(18.53	(13.25
Female head	0.23	0.23
	(0.42	(0.42
Head not agriculture	0.16	0.14
	(0.37	(0.35
Head school	0.25	0.30
	(0.43	(0.46
HH has many plots	0.57	0.53
	(0.49	(0.50
Cattle	0.66	0.67
	(0.47	(0.47
Financial access	0.25	0.50
	(0.43	(0.50
Distance to town	326.8	378.6
	(229.93	(285.74
Dependency ratio	0.49	0.50
	(0.24	(0.23
Death shock	0.08	0.02
	(0.27	(0.13
Illness shock	0.23	0.08
	(0.42	(0.27
Cropdamage shock	0.10	0.03
	(0.30	(0.17
Livestock shock	0.09	0.05
	(0.29	(0.21
Jobloss shock	0.01	0.00
	(0.09	(0.04
Price shock	0.02	0.18
	(0.14	(0.38
PSNP beneficiary	0.00	0.15
	(0.00	(0.36
Highlands-drought	0.39	0.37
	(0.49	(0.48
Highlands-reliable	0.38	0.34
	(0.48	(0.47
Lowlands-reliable	0.03	0.11
	(0.18	(0.3)
Lowlands-enset	0.19	0.18
	(0.40	(0.39
Number obs	8431	8807

Table 3.2 Main regression results

	Baseline	Specparsi	Spec3
(Intercept)	7.794***	7.800***	7.815***
	(0.018)	(0.018)	(0.020)
Drought shock	−0.015***	−0.020***	−0.058***
	(0.002)	(0.003)	(0.007)
boot.se	0.0025	0.0028	0.0078
boot.ci	(-0.020, -0.0104)	(-0.0242, -0.0130)	(-0.0728, -0.0425)
Drought* Head school		0.000	0.000
		(0.001)	(0.001)
Drought*Femalehead		0.000	
		(0.000)	
Drought* PSNP		0.005***	0.005***
		(0.001)	(0.001)
Drought*non-agri hh			−0.000
			(0.001)
Drought*dist. market			0.000***
			(0.000)
Drought* dependency			0.001
			(0.001)
Drought* anycattle			0.018***
			(0.004)

Drought* Fin. access			0.001**
			(0.000)
Drought* Illness			0.000
			(0.001)
Drought* Cropdamage			−0.001*
			(0.001)
Drought* livestock shock			−0.000
			(0.001)
D* highlands reliable			0.005***
			(0.001)
D* lowlands reliable			−0.004***
			(0.001)
D* lowlands enset			0.004***
			(0.001)
R-squared	0.245	0.246	0.251
Adj. R-squared	0.244	0.245	0.249
AIC	20047.21	20022.04	19937.65
BIC	20271.92	20270.00	20263.10
Num. obs.	17134	17134	17134
boot.cv.cropleap	0.249	0.248	0.248
boot.cv	0.170	0.171	0.169
RMSEtest05	0.0232	0.0281	0.0291
RMSEtest11	0.0438	0.0431	0.0490

RMSEtestHRE	0.0514	0.0546	0.0602
RMSEtestHDR	0.0302	0.0297	0.0544
RMSEtestLOE	0.0042	0.0060	0.0055
RMSEtestLOR	0.0602	0.0080	0.0013

Notes: s.e.=standard error cv=cross validation. R2=r-squared; AIC=Aikike information Criterion; BIC= ; boot=bootstrap; RMSE=root mean squared error. Sample size = 17238. Full results in appendix.

Figure 3.1

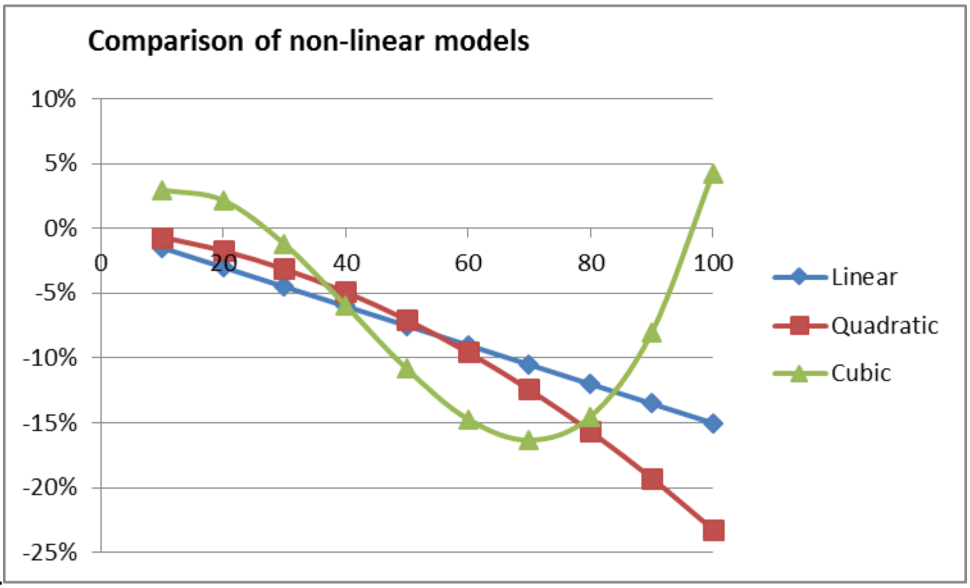


Table 5.1: Testing the models with 2005 and 2011 as training and "holdout" data

	Baseline	Model 1	Model 2	Model 3	Model 4
Original R^2	0.245	0.246	0.247	0.247	0.251
Original Adj. R^2	0.244	0.245	0.245	0.246	0.249
Original AIC	20047.21	20022.04	20072.26	20065.60	20001.73
Original BIC	20271.92	20270.00	20335.72	20360.06	20319.43
Num. obs.	17134	17134	17134	17134	17134
boot.cv	0.170	0.171	0.170	0.170	**0.169**
RMSEtest05	**0.0232**	0.0281	0.0338	0.0272	0.0291
RMSEtest11	0.0438	0.0431	**0.0427**	0.0429	0.0490

Notes: cv=cross-validation, R2=r-squared; AIC=Aikike information Criterion; BIC= ; boot=bootstrap; RMSE=root mean squared error. HRE=Highlands, Reliable Region; HDR=Highland, drought-prone Region; LOE=Lowlands Enset growing region; LOR=Lowlands reliable region. 05,11 refer to the datasets collected in 2005, 2011 respectively (and defined above).

Table 5.2: Testing the models with 2005 and 2011 as training and “holdout” data – squared model

	Baseline	Model 1	Model 2	Model 3	Model 4
R^2	0.245	0.246	0.244	0.245	0.248
Adj. R^2	0.244	0.245	0.243	0.243	0.246
AIC	20044.22	20020.87	20068.66	20065.60	20003.63
BIC	20276.69	20276.58	20339.87	20360.06	20329.08
Num. obs.	17134	17134	17134	17134	17134
boot.cv	0.170	0.171	0.170	0.170	**0.169**
RMSEtest05	0.0297	0.0386	0.0338	0.0344	0.0348
RMSEtest11	0.0412	0.0408	0.0422	0.0424	0.0453

Notes: cv=cross validation. R2=r-squared; AIC=Aikike information Criterion; BIC= ; boot=bootstrap; RMSE=root mean squared error.

Table 5.3: Results of excluding regions

	Baseline	Model 1	Model 2	Model3
RMSEtestHRE	0.0514	0.0546	0.0595	0.0602
RMSEtestHDR	0.0302	0.0297	0.0519	0.0544
RMSEtestLOE	0.0042	0.0060	0.0055	0.0055
RMSEtestLOR	0.0602	0.0080	0.0034	0.0013

Notes: cv=cross validation. R2=r-squared; AIC=Aikike information Criterion; BIC= ; boot=bootstrap; RMSE=root mean squared error. HRE=Highlands, Reliable Region; HDR=Highland, drought-prone Region; LOE=Lowlands Enset growing region; LOR=Lowlands reliable region.

Figure 6.1

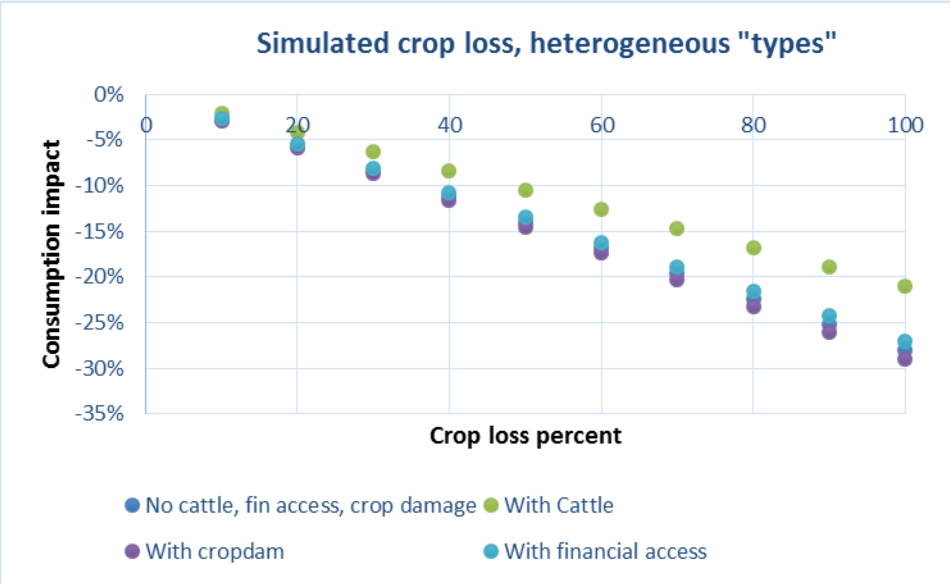
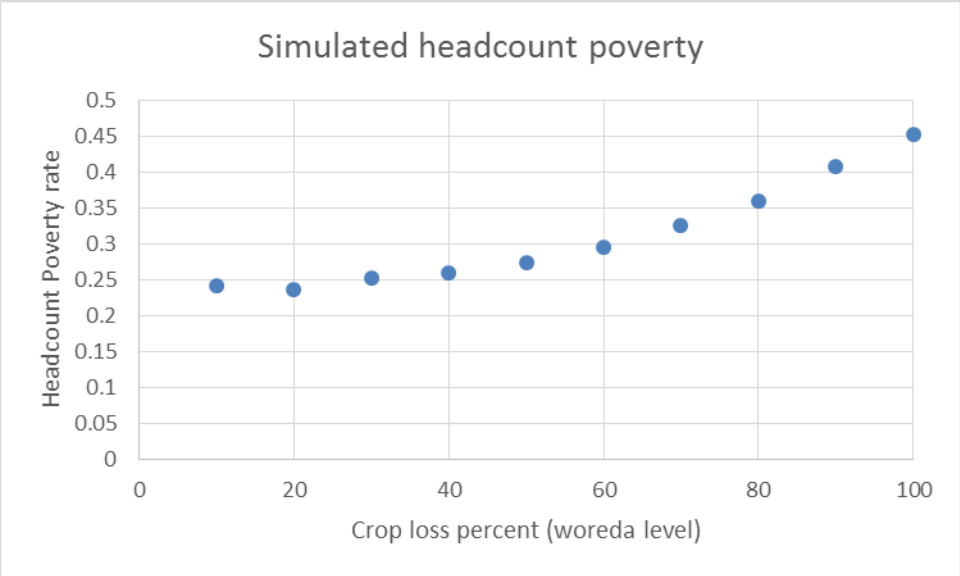

Simulated crop loss, heterogeneous "types"
0%
-5%
-10%
-15%
-20%
-25%
-30%
-35%
0
20
40
60
80
100
Consumption impact
Crop loss percent
No cattle, fin access, crop damage
With Cattle
With cropdam
With financial access

Figure 6.2


Simulated headcount poverty
Headcount Poverty rate
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
20
40
60
80
100
Crop loss percent (woreda level)