Spillovers in Village Consumption: Testing the Extent of Partial Insurance

Tavneet Suri[1]

Abstract

This paper contributes to the literature on risk sharing by proposing a empirical strategy that allows a joint test of both the implications of perfect risk-sharing, i.e. that individuals can smooth away all idiosyncratic shocks and their consumption co-moves one for one with aggregate (group) consumption. This test is more general than earlier approaches and also allows me to estimate the *extent* of risk-sharing within groups in an economy. I study the consumption of maize and food in rural Kenya using this approach. I find that households have the ability to smooth their consumption of maize and other food extremely well, even though I reject complete risk sharing in a few cases. I then look at how well villages are able to pool risk within districts and find that there is significant partial (not complete) insurance amongst villages within districts. These findings have policy implications for government food subsidies, food security and other insurance programs in terms of the level of aggregation at which governments should target such safety net policies.

JEL Classification – C23, O17, Q18

Keywords – Spillovers, risk sharing, consumption variation, Kenya

[1]MIT Sloan School of Management. Email: tavneet@mit.edu. I am extremely grateful to Michael Boozer and Robert Townsend for extensive discussions on this project. I would also like to thank Joe Doyle, Markus Goldstein, Thomas Jayne, Gustav Ranis, Paul Schultz, Tom Stoker, Christopher Udry and seminar audiences at Yale, Chicago and NEUDC for comments. A special thanks to Professor Thomas Jayne and Margaret Beaver at Michigan State for making the data easily available and for all their help with using the survey. The data come from the TAMPA project, between Tegemeo Institute at Egerton University, Kenya and Michigan State University, funded by USAID, April 1996 - June 2005. The rainfall data come from the Climate Prediction Center and are part of the USAID/FEWS project. A special thanks to Tim Love for all his help with these data.

1 Introduction

Developing economies are often described as high-risk environments, where weather shocks (e.g. El Nino, monsoons, etc.), human disease and crop disease are prevalent. These economies are often largely dependent on agriculture as a source of income, so households face extremely uncertain and variable incomes. In addition, average incomes in these economies tend to be extremely low, with a large percentage of households being at or close to subsistence levels[1]. All this, together with the fact that in several low-income countries certain markets either do not exist, or if in fact they exist, they often work imperfectly, means that institutions (often informal) arise to help individuals with income uncertainties. A question that has interested development economists and policy makers for years now has been not only the extent of formal and informal mechanisms used to deal with risk, but how efficient such mechanisms are, especially with regards to food and staple consumption. These mechanisms include, for example, the purchase or sales of assets, mostly animals (see Rosenzweig and Wolpin (1993) and Fafchamps, Udry and Czukas (1998)), the use of grain storage, borrowing from village money lenders, transfers from relatives (see Rosenzweig (1988)) and government policies, such as food subsidies, etc.[2].

The aim of this paper is to rigorously test the effects of all such institutions in a given economy, similar in spirit to Townsend's (1994) general equilibrium test. However, this paper contributes to the existing literature in several ways. First, it allows a joint test of *both* of the implications of perfect risk sharing, namely that individual consumption co-moves one for one with aggregate group (here, village) consumption, and that all idiosyncratic shocks can be completely smoothed away by the household and should therefore have no effect on individual consumption (conditional on aggregate group consumption). I am able to test both these implications jointly by adapting a contrast estimator derived in the peer effects literature (see Boozer and Cacciola (2001)) to non-experimental data. Second, the contrast estimator actually estimates the *extent* of risk sharing within a group. This contrast estimator describes the risk sharing problem in terms of a spillover. The idea of a spillover is a natural way to think about risk sharing since, theoretically, only aggregate shocks to income, and not idiosyncratic shocks, should have an effect on household consumption. By comparing how a household's consumption responds to an idiosyncratic shock to how the average village responds to an aggregate shock, I can estimate this "spillover" and use this as a measure of the extent of risk sharing. Third, this comparison of within village responses to between village responses is a crucial part of the estimation strategy as it enables us to build the correct "counterfactual". I describe in detail

[1] For example, the World Development Indicators estimates that in the least developed countries, real GDP per capita levels were at \$305 for 2002, with about 76% of the population of these countries living in rural areas. The comparable figures for Sub-Saharan Africa were \$575 and 67%. The UN Human Development Report estimates that for the period 1990-2001, an average of 23% of Kenya's population lived on less than \$1 a day and 59% on less than \$2 a day.

[2] In developed economies, there are a number of other insurance mechanisms, like unemployment insurance, disability and medical insurance; welfare and other social government programs; intergenerational altruism has also been studied in this context (see Altonji, Hayashi and Kotlikoff (1989)); all state-contingent government transfers, like natural disaster relief, etc.

why the comparison of individual responses to aggregate village responses provides a benchmark against which to measure the effect of risk sharing, and why this benchmark is needed. Looking only at how individual shocks within a village affect consumption misses an important part of the story: what the relevant aggregate response is. Finally, this test avoids the mechanical estimation issues that have plagued this literature.

The policy value of the question at hand regards the importance of government safety nets in this economy, and their role in ensuring adequate food supplies, especially during times of adverse income shocks. What is the most efficient way to target government food security and insurance policies? The idea is to look at risk sharing across different levels of aggregation and analyze where risk–sharing breaks down enough to warrant safety nets. So, in addition to understanding how well households pool risk within villages, I look at how well villages pool risk within districts. The policy relevance is clear: if consumption can be smoothed well by both households and villages, then governments should be concerned with providing food security and assistance programs when income shocks are aggregate to the district or province. However, if even households are unable to pool consumption, then individual shocks can have serious consequences for these households and there is an immediate and urgent need for a government role in food security. This research also relates closely to the literature on poverty, vulnerability and the role of safety nets in developing economies[3].

I test the null of full insurance using this contrast estimator under various specifications. In most cases, I am unable to reject Pareto efficiency. Where I can reject the null, there is evidence of a large extent of partial insurance - households are not far from Pareto efficiency. When I use the contrast estimator to test how well villages pool maize consumption within districts, I find a significant amount of partial risk sharing within districts, but can reject perfect smoothing. These results have clear policy implications. Individual household level shocks are reasonably well dealt with in my sample of households, at least with respect to maize consumption. The results for villages, on the other hand, show that there is a role for governments when shocks are aggregate to villages, i.e. at the district level and in all likelihood much more so at the province level.

The rest of this paper is structured as follows. Section 2 outlines a stylized theoretical model illustrating the general equilibrium implications of complete insurance. In Section 3, I briefly review the relevant literature in the field. Section 4 outlines the empirical strategy. I discuss in detail what partial insurance is and how the contrast estimator is able to estimate the extent of partial insurance. I also describe how the contrast estimator avoids the specification issues that have plagues this literature (a detailed discussion is left to the Appendix). Section 5 discusses the data, Section 6 the results and Section 7 concludes.

[3] See Morduch and Kamanou (2001), Dercon and Krishnan (2003), Barrett, Holden and Clay (2002) and Udry (1999).

2 Theoretical Implications of Full Insurance

This section briefly describes the basic theory behind optimal risk allocations and how it relates to the empirical tests that have been implemented in the literature. The classic works in this literature are Wilson (1968) and Diamond (1967). As I show below, the principal implication of complete risk sharing is that individual consumption responds only to aggregate risk (aggregate shocks) and not to any idiosyncratic risk (individual shocks).

As Cochrane (1991) states, this implication of complete risk sharing or insurance is analogous to "... a cross-sectional counterpart to the permanent income hypothesis: full insurance implies that consumption should not vary across individuals in response to idiosyncratic shocks, just as constant borrowing and lending opportunities imply that consumption should not vary over time in response to foreseeable shocks." This analogy has motivated a lot of the empirical literature on testing for full insurance and consumption smoothing, in line with Hall's (1978) test for a random walk in consumption. However, it is important to note that the permanent income hypothesis and consumption insurance are very distinct propositions. The former is a panel concept and deals with individuals smoothing consumption over their lifecycle. The latter is cross-sectional with the village as the provider of insurance. This paper deals strictly with the latter.

As per Bardhan and Udry (1999)[4], I look at an economy with just one village with no credit markets. I assume throughout that the village is the relevant risk sharing group, i.e. the appropriate grouping amongst which households pool risk. This is a common assumption in the literature as there is usually no available information on the composition of the relevant group. Some researchers have looked at other assumptions on the group, such as Goldstein (1999) who looks across ethnic groups, as does Grimard (1997). Morduch (2001) looks within castes, and Munshi and Rosenzweig (2005) within sub-castes. Note that in a cross-section of data, the relevant risk sharing group cannot be inferred from the data unless data on group composition is collected, see Manski (1993, 2000)[5].

I also assume there is just one time period but multiple possible states of nature (without loss of generality). This is not an important simplification as all the following easily generalizes. Say there are N households in the village, indexed by i and S states of nature, indexed by s, each with a probability of occurrence of π_s. Without loss of generality, by definition of state, the probabilities of occurrence of each state of nature do not vary by household. Income is exogenously given in each state of nature for each household, i, by y_{is}. Say the utility for household i is given by[6]:

$$U_i = \sum_s \pi_s u_i(C_{is}) \tag{1}$$

[4] This is a stylized model of Pareto efficient risk sharing, similar to the one presented in Townsend (1993, 1994, 1995), Morduch (2001), etc. Also see Scheinkman (1984).

[5] Boozer and Suri (2008) look at how to identify risk sharing groups using long panel data.

[6] With multiple time periods and a discount factor, β, this utility function generalizes to $U_i = \sum_t \beta^t \sum_s \pi_s u_i(C_{ist})$.

where C_{is} is the consumption for household i in state s. Equation (1) is the weighted sum of the utilities for each state of nature, where the weights are given by the probability of the state, π_s. $u_i(.)$ is household i's utility function for each state of nature and it depends only on the consumption in that state. I assume $u_i(.)$ is twice differentiable with $u'(.) > 0$ and $u''(.) < 0$. To find the Pareto efficient allocation of risk in this economy, I maximize a weighted sum of household utilities subject to constraints of non-negative consumption and total consumption being equal to total income in each state, i.e.:

$$Max \sum_i \omega_i U_i \tag{2}$$

$$s.t. \sum_i C_{is} = \sum_i y_{is} \ \forall s \tag{3}$$

$$and \quad C_{is} > 0 \ \forall s \tag{4}$$

where ω_i are the household specific weights on their utility functions. From the first order conditions for this problem for two households, i and j,

$$\frac{u_i'(C_{is})}{u_j'(C_{js})} = \frac{\omega_j}{\omega_i} \ \forall i, j, s \tag{5}$$

I impose a constant absolute risk aversion (CARA) utility function[7] of the following form:

$$u_i(x) = -\frac{1}{\sigma} e^{-\sigma x} \ \forall i \tag{6}$$

Substituting into the first order conditions in equation (5),

$$\frac{e^{-\sigma C_{is}}}{e^{-\sigma C_{js}}} = \frac{\omega_j}{\omega_i} \tag{7}$$

Taking logs gives the following expression for household consumption:

$$C_{is} = C_{js} + \frac{1}{\sigma}[ln(\omega_i) - ln(\omega_j)] \tag{8}$$

Equation (8) holds for all households in the village. It illustrates clearly an important implication of complete risk sharing that can be empirically tested directly with suitable panel data as it is overidentified[8]. Instead, I follow the previous research and aggregate up all the N equali-

[7] Mace (1991) outlines two different utility functions and what they imply in terms of the empirical specification. A constant absolute risk aversion (CARA) utility function implies a specification in logs of consumption, while an exponential utility function implies a specification in levels.

[8] With long panel data, it is possible to empirically test equation (8) of the theory directly. To make this point clearer, consider equation (8) again $C = C_{js} + \frac{1}{\sigma}[ln(\omega_i) - ln(\omega_j)]$. With panel data, this equation is overidentified. For each household, with panel data on its consumption behavior, we can use both the cross-sectional and time series aspects of the data to test all the implications of complete risk sharing. This is not the

ties like equation (8) (one for each household) to derive the following expression for household consumption in terms of average village consumption:

$$C_{is} = \bar{C}_s + \frac{1}{\sigma}[ln(\omega_i) - \frac{1}{N}\sum_j ln(\omega_j)] \tag{9}$$

where $\bar{C}_s = \frac{1}{N}\sum_i C_{is}$ i.e. the average consumption in the village.

The first term in brackets in equation (9) varies by household and the second term by village. The term in brackets therefore corresponds to a household fixed effect which I label α_i. Rewriting equation (9),

$$C_{is} = \bar{C}_s + \alpha_i \tag{10}$$

Equation (10) illustrates clearly the two main theoretical implications of an economy where there is complete risk sharing:

1. No matter what the history of shocks, income, etc. household consumption will co-move one for one with average village consumption.

2. Household level shocks or incomes should play no role in determining household consumption, once average village consumption has been controlled for.

In a cross-section of data, therefore, idiosyncratic shocks (or incomes) should not affect individual consumption as long as average village consumption is controlled for. The test for complete risk sharing is therefore a joint test for the co-movement of household consumption with average village consumption and individual incomes or shocks not affecting individual consumption.

3 Related Empirical Literature

The literature on consumption smoothing and risk sharing mechanisms is large. I focus specifically on reviewing the main general equilibrium-type tests of complete insurance and do not include a discussion of the empirical literature on specific risk pooling mechanisms[9].

I start with a discussion of Townsend (1994), the main paper in this literature that derives the general equilibrium implications for perfect risk sharing and tests it for a developing country. A lot of the literature that has followed has been similar, extending Townsend's basic ideas.

Townsend first runs household by household regressions as follows:

$$C_{it} = \alpha_i + \beta_i\bar{C}_t + \delta_i X_{it} + \psi_i A_{it} + \epsilon_{it} \tag{11}$$

aim of this paper, so I do not discuss exactly how. See Altug and Miller (1990) and Boozer and Suri (2008) for examples.

[9] Good reviews can be found in Besley (1995), Rosenzweig and Wolpin (1993) and Alderman and Paxson (1992).

where C_{it} is consumption of household i at time t; $\bar{C}_t$ is the average village consumption at time t; X_{it} are household specific covariates, including income or shock measures; A_{it} refers to demographic characteristics of household i at time t (the age-sex composition of the household). Under the null of full insurance, for a given household over time, we would expect each β_i coefficient on average village consumption to be unity and each δ_i coefficient on individual incomes or shocks to be zero. Testing whether the β_i's are one is a test for whether individual household consumption co-moves one for one with average village consumption. And, testing if the δ_i's are zero is a test for whether household specific income or shocks affect household consumption, once we control for aggregate consumption.

Townsend also estimates the model for all households by looking at the following pooled version of equation (11) which assumes no household level heterogeneity in the β, ψ and δ coefficients, but retains the household level fixed effects (α_i):

$$C_{ijt} = \alpha_i + \beta \bar{C}_{jt} + \psi A_{ijt} + \delta X_{ijt} + \epsilon_{ijt} \tag{12}$$

where C_{ijt} is the consumption for household i in village j at time t. Econometrically, estimating equation (12) by OLS is an issue. In the simplest case with no covariates, OLS will mechanically give a $\widehat{\beta}$ coefficient[10] of one since OLS fits the mean. The estimated $\widehat{\beta}$ will therefore not have a behavioral interpretation. We could change all the underlying household consumptions around and we would still get an estimated $\widehat{\beta}$ of one. This "problem"[11] has been discussed by Townsend (1994) and Deaton (1990) and in detail by Boozer and Cacciola (2001). Appendix A discusses in detail the mechanical aspects of specifications like equation (12) and why they often do not have a behavioral interpretation.

To avoid this mechanical specification issue, Townsend first uses (long) panel data to run household-specific regressions as in equation (11), which allows him to estimate a $\widehat{\beta}$ for each household. However, he states: "As Deaton (1990) has pointed out, the coefficients on the average consumption variable must average to unity across households if the sample is sufficiently large and if no other terms are included in the regression, even if household regressions are run one at a time. In this sense the average value for the coefficients tells us nothing." So, he estimates equation (11) household by household and looks at the variance of the estimated household specific $\widehat{\beta}_i$'s around one.

In the cross-section, to avoid the specification problems mentioned above, Townsend looks at the following regression that avoids estimating the coefficient on average village consumption:

$$C_t^i - \bar{C}_t = \alpha_i + \gamma A_t^i + \delta_W \hat{X}_t^i + u_t^i - \delta_W v_t^i \tag{13}$$

He imposes a β coefficient of one and then tests whether $\widehat{\delta}$ is significantly different from zero[12].

[10] Throughout the β coefficient refers to the coefficient on average village consumption in a pooled household level OLS regression specification, like equation (12).

[11] This is an issue entirely aside from the reflection problem that Manski (1993, 2000) discusses.

[12] Remember that the test for full insurance is a joint test of the null hypothesis that $\widehat{\beta} = 1$ and $\widehat{\delta} = 0$.

This is close to (but not quite) a village fixed effects specification as he subtracts out average village consumption from individual consumption and uses this as the dependent variable. Townsend estimates equation (13) via household fixed effects and looks at a Griliches-Hausman (1986) measurement error correction[13]. He rejects full insurance, but concludes: "Although the model is rejected statistically, it does provide a surprisingly good benchmark. Household consumptions co-move with average village consumption. More clearly, household consumptions are not much influenced by contemporaneous own income, sickness, unemployment, or other idiosyncratic shocks, controlling for village consumption (i.e. for village level risk)."

Cochrane (1991) has a similar test of consumption insurance. He looks at the effect of idiosyncratic exogenous variables on consumption growth. The exogenous income shocks he looks at are long illness, involuntary job loss and spells of unemployment. He finds that long illnesses and involuntary job losses do affect consumption growth, which allows him to reject the null of full insurance. But, he is unable to reject the null for loss of work due to strikes and spells of unemployment. Mace (1991) looks at the effects of changes in aggregate consumption, individual income and employment status on changes in consumption. She finds that individual income changes have small though significant effects on consumption changes. She concludes that full insurance holds up quite well. Deaton (1990, 1992, 1997) tests whether individual income affects consumption while controlling for village fixed effects in Cote d'Ivoire. The idea behind the village fixed effects (village dummies) is to control for aggregate village consumption. Under the null of full insurance, individual incomes should play no role in determining consumption. He finds that individual incomes do matter and he is therefore able to reject the null of full insurance.

There are a number of papers that are in a similar vein that use similar approaches to these to test models of risk sharing. I don't discuss these in detail. Examples include Besley (1995), Goldstein (1999), Grimard (1997), Morduch (1999, 2001), Dercon and Krishnan (2003) and Ravallion and Chaudhuri (1997). I now go on to discuss the test I use.

4 Econometric Specification

This section outlines my empirical specification that allows a joint test of the both implications of risk sharing models. I use a contrast estimator from the peer effects literature (see Boozer and Cacciola (2001) for more detail). This contrast estimator takes a different approach to estimating models of consumption smoothing. I think of risk sharing as a spillover effect that arises from a given household being part of a village. Theoretically only aggregate shocks to income, and not idiosyncratic shocks, should have an effect on household consumption. Any

[13] Since this within estimate, δ_W, is not consistent because of measurement error, Townsend also estimates the first differenced version, which is also inconsistent because of the measurement error:

$$(C_t^j - \bar{C}_t) - (C_{t-1}^j - \bar{C}_{t-1}) = \gamma(A_t^j - A_{t-1}^j) + \delta_D(\hat{X}_{t-1}^j - \hat{X}_{t-1}^j) + (u_t^j - u_{t-1}^j) - \delta_D(v_t^j - v_{t-1}^j)$$

A combination of δ_W and δ_D are used to give a consistent estimate of δ, the coefficient on household income. See Griliches and Hausman (1986).

shock a household faces can have both idiosyncratic as well as aggregate components. So, a household will respond to a given shock in various ways. I look at how households within a village respond to a shock with respect to their consumption decisions. Econometrically, this looks at a village fixed effects specification to see how shocks affect consumption. The village fixed effects control for the aggregate consumption as well as aggregate components of the shock. I then compare this within village response to how the average village responds. Econometrically, this average response comes from an aggregate village level regression of shocks on consumption, i.e. the between village regression. This comparison between the within and between effects of shocks, in fact, estimates a measure of the spillover at the village level, i.e. the insurance against shocks (or extra consumption smoothing) that comes from a household being part of a given village.

As I describe in more detail below, my test for risk sharing therefore accomplishes the following:

1. It allows an omnibus test of Pareto efficient allocations of risk that tests both the theoretical implications of perfect markets.

2. It provides an estimate of the extent of partial village shared insurance.

3. Conceptually, it adds a benchmark to the standard test that acts as a "counterfactual", something that has not been accounted for in earlier work.

4. It avoids the mechanical specification issues other researchers have encountered when trying to estimate such models via OLS. This mechanical issue is described briefly above and in more detail in Appendix A.

I first describe the contrast estimator specification and then I go on to discuss in detail what I mean by partial insurance and why this give an appropriate measure of the extent of insurance.

4.1 The Contrast Estimator Specification

This section outlines the contrast estimator (derived in detail by Boozer and Cacciola (2001)). The contribution of this analysis is two fold: first, to apply this estimator to the context of insurance and, second, to extend it to the case of non-experimental data. In the context of insurance, the contrast estimator is simply a comparison of how individual households in a given village respond to a shock with respect to their consumption decisions, relative to how the average village responds.

To illustrate how the contrast estimator works, consider a rewrite of equation (12):

$$C_{ij} = \beta \bar{C}_j + \delta S_{ij} + \phi X_{ij} + \gamma Z_j + \epsilon_{ij} \tag{14}$$

where C_{ij} is the consumption for household i in village j, $\overline{C_j}$ is the average consumption for village j, S_{ij} is the shock experienced by household i in village j, X_{ij} are household level covariates (including demographics) and Z_j are village level covariates.

For simplicity, just to explain the estimation procedure, I have ignored the household fixed effects. I do have panel data and will account for the household fixed effects in the empirical work.

The within village (or village fixed effects) estimate is then given by:

$$C_{ij} - \bar{C}_j = (S_{ij} - \bar{S}_j)\delta^W + (X_{ij} - \bar{X}_j)\phi^W + \epsilon_{ij} - \bar{\epsilon_j} \tag{15}$$

where S_{ij} is a measure of the household specific shock, $\bar{S}_j$ the average village shock and δ^W the within village estimate of how consumption responds to shocks. The covariates in this specification are all deviations from their village mean, hence the absence of village level covariates.

I also estimate the between village effect, which measures how the average village responds to aggregate village shocks. It is easily run by averaging up equation (14) to the village level and then running OLS. This between village regression is given by the following equation where I call δ^B the between village estimate of how consumption responds to aggregate shocks:

$$\bar{C}_j = \bar{S}_j\delta^B + \bar{X}_j\phi^B + Z_j\gamma^B + \bar{\epsilon_j} \tag{16}$$

I can then estimate the extent of partial insurance or the spillover as per the contrast estimator. I estimate the $\hat{\beta}$ in equation (14) using the $\widehat{\delta}^W$ and the $\widehat{\delta}^B$ from equations (15) and (16) respectively[14].

$$\hat{\beta} = 1 - \frac{\delta^W}{\delta^B} \tag{17}$$

Relating this back to the theory, it is clear that in the case where $\widehat{\delta}^W = 0$, there is full insurance at the village level (as long as $\widehat{\delta}^B > 0$). Why? When $\widehat{\delta}^W = 0$, $\hat{\beta} = 1$ from equation (17) above, so there is perfect co-movement of household consumption with village consumption. In addition, when $\widehat{\delta}^W = 0$, it means that individual shocks do not affect individual consumption within villages. What makes this test novel is that it is a combination of the responses to shocks that measures the co-movement of consumption within a village. A test of whether the estimated β coefficient is significantly different from one not only tests for co-movement in consumption, but also incorporates an estimate of whether individual shocks matter within a village. Since $\widehat{\beta}$ is computed as in equation (17) above, it compares the estimates of consumption responses to individual and aggregate shocks. It therefore combines both the implications of full insurance into one estimate. If I reject the null of perfect risk sharing, I have a measure of the extent of consumption smoothing, a measure of the extent of partial insurance, from the estimated β

[14] Boozer and Cacciola (2001) show that this method of moments estimator computed as a combination of standard panel estimators is equivalent to a particular IV estimator of β in equation (14). In my case this would amount to equation (14) being specified in terms of the leave-out mean village consumption (mean of the village excluding the specific household) instead of average village consumption. The IV specification would then instrument for this leave out mean village consumption using the leave out mean village shock while controlling (in both stages of the IV regression) for the individual shock.

coefficient.

4.2 What is Partial Insurance?

Say I can credibly estimate the coefficient on average village consumption (β) in equation (14) and that it is significantly different from one, i.e. I reject the null hypothesis of full insurance. What does the value of the estimated $\widehat{\beta}$ coefficient mean? What do I mean by the term "partial insurance"? To answer these questions, we must think about what the possible alternatives to full insurance are. There is no clear theoretical alternative; an estimated $\widehat{\beta}$ coefficient different from one could imply any or all of the following:

1. The relevant risk sharing group is not the village but some other group, like ethnicity, caste, sex or a combination of these.
2. The relevant risk sharing group is the village, but individuals in the risk sharing groups are not fully insuring each other. Instead, when they experience a shock, transfers to smooth the shock may take place, but such transfers are not enough for households to completely insure themselves in response to individual level shocks. The mechanisms that we think of as being used to smooth consumption either do not exist or do not work efficiently.
3. Individuals are in fact not smoothing consumption over space, but instead they are insuring with their future selves. Individual households are saving and borrowing to smooth consumption over their lifecycle in the face of individual shocks. The basic theory in Section 2 assumes no credit markets, but in practice some credit markets do exist, though they may not be perfect.

The phrase "partial insurance" has been used in the literature to refer to situations where there is evidence of at least some consumption smoothing, even though consumption allocations may not be fully Pareto optimal. What the form of this partial insurance is has never been quantified. My test for partial insurance has two components. It is an omnibus test, like Benjamin's (1992) test for separation in agricultural household models. If I reject full insurance, the specific alternative hypothesis is not clear. However, I can estimate the extent of insurance (given a maintained hypothesis as to what the relevant insurance group is). My test involves estimating the extent to which individual household consumptions co-move with aggregate village consumption so that it has a behavioral interpretation. In addition, I can simultaneously look at the coefficients on individual incomes or shocks so I can jointly test both implications of the theory.

Studies that look at the importance of individual income (the δ coefficient) are also omnibus tests. However, the ideal test for complete risk sharing is very much a joint test that $\widehat{\beta} = 1$ and $\widehat{\delta} = 0$ in a specification like equation (14). I interpret the magnitude of β (and not the estimated coefficient on individual income) as a measure of the extent of partial insurance. This comes directly from thinking of the extent of insurance as a spillover. It measures the extent of consumption smoothing that takes place at the village level or, more intuitively, it is the extent of co-movement of household consumption with average village consumption. Insurance provides a means by which households can reduce the variance in their consumption profiles: it is about smoothing consumption.

The theory of Pareto efficient risk pooling outlined in Section 2 has very specific implications for the distribution of shocks: individual shocks and not aggregate shocks, to income should affect household consumptions within a village. The estimator of β I use therefore contrasts the extent to which a given shock affects individual household consumptions within a village, relative to between villages. I show that the estimate of β is a combination of the responses to individual vs. aggregate level shocks. This makes it a good measure of the extent of insurance, as it accounts clearly for the distribution of shocks under complete risk sharing. It is the comparison of individual responses to aggregate responses to shocks that adds a unique conceptual notion to how insurance is measured. This comparison allows us to build the correct counterfactual, with the between village variation in consumption as a benchmark against which to measure the spillover insurance effect. I discuss this in more detail when I outline the estimator below.

Say I estimate a β less than one, what the β is measuring is still a question. Going back to the theoretical model presented in Section 2, we assumed the economy consisted of just a single village and that the village was the relevant risk sharing group. However, it is entirely possible that this is not the case - that indeed some other group, defined along the lines of, say, ethnicity or sex, is the relevant risk pooling group. What theoretical implications does this have? I look at the easiest case: say the economy consists of just one village, but the relevant risk pooling group is only a subgroup of the village, while the rest of the households in the village do not pool risk at all. For those individuals belonging to the risk pooling group, their consumption is determined by average consumption of the group (not of the village). For those individuals in the village who do not belong to the risk pooling group, they do not smooth consumption at all and therefore for them, $C_{ks} = y_{ks}$ for all states, s. I can then derive the following expression for the consumption of individuals who belong to the risk pooling group:

$$C_{is} = \frac{N}{n}\bar{C}_s - \frac{1}{n}\sum_{k \notin J} y_{ks} + \alpha_i \tag{18}$$

where $\alpha_i = \frac{1}{\sigma}[ln(\omega_i) - \frac{1}{n}\sum_{i \in J} ln(\omega_j)]$, $\notin J$ indicates that the household is not a member of the risk pooling group, N is the total number of households in the village, n is the number of households that belong to the insurance group and $\bar{C}_s$ is the average village consumption. For households that do belong to the risk sharing group, individual consumption will no longer co-move with the village average one for one. Some individual incomes will play a role in determining consumption - for individuals in the village that are not a part of this group, consumption in each state is simply their income. So, consumption trends in the village will depend on the covariance of village income with consumption among those insuring. It is not clear what implications for the coefficient on average village consumption this will have, without assuming a specific income process. But, β will no longer be one. Equation (18) also illustrates that, in the unrealistic case where $\sum_k y_{ks}$ is uncorrelated with $\bar{C}_s$, then the estimated β will simply be the inverse of the fraction of households that belong to the risk sharing group, i.e. $\frac{N}{n}$ so that $\beta > 1$.

This thought experiment was the simplest: instead of the whole village insuring, what if we have just a subgroup? It is easy to imagine multiple villages in an economy with overlapping subgroups being the relevant risk sharing groups. This complicates any attempts at constructing theoretical models of such situations. There may also be cases where groups are unable to fully smooth all idiosyncratic shocks. Once again, deriving the theoretical implications of such situations is unnecessarily complicated for the purposes of this paper. The important thing to note is that any situation of rejecting full insurance could imply a number of alternatives.

5 Data

The data requirements here are not overly stringent. At a minimum, I require a single cross-section with a large number of groups and household level shock information, but short (two time periods) panel data is ideal. The data I use is from the Tegemeo Agricultural Monitoring and Policy Analysis (TAMPA) Project, between Tegemeo Institute at Egerton University, Kenya and Michigan State University, funded by the US Agency for International Development, Kenya. This is a household level survey aimed at monitoring smallholder production patterns, consumption and incomes, as well as identifying important policy agendas for farmers. The survey covers a large part of rural Kenya and collects food consumption, income, crop, production and credit data as well as some village/community level information.

The shock measure I use is constructed using rainfall data. The rainfall data comes from The Climate Prediction Center as part of the USAID/FEWS (Famine Early Warning System) project. It is available by latitude and longitude across the country. The rainfall estimates I use are constructed using actual rain station data and incorporate METEOSAT 5 satellite data, GTS (Global Telecommunication System) rain gauge reports, data and models on wind and relative humidity, and orography[15]. Note that for the empirical strategy here, I need a household specific shock, but rainfall is at the village level. Even though the simple rainfall measure is the same for a number of households, each household has very different agricultural characteristics. So, a given amount of rainfall is likely to affect different households differently. I use this idea to construct a household level shock measure.

I first find the village level rainfall shock which is just the deviation of the main season rainfall from its long term mean. I take the deviation from mean rainfall as the appropriate measure of the shock is that mean rainfall is just a description of households' expectations as regards rainfall. There are two rainy seasons in most parts of Kenya, though the timing of these varies across the country. The main season is by far the more prominent in terms of maize cropping. For example, in the 2000 sample of my data, the average acreage cropped during the

[15] See Herman, Kumar, Arkin and Kousky (2002) for more detail or contact the author.

main season is about three and a half times that during the short season (about three times in 1997). Similarly, about 33% of households in this sample (and 37% in 1997) do not report any acreage farmed during the short season. So, I focus on main season rainfall shocks. To get a measure of how this rainfall shock affects households within villages, I interact it with the total number of crop acres farmed by the household (irrespective of whether it is owned or rented). I use the total acreage farmed as opposed to acreage harvested or quantity of crop harvested as these alternatives would include the effects of the shock.

Rainfall shocks have strong non-linear effects on yields. Very little rainfall is bad (i.e. results in a poor harvest) for households, as is too much rainfall. It is quite possible that a little bit of rainfall below mean rainfall has negative effects on harvests of crops but a symmetric case with just a little bit of rainfall above the mean has positive impacts. Figure 1 shows the relationship between maize harvest and my shock measure for the 1997 sample. This figure illustrates clearly the fact that in my sample of data, there are no shocks above mean rainfall that are large enough to have strong negative impacts. So, I just treat the shocks as having linear effects[16].

In terms of consumption data, I focus primarily on maize consumption. The survey does not collect complete consumption/expenditure modules so I cannot look at total household consumption. I also look at total crop (i.e. non-purchased food) consumption. Table 1 describes the households in the two samples of my data, 2000 and 1997. Average per capita maize consumption is KShs 2,081.76 in 2000 and KShs 1,681.62 in 1997[17]. Similarly, the figures for per capita crop consumption are KShs 4,242.51 and KShs 2,955.16 respectively. Average acreage farmed over these two periods increases from 4.319 in 1997 to 5.367 in 2000. And, in both years, the value of the shock (the pure rainfall deviation as well as the interaction) is negative on average, implying less rainfall than average in both these periods. The value for the shock (interaction) is -0.115 in 2000 and -0.418 in 1997.

6 Results

This section describes my results for the contrast estimator, computed as a method of moments estimator as above. The estimate of $\widehat{\beta}$ incorporates both components of the test for perfect risk sharing. This makes the joint tests for full insurance somewhat more complicated, though the overall estimate of β and whether it is significantly different from one is a good test in of itself. The drawback of just testing whether $\widehat{\beta} = 1$ is for one scenario: if I am unable to reject the null based on this estimate of β, it may be because the between estimate ($\widehat{\delta}^B$) is small and not significant such that I do not have enough power to be able to reject the null. The better test for perfect risk sharing in the contrast estimator case is a joint test of $\beta = 1$ and $\delta^W = 0$, which amounts to a joint test that $\delta^W = 0$ and $\delta^B \neq 0$. These joint tests are slightly

[16]I have split the rainfall shocks to differentiate between shocks with rainfall above the mean ("Too Much Rain") and shocks with rainfall below the mean ("Too Little Rain"). I looked at contrast estimator specifications across these types of shocks and found that the positive shocks have very little power and the negative shocks drive the results. These results are available from the author upon request.

[17]The exchange rate at the end of 2000 was about 78.4 KShs (Kenyan shillings) to the US dollar.

complicated, though one advantage is that the two tests are independent since the within and between estimates are orthogonal to each other. To account for the fact that the two tests are on the same sample, I use a Bonferroni correction. I report the overall results of the joint tests here, with the corrected p-values. If I am able to reject full insurance, the estimated β is a measure of the extent of partial insurance.

With regard to the model specifications, I look at results for both maize consumption and crop consumption. With panel data, I can also account for the household fixed effects that the theory implies. In addition to these specifications, I look at how well villages are insured within districts. There are only twenty two districts in my data, so these results should be interpreted with caution. Finally, I test for complete risk sharing using a Townsend (1994) test and compare these results to the contrast estimator results. The Deaton (1990, 1992, 1997) test is equivalent to my within village specification.

One final concern with the contrast estimator results is measurement error, which could account for the differences in the estimated within and between slopes because the between specification just aggregates up the individual level data and may therefore just be eliminating noise. The standard panel data solution to this problem is to use a Griliches and Hausman (1986) correction, as per Townsend (1994). However, such a correction cannot be applied here for two reasons. First, it requires at least three periods of data. More importantly, I am looking at “cross-sectional panels”, and implementing a Griliches-Hausman procedure would involve making restrictive assumptions on the spatial distribution of shocks. So, I look at another less formal way of accounting for measurement error and still find a large extent of risk sharing.

6.1 Basic Contrast Estimator Results

Tables 2 and 3 show the within (δ^W from equation (15)) and between (δ^B from equation (16)) estimates for the effects of shocks on maize and crop consumption for 1997 and 2000. A question that arises in the empirical work is whether to look at levels or logs of consumption. Theoretically this depends on the functional form of the utility function chosen, as mentioned in Section 2. Tables 2 and 3 report results using both levels and logs of consumption. For there to be evidence of an insurance spillover, the within estimate has to be less than the between, i.e. there is an extra benefit in terms of insuring away the impacts of shocks at the village level that isn't there at the individual level.

Table 2 looks at maize consumption for the two periods, both the level and log specifications. For the 2000 sample, the level specification shows a consumption spillover estimate of β of 0.608 (from the first two columns of this table), which, given the standard error of 0.310, is not significantly different from one at the 5% value (the p-value on the t-test is 0.21). On the joint test of $\widehat{\delta}^W = 0$ and $\widehat{\delta}^B \neq 0$, however, I can reject complete risk sharing at the 5% level (the p-value is 0.03). Note that on the joint test, if either the δ^W estimate is very significant or the δ^B estimate is not at all significant, I will tend to reject the joint null. The graphical version of this estimate is shown in Figure 2, which shows the relationships between the rainfall shock and household maize consumption, both within and between villages. The difference in the two slopes is a measure of the extent of insurance.

The next two columns of Table 2 show similar results for the 1997 cross-section of data: a spillover estimate of 0.677, with a standard error of 0.217, again not significantly different from one (with a corresponding p-value of 0.14). On the joint test, I cannot reject complete risk sharing (p-value is 0.08). Moving onto the log specifications for maize consumption, the results are somewhat stronger. The 2000 data show a spillover estimate of β of 0.759 with a standard error of 0.165, not significantly different from one, with a p-value on the t-test of 0.14. In terms of the joint test, I cannot reject the null (p-value of 0.12). Similarly, the 1997 data show a β estimate of 0.711 with a standard error of 0.142, giving a p-value on the t-test of 0.11. With the joint test, I cannot reject the null (p-value on the joint test is 0.08). These results in logs are interesting to compare to a Deaton test that just looks at the δ^W estimates. Under a Deaton test, I would in fact reject the null for 1997, but not for 2000 (also see Table 8).

Table 3 shows a similar analysis, this time looking at crop consumption. Here, the results are slightly different across the two years. For the level specification, the 2000 data shows a spillover of 0.933 (standard error of 0.316), which is not significantly different from one (p-value on the t-test of 0.83). In the case of the joint test, I cannot reject the null of complete risk sharing (a p-value of 1.00). Similar results are borne out in the logs specification for 2000 (the p-value on the t-test for $\widehat{\beta} = 1$ is 0.88). However, the results for the 1997 data are quite different from 2000. In both the case of the level specification as well as the log specification, I can clearly reject perfect risk sharing. In the levels case, the spillover estimate is 0.555 (the standard error is 0.198 and the p-value on the t-test for $\widehat{\beta} = 1$ is 0.02) and the joint test has a p-value of 0.001. In the logs case, the spillover estimate is 0.684 (with a standard error of 0.111 and a p-value on the t-test for $\widehat{\beta} = 1$ of 0.005) and the joint test has a p-value of 0.0002. So, the 1997 sample is the one clear case where I am able to reject complete risk sharing. But, there is strong evidence of significant smoothing of crop consumption, with a spillover estimate of about 0.6.

The specifications in Tables 2 and 3 should not be taken too literally. The theory implied a household fixed effect in the estimation. So, Table 4 looks at a specification in changes in consumption to account for the household fixed effects, showing the results for both the non-log and log versions of maize consumption. The first two columns show a spillover of 0.778, with a standard error of 0.104, which is significantly different from one; the p-value on the t-test is 0.03. On the joint test of risk sharing, I can reject complete risk sharing (the p-value is 0.02). The log specification is, however, stronger, with a spillover estimate of 0.907 (with a standard error of 0.064 and a p-value on the t-test for $\widehat{\beta} = 1$ of 0.15). Likewise, on the joint test (with a p-value of 0.24), I cannot reject the null. Figure 3 shows the graphical version of this spillover estimate.

Moving on to Table 5 and the results for crop consumption, the non-log specification shows a β estimate of 0.777 (standard error of 0.135), which is not significantly different from one (t-test p-value is 0.10). The joint test does not allow me to reject perfect risk sharing (with a p-value of 0.13). The results are very similar for the log specification (a spillover estimate of 0.927 that is not significantly different from one) and I am unable to reject the joint null of complete risk sharing (p-value on the joint test is 0.515).

To summarize all these results, it is clear that households are able to smooth maize and other crop consumption rather well in my sample. In a lot of cases, I am unable to reject perfect risk sharing, more so in the fixed effects specifications which align more closely with the theory. This may seem somewhat surprising, but I am looking only at maize and other crop consumption. Maize is the main staple, so it is likely that when households face shocks, they will first try to smooth their consumption of staple food. It is not surprising that the results are less strong for the case of crop consumption. And, it is unlikely that total household consumption would be smoothed as effectively as maize consumption.

6.2 How Well do Villages Pool Risk Within Districts?

This section extends the basic specification, with a more policy motive behind the analysis. The above results indicated that households in rural Kenya manage to smooth maize consumption and other food consumption reasonably well. This then begs the policy question of what the government's ideal role is for food security in this economy. Under complete markets, households can smooth out individual shocks and not aggregate shocks. The question is therefore at what level of aggregation should governments target their food subsidy/food security policies, especially at the time of adverse shocks. To answer this, we need to know how well villages are insured within, say districts, a higher level of geographic aggregation than villages. If villages are well insured within districts, then it points to government policy playing a role at a higher level of aggregation than villages, such as districts and perhaps even provinces. Given there are only eight provinces in Kenya, this would be hard to test empirically.

These specifications use a similar analysis as above, but just look at how well villages in districts smooth consumption (as opposed to households in villages). Table 6 shows the within and between district contrast estimates for the 2000 sample and the fixed effects sample, both just for the logs of consumption. The most important thing to note is the wide variance in estimates of the spillover effect. This is mostly the result of a lack of precision due to the much smaller samples. A quick summary of the 2000 results does not really show strong evidence of consumption smoothing – the estimates of the spillover vary widely from -0.56 to 0.126, with extremely large standard errors. However, the more appropriate and relevant fixed effects specifications show a β estimate of 0.565 (with a standard error of 0.232), which is significantly different from one only at the 10% level (the p-value on the t-test is 0.08). In terms of the joint test for risk sharing, though, I am able to reject the null (with a p-value of 0.002). Looking at crop consumption, the spillover estimate on the fixed effects specification is 0.718, with a standard error of 0.142, which is significantly different from one (p-value is 0.05). The joint test rejects with a p-value of 0.01. So, I am able to clearly reject complete consumption smoothing, but there is strong evidence of partial risk sharing, even at the district level. Villages are able to smooth shocks such that their maize consumption co-moves with district consumption with a factor of between 0.6 and 0.7. This implies that there is a role for governments to help villages deal with negative aggregate shocks, albeit small. But, province level shocks are where government food security programs have a very strong role to play. In the face of large aggregate shocks to the district and/or province, households are unable to pool this risk completely via formal/informal mechanisms and there is therefore a role for government policy here.

6.3 Townsend (1994) Specifications

As a final test, I can run specifications similar to those in Townsend (1994) as follows:

$$C_{ij} - \bar{C}_j = \alpha_i + \delta^T S_{ij} + \phi^T X_{ij} + \gamma^T Z_j + \epsilon_{ij} \tag{19}$$

I look at both the single cross-section results (without the α_i in equation (19)) as well as those with the household fixed effects. The Townsend test is whether the estimated δ^T coefficient is significantly different from zero or not. Table 7 reports the results for three cases: 1997, 2000 and the household fixed effects specifications, for just the logs of maize and crop consumption. For all three maize consumption specifications, I am unable to reject complete risk sharing at the 5% level of significance. For crop consumption, I can reject complete risk sharing for just the 1997 sample.

Table 8 illustrates how the contrast estimator results compare to results from Deaton-type and Townsend-type tests. This table does illustrate some of the usefulness of my test. Not only does Table 8 show whether the three types of tests (contrast, Deaton and Townsend) reject the null of perfect risk pooling or not, but it also presents the p-values on the null hypothesis in each case. There are instances where the contrast estimator refines the test (i.e. it is a stronger test), given it is a joint test of all the implications of optimal risk sharing, in contrast with the Deaton and Townsend tests. In addition, the estimated coefficient of interest, β, is a measure of the extent of risk sharing in the economy and helps to understand how far from Pareto optimality the economy is.

6.4 Measurement Error

A final concern with the above results is that measurement error is driving the difference between the within and between estimates. Since the between regression just averages up the individual level regression, the averaging process may eliminate some of the noise. It is not possible to use a Griliches-Hausman panel data measurement error correction strategy here. Instead, I look at how important measurement error may be in a less formal way.

The way I do this is to aggregate up the household level observations in each village into sub-groups. I then treat each of these sub-groups as individual observations and test complete risk pooling using the contrast estimator specification. The important part of this is how I aggregate up households in each village into these sub-groups. Ideally, I would have the location of each household with respect to other households in the village. Unfortunately, household level GPS information is not available for my sample. If I did have the location of each household, then I could actually look at the spatial distribution of the shocks and try to adapt the Griliches-Hausman procedure to the scenario of the "cross-sectional panels" I have here.

Instead, I look at a distance measure that I have data on at the household level. I look at how far households are from piped water. I could have used other distance measures in the data, like distance to the nearest matatu stop (the main form of public transport). However, the water measure seems the most appropriate, as all households need and use water (a lot of households have bicycles and therefore may never use the matatu). I divide up the households in each village into sub-groups, based on quintiles of the distribution of this distance to piped water variable. I aggregate up consumption to the level of these sub-groups and then treat these as distinct observations and estimate the model.

Table 9 shows the results. Note that the number of observations is much lower now, whereas the number of villages is still the same. I report the results only for the logs of maize and crop consumption, for both the 2000 and the household fixed effects samples. In the case of maize consumption, the 2000 sample shows a spillover estimate of 0.541 (I can reject the joint null). The fixed effects sample shows an estimate of the spillover of 1.013 (I cannot reject the joint null). Similarly, the two estimates for crop consumption are 0.592 and 0.966, respectively and I cannot reject the joint null for the fixed effects sample once again. These results, at least for the fixed effects case, are very similar to those in earlier tables, which points to measurement error perhaps not being a dominant issue.

7 Policy Implications and Conclusion

The basic theoretical implications of full insurance have led economists to analyze the effect of a group outcome (here average village consumption) on individual outcomes (here household consumption). This paper draws on some of the empirical advances in the peer effects literature to better understand the extent of risk sharing across villages in Kenya. In particular, I use the contrast estimator derived by Boozer and Cacciola (2001) to estimate the extent of insurance. I extend their estimator in two significant ways: (i) I apply it to the case of insurance and illustrate how the extent of risk sharing can be thought of very intuitively as a spillover and (ii) I describe how their estimator works in the case of non-experimental data.

I find strong evidence of a significant spillover at the village level, which I interpret as the extent of partial insurance. The estimates range from a β coefficient of 0.56 to 0.97 in the cross-sectional specifications and between 0.78 and 0.93 in the household fixed effects specifications, depending on whether the dependent variable is maize or crop consumption and whether it is in levels or logs. The test here is an omnibus test: it does not distinguish between the reasons that may account for the failure of the complete risk sharing hypothesis. Overall, the results presented here look extremely favorable for perfect risk sharing. Only in a few cases am I able to reject the null of Pareto efficient consumption smoothing, a striking result given the study area. This result also holds up when I account for measurement error in consumption. What makes this analysis especially interesting is that I look at the consumption of the main staple crop in Kenya, maize. With more detailed consumption data, it would be interesting to look at whether the consumption of other goods shows evidence of such strong smoothing in Kenya.

These results have important policy relevance for rural households in Kenya. The results show that households in my sample are able to, through various formal and/or informal means, smooth well consumption of food, in particular, maize, their main staple. Households facing uncertain, variable incomes are, at a minimum, able to insure away enough of the risk so as to smooth their consumption of food. This relates to the role of food subsidies and the role of governments in providing safety nets to account for the basic consumption needs of households in such economies. My results point to a limited need for such subsidies aimed at the household level during adverse income shocks, perhaps more of a focus on whether subsistence consumption levels are being met in the first place, especially for the households that are in the lower tails of the income distribution. The aggregated results show that villages are able to smooth away some of their negative shocks, but not perfectly. Governments therefore have a strong role in helping groups smooth out aggregate shocks, somewhat at the village level and far more so at the district and province level.

On a broader scale, this work has strong relevance for economics in general. Group or social interactions are extremely widespread in the literatures on externalities and spillovers, yet the credible identification of such effects is limited. Since the contrast estimator allows credible identification of such effects, it has widespread significance across several fields. For example, Boozer and Cacciola (2001) study peer effects in the US class size debate. Kremer and Miguel (2001) look at the spillovers of de-worming health projects on education in Kenya. Conley and Udry (2002) study the effects of networks on learning in the pineapple industry in Ghana. Forbes and Rigobon (1999) survey some of the conceptual and empirical issues involved in measuring financial contagion. There has been some more recent work in the field of urban economics on spillovers and externalities in cities (see Lucas (2001)) which is also related to the empirical growth literature (see Durlauf (1999)). This drives home the importance of the contrast estimator given its easy application to non-experimental data.

8 Appendix A: Econometric Issues for Grouped Data

Previous researchers in this literature have used regression as a tool of analysis and have focused on the aggregate implications of complete risk sharing as illustrated by equation (14). Consider the following econometric specification[18]:

$$C_{ij} = \beta \bar{C}_j + \delta X_{ij} + \gamma Z_j + \epsilon_{ij} \tag{20}$$

where C_{ij} is consumption for household i in village j; $\bar{C}_j$ is the average consumption for village j; X_{ij} are household specific covariates and Z_j are village specific covariates. Equation (20) comes directly from the theoretical model in Section 2, except I have eliminated the household fixed effect for simplification. I could specify equation (20) in changes instead of levels and that would still have the same issues. Under the null of full insurance, β should be one and δ (the coefficient on individual incomes) should be zero. There are two issues with a specification like (20) above: one mechanical and one conceptual. This Appendix is devoted to discussing the former in some detail.

Estimating equation (20) by OLS is problematic. Think of the simplest case with no covariates, OLS will mechanically estimate β to be one. OLS fits the mean and the mean of individuals' consumptions across villages in the sample is the same as the mean of average village consumptions. In this case, the estimated β coefficient will not have a behavioral interpretation. We could change all the underlying household consumptions and would still get a β coefficient of one. This "problem" has been discussed by Townsend (1994), Deaton (1990) and in detail by Boozer and Cacciola (2001).

This section draws on Boozer and Cacciola (2001), who derive the mechanical aspects underlying specifications like equation (20). I outline four relevant cases to understand why such specifications mostly have no behavioral interpretation. Some researchers (see Goldstein (1999)) opt to use the leave out group mean (i.e. the mean of the group excluding the specific household) instead of the regular mean, but this still does not solve the issue. The four cases I discuss are: (i) using the full mean as the independent variable without any covariates, (ii) using the leave out mean without covariates, (iii) using the full mean with covariates, and (iv) using the leave out mean with covariates.

I briefly discuss the econometric implications of the first two of these and then summarize how covariates may complicate matters but do not solve the issue. The first two cases deal with a simplified version of equation (20) where there is a single independent variable - the mean or the leave out mean of consumption - and no covariates. In the case of the mean, I can re-write equation (20) as:

$$C_{ij} = \beta \bar{C}_j + \epsilon_{ij} \tag{21}$$

[18] I begin by looking at a single cross-section and not worrying about a time index. This is not an important simplification.

I can then estimate $\hat{\beta}$ as:

$$\hat{\beta} = \frac{\sum_j \sum_i C_{ij}\bar{C}_j}{\sum_j \sum_i \bar{C}_j} = \frac{\sum_j N\bar{C}_j^2}{\sum_j N\bar{C}_j^2} = 1 \tag{22}$$

In the case of the leave out mean, Boozer and Cacciola (2001) show that the β coefficient can be written as follows:

$$\hat{\beta} \approx 1 - \frac{WSS}{(N-1)BSS} \tag{23}$$

where WSS and BSS are the within and between sum of squares respectively. Extending this last case to account for covariates, equation (23) becomes

$$\hat{\beta} \approx 1 - \frac{(WSS|x'_{ij} - \bar{x'}_j)}{(N-1)(BSS|\bar{x'}_j)} \tag{24}$$

This expression for the estimate of β using the leave out mean has a nice interpretation. The estimated β depends on the type of covariates included in the regression. If we included covariates that vary only at the village level, it will drive the estimated β down. This is because a covariate that varies only at the group level affects only the between variation, hence the BSS. It is orthogonal to the WSS and therefore has no effect on the conditional WSS. Including village level covariates therefore drives the estimated β down. For covariates that vary both within and between villages, the effect on the estimated β is ambiguous. Covariates that explain relatively more of the within variation (relative to the between) will result in an estimated β closer to one. As Boozer and Cacciola (2001) point out, a consequence of this is that identification of β can be "fragile" as estimates will be sensitive to the inclusion or exclusion of certain covariates. The estimated β therefore does not necessarily have a behavioral interpretation.

The β instead tends to be a "catch-all" for why consumption may vary within and between villages. On the behavioral side, any sorting into villages without insurance could mean a β equal to one. On the researcher's side, adding variables that explain within and not between village variation would drive the estimated β closer to one. The important thing to take away from this discussion is that the dispersion of the β coefficient around one in the general case (i.e. with covariates) tells us how much covariates explain the within village vs. the between village variation. It tells us little about underlying insurance in the villages and therefore little about the underlying behavioral model. We could just as easily change around households' consumption allocations and still get a coefficient close to one on average consumption[19].

How is this mechanical specification issue avoided? Townsend uses panel data to estimate household specific β coefficients. And, in the cross-section, his dependent variable is the deviation of individual consumption from average village consumption. Deaton controls for village fixed effects. I use the contrast estimator which avoids this mechanical issue very cleanly.

[19] See Boozer (2002) for the econometric issues that arise in the more general case of grouped data.

References

Alderman, Harold, and Christina Paxson, (1992), "Do the Poor Insure? A Synthesis of Literature on Risk and Insurance in Developing Countries".

Altonji, Joseph, Fumio Hayashi, and Laurence Kotlikoff, (1989), "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data", Working Paper Number 3046, Cambridge, Mass., NBER, July 1989.

Altug, Sumru and Robert Miller, (1990), "Household Choices in Equilibrium", *Econometrica*, 58(3), 543-570.

Bardhan, Pranab and Christopher Udry, (1999), "Development Microeconomics", Oxford University Press.

Barrett, Christopher, Stein Holden, and Daniel Clay, (2002), "Can Food-For-Work Programs Reduce Vulnerability", in Stefan Dercon, ed., *Insurance Against Poverty*, Oxford: Oxford University Press.

Benjamin, Dwayne, (1992), "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models", *Econometrica*, 60(2), 287-322.

Besley, Timothy, (1995), "Non-market Institutions for Credit and Risk Sharing in Low-Income Countries", *Journal of Economic Perspectives*, 9(3), Summer 1995, 115-127.

Boozer, Michael, (2002), "Identification of Group Effects in Individual Data with Clustering", Working Paper, Yale University, July 2002.

Boozer, Michael and Tavneet Suri, (2004), "Testing for Consumption Insurance and the Identification of Risk Sharing Groups", Working Paper, Yale University.

Boozer, Michael, and Stephen Cacciola, (2001), "Inside The 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data", Economic Growth Center Discussion Paper 832, Yale University, June 2001.

Cochrane, John, (1991), "A Simple Test of Consumption Insurance", *Journal of Political Economy*, 99(5), 1991, 957-976.

Conley, Timothy and Christopher Udry, (2002), "Learning About A New Technology: Pineapple in Ghana", Working Paper, University of Chicago and Yale University.

Deaton, Angus, (1997), "Analysis of Household Surveys: A Microeconometric Approach to Development Policy", John Hopkins University Press for the World Bank, 1997.

Deaton, Angus, (1992a), "Saving and Income Smoothing in Cote d'Ivoire", *Journal of African Economies*, 1(1), March 1992, 1-24.

Deaton, Angus, (1990), "On Risk, Insurance, and Intra-Village Consumption Smoothing", Manuscript, Princeton University, November 1990.

Dercon, Stefan, and Pramila Krishnan, (2003), "Food Aid and Informal Insurance", WIDER Discussion Paper No. 2003/09, United Nations University.

Diamond, Peter, (1967), "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty", *American Economic Review*, 57(3), September 1967, 759-776.

Durlauf, Steven, (1999), "Notes on Group Membership and Poverty", Prepared for World Bank 1999 Summer Research Workshop on Poverty.

Fafchamps, Marcel, Chris Udry and Katie Czukas, (1998), "Drought and Saving in West Africa: Are Livestock a Buffer Stock?", *Journal of Development Economics*, 55(2), pp. 273-306.

Forbes, Kristin and Roberto Rigobon, (1999), "Measuring Contagion: Conceptual and Empirical Issues", Paper prepared for the UNU-WIDER Conference on Financial Contagion in Helsinki, Finland.

Goldstein, Markus, (1999), "Chop Time, No Friends: Examining Options for Individual Insurance in Southern Ghana", Manuscript, Department of Agricultural and Resource Economics, University of California, Berkeley.

Griliches, Zvi, and Jerry Hausman, (1986), "Errors in Variables in Panel Data", *Journal of Econometrics*, 31, 93-118.

Grimard, Franque, (1997), "Household Consumption Smoothing Through Ethnic Ties: Evidence from Cote d'Ivoire", *Journal of Development Economics*, 53, 391-422.

Hall, Robert, (1978), "Stochastic Implications of the Life-Cycle Permanent Income Hypothesis: Theory and Evidence", *Journal of Political Economy*, 86(6), December 1978, 971-87.

Herman, Alan, Vadlamani Kumar, Phillip Arkin, and Jamie Kousky, (2002), "Objectively Determined Ten Day African Rainfall Estimates Created for Famine Early Warning Systems", Climate Prediction Center.

Lucas, Robert, (2001), "Externalities and Cities", *Review of Economic Dynamics*, 4(2), April 2001, 245-274.

Mace, Barbara, (1991), "Full Insurance in the Presence of Aggregate Uncertainty", *Journal of Political Economy*, 99(5), 1991, 928-956.

Manski, Charles, (2000), "Economic Analysis of Social Interactions", *Journal of Economic Perspectives*, 14(3), Summer 2000, 115-36.

Manski, Charles, (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, 60, 1993, 531-542.

Miguel, Edward and Michael Kremer, (2001), "Worms: Education and Health Externalities in Kenya", NBER Working Paper Number 8481, July 2001.

Morduch, Jonathan and Gisele Kamanou, (2001), "Measuring Vulnerability", forthcoming in Stefan Dercon (ed.), *Insurance Against Poverty*, Oxford University Press, 2003.

Morduch, Jonathan, (2001), "Consumption Smoothing Across Space: Testing Theories of Risk Sharing in the ICRISAT Study Region of South India", Working Paper, NYU, Presented at the UNU/WIDER Meeting on Insurance Against Poverty, June 2001.

Morduch, Jonathan, (1999), "Between the Market and State: Can Informal Insurance Patch the Safety Net?", *World Bank Research Observer*, 14(2) December 1999, 187-207.

Ravallion, Martin and Shubham Chaudhuri, (1997), "Risk and Insurance in Village India: Comment", *Econometrica*, 65/1 (January): 171-184.

Rosenzweig, Mark, and Kenneth Wolpin, (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.*

Rosenzweig, Mark, (1988), "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries", *Economic Journal*, 98, 1148-1170.

Scheinkman, Jose, (1984), "General Equilibrium Models of Economic Fluctuations: A Survey of Theory", Manuscript, University of Chicago, June 1984.

Townsend, Robert (1995), "Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies", *Journal of Economic Perspectives*, 9(3), Summer 1995, 83-102.

Townsend, Robert (1994), "Risk and Insurance in Village India", *Econometrica*, 62(3), May 1994, 539-591.

Townsend, Robert (1993), "The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models", Princeton University Press, 1993.

Udry, Christopher, (1999), "Poverty, Risk and Households", Prepared for World Bank 1999 Summer Research Workshop on Poverty.

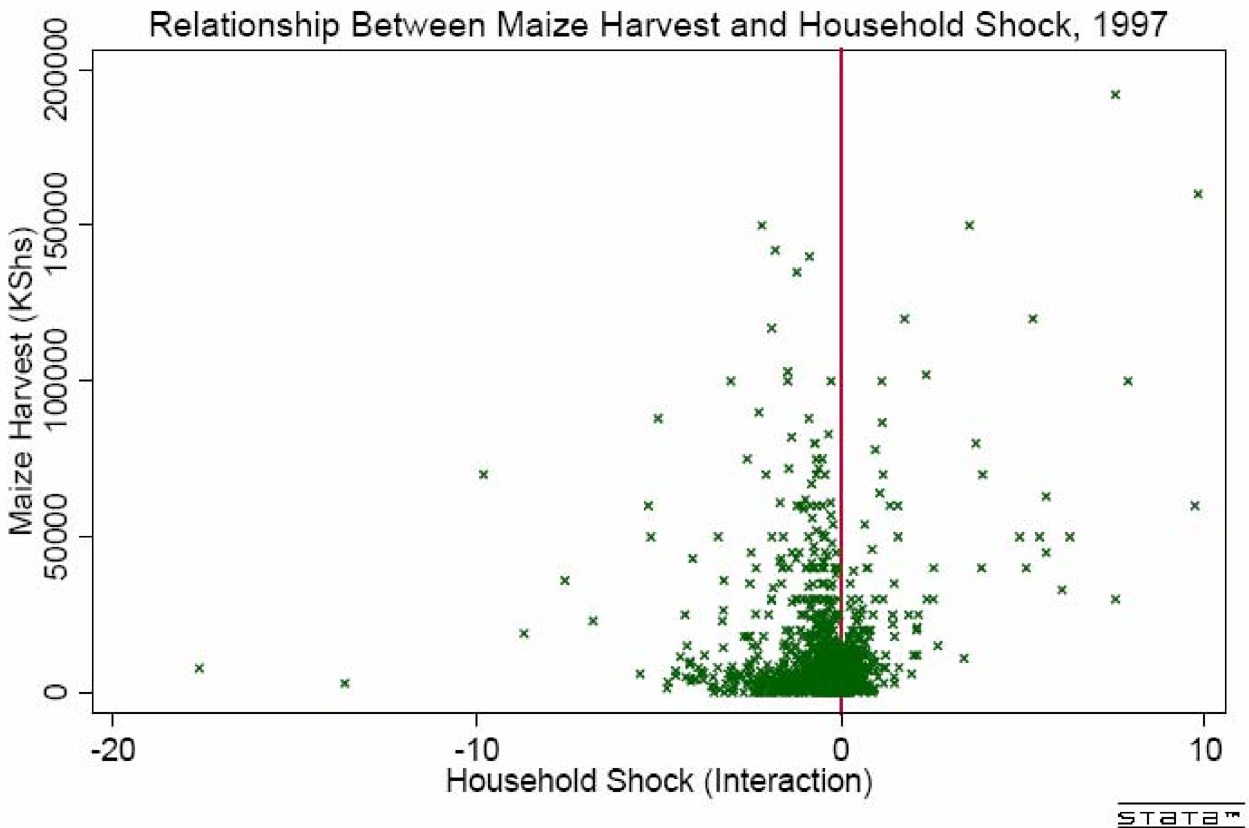
United Nations, *Human Development Report*, Various Issues.

Wilson, Robert, (1968), "The Theory of Syndicates", *Econometrica*, 36, 1968, 119-132.

Wooldridge, Jeffrey, (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press, 2002.

World Bank, *World Development Indicators*, Online version.

Figure 1



Relationship Between Maize Harvest and Household Shock, 1997
Maize Harvest (KShs)
200000
150000
100000
50000
0
-20
-10
0
10
Household Shock (Interaction)
stata™

Figure 2
Relationship Between Shock and Household Consumption
Within and Between Villages
(2000 Sample in Levels)

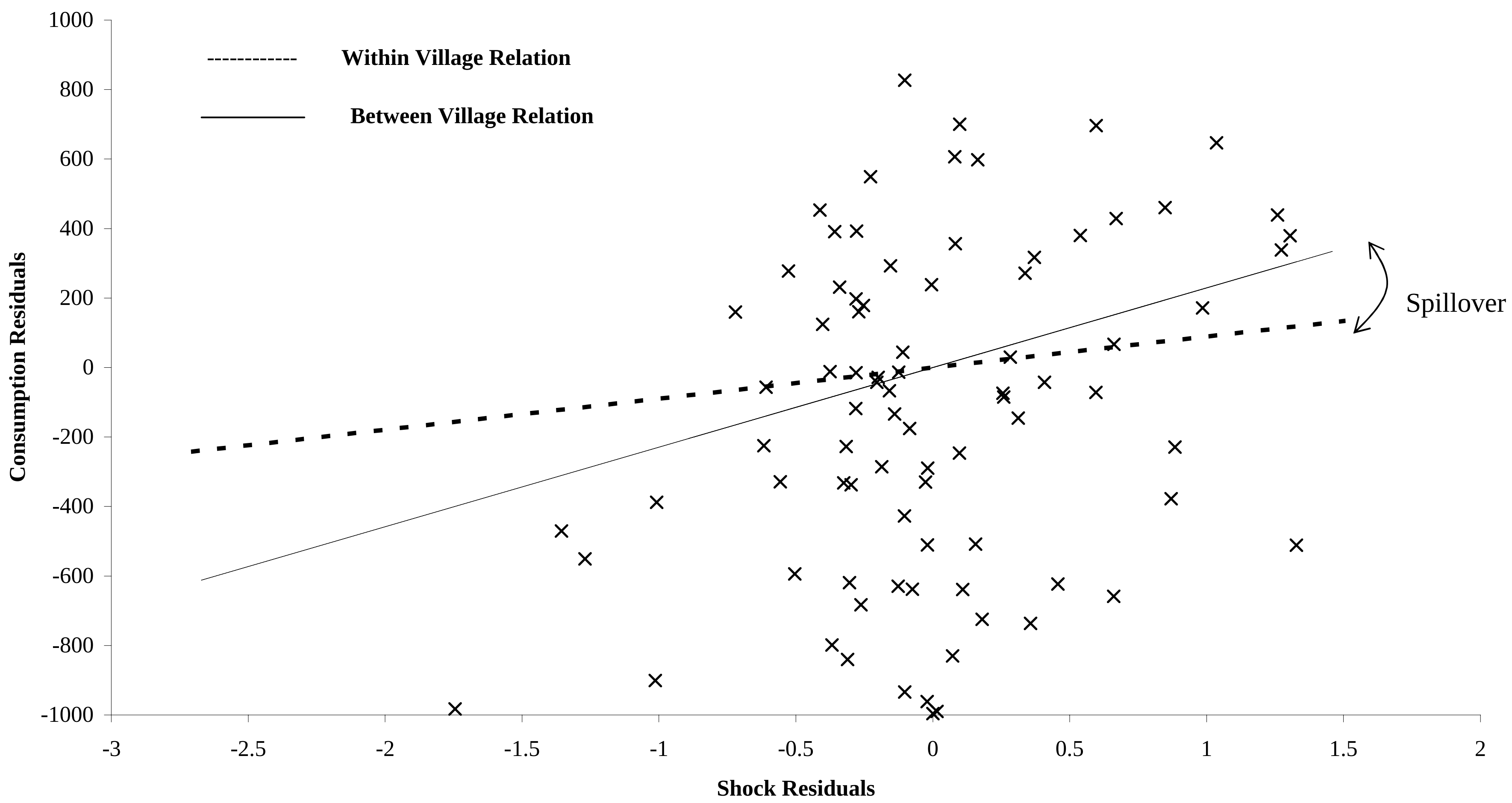


Figure 3
Relationship Between Shock and Household Consumption
Within and Between Villages
(Household Fixed Effects Sample in Logs)

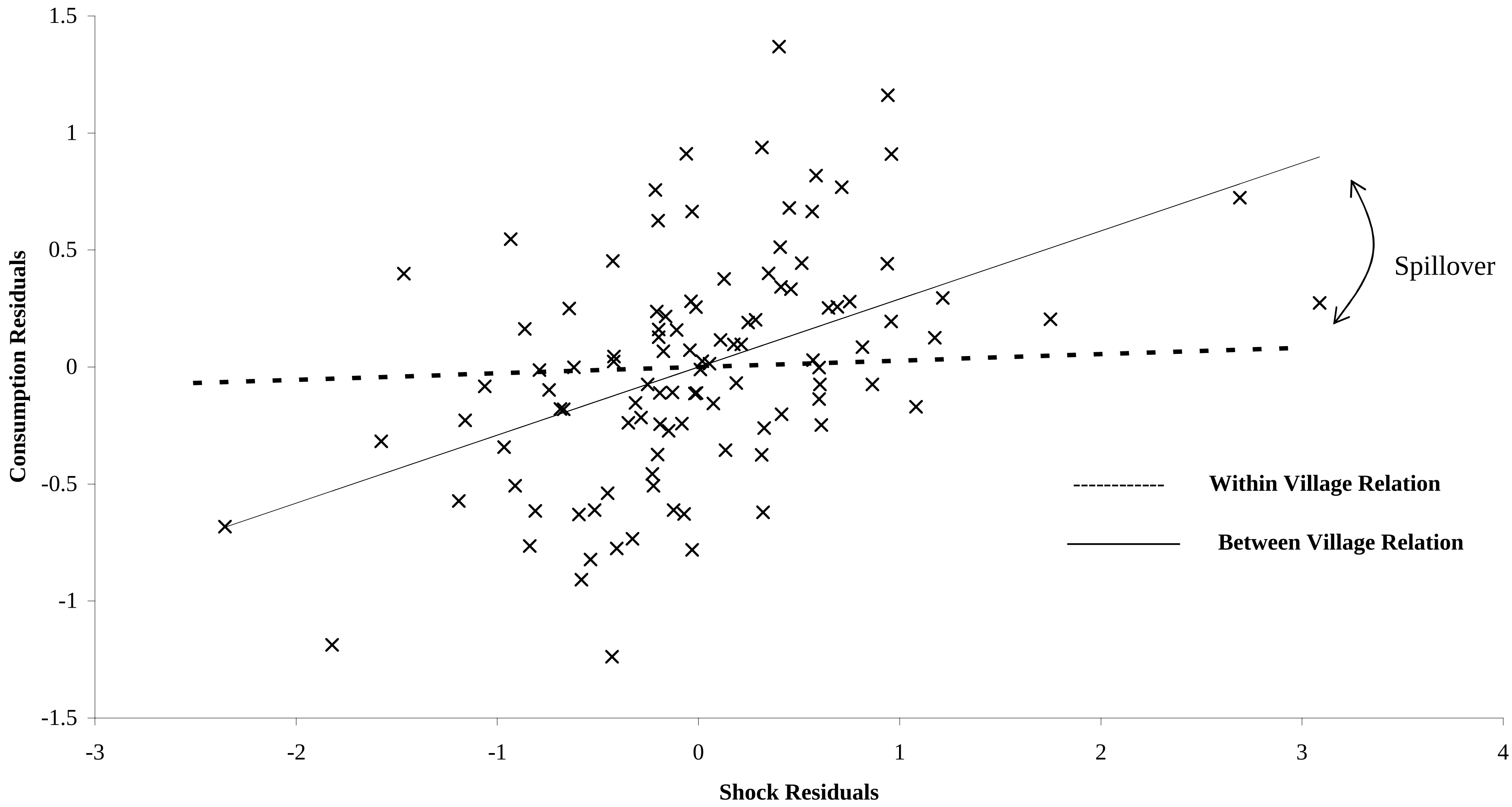


Table 1: Summary Statistics

Variable	2000 Sample 1438 Households	1997 Sample 1496 Households
Average Household Per Capita Maize Consumption	2081.76 (1992.73)	1681.62 (2011.94)
Average Household Per Capita Crop Consumption	4242.51 (3520.65)	2955.16 (3011.02)
Average Value of Household Harvest	77739.34 (123609.6)	50560.96 (97164.53)
Average Total Acreage Farmed	5.367 (9.516)	4.319 (5.664)
Average Rainfall Shock (Current Seasonal Rainfall Minus Seasonal Mean, Normalized by the Seasonal Mean)	-0.053 (0.287)	-0.081 (0.221)
Average Value of Household Specific Shock (Average Rainfall Shock*Total Acreage)	-0.115 (3.256)	-0.418 (1.616)
Household Demographics:		
Average Household Size	8.373 (3.308)	6.996 (2.646)
Average Percent of Household that is Male and Between 6 and 16 years of age	0.150 (0.138)	0.162 (0.152)
Average Percent of Household that is Female and Between 6 and 16 years of age	0.145 (0.137)	0.164 (0.155)
Average Percent of Household that is Male and Between 17 and 39 years of age	0.197 (0.152)	0.179 (0.160)
Average Percent of Household that is Female and Between 17 and 39 years of age	0.195 (0.125)	0.163 (0.133)
Average Percent of Household that is Male and 40 years of age or Above	0.109 (0.103)	0.102 (0.113)
Average Percent of Household that is Female and 40 years of age or Above	0.107 (0.117)	0.113 (0.133)

Notes: Standard deviations are in parentheses. Consumption and harvest data is in Kenyan shillings.

Table 2: Contrast Estimator: Within and Between Village Regressions

Variable	Dependent Variable: Per Capita Maize Consumption				Dependent Variable: Ln Per Capita Maize Consumption			
	2000		1997		2000		1997	
	Within	Between	Within	Between	Within	Between	Within	Between
Household Specific Linear Shock (Interaction)	89.80 (38.12)	229.06 (152.73)	116.20 (57.60)	359.71 (162.26)	0.027 (0.014)	0.112 (0.050)	0.054 (0.026)	0.236 (0.092)
Total Acreage	84.12 (8.58)	146.77 (33.43)	108.32 (12.72)	257.23 (34.89)	0.029 (0.003)	0.030 (0.011)	0.046 (0.006)	0.120 (0.020)
Village Level Main Effect of Linear Shock	-	-1589.6 (849.07)	-	-645.10 (872.76)	-	-0.541 (0.279)	-	-0.304 (0.494)
Household Size	-108.91 (16.68)	154.80 (76.82)	-130.68 (20.27)	56.91 (80.16)	-0.056 (0.006)	0.021 (0.025)	-0.076 (0.009)	0.039 (0.046)
Constant	3878.85 (397.95)	4628.95 (2699.65)	2328.00 (349.77)	-3578.93 (1976.81)	8.092 (0.148)	7.627 (0.885)	7.772 (0.161)	4.567 (1.092)
R-squared	0.139	0.386	0.113	0.445	0.155	0.309	0.142	0.393
Number of Groups	107		107		107		107	
Number of Observations	1437		1495		1436		1477	
Contrast Estimate ($1-\delta^W/\delta^B$)	0.608 (0.310)		0.677 (0.217)		0.759 (0.165)		0.771 (0.142)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. Within and between regressions are given by equations (16) and (17) respectively.

Table 3: Contrast Estimator: Within and Between Village Regressions
Cross Sectional Results for Crop Consumption Data

Variable	Dependent Variable: Per Capita Crop Consumption				Dependent Variable: Ln Per Capita Crop Consumption			
	2000		1997		2000		1997	
	Within	Between	Within	Between	Within	Between	Within	Between
Household Specific Linear Shock (Interaction)	14.69 (67.01)	218.21 (240.30)	302.24 (88.42)	679.03 (226.58)	0.0024 (0.013)	0.058 (0.051)	0.101 (0.025)	0.320 (0.080)
Total Acreage	187.61 (15.09)	184.90 (51.85)	215.72 (19.53)	338.81 (48.67)	0.037 (0.003)	0.022 (0.011)	0.060 (0.006)	0.110 (0.017)
Village Level Main Effect of Linear Shock	-	-1261.65 (1328.76)	-	-2735.28 (1218.80)	-	-0.188 (0.283)	-	-1.244 (0.428)
Household Size	-246.26 (29.34)	43.90 (118.71)	-244.02 (31.10)	-48.61 (112.03)	-0.060 (0.006)	-0.0047 (0.025)	-0.077 (0.009)	0.001 (0.040)
Constant	9569.55 (699.49)	13509.18 (4233.25)	4825.79 (536.58)	3788.07 (2766.72)	8.984 (0.132)	9.203 (0.903)	8.344 (0.152)	7.904 (0.962)
R-squared	0.218	0.259	0.172	0.396	0.241	0.196	0.182	0.393
Number of Groups	107		107		107		107	
Number of Observations	1437		1496		1435		1485	
Contrast Estimate $(1-\delta^W/\delta^B)$	0.933 (0.316)		0.555 (0.198)		0.966 (0.226)		0.684 (0.111)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. Within and between regressions are given by equations (16) and (17) respectively.

Table 4: Household Fixed Effects Contrast Estimator
Within and Between Regressions

Variable	Dependent Variable: Per Capita Maize Consumption		Dependent Variable: Ln Per Capita Maize Consumption	
	Within Estimate	**Between Estimate**	**Within Estimate**	**Between Estimate**
Household Specific Shock (Interaction)	105.20 (41.66)	473.34 (117.39)	0.027 (0.018)	0.291 (0.057)
Total Acreage	52.44 (12.82)	331.02 (48.15)	0.031 (0.005)	0.096 (0.023)
Village Level Main Effect of Shock	-	-1869.02 (696.70)	-	-1.192 (0.337)
Household Size	-149.13 (28.74)	-189.49 (114.32)	-0.073 (0.012)	-0.248 (0.055)
Constant	576.79 (78.09)	878.49 (200.36)	0.386 (0.033)	0.793 (0.098)
R-squared	0.052	0.516	0.083	0.462
Number of Groups	107		107	
Number of Observations	1409		1391	
Contrast Estimate ($1-\delta^W/\delta^B$)	0.778 (0.104)		0.907 (0.064)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. Within and between regressions are given by equations (16) and (17) respectively. All variables for the household fixed effects specifications are in changes (to account for the household fixed effects).

Table 5: Household Fixed Effects Contrast Estimator
Within and Between Regressions

Variable	Dependent Variable: Per Capita Crop Consumption		Dependent Variable: Ln Per Capita Crop Consumption	
	Within Estimate	Between Estimate	Within Estimate	Between Estimate
Household Specific Shock (Interaction)	169.13 (91.83)	759.26 (203.57)	0.018 (0.016)	0.246 (0.047)
Total Acreage	103.09 (28.27)	267.94 (83.50)	0.031 (0.005)	0.080 (0.019)
Village Level Main Effect of Shock	-	-2592.51 (1208.13)	-	-1.019 (0.278)
Household Size	-378.24 (63.35)	-584.29 (198.07)	-0.079 (0.011)	-0.224 (0.046)
Constant	1932.44 (171.96)	2779.37 (347.39)	0.538 (0.030)	0.881 (0.080)
R-squared	0.059	0.424	0.096	0.477
Number of Groups	107		107	
Number of Observations	1412		1402	
Contrast Estimate ($1-\delta^W/\delta^B$)	0.777 (0.135)		0.927 (0.067)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. Within and between regressions are given by equations (16) and (17) respectively. All variables for the household fixed effects specifications are in changes (to account for the household fixed effects).

Table 6: How Well Are Villages Pooled Within Districts?
Contrast Estimators: Within and Between District Regressions
Results for 2000 (Levels) and Household Fixed Effects (Changes) Samples

Variable	Ln Per Capita Maize Consumption				Ln Per Capita Crop Consumption			
	2000		HH FE		2000		HH FE	
	Within	Between	Within	Between	Within	Between	Within	Between
Household Specific Shock	0.083	0.095	0.134	0.308	0.039	0.025	0.092	0.326
	(0.026)	(0.095)	(0.038)	(0.139)	(0.025)	(0.092)	(0.030)	(0.125)
Total Acreage	0.045	-0.19	0.040	-0.137	0.035	-0.030	0.037	-0.222
	(0.008)	(0.031)	(0.017)	(0.156)	(0.007)	(0.030)	(0.013)	(0.141)
Village Level Main Effect of Shock	-1.004	-0.798	-1.594	-1.381	-0.634	0.034	-1.320	-1.703
	(0.395)	(0.643)	(0.451)	(0.851)	(0.381)	(0.623)	(0.352)	(0.767)
Household Size	-0.058	0.110	-0.088	-0.220	-0.052	0.028	-0.110	-0.107
	(0.022)	(0.076)	(0.044)	(0.217)	(0.021)	(0.074)	(0.034)	(0.195)
Constant	7.442	3.296	0.438	1.119	8.816	10.283	0.631	1.112
	(0.651)	(6.460)	(0.082)	(0.388)	(0.628)	(6.255)	(0.064)	(0.350)
R-squared	0.424	0.664	0.234	0.714	0.342	0.567	0.302	0.705
Number of Groups	22		21		22		21	
No. of Observations	107		103		107		103	
Contrast Estimate $(1-\delta^W/\delta^B)$	0.126		0.565		-0.56		0.718	
	(0.916)		(0.232)		(5.87)		(0.142)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. Within and between regressions are given by equations (16) and (17) respectively. All variables for the household fixed effects specifications are in changes (to account for the household fixed effects).

Table 7: Townsend Type Regressions

Dependent Variable is Household Consumption Minus Leave Out Village Mean Consumption

Results for Only Logarithmic Specifications are Reported

Variable	Maize Consumption			Crop Consumption		
	2000	1997	HH FE	2000	1997	HH FE
Household Specific Shock	0.020 (0.013)	0.036 (0.024)	0.021 (0.016)	-0.001 (0.012)	0.068 (0.022)	0.012 (0.015)
Total Acreage	0.026 (0.003)	0.037 (0.005)	0.028 (0.005)	0.034 (0.003)	0.048 (0.005)	0.028 (0.005)
Village Level Main Effect of Shock	-0.012 (0.088)	-0.035 (0.147)	-0.140 (0.113)	0.084 (0.079)	-0.168 (0.140)	-0.060 (0.103)
Household Size	-0.049 (0.006)	-0.065 (0.009)	-0.071 (0.012)	-0.053 (0.005)	-0.065 (0.009)	-0.077 (0.011)
Constant	0.707 (0.148)	0.692 (0.160)	0.072 (0.034)	0.868 (0.133)	0.655 (0.153)	0.090 (0.031)
R-squared	0.128	0.115	0.073	0.202	0.146	0.087
No. of Observations	1436	1477	1391	1435	1485	1402

Notes: Standard errors are in parentheses. Regressions all control for household age-sex composition. Equation (19) shows the specifications run. All variables for the household fixed effects specifications are in changes (to account for the household fixed effects).

Table 8: Comparison of Estimates to Deaton (1990) and Townsend (1994)

Specification	Contrast Estimator Test		Deaton (1990) Test		Townsend (1994) Test	
	Reject?	p-value	Reject?	p-value	Reject?	p-value
Per capita maize consumption, 2000	*Reject*	*0.03*	*Reject*	*0.02*	*Cannot Reject*	*0.07*
Per capita maize consumption, 1997	*Cannot Reject*	*0.08*	*Reject*	*0.04*	*Cannot Reject*	*0.30*
Per capita maize consumption, FE	Reject	0.02	Reject	0.01	Reject	0.05
Ln (Per capita maize consumption), 2000	Cannot Reject	0.12	Cannot Reject	0.054	Cannot Reject	0.14
Ln(Per capita maize consumption), 1997	*Cannot Reject*	*0.08*	*Reject*	*0.04*	*Cannot Reject*	*0.13*
Ln(Per capita maize consumption), FE	Cannot Reject	0.24	Cannot Reject	0.13	Cannot Reject	0.19
Per capita crop consumption, 2000	Cannot Reject	1.00	Cannot Reject	0.83	Cannot Reject	0.95
Per capita crop consumption, 1997	Reject	0.001	Reject	0.0006	Reject	0.03
Per capita crop consumption, FE	Cannot Reject	0.13	Cannot Reject	0.07	Cannot Reject	0.12
Ln (Per capita crop consumption), 2000	Cannot Reject	0.635	Cannot Reject	0.85	Cannot Reject	0.91
Ln(Per capita crop consumption), 1997	Reject	0.0002	Reject	0.0001	Reject	0.002
Ln(Per capita crop consumption), FE	Cannot Reject	0.515	Cannot Reject	0.26	Cannot Reject	0.41

Notes: This table summarizes all the results, computed via the contrast estimator, a Deaton test (just a test on the significance of the within estimate) and a Townsend test (reported in Table 7). The rows of the table describe the specification. The specifications that are in italics are those where the contrast estimator gives different results to either of the other two types of tests.

Table 9: Dealing With Measurement Error
Within and Between Village Regressions
Unit of Observation is Averaged Quintiles Within a Village

Variable	**Ln Per Capita Maize Consumption**				**Ln Per Capita Crop Consumption**			
	2000		**FE**		**2000**		**FE**	
	Within	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**
Household Specific Shock	0.045	0.098	-0.002	0.156	0.020	0.049	0.004	0.118
	(0.036)	(0.050)	(0.050)	(0.056)	(0.031)	(0.045)	(0.049)	(0.051)
Total Acreage	0.026	0.015	0.018	0.054	0.044	0.010	0.045	0.048
	(0.010)	(0.012)	(0.019)	(0.031)	(0.009)	(0.011)	(0.018)	(0.025)
Village Level Main Effect of Shock	-	-0.407	-	-0.503	-	-0.106	-	-0.293
		(0.308)		(0.363)		(0.280)		(0.325)
Household Size	-0.059	0.038	-0.089	-0.255	-0.053	0.026	-0.089	-0.189
	(0.021)	(0.030)	(0.043)	(0.061)	(0.018)	(0.028)	(0.042)	(0.053)
Constant	8.063	7.342	0.386	0.747	8.904	9.073	0.481	0.715
	(0.418)	(0.661)	(0.091)	(0.114)	(0.354)	(0.601)	(0.089)	(0.100)
R-squared	0.121	0.177	0.124	0.359	0.199	0.179	0.115	0.349
Number of Groups	107		107		107		107	
No. of Observations	280		274		280		274	
Contrast Estimate	0.541		1.013		0.592		0.966	
($1-\delta^W/\delta^B$)	(0.436)		(0.321)		(0.735)		(0.416)	

Notes: Standard errors are in parentheses. Standard errors for the contrast estimators are computed using the delta method. Regressions all control for household age-sex composition. A unit of observation here is the averaged quintile of distance to piped water within a village.