RISK SHARING TESTS AND COVARIATE SHOCKS: DROUGHT, FLOODS, AND PESTS IN UGANDA

ETHAN LIGON

Abstract. Full risk-sharing implies that marginal utilities of expenditure (MUEs) have a simple factor structure; Pareto weights divided by a common price. But MUEs are inferred from expenditures. The standard approach to inferring MUEs and testing risk-sharing assumes homothetic preferences, unitary income elasticities, and identical price elasticities. This works for idiosyncratic shocks, which affect budgets but not prices, but not "covariate" shocks which change prices. We obtain the class of preferences which allows one to infer MUEs from expenditures. We estimate these using Ugandan data. With these MUEs risk-sharing tests of covariate shocks deliver sensible results; the standard tests do not.

ligon@berkeley.edu
University of California
Berkeley, CA 94720–3310

Keywords. Risk sharing, Two-way fixed effects, Non-homothetic demands, Uganda, Marginal utility of expenditures, aggregate shocks

Date: August 5, 2025.

1. Introduction

It is well known that in a setting with risk-averse households and uncertainty, an efficient allocation will eliminate any "idiosyncratic" risk, in the sense that given this allocation the ratio of any two households' marginal utilities will vary only in fixed proportion, regardless of the realization of the uncertain state. Suppose for example that Farmer A is unlucky and a hailstorm damages his rice crop. Aggregate resources are reduced by the size of the damage, even if all the damage accrued to A's crops. So Farmer A's marginal utility of rice will increase, but if efficiency prevails, then everyone else's marginal utility of rice will increase proportionally, even if the shock was "idiosyncratic" in that it only affected the crops of one farmer.

The risk-sharing problem is often framed in terms of a planning problem, with different households assigned some Pareto weights *ex ante* and the planner choosing allocations such that ratios of households' marginal utilities of different goods (e.g., A's marginal utility of rice to B's marginal utility of rice) are equal to the ratio of their Pareto weights. But the problem immediately decentralizes in the sense that we can describe it in terms of prices and expenditures. In particular, we can say that an efficient allocation (a) keeps ratios of any two households' marginal utilities of expenditures (MUEs) constant across date-states; and (b) that given these constant ratios, allocations (or household expenditures) depend *only* on common prices, faced by all the households efficiently sharing risk.[1]

Townsend (1994) pioneered the idea of testing this "efficient risk-sharing" hypothesis in the development literature, and Deaton (1990) showed that one can implement the test as a two-way fixed effects regression. The general approach involves regressing a measure of the (log) marginal utility of expenditures on the (log) average of households' marginal utilities of expenditures, allowing for household-specific intercepts. The idea is to directly test the proportionality of marginal utilities of expenditures implied by efficiency. And in particular this proportionality implies a testable exclusion restriction: the event of the "idiosyncratic" shock should *not* affect individual MUEs after controlling for its effect on the average MUE. Tests along these lines

[1]There are many different ways in which such efficient allocations could be implemented. Some of these would involve formal markets for contingent claims (Arrow & Debreu, 1954) or repeated exchange of some set of securities (Arrow, 1964; Arrow & Hahn, 1983) or forward markets (Townsend, 1978). These market may not even need to be complete (Levine & Zame, 2002). Neither must they be formal; various forms of informal exchange (Platteau & Abraham, 1987) or reciprocity (Cashdan, 1985) may well constitute a system of informal insurance approaches full efficiency, even in the complete absence of formal markets or legally enforceable contracts (Fafchamps, 1992; Ligon et al., 2002).

have become the stuff of textbooks (Bardhan & Udry, 1999; Ray, 1998) and have been conducted in settings all over the world, in hundreds of different studies.[2] When the "idiosyncratic shock" is some measure of household income (or often deviations from mean income) such tests reliably reject the efficient risk-sharing hypothesis—household-level variation in income has a statistically significant effect on household consumption expenditures. Despite this rejection, many economists would say that the *magnitude* of this effect is not typically very large—empirically, the elasticity of consumption expenditures with respect to income is in the ballpark of 5-20%. So one might take the position that most idiosyncratic risk is shared.

But what about risk that is not idiosyncratic? After all, that hailstorm could damage the crops of not just one farmer in the village, but the crops of most farmers. Such a shock is sometimes said to be "covariate," capturing the idea that if we condition on the shock affecting Farmer A then we'd expect some damage also to Farmer B. Notice that a covariate shock need not be universal—some households may be more affected than others.

If risk is shared efficiently, then how would we expect a covariate shock to affect outcomes? The central precept that marginal utilities of expenditure should vary in proportion will still hold, so if we were to regress log MUEs on the village average, allowing for household fixed effects, then efficient risk-sharing implies that the exclusion restriction should still hold. However, if the shock affects enough production in the village, then we might also expect the shock to have an effect on marginal utilities of expenditure because of its effect on local *prices*. Viewed from this perspective, a covariate shock is one that affects prices, while an idiosyncratic shock does not.[3]

This brings us to a practical difficulty: how do we measure MUEs? Existing approaches assume some form of homothetic preferences, which have the feature that MUEs can be written as a function of total expenditures. The most commonly encountered case is preferences exhibiting constant relative risk aversion (CRRA). We

[2]Some recent prominent examples in the development literature include Angelucci and De Giorgi (2009), Chiappori et al. (2014), Jack and Suri (2014), Kinnan (2022), Mazzocco and Saini (2012), Munshi and Rosenzweig (2016), Samphantharak and Townsend (2018), and Santàeulalia-Llopis and Zheng (2018), while earlier studies are covered in surveys such as Alderman and Paxson (1994), O. P. Attanasio and Weber (2010), Dercon (2005), Morduch (1995), and Townsend (1995). The technique we attribute to Townsend arguably started in macro or international (O. Attanasio & Davis, 1996; Cochrane, 1991; De Giorgi et al., 2020; Lee & Lee, 2025; Mace, 1991; Obstfeld, 1993) and is still used. Though testing for risk-sharing across households is the classic approach, a similar setup is used in other settings, for example to test for sharing across people or generations in the unitary household model (e.g., Altonji et al., 1992; Hayashi et al., 1996; Dercon & Krishnan, 2000; Lise & Yamada, 2019; Theloudis et al., 2025).

[3]Appendix C.5 documents such a positive association between reported shocks and farmgate prices.

demonstrate in Section 2 that in the CRRA case this leads to a test involving a two-way fixed effects (TWFE) regression of log total expenditures on a set of household fixed effects and a set of market-time effects. If preferences are correctly specified, then the market-time effects will capture the effects of changes in prices.

But CRRA preferences are very restrictive: among other things all goods must have unitary income elasticities, and expenditure shares for different consumption goods will not vary with income. But Engel (1857) and a host of more recent evidence emphatically asserts (e.g., Jensen & Miller, 2008) that utility functions are *not* homothetic,[4] and so the usual risk-sharing regression is mis-specified. In particular, without homothetic preferences prices affect total expenditures in a non-separable way, so that the market-time effects of the TWFE regression cannot control for variation in prices, and the disturbance term in the TWFE regression will depend on prices. Using data from Uganda, we show that this mis-specification can lead to surprising results: in the usual specification covariate shocks such as droughts, floods, or pests appear to lead to welfare improvements (Section 7).

Can we fix this problem? Is it possible to use data on expenditures to construct more general measures of non-homothetic MUEs, and use these to obtain a valid test of risk-sharing against covariate shocks? To answer this, we go back to some basic consumer theory to find the broader class of preferences which permits us to infer MUEs using only data on expenditures (and household characteristics). We are able to completely describe this class by showing the expenditure system can be expressed in a form known as the generalized Pexider functional equation (Aczél & Dhombres, 1989), and exploiting results from the theory of functional equations to describe the complete set of solutions to these equations which comprise valid demand systems. The solutions take one of two forms: (i) a family of semiparametric demands which generalize demands associated with CRRA utility; and (ii) a second family of semiparametric demands which generalize Stone-Geary. Both families are non-homothetic, and together exhaust the class of demands from which MUEs can be constructed using expenditure data. We use the first family, which is amenable to estimation using linear methods; the second is not.

The problem which concerns us in this paper involves three distinct elements. First, the structure of actual preferences isn't homothetic, so total expenditures depend on prices in a more complicated way than is assumed by the usual risk-sharing

[4]We present additional evidence on this point in Section C.2.

regression—in particular not all goods have unitary income and price elasticities. Second, covariate shocks cause changes in prices. These first two elements are enough to generate correlations between covariate shocks and the consumption aggregate, even if risk is perfectly shared. Third, in practice the constructed sum of consumption expenditures is not complete, and excludes certain elastic goods or services, so that the constructed sum taken together has an income elasticity less than one. The consequence is that shocks that increase prices for observed inelastic goods will tend to increase observed expenditures.

Section 2 discusses full risk-sharing in the general case. Here the theory is remarkably clear on two basic points, both of which I believe to be novel. First, one cannot sum up expenditures on goods that have different income elasticities and use this sum to construct MUEs. Second, in order to use time effects to handle the effects of (possibly unobserved) prices on item-level expenditures, we must be able to express the system of expenditures as an additively separable function of the MUE and prices. In Section 3 we show that the requirement that (some transformation of) item-level expenditures have this separability property (along with some standard regularity conditions) is equivalent to the utility function taking one of two particular semi-parametric forms, but only one of these forms can be estimated using linear methods. Importantly, this class of utility functions nests the homothetic forms previously used in tests of risk sharing. This class of utility functions then dictates the form of demands and the risk-sharing test. Each of these two points is independent of actual risk-sharing arrangements.

We next turn to empirics: Given the form demands must take if we're to control for prices using time effects, how can we construct estimates of MUEs? The equations for log expenditures take a form that Hansen (2022) calls a factor model with additional regressors. Section 3.4 discusses the estimation of MUEs using expenditure data. We use the data from Uganda to construct estimates of MUEs, which are the natural objects to use in a test of risk-sharing. We conduct this test, and demonstrate that this more general demand system delivers much more sensible results in a test of risk-sharing than does the standard CRRA approach.

Finally, our construction of MUEs is independent of the risk-sharing hypothesis, and so these objects could be used in many other settings, or in the estimating and testing of the many dynamic models which put structure on the evolution of MUEs over time. We offer some thoughts and suggestions about ways in which these might proceed.

2. Risk sharing (The General Case)

Suppose that preferences are "regular" (i.e., can be represented by an increasing, concave, continuously differentiable utility function). More restrictively, assume that preferences are von Neumann-Morgenstern and intertemporally separable.

Following Townsend (1994), consider the problem facing a social planner in an environment with uncertainty; at date t state $s \in \{1, 2, \ldots, S\}$ is realized with probability $\Pr_t(s)$. The planner maximizes a weighted sum of households' utilities, with the (Pareto) weight θ_i associated with household i's utility. Utility at date t is discounted by some $\beta_t \in (0, 1)$. At each date-state (t, s) the planner allocates a consumption *budget* to each household, but has to respect the aggregate resource constraint that the sum of all expenditures must be less than some given quantity $\bar{x}_t(s)$. For household i at date t in state s let the budget allocated be $x_{it}(s)$. Given this budget and taking prices as given the household then solves the usual consumer problem.

In general, indirect utility within the period will depend not only on the size of the budget, but also a complete vector of prices $p_t(s)$ and household characteristics z_{it}. Then the planner's intertemporal problem can be written

$$\max_{x_{it}(s)} \sum_i \theta_i \sum_t \beta_t \sum_s \Pr_t(s) V(x_{it}(s), p_t(s); z_{it}(s)) \tag{1}$$

subject to the aggregate budget constraint $\sum_i x_{it}(s) = \bar{x}_t(s)$ for all t, s.

Let $\nu_t(s) = \mu_t(s)\beta_t \Pr_t(s)$ be the multiplier on the aggregate resource constraint at time t in state s. First order conditions associated with the assignment of $x_{it}(s)$ are

$$\theta_i \frac{\partial V}{\partial x}(x_{it}(s), p_t(s); z_{it}(s)) = \theta_i \lambda(x_{it}(s), p_t(s); z_{it}(s)) = \mu_t(s) \qquad \text{for all } (i, t, s),$$

where $\lambda(x, p, z)$ is a function that can be interpreted as the household's marginal utility of expenditures (MUE), and where $\mu_t(s)$ is the shadow price associated with the aggregate budget within date-state (t, s). Taking logarithms and rearranging,

$$\log \lambda_{it}(s) = \log \mu_t(s) - \log \theta_i. \tag{2}$$

Equation (2) is the hallmark of full risk-sharing, expressing the simple factor structure of optimal allocations. It also immediately lends itself to testing: The right hand side can be estimated using panel data with two-way fixed effects (time and household), assuming only that prices are common.

Everything up to this point is standard and implied by Pareto optimality, provided only that agents are risk averse and have well-behaved preferences which are separable across dates and states.

However, before taking the predictions of full risk sharing to data one must take a stand on how to construct the MUE function $\lambda(x, p, z)$. Townsend adopted a representation of momentary household utility which depends only on the consumption aggregate, normalized by a price index and a scalar function of household characteristics; this is equivalent to assuming a homothetic utility function (Blackorby & Donaldson, 1988). The empirical literature has mostly followed his example. Typical risk-sharing tests following Townsend assume homothetic or (less often) quasi-homothetic (e.g., Ogaki & Zhang, 2001; Zhang & Ogaki, 2004) preferences, for example assuming the household indirect utility function takes the "Constant Relative Risk Aversion" (CRRA) form

$$V(x, p, z) = \frac{(x/(\pi(p)g(z)))^{1-\gamma} - 1}{1-\gamma},$$

where $\pi(p)$ is a scalar price index, and where $g(z)$ is a scalar function mapping household characteristics into "adult equivalents".[5] Then the household's marginal utility of expenditures is given by

$$\lambda(x, p, z) = \frac{x^{-\gamma}}{(\pi(p)g(z))^{1-\gamma}}.$$

Substituting this expression into (2) and re-arranging yields an estimating equation of the form exploited by Townsend and Deaton,

$$\log x_{it} = \frac{1}{\gamma}\log\theta_i - \left[\frac{1}{\gamma}\log\mu_t + \frac{1-\gamma}{\gamma}\log\pi(p_t)\right] - \frac{1-\gamma}{\gamma}\log g(z_{it}). \tag{3}$$

So we estimate this by regressing the log of the consumption aggregate on household fixed effects (which identify $\gamma^{-1}\log\theta_i$) and time (or perhaps village-time) effects (which identify the term in square brackets involving only prices), and some known function g of observed household characteristics.

From the comparison of the two equations (2) and (3) three important points emerge. First, that the Townsend/Deaton risk-sharing regression is a special case of the more general (2). Second, that the marginal utility of expenditures λ already automatically incorporates information on household characteristics that may affect demand, and does so much more flexibly than does the $g(z)$ function that appears in the Townsend approach (c.f., Lewbel, 2010), as $g(z)$ cannot vary across different goods. Third, in general λ also depends on the entire vector of prices p, while in the CRRA case prices affect total expenditures only via a single scalar price index $\pi(p)$.

[5]Townsend (1994) actually works principally with exponential or CARA utility, with $V(x, p, z) = -\frac{1}{\sigma}\exp[-\sigma(x/\pi(p) - g(z))]$ which delivers a regression specified in levels rather than logs of total expenditures, but the subsequent literature has generally adopted the CRRA specification.

Townsend-style risk-sharing tests are generally implemented by adding some measure of a "shock" to (3), and testing the exclusion restriction (idiosyncratic shocks shouldn't affect MUEs). But even if perfect risk-sharing doesn't hold (we can think of this as the Pareto weights θ_i varying with the state), total expenditures in this framework depend only on Pareto weights (reflecting households' relative wealths), on prices p (capturing aggregate shocks to demand and supply), and on characteristics z (which may function as demand "shifters"). So "shocks" can affect total expenditures only via one of these three channels.

Under a maintained hypothesis of full risk-sharing, idiosyncratic income variation will be insured, but variation in prices will still affect expenditures, implying a regression of the form

$$\log \lambda_{it} = \log \mu_t - \log \theta_i + \delta \text{Shock}_{it} + e_{it}. \tag{4}$$

When we observe λ_{it}, we can implement this regression simply by adding two-way fixed effects—time effects account for $\log \mu_t$ and prices, while household fixed effects account for $\log \theta_i$. This leaves a disturbance term e_{it} which can be interpreted as either (or both of) measurement error in the dependent variable, or the effects of time-varying unobserved household characteristics on expenditures. In either case full insurance implies the exclusion restriction $\delta = 0$. (Here we also assume that the shock doesn't affect unobserved household characteristics.) Note that in this case full insurance implies $\delta = 0$ regardless of whether the shock is idiosyncratic (doesn't affect prices) or covariate (affects prices).

What about the case in which we observe λ_{it}, but there's not full insurance? In this framework a covariate shock can affect households in two different ways. First, via a common price channel—a drought may induce a local supply shock, and changes prices (and MUEs) for all households. And second, via an idiosyncratic income channel—farmers who grow, e.g., more or less maize will have their incomes affected to a greater or lesser extent, and if their incomes are not fully insured this will induce idiosyncratic variation in their MUEs. This correlation implies $\delta \neq 0$, in which case we would reject full insurance, and interpret δ as the average of the idiosyncratic response.

Now, consider the special case of CRRA preferences, for which MUEs can be expressed in the separable form $\log \lambda(x, p, z) = -\gamma \log \left(\frac{x}{\pi(p) g(z)} \right) - \log \pi(p) - \log g(z)$.

Adding a "shock" to the right-hand side of (3) yields the CRRA risk-sharing regression

$$\log x_{it} = \frac{1}{\gamma}\log\theta_i - \frac{1}{\gamma}\left[\log\mu_t + (1-\gamma)\log\pi(p_t)\right] + \delta\text{Shock}_{it} - \frac{1-\gamma}{\gamma}g(z_{it}) + \frac{1}{\gamma}e_{it}, \tag{5}$$

where now the *joint* hypothesis of full risk-sharing and CRRA preferences implies $\delta = 0$. Suppose that there's full risk-sharing but preferences are *not* CRRA. If the shock is idiosyncratic (and so doesn't affect prices, and also doesn't affect characteristics z) then we would still expect $\delta = 0$. But if the shock affects prices then the exclusion restriction fails, because in this case the single price index $\pi(p)$ can't account for the effects of changes in prices on the composition of expenditures. It follows that the disturbance term e_{it} must be a function of those prices.

More particularly, in the face of increased prices expenditures on inelastic goods such as food will *increase*. And since most of the expenditure items we have household-level data on are different sorts of food, we might expect a measure of total food expenditures to be positively correlated with "covariate" negative shocks which increase local prices, such as drought, floods, pests, or changes in prices for agricultural inputs. To put a finer point on it, if the measure of x_{it} is not expenditures on *all* non-durable goods and services, but is expenditures on a subset of goods which have inelastic demands, then by definition expenditures on that subset will tend to increase with increases in prices of those goods. Per Engel (1857) the subset "food" would be a good example, with more recent evidence including McKenzie (2003) and Thomas and Frankenberg (2007). And in this case we would predict that any shock that causes increases in food prices will be associated with $\delta > 0$.

So, to sum up: Covariate shocks affect prices and hence expenditures in ways that aren't properly accounted for in a specification of the risk-sharing regression that assumes homothetic preferences (i.e., that prices have no effect on expenditures). Thus, even if there *is* full insurance covariate shocks may be correlated with total expenditure. Further, if the constructed consumption aggregate excludes some elastic goods in particular, then we would expect the correlation between covariate shocks and the constructed consumption aggregate to be positive.

3. Inferring MUEs from Expenditure Data

The problem we've identified is that the usual risk-sharing regression provides a joint test of full risk-sharing *and* CRRA utility. Assuming CRRA utility allows us to use nothing more than panel data on expenditures (and perhaps household demographics) to construct risk-sharing tests based on panel estimation of two-way fixed

effects. In practice these are highly desirable properties. Are there non-homothetic utility structures which could more flexibly account for demand responses to changes in prices while at the same time preserving these desirable properties?[6] In this section we establish that the answer is "yes", and obtain the complete class of demand systems which (i) are implied by some regular utility function; and (ii) can be constructed using (just) data on consumption expenditures (and perhaps household demographics).

3.1. **Preliminaries.** Here we begin to develop the basic theory necessary to describe the class of demands that interests us. As a first step we set aside any explicit consideration of household characteristics, and focus on demand systems that are functions of just budget and prices. There's no loss of generality in this. With time-separable von-Neumann-Morgenstern preferences we can simply describe utility functions and demand implicitly conditioning on those characteristics; we reintroduce those characteristics as demand-shifters when we turn to empirics.

Thus, let $\mathcal{U}_n$ be the set of strictly increasing, strictly concave, twice-continuously differentiable functions mapping $\mathbb{R}^n_+$ into $\mathbb{R}$, and call $\mathcal{U}_n$ the set of *regular* utility functions over $\mathbb{R}^n_+$.

For a household with a utility function $U \in \mathcal{U}_n$ with a total budget $\bar{x} > 0$ facing prices $p \in \mathbb{R}^n$, a Lagrangian formulation of the *consumer's problem* is to solve $\max_{c \in \mathbb{R}^n_+} U(c) + \lambda(\bar{x} - p \cdot c)$, with λ the Lagrange multiplier or MUE that we would like to obtain for our risk-sharing test.

We can express a solution to the consumer's problem in terms of demand functions which depend on prices and λ. This form of demands was advocated by Ragnar Frisch, so we might say that Frischian demands map the product of positive quantity λ and n prices into n quantities demanded. We say that

Definition 1. An n-vector of Frischian demands $f(p, \lambda)$ is *rationalized* by U if there exists a $U \in \mathcal{U}_n$ such that

$$u_j(f(p,\lambda)) := \frac{\partial U}{\partial c_j}(f(p,\lambda)) = p_j \lambda \tag{6}$$

for all p in any open subset of $\mathbb{R}^n_+$ and $\lambda > 0$ for $j = 1, \ldots, n$.

Similarly, we say a given f is *rationalizable* if there exists a $U \in \mathcal{U}_n$ which rationalizes f.

[6] A complementary approach taken by Atkin et al. (2024) exploits non-homothetic utility structures which can flexibly respond to changes in relative prices, but delivers money-metric utility changes, rather than MUEs, and so is inappropriate for measuring risk-sharing.

Definition 1 basically[7] requires that demands be interior solutions to the problem of maximizing some regular utility function subject to a budget constraint. If a consumer has a utility function U, and solutions to that consumer's problem are characterized by the first order conditions (6), then these demands will also be solutions to this consumer's problem.

From the form of the right-hand side of (6) it's apparent that Frischian demands must be homogeneous of degree zero in $(p, 1/\lambda)$. It follows that we can simplify the expression of Frischian demands, as we have $f(p, \lambda) = f(p\lambda, 1)$, which can then simply be written as $f(p\lambda)$. This same point follows from the observation that at the optimum the vector of marginal utilities u can be inverted to yield $f(p, \lambda) = u^{-1}(p\lambda)$, so that we know we can write Frischian demands as simply $f(p\lambda)$.

3.2. **MUEs for risk-sharing tests.** The problem: we observe consumption expenditures $\{x_j\}$ for some (but perhaps not all) goods j. From these data we wish to infer values for MUEs; further, we want to be able to use these MUEs in a risk-sharing test that can be implemented using two-way fixed effects.

The fundamental risk-sharing equation is

$$\log \lambda = \log \mu(p) - \log \theta,$$

but we don't directly observe λ, only expenditures. So if we're to preserve the risk-sharing test we need to be able to write some transformation of expenditures as an additively separable function of prices and $\log \lambda$. Note that CRRA utility does exactly this as expenditures for good j in the CRRA case satisfy $-\gamma \log x_j = \log \lambda + (1-\gamma) \log p_j$, or (summing over goods) $\log x = \frac{-1}{\gamma} \log \lambda + \log \pi(p)$ for some linearly homogeneous price index $\pi(p)$. However, other more general preferences will also work. What are these preferences?

In general, for any good j we need functions (ϕ_j, a_j) such that

$$\phi_j(x_j) = \log \lambda + a_j(p). \tag{7}$$

If there exist functions (ϕ_j, a_j) satisfying (7), then we can substitute into the fundamental risk-sharing equation, obtaining

$$\phi_j(x_j) = [\log \mu(p) + a_j(p)] - \log \theta.$$

[7]The definition doesn't seem to allow for important real-world cases, such as two goods which are perfect substitutes (the rationalizing utility function wouldn't be strictly concave). However, extending the consumer's problem to allow for household production allows us to cover this and a broad class of other cases (Ligon, 2020).

Since the term in brackets varies only with prices this can serve as the basis of the kind of risk-sharing test that we're after, with time-effects identifying $\log \mu(p) + a_j(p)$ and household fixed effects identifying $\log \theta$. The key property is (7); this is a special case of a more general property we'll call λ-separability.[8]

Definition 2. The Frischian expenditures on good j, $x_j(p,\lambda) \equiv p_j f_j(p\lambda)$ are λ-*separable* if there exist functions (ϕ_j, a_j, b_j) such that

$$\phi_j(x_j(p,\lambda)) = a_j(p) + b_j(\lambda), \tag{8}$$

with ϕ_j continuously differentiable and a_j either varies with p or is equal to zero.

Note that while rationalizability is a property of the entire system of demands and expenditures, (λ-) separability is a property of a particular good. In particular it's possible that some but not all demands or expenditures are (λ-) separable.

3.3. **Demands and utilities when expenditures are λ-separable.** Exploiting the fact that expenditures must be linearly homogeneous, it turns out that one can write any rationalizable λ-separable expenditures in the form

$$k(p+\lambda) = g(\lambda)\ell(p) + h(p),$$

which is called the *generalized Pexider* equation (Appendix A). This gives us a single functional equation in two variables, which can be solved for the four functions g, h, k, and ℓ. Exploiting this allows us to describe all rationalizable demands and utilities when expenditures are λ-separable:

Theorem 1. *If demands are rationalizable (Definition 1) and expenditures on good j are λ-separable (Definition 2) with ϕ_j increasing; $a_j(p)$ either non-constant or zero, and continuous at a point; and with b_j continuous at a point, then transformation functions ϕ_j, Frischian demands f_j and rationalizing marginal utility u_j must belong to one of the following two cases of demands for positive constants α_j and β_j, and constants σ_j:*

1. *(Constant Frisch Elasticity): $\phi_j(x_j) = \log(x_j)$; $f_j(p,\lambda) = (\alpha_j/(\lambda p_j))^{\beta_j}$; and $u_j(c) = \alpha_j c_j^{-1/\beta_j}$.*
2. *(Generalized Stone-Geary): $\phi_j(x_j) = x_j^{\beta_j}$; $f_j(p,\lambda) = [(\alpha_j/(\lambda p_j))^{\beta_j} + \sigma_j]^{1/\beta_j}$; and $u_j(c) = \alpha_j \left(c_j^{\beta_j} - \sigma_j\right)^{-1/\beta_j}$.*

Proof. See Appendix A. □

[8]This property generalizes what Browning et al. (1985) calls "Case 2" demands, discussed in detail in Ligon (2016b).

3.3.1. *Rationalizing Utility Functions.* The labels of the different cases in Theorem 1 indicate names for the rationalizing utility function U having marginal utilities $u_j(c)$; an example of "Constant Frisch Elasticity" (CFE) utility can be written as $U(c) = \sum_{j=1}^{n} \alpha_j \beta_j \frac{c_j^{1-1/\beta_j}-1}{\beta_j - 1}$.[9] The CFE system generalizes the Constant Elasticity of Substitution (CES) system (take $\beta_j = \beta$), of which the Cobb-Douglas system is a limiting case (take $\beta \to 1$, applying L'Hôpital's rule). Both the CES and Cobb-Douglas cases are homothetic, and consistent with the CRRA indirect utility function. Finally, the "Generalized Stone-Geary" case gives what is, to the best of my knowledge, a marginal utility function which has not previously appeared in the literature. This case gives demands which are not linear in parameters, which may limit its usefulness in applied empirical work. However, when $\beta_j = 1$ one obtains the quasi-homothetic Stone-Geary utility function, which suggests that it could be used to explore the behavior of Engel curves, perhaps exploiting a Box-Cox approach to estimation.

3.4. **Estimating** $\log \lambda$**.** Theorem 1 gives us Frischian quantities demanded as a function of prices and MUEs, but our present motivation for using these demands is to obtain estimates of MUEs from these demand systems using only expenditures. Other practical concerns are also important: we do not wish to claim that we observe expenditures for all goods, just a subset; the expenditure system should allow for rich and unknown patterns of substitution; the demand system should allow for expenditures on some goods to be zero.

To this end, imagine extending the demand system by imagining that demands of Theorem 1 to be for goods which may be purchased at a price p_j, or it may be produced at a unit cost $P_j(p)$ by using a vector of inputs having prices p, with $P_j(p)$ linearly homogeneous. The expenditure systems of the theorem are λ-separable by construction, so we know that after transforming expenditures by taking logarithms (CFE) or raising expenditures for j to a power β_j (Generalized Stone-Geary) we will obtain an expression which is additively separable in λ and p. An exact parametric form governing the relation between λ and expenditures is given by the theorem, but in the extended demand system prices can influence expenditures via any linearly homogeneous function $\tilde{a}_j(p)$.

[9]Ligon (2020) gave this name to a general form of this utility function, but special cases include the "direct addilog" of Houthakker (1960) or the "constant relative income elasticity" form of Caron et al. (2014).

Accordingly, using Theorem 1, the condition that expenditures be λ-separable implies that expenditures must take one of two forms:[10]

(1) $\log x_j = \log \tilde{a}_j(p) - \beta_j \log \lambda$; or
(2) $x_j = \left[\left(\frac{\alpha_j}{\lambda}\right)^{\beta_j} + a_j(p) \right]^{1/\beta_j}$.

The second form (generalized Stone-Geary) does not easily allow us to estimate log MUEs using data on expenditures. But the first does, using standard methods of estimation with panel data. Here, instead of the panel dimensions varying over households and time periods, they vary over households and items of consumption expenditure. In particular, for notational convenience let $w_{it} = -\log \lambda_{it}$, and index goods by j and households by i. Then we have for *each* period t

$$\log x_{it}^j = a_j(p_t) + g_j(z_{it}) + \beta_j w_{it} + \epsilon_{it}^j, \tag{9}$$

$a_j(p_t)$ measures the effect of common time t prices p_t on expenditures, where z_{it} are observable household characteristics, such as household size and composition, where β_j gives us (minus) the elasticity of expenditures on good j with respect to MUE, and where ϵ_{it}^j can be regarded as measurement error in expenditures, or perhaps the effects of unobserved household characteristics on those expenditures. As prices are assumed to be common, the term $a_j(p_t)$ we can account for using good-time effects (or good-time-market effects, if it is thought that prices vary across spatially distinct market regions, in which case t should be thought of as indexing different market-periods).[11]

We call the system (9) the *Constant Frisch Elasticity* (CFE) expenditure system, as it is an example of the systems considered by Frisch (1959) with its chief distinguishing characteristic the fact that the elasticities (the β_j) of expenditures with respect to MUE are constant.

Different strategies may be employed to estimate the functions g_j; our preference here is perhaps the simplest, which is to simply assume a linear form $g_j(z) = \gamma_j' z$ for some vector of parameters γ_j which may vary across goods. Then estimation proceeds

[10]Further, Ligon (2020) demonstrates that expenditures on an input k used to produce a top-level good j with CFE demand will itself take the form $\log x_k = \log a_{jk}(p) - \beta_j \log \lambda$.

[11]In principle rather than assuming that all households face the same prices and using latent variables one could use data on the prices (or marginal costs, if pricing was non-linear) to estimate the $a_j(p_t)$ if it was thought that the law of one price was unlikely to hold. But this would be a high-dimensional problem, and one of the chief attractions of the CFE demand system is that it can be estimated *without* observing prices. Further, some heterogeneity in unobserved prices can be accommodated. If, for example, differently situated households faced proportional iceberg costs for any goods acquired this would lead to observationally equivalent outcome of facing the same prices, but having a larger MUE.

in two steps. First, estimate the regression

$$\log x_{it}^j = a_t^j + \gamma_j' z_{it} + e_{it}^j, \tag{10}$$

with a_t^j a set of good-time-market effects, and e_{it}^j a composite error term, since from (9) we have $e_{it}^j = -\beta_j w_{it} + \epsilon_{it}^j$.

We consider three different possibilities for our second step. First, summing the error terms e_{it}^j over different goods j for each household i at t we obtain

$$w_{it} = \frac{\sum_j e_{it}^j}{\sum_j \beta_j} - \frac{\sum_j \epsilon_{it}^j}{\sum_j \beta_j}.$$

We can assume that $\frac{1}{J}\sum_j \beta_j = 1$ without loss of generality, as this simply amounts to choosing the scale on which we measure w_{it}. Then a natural analog estimator of w_{it} is simply the average of the estimated residuals $\hat{e}_{it}^j$ from (10), or

$$\tilde{w}_{it} = \frac{1}{J}\sum_j \hat{e}_{it}^j.$$

This will be an unbiased estimator under standard conditions (principally that (w_{it}, z_{it}) are orthogonal to the disturbance ϵ_{it}^j in (9)). Call this the "average" estimator of w_{it}.

Second, if the variance of ϵ_{it}^j varies across goods, then (using the usual logic of generalized least squares) we can construct a more efficient estimator. Let σ_j^2 denote the variance of the error for the jth good. In this heteroskedastic case the weighted average

$$\hat{w}_{it} = \frac{1}{J}\sum_j \hat{e}_{it}^j/\sigma_j$$

will be a more efficient estimator of w_{it}. The standard deviation of the estimated residuals $\hat{e}_{it}^2$ computed from the sample of all households in all periods is a feasible estimator for σ_j. Call this the "weighted average" estimator of w_{it}.

But if accounting for variance improves efficiency, then the GLS logic encourages us to further consider covariance. This leads to a model similar to that of Jöreskog and Goldberger (1975) and our third estimator. Rather than averaging, this "factor analysis" approach focuses on the sample covariance matrix of the residuals. This has several benefits, including increased efficiency, but also because it yields estimates of the β_j as well as the MUEs, and because it allows us to compare ws over time. This comes at the cost of requiring some additional structure on the covariance of the ϵ_{it}^j disturbances. The factor analysis estimator is described in Appendix B. [12]

[12]Additional details of assumptions and methods are given in Ligon (2020), while code to compute estimates is provided by Ligon (2023).

Here we adopt the factor analysis approach estimator for our main results. This allows us to present estimates of β, and of MUEs that are comparable over time. But as it happens, in our present application the choice of estimator turns out to matter very little, as the correlations between the factor analysis estimates of the w_{it} and the average estimates are 0.97, while the correlation between the two averaging estimators is 0.99. Almost none of the results we present on MUEs vary importantly with the choice of estimator.[13]

4. Data from Uganda

A goal of this paper is to provide tools to estimate MUEs with only modest data requirements. In particular, we've devised methods which allow us to construct estimates of MUEs from a rich class of non-homothetic preferences using nothing more than data on expenditures and basic household demographics. These are the same data that are often used to construct MUEs for homothetic preferences, typically some measure of total consumption expenditures per capita or adult equivalent.

4.1. **Sample Data From Uganda.** Household MUEs can be estimated using a single cross-section, but our test of risk-sharing requires panel data on households over time. We use panel survey data from Uganda, covering eight waves which span the years 2005–2020. These surveys are called the Ugandan National Panel Survey, and are conducted by the Ugandan Bureau of Statistics part of the World Bank's *Living Standards Measurement Surveys* (LSMSs), which have now been conducted in many countries across many years (Deaton, 1997). Waves were not always collected at regular intervals, however, and only a subset of households appear multiple times. In particular, while we observe usable data from 5601 distinct households over this period, 1007 of these households appear only once, and so contribute nothing of value to the risk sharing test (though data on these households' consumption expenditures is still useful for estimating the CFE demand system). Table 1 describes the number of households observed in each round (after dropping those households missing the data to necessary to estimate the demand system), and how many households observed in a particular round also turn up in another. For example, looking at the first row of the table, we have data for 2911 households in 2005, and of these 1187 are also observed in 2019-20.

Overall, we have an eight year unbalanced panel with a total of 22,791 usable household-year observations. The unconditional probability of a sample household

[13]The sole exception is Figure 4, which takes as an input ws which are comparable over time.

TABLE 1. Households' Attrition in the Ugandan LSMS. Includes only households with data necessary to estimate demand system.

Round	2005	2009	2010	2011	2013	2015	2018	2019
2005-06	2911	2313	2094	2139	1444	1354	1252	1187
2009-10	—	2760	2295	2336	1532	1441	1319	1253
2010-11	—	—	2457	2206	1444	1355	1251	1188
2011-12	—	—	—	2687	1575	1477	1364	1291
2013-14	—	—	—	—	2974	2638	2415	2265
2015-16	—	—	—	—	—	3115	2726	2546
2018-19	—	—	—	—	—	—	3010	2708
2019-20	—	—	—	—	—	—	—	2877

observed in one wave not being observed in the next is about 0.14. The lack of balance in the sample is driven less by households attriting, and more because of the sample design, as some households were intentionally dropped and new "refreshment" samples were added in certain waves. Most particularly, inl the 2013–14 wave a third of the sample was replaced (this can be seen in Table 1). These sampling decisions are much more important to determining the probability of a random household remaining in the panel across rounds than the decisions or experience of the household itself, and in years where no part of the sample was intentionally dropped attrition is quite modest by the standards of panel surveys in low income countries (e.g., in 2019–20 attrition rates are less than 4%).

4.2. **Data for Estimating MUEs.**

4.2.1. *Expenditure Data.* Each wave of the LSMS includes a detailed elicitation of different kinds of food, plus two non-food items, "cigarettes" and "other tobacco." If one takes the union of all items across all eight waves there are 175 distinct items; however, many of these simply involves differences in language (e.g., "sweet bananas" vs "bogoya"), or what are actually differences in units (e.g., "matoke (heap)" vs "matoke (cluster)"). Finally, some simply involve mis- or variant spelling ("yogurt" vs "yogurt"). Reconciling these leads to a list of 130 distinct items. Eliminating items which are not observed in every round, or combining closely related items ("ground nuts (shelled)" and "ground nuts (pounded)") gives us a final list of 41 goods, listed in Figure C.1.

4.2.2. *Household Characteristics.* Table 2 reports on the characteristics of the sample households we use in our demand estimation, by year. These characteristics are chosen

because the household-level marginal utility of expenditures seems likely to depend not only on household size, but also on its composition.

TABLE 2. Mean characteristics of households in Uganda by year. Figures in parentheses are standard deviations. Households missing data on any of these characteristics or with inadequate data on food expenditures are excluded from the analysis.

(t/N_t)	Girls	Boys	Women	Men	Rural	log HSize
2005-06	1.62	1.61	1.22	1.06	0.73	1.53
2911	(1.48)	(1.49)	(0.80)	(0.83)	(0.44)	(0.64)
2009-10	1.79	1.83	1.29	1.15	0.75	1.65
2760	(1.53)	(1.60)	(0.81)	(0.91)	(0.43)	(0.60)
2010-11	1.91	1.91	1.34	1.18	0.78	1.70
2457	(1.60)	(1.61)	(0.86)	(0.94)	(0.41)	(0.59)
2011-12	1.85	1.83	1.31	1.16	0.80	1.67
2687	(1.57)	(1.58)	(0.80)	(0.91)	(0.40)	(0.58)
2013-14	1.62	1.62	1.27	1.13	0.74	1.57
2974	(1.50)	(1.50)	(0.76)	(0.87)	(0.44)	(0.61)
2015-16	1.32	1.33	1.20	1.06	0.75	1.38
3115	(1.39)	(1.39)	(0.76)	(0.89)	(0.43)	(0.71)
2018-19	1.51	1.46	1.26	1.13	0.75	1.53
3010	(1.40)	(1.36)	(0.71)	(0.87)	(0.43)	(0.60)
2019-20	1.49	1.43	1.27	1.11	0.77	1.51
2877	(1.40)	(1.32)	(0.70)	(0.86)	(0.42)	(0.60)
Pooled	1.63	1.62	1.27	1.12	0.76	1.56
22791	(1.49)	(1.49)	(0.78)	(0.88)	(0.43)	(0.63)

5. Estimates of the CFE Expenditure System

With the Ugandan data described above we estimate the β_j parameters and the values of $\log \lambda_{it}$ for every household-year. There are some details of our preferred specification. We allow for demands to depend on the observed household characteristics z_{it} which appear in Table 2. We also need to take a stand on what constitutes a "market," or the geographical area within which households are assumed to face the same prices. Uganda is divided into four regions (central, east, west, north), and we assume that each region is a distinct market, thus allowing prices to vary by market-year. To square this with the notation of the paper one should thus regard t as an index of market-years. In our estimation of (9) this implies that t indexes

$4 \times 8 = 32$ market-years, and that our specification thus includes $32 \times 41 = 1312$ market-year-good dummy variables.

Figure 1 shows our estimates of the β_j elasticities; note that we can easily reject the hypothesis that these are all equal, as they would be in the special case of CRRA—there's a wide range of income elasticities across different goods, even for the same household, and the least elastic (those with low values of β_j) are what we might expect (starchy staples, salt). Similarly, goods such as fresh milk, sweet bananas, coffee, oranges, and passion fruit all exhibit high elasticities. The ratios of the β coefficients are equal to ratios of income elasticities, implying that passion fruit is roughly three times as income elastic as cassava.

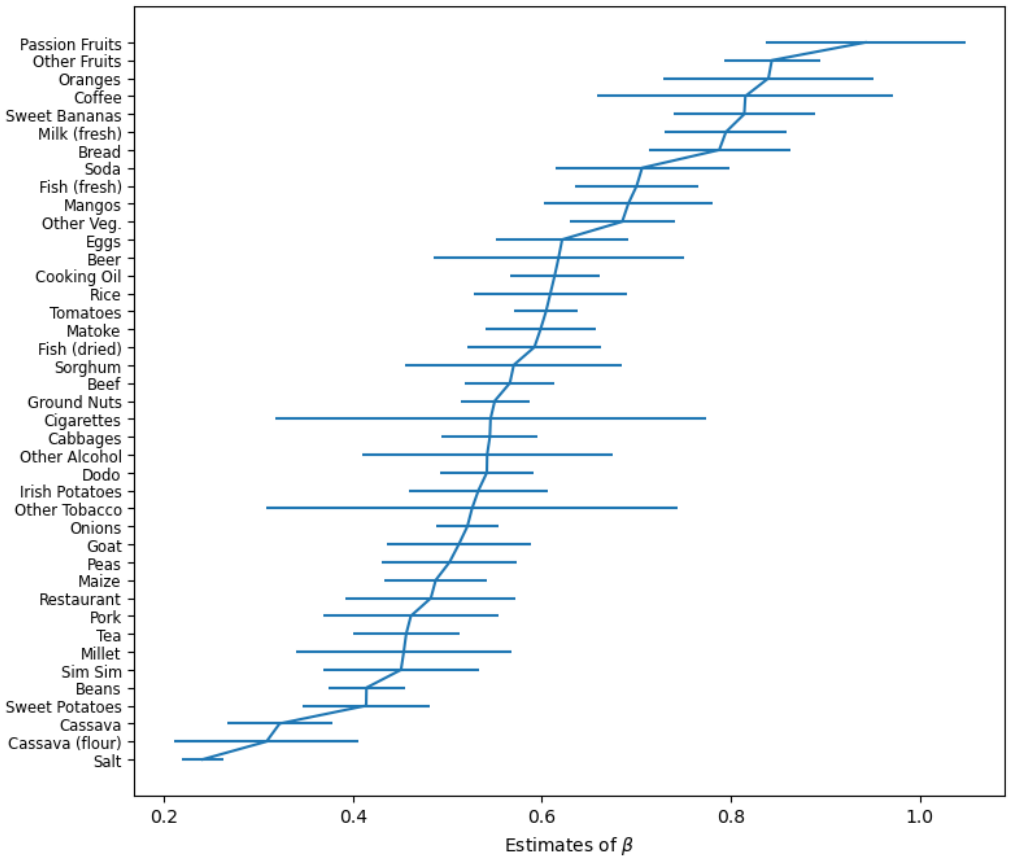


FIGURE 1. Estimates of Frishian elasticities β_j. These are proportional to income elasticities. Horizontal bars are 95% confidence intervals.

We estimate $w = -\log \lambda_{it}$ (we use the negative of the log MUE to make it easier to interpret the sign of the result—bigger is better) for every household in every year. Figure 2 presents kernel density estimates of the distribution of w for every

year. There is some evidence of this distribution shifting across years. And from these cross-sectional distributions we can't tell anything about how the position of a particular household changes over time.

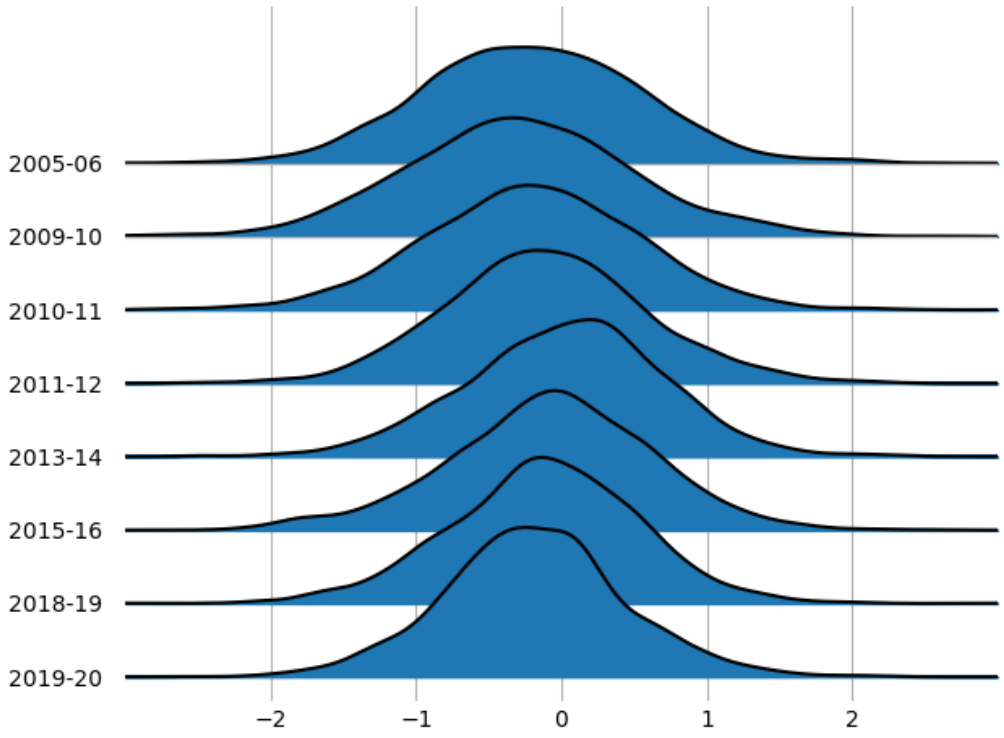


FIGURE 2. Distribution of w by year.

6. Direct Evidence on the Effects of Covariate Shocks

To use estimated MUEs to test whether households are insured against covariate shocks, the only additional data required is data on those shocks, which we will use below to test whether covariate shocks affect differently situated households differently, or whether the effects of these shocks are captured entirely by changes in the prices faced by all households, consistent with the hypothesis of full insurance.

Yet this approach to measuring the effects of shocks is indirect. This is meant to be a virtue, because of the limited data requirements and simplicity of the test. However, for the Ugandan panel we use we also have additional data on (i) subjective reports on the effects of particular shocks; (ii) how households coped with those shocks; (iii) farm-gate prices for some locally produced goods; (iv) local market prices for a variety of foods; and (v) information on the quantities of food households consume.

We use these additional data to directly trace out the effects of covariate shocks in Appendix C. Data limitations mean that such an account will be somewhat narrative and more informal than our main analysis. Nevertheless, in the course of this narrative we present evidence that in Uganda:

§C.2: Preferences are non-homothetic, with expenditure shares that vary with budget and prices. It follows that changes in income or prices (whether induced by shocks or not) can be expected to induce changes in the composition of households' food expenditures.

§C.3: Households report being affected by droughts, floods, agricultural pests, and agricultural prices with some frequency, and these shocks were indeed *covariate*, in the sense that the probability of such reports covaries across space within a given year.

§C.4: Covariate shocks such as droughts, floods, and pests had a primary effect on agricultural production, which in turn led to reported declines in household income and consumption. For example, 92% of households who reported drought said it affected their production, 80% said it affected their income, and 50% said it affected their consumption.

§C.5: Drought, pests, and adverse prices (but not floods) tended to lead to statistically significant increases in local farm-gate prices.

§C.6: Households which experienced covariate shocks responded primarily through methods of self-insurance: reducing consumption, drawing on savings, and increasing labor supply, all strategies one might employ in the face of increased prices. In contrast, those who experienced idiosyncratic shocks most often relied on help from friends and family.

§C.7: Most relative *market* prices for different foods were significantly different in places that experienced covariate shocks (including floods) versus those that did not.

§C.8: The recent experience of drought, floods, or pests had significant effects on the composition of diet, with negative consequences for dietary diversity and for nutrition.

7. Tests of Risk Sharing in the Face of Covariate Shocks

In this section we want to actually conduct the risk-sharing tests described above. We'll use two different approaches. The first is the classic test assuming CRRA utility: for these the dependent variable is the log of total expenditures. The second is the

test assuming our more general CFE preferences; for these the dependent variable is $w = -\log\lambda$.

7.1. **Effects of shocks on welfare last year.** The risk-sharing test we're concerned with is meant to provide a test of the null hypothesis that $\mathrm{E}(\log\lambda|p,z) = \mathrm{E}(\log\lambda|p,z,\mathrm{Shocks})$; that is, whether shocks affect log MUEs after we control for possible changes to prices or household characteristics. If demands are consistent with the CFE specification, then the welfare measures w we've constructed are estimates of $-\log\lambda$, and we are justified in using w as the dependent variable in the two-way fixed effects regressions we've described. When $\beta_j = \bar{\beta}$ for all goods j, then CFE demands will coincide with the special case of CRRA, and we will have $-\log\lambda$ proportional to $\log x$.

Table 3 describes the effects of different shocks on both w (equation 4, the CFE risk-sharing regression) and on $\log x$ (equation 5, the CRRA risk-sharing regression). Reported standard errors here and elsewhere are robust (Arellano, 1987).

Recall that if preferences are *not* CRRA then we'd expect the disturbance term in the TWFE regression to depend on prices. In particular, if income and price elasticities are not equal to one for all goods then we would expect that a regression of log total food expenditures on shocks which increased food expenditures would yield positive estimated values of δ. Similarly, if Frischian elasticities vary across goods but are constant as in the CFE system then we would expect the time-market effects in the regression to correctly capture the effects of prices on expenditures, so that estimated values of δ should be zero (if there's full risk-sharing).

The results of Table 3 confirm this reasoning rather dramatically. In fact, *every* covariate shock has a significant positive coefficient in the CRRA regressions, while *no* covariate shock has a significant coefficient in the CFE regressions, just as we would predict if there's full insurance and demands are CFE. This is strong evidence in favor of the view that these negative covariate shocks increase local consumption expenditures by causing an increase in some local prices,[14] and that inclusion of time-market effects in the CRRA specification does not control for these price effects correctly.

[14]See Appendix C.5 for direct evidence on this point.

TABLE 3. Effects of different shocks on welfare measures w and $\log x$.^{N,T,H,X,FE}

Shock	w (CFE)	$\log x$ (CRRA)
Health	-0.057^{**}	-0.028
	(0.024)	(0.019)
Theft	0.023	0.022
	(0.033)	(0.026)
Death (non-earner)	0.023	0.020
	(0.050)	(0.036)
Death of earner	0.033	0.007
	(0.089)	(0.065)
Income^{I}	0.056^{***}	0.075^{***}
	(0.007)	(0.006)
Drought	0.007	0.043^{***}
	(0.015)	(0.011)
Floods	0.030	0.091^{***}
	(0.033)	(0.027)
Pests	0.042	0.094^{***}
	(0.034)	(0.024)
Prices	-0.044	0.099^{***}
	(0.038)	(0.028)

* $p<0.10$;
** $p<0.05$.
*** $p<0.01$.

N Total observations are 22791 unless otherwise noted.

T All regressions involve 8 rounds and 4 markets.

H A total of 5604 distinct households, unless otherwise noted (but not all households appear in all rounds).

X All regressions include the number of boys, girls, men, women, log of household size, and a "rural" dummy as additional covariates.

FE Household Fixed Effects: Yes; Market-year Effects: Yes

I The regression on income has $N = 9721$, and $H = 3851$.

Among the idiosyncratic shocks, income is significant, has the expected sign, and has a magnitude similar to that found in other Townsend-style tests of full insurance. No other coefficients in either specification are significant, with the sole exception of Health for CFE. This last should not be a surprise, and neither should it be interpreted as a rejection of the risk-sharing model, as our estimation of the CFE demands did not include any information on health as a household characteristic. If health affects demands, in our specification it can do so either via a shock to the budget and thus w (which would be at odds with full insurance, but not the CFE demand specification) or via the disturbance term in the demand equations, which explicitly depends on unobserved household characteristics such as health. The data we have on expenditures is for the past week; the data we have on shocks is for the past year. Could this somehow cause the apparently abberant effects of positive shocks on expenditures, rather than something to do with prices? Perhaps spending increases as people attempt recovery from the shock?[15] We examine this by considering shocks at different temporal proximities to the expenditures. We construct a dummy variable which takes the value one if there's a reported covariate shock within the m months prior to the interview, and then re-estimate our risk-sharing regressions allowing m to vary from zero months up to twelve months.

Results are reported in Figure 3. This figure provides even stronger evidence favoring the CFE over the CRRA risk-sharing specifications. For the CRRA ($\log x$) specification, having had any covariate shock within the last year has a significant positive effect on log consumption, in every month. In contrast, for the CFE (w) specification, having any covariate shock within the last month has a rather large and significant negative effect on w. When the shock occurs in earlier months the magnitudes of the effects of shocks on w are uniformly smaller than the CRRA case, though some are positive and statistically significant. The full risk-sharing hypothesis holds that covariate shocks should influence w only through prices, so we could interpret the significant negative effect in the first month as evidence against the full risk-sharing hypothesis.

To try to get a better understanding of these patterns, Table 4 reports the same kind of regressions described by Figure 3 which allow for shock windows of zero through twelve months, but keeps the different kinds of shocks separate, as in Table 3 (except that income is dropped, as we don't have detailed data on the timing of

[15]Though here it's worth remembering that we measure non-durable expenditures, mostly food, as distinct from investment.

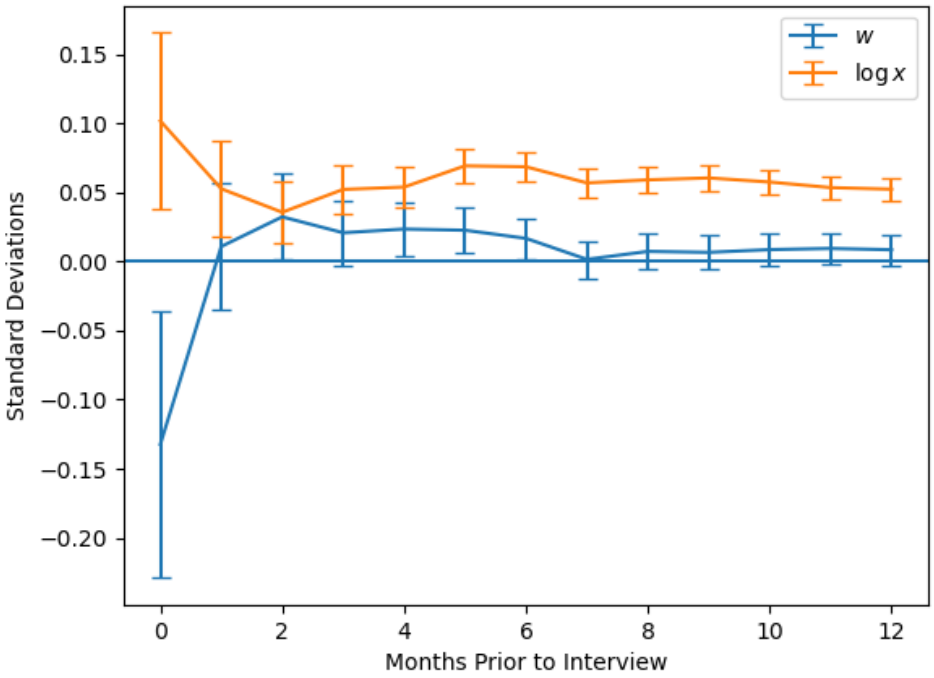


FIGURE 3. Effects of any covariate shock within the last m months on welfare measures, from a two-way panel regression otherwise specified as in Table 3. Scale is in standard deviations of the dependent variable (w or $\log x$). Error bars cover a span of two standard errors about the point estimate.

its receipt). As in that table, most of the covariate shock coefficients (the last four columns) are significant and positive in the CRRA specification at a 95% percent level of confidence, while only two are positive and significant at this level for the CFE specification reported in the subsequent Table 5.

TABLE 4. Effects of different shocks within the last m months on current log consumption expenditures.$^{N,T,H,X,\text{FE}}$

Months	Health	Theft	Death (non-earner)	Death (earner)	Drought	Floods	Pests	Prices
0	0.013	0.058	0.088	0.779***	−0.068	0.283***	0.291**	−0.220
	(0.079)	(0.092)	(0.202)	(0.218)	(0.097)	(0.102)	(0.120)	(0.198)
1	−0.008	0.007	−0.024	0.216	−0.015	0.104	0.203***	0.068
	(0.039)	(0.051)	(0.103)	(0.136)	(0.047)	(0.086)	(0.075)	(0.105)
2	−0.030	−0.010	0.015	0.345**	−0.022	0.119*	0.177***	0.090
	(0.032)	(0.040)	(0.076)	(0.175)	(0.030)	(0.061)	(0.057)	(0.064)
3	−0.019	−0.011	0.042	0.294**	−0.000	0.172***	0.156***	0.104**
	(0.028)	(0.036)	(0.061)	(0.142)	(0.023)	(0.049)	(0.046)	(0.053)
4	−0.031	−0.021	0.042	0.104	0.016	0.137***	0.134***	0.122***
	(0.026)	(0.033)	(0.052)	(0.124)	(0.019)	(0.041)	(0.041)	(0.046)
5	−0.034	−0.022	0.048	0.046	0.050***	0.116***	0.158***	0.097**
	(0.024)	(0.031)	(0.048)	(0.109)	(0.016)	(0.037)	(0.037)	(0.040)
6	−0.037*	−0.007	0.026	0.037	0.059***	0.080**	0.142***	0.120***
	(0.023)	(0.028)	(0.044)	(0.095)	(0.014)	(0.034)	(0.033)	(0.038)
7	−0.023	0.012	0.005	0.021	0.046***	0.061*	0.131***	0.109***
	(0.022)	(0.028)	(0.043)	(0.088)	(0.013)	(0.031)	(0.031)	(0.035)
8	−0.022	−0.001	0.018	0.048	0.052***	0.068**	0.129***	0.102***
	(0.021)	(0.027)	(0.041)	(0.079)	(0.012)	(0.030)	(0.029)	(0.033)
9	−0.019	0.004	0.026	0.030	0.056***	0.065**	0.121***	0.097***
	(0.020)	(0.027)	(0.039)	(0.074)	(0.012)	(0.029)	(0.028)	(0.031)
10	−0.018	0.010	0.028	0.020	0.049***	0.074***	0.118***	0.107***
	(0.019)	(0.027)	(0.037)	(0.069)	(0.011)	(0.028)	(0.027)	(0.029)
11	−0.025	0.014	0.032	0.026	0.046***	0.072***	0.104***	0.098***
	(0.019)	(0.026)	(0.037)	(0.067)	(0.011)	(0.027)	(0.025)	(0.029)
12	−0.028	0.022	0.020	0.007	0.043***	0.091***	0.094***	0.099***
	(0.019)	(0.026)	(0.036)	(0.065)	(0.011)	(0.027)	(0.024)	(0.028)

* $p<0.10$;
** $p<0.05$.
*** $p<0.01$.
N Total observations are 22791.
T All regressions involve 8 rounds and 4 markets.
H A total of 5604 distinct households (but not all households appear in all rounds).
X All regressions include the number of boys, girls, men, women, log of household size, and a "rural" dummy as additional covariates.
FE Household Fixed Effects: Yes; Market-year Effects: Yes

TABLE 5. Effects of different shocks within the last m months on current $w = -\log \lambda$.$^{N,T,H,X,\text{FE}}$

Months	Health	Theft	Death (non-earner)	Death (earner)	Drought	Floods	Pests	Prices
0	−0.057	0.239**	−0.072	1.013**	−0.367**	0.357*	−0.087	−0.528***
	(0.092)	(0.116)	(0.309)	(0.454)	(0.152)	(0.210)	(0.152)	(0.172)
1	0.014	0.045	−0.079	0.406**	−0.061	0.095	0.125	0.059
	(0.050)	(0.065)	(0.147)	(0.204)	(0.063)	(0.104)	(0.098)	(0.153)
2	0.018	0.060	−0.087	0.485**	0.005	0.062	0.114	0.057
	(0.042)	(0.053)	(0.107)	(0.214)	(0.042)	(0.074)	(0.078)	(0.115)
3	0.015	0.005	0.037	0.367**	−0.002	0.102*	0.111*	−0.086
	(0.038)	(0.048)	(0.090)	(0.176)	(0.031)	(0.059)	(0.062)	(0.087)
4	−0.027	−0.008	0.020	0.107	0.008	0.115**	0.076	−0.064
	(0.035)	(0.045)	(0.075)	(0.159)	(0.025)	(0.049)	(0.056)	(0.068)
5	−0.033	−0.014	0.027	0.082	0.018	0.077*	0.104**	−0.096*
	(0.032)	(0.041)	(0.069)	(0.138)	(0.022)	(0.044)	(0.051)	(0.056)
6	−0.042	−0.007	0.036	0.066	0.015	0.038	0.089*	−0.077
	(0.030)	(0.037)	(0.064)	(0.123)	(0.019)	(0.041)	(0.047)	(0.051)
7	−0.036	0.006	0.015	0.060	−0.000	0.014	0.053	−0.069
	(0.029)	(0.037)	(0.062)	(0.116)	(0.018)	(0.038)	(0.043)	(0.048)
8	−0.036	0.002	0.033	0.102	0.009	0.007	0.065	−0.071
	(0.028)	(0.035)	(0.058)	(0.108)	(0.016)	(0.037)	(0.042)	(0.044)
9	−0.036	0.003	0.073	0.084	0.006	0.020	0.065*	−0.072*
	(0.027)	(0.034)	(0.055)	(0.101)	(0.016)	(0.036)	(0.039)	(0.042)
10	−0.034	0.021	0.057	0.053	0.006	0.021	0.068*	−0.052
	(0.026)	(0.034)	(0.053)	(0.096)	(0.015)	(0.035)	(0.037)	(0.040)
11	−0.049*	0.019	0.049	0.078	0.008	0.021	0.054	−0.043
	(0.025)	(0.033)	(0.051)	(0.093)	(0.015)	(0.034)	(0.036)	(0.038)
12	−0.057**	0.023	0.023	0.033	0.007	0.030	0.042	−0.044
	(0.024)	(0.033)	(0.050)	(0.089)	(0.015)	(0.033)	(0.034)	(0.038)

* $p<0.10$;
** $p<0.05$.
*** $p<0.01$.
N Total observations are 22791.
T All regressions involve 8 rounds and 4 markets.
H A total of 5604 distinct households (but not all households appear in all rounds).
X All regressions include the number of boys, girls, men, women, log of household size, and a "rural" dummy as additional covariates.
FE Household Fixed Effects: Yes; Market-year Effects: Yes

Using a two-way fixed effects regression we've seen that covariate shocks have at most a small effect on w, save for the first month. We've argued that this is because the main mechanism by which covariate shocks affect MUEs is via prices, and the two-way fixed effects regression controls for these if demands are CFE. What if we *don't* control for prices? A panel regression incorporating only household fixed effects

would then reveal the effects that shocks have on MUEs via prices. Of course, attribution becomes problematical, since the incidence of covariate shocks across years and markets will be correlated not only with prices, but possibly other shocks. Accordingly, Figure 4 shows the results from a series of regressions of w and $\log x$ on the number of reported covariate shocks within the last m months, just as in Figure 3, but in this case only as a "one-way" panel estimator, controlling for household fixed effects but *not* time-market effects.

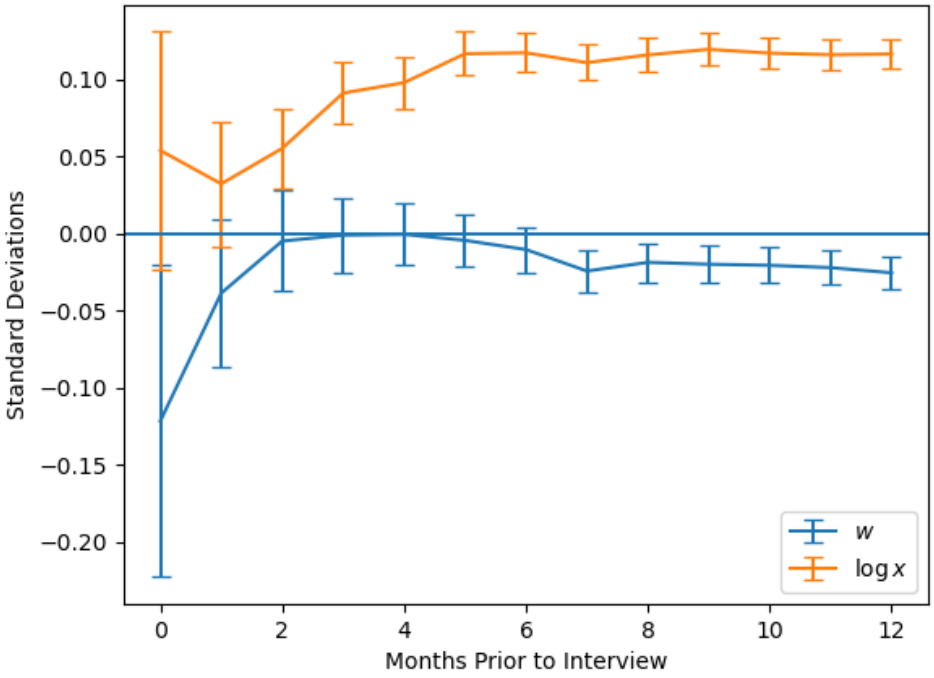


FIGURE 4. Effects of any covariate shock within the last m months on welfare measures with household fixed effects, but otherwise specified as in Table 3. Scale is in standard deviations of the dependent variable (w or $\log x$). Error bars cover a span of two standard errors about the point estimate.

Here w falls with reported "negative" covariate shocks (drought, floods, pests, prices) in the way we would expect, with estimated coefficients significantly negative for both the month of the shock and at more distant intervals (anything from seven to twelve months). In contrast, "negative" covariate shocks have a uniformly *positive* effect on expenditures, significant at all but the shortest intervals. This is once again consistent with the view that these shocks increase local prices of goods with inelastic

demands, leading to an increase in expenditures but a decrease in consumption of those goods.

8. Conclusion

8.1. **Implications for Tests of Full Risk Sharing.** Using standard risk-sharing regressions to test for insurance against covariate shocks (shocks that affect relative prices) can yield very surprising results—in the example of Uganda, droughts, floods, pests, and adverse changes in prices appear to be at least partially uninsured, but to *improve* welfare as measured by "real" consumption expenditures.

We argue that these surprising results are a consequence of the standard risk-sharing regressions actually being a *joint* test of full insurance along with preferences being homothetic—in this case changes in relative prices affect welfare only via a single scalar price index. There is extremely strong evidence against utility being homothetic, starting with Engel (1857), since homothetic utility implies that all demands must have an income elasticity of one. This doesn't matter much when one tests risk-sharing with respect to idiosyncratic shocks, since almost by definition these won't affect prices. But it can matter very much when prices change, since increases in prices can very easily increase total "real" expenditures while utility actually falls.

There are two great virtues of the risk-sharing regression framework. The first is its theoretical simplicity—it's really a model of households' marginal utilities of expenditure (MUEs), and in a world with full insurance these have a simple factor structure which is easily tested using panel methods with two-way fixed effects. The second is that assuming homothetic utility allows us to write the MUE as a simple function of nothing but total expenditures (typically $x^{-\gamma}$). And there is plenty of carefully collected data on household expenditures which one can use to construct these MUEs for dozens of countries over many years.

However, if we want to understand insurance against covariate shocks or price changes, the evidence that we need to abandon homothetic utility is overwhelming. So are there other ways to construct estimates of MUE that depend only on expenditures? We show that there are. If we use data on *item*-level expenditures, instead of adding these up to obtain a total, there's information based on the composition of these expenditures which can be used to estimate the MUE λ. A condition for being able to use expenditure data to estimate λ is that the expenditure system must be separable in λ and prices. When we impose this separability and exploit the homogeneity of

expenditures in prices, we show that the expenditure system can be written in a form called a generalized Pexider functional equation.

We exploit results from the theory of functional equations to obtain the entire class of possible demand systems consistent with inferring the MUE from nothing more than expenditures. These solutions fall into two families of semiparametric demands—one corresponds to a generalization of CRRA utility we call "Constant Frisch Elasticity" (CFE), while the other corresponds to a generalization of Stone-Geary utility. Both admit non-homothetic preferences and very flexible responses to changes in relative prices, but only the CFE system is easily estimated.

We use an eight round panel dataset from Uganda to estimate MUEs from expenditure data. We also obtain estimates of elasticities which emphatically reject the hypothesis of unitary income elasticities (a feature of CRRA demands). Using self-reported data on both covariate and idiosyncratic shocks, we estimate the risk-sharing regressions, using as dependent variables (a) the logarithm of total expenditures, as in the usual CRRA case; and (b) the values of $w = -\log\lambda$ estimated from the CFE expenditure system in a two-way panel regression (which also includes household demographics). As theory predicts, in this specification covariate shocks and adverse price changes have significant effects on log total expenditures (even if perfectly observed)[16], and *no* significant effect on the CFE estimates of MUE, as these are constructed to account for any changes in prices.

8.2. **Coda: MUEs are Everywhere.** This paper has focused on the case of full risk-sharing. This application may seem rather narrow, and so of limited interest. However, the construction of estimates of MUEs is independent of the efficient risk-sharing hypothesis, so estimated MUEs could be used to estimate and test any of a large variety of dynamic life-cycle models, following a program suggested by Blundell (1998). The usual consumption Euler equation is, after all, a statement about MUEs across time, and ratios of MUEs across periods give us a way to calculate intertemporal marginal rates of substitution, free of the usual homotheticity assumptions.

The practice of using CRRA preferences and measuring MUEs as an expenditure aggregate raised to some negative power is extremely common across many fields of economics. The novel tool developed in this paper is an improved way to measure MUEs. We've constrained ourselves to using nothing more than the same expenditure

[16]If in addition observed expenditures disproportionally omit certain goods or services with high income elasticities, then covariate shocks will tend to have a positive effect on total expenditures, as observed here.

data necessary to construct the expenditure aggregates typically employed by those working with CRRA or other homothetic preferences. The central idea is that CFE MUEs can be easily computed from the same data used to construct CRRA MUEs, and so serve as a more general drop-in replacement.

Consider first applications in other fields that use Townsend-style tests of risk-sharing. Though testing for risk-sharing across households is the classic approach, a similar setup is used in other settings, for example to test for sharing across people or generations in the unitary household model (e.g., Altonji et al., 1992; Hayashi et al., 1996; Dercon & Krishnan, 2000; Lise & Yamada, 2019; Theloudis et al., 2025).[17] The technique we attribute to Townsend arguably started in macro or international (O. Attanasio & Davis, 1996; Cochrane, 1991; De Giorgi et al., 2020; Lee & Lee, 2025; Mace, 1991; Obstfeld, 1993) and is still used.

Then there's the development of a literature on "partial insurance" (Blundell et al., 2008), which models dynamics in a permanent income-life cycle world where insurance is imperfect, but households have access to (at least) some form of credit market or intertemporal storage. The derivation of the key estimating equation is somewhat involved, but delivers an estimating equation which looks just like a differenced version of the Townsend regression with changes in log consumption as a dependent variable, and measures of each of permanent and transitory income as "shocks." This has many of the same features as the Townsend risk-sharing test—a possibly mis-measured MUE, with mis-measurement which can contaminate the interpretation of the time dummy as controlling for changes in prices (returns). Most examples of the use of this setup come from the macro literature (e.g., Kaplan & Violante, 2010; Arellano et al., 2017), but there are also applications in development (Bazzi et al., 2015; Santàeulalia-Llopis & Zheng, 2018).

Finally, there's scope for use of MUEs from the CFE demand system even in applications where the author may not be explicitly thinking in terms of MUEs. For example, there are many applications that start with a Rubin causal model wiith the intention of estimating treatment effects on (log) consumption. But this often leads to a TWFE estimator that also has a structural interpretation. This is particularly the case in a large number of RCTs in development, where it's very common to use a double-difference (a special case of TWFE) to estimate a treatment effect on total consumption expenditures (in levels or in logs). In these applications what

[17]MUEs also play a central role in the literature on the collective household (e.g., Browning et al., 1994). This literature also suggests a way in which one could extend preferences to incorporate other-regarding motives or public goods.

we're calling "risk-sharing" might be interpreted as "spillovers," and the use of fixed effects might be interpreted as a test of "parallel trends," but it's essentially the same econometric specification that Townsend employed. In this case there will also be a MUE & risk-sharing interpretation, with the benefit that if assignment is random there's also a causal interpretation. See, e.g., (Angelucci et al., 2018; Banerjee et al., 2015; Haushofer & Shapiro, 2016; Karlan et al., 2014).

9. References

Aczél, J., & Dhombres, J. (1989). *Functional equations in several variables* (Vol. 31). Cambridge University Press.

Alderman, H., & Paxson, C. H. (1994). Do the poor insure? A synthesis of the literature on risk and consumption in developing countries. In *Economics in a changing world: Volume 4: Development, trade and the environment* (pp. 48–78). Springer.

Altonji, J., Hayashi, F., & Kotlikoff, L. (1992). Is the extended family altruistically linked? Direct tests using micro data. *The American Economic Review*, *82*, 1177–1198. https://www.jstor.org/stable/pdf/2117473.pdf

Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability: Held at the Statistical Laboratory, University of California, December, 1954, July and August, 1955*, *1*, 111.

Angelucci, M., & De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect non-eligibles' consumption. *The American Economic Review*, *99*(1), 486–508.

Angelucci, M., De Giorgi, G., & Rasul, I. (2018). Consumption and investment in resource pooling family networks. *The Economic Journal*, *128*(615), 2613–2651.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, *49*(4), 431–34.

Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, *85*(3), 693–734.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk bearing. *Review of Economic Studies*, *31*, 91–96.

Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, 265–290.

Arrow, K. J., & Hahn, F. (1983). *General competitive analysis*. Elsevier.

Atkin, D., Faber, B., Fally, T., & Gonzalez-Navarro, M. (2024). Measuring welfare and inequality with incomplete price information. *The Quarterly Journal of Economics*, *139*(1), 419–475.

Attanasio, O., & Davis, S. J. (1996). Relative wage movements and the distribution of consumption. *Journal of political Economy*, *104*(6), 1227–1262.

Attanasio, O. P., & Weber, G. (2010). Consumption and saving: Models of intertemporal allocation and their implications for public policy. *Journal of Economic Literature*, *48*, 693–751.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *American economic journal: Applied economics*, *7*(1), 22–53.

Bardhan, P., & Udry, C. (1999). *Development microeconomics*. Oxford University Press.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach*. John Wiley & Sons.

Bazzi, S., Sumarto, S., & Suryahadi, A. (2015). It's all in the timing: Cash transfers and consumption smoothing in a developing country. *Journal of Economic Behavior & Organization*, *119*, 267–288.

Bentze, T. P., & Wollburg, P. R. (2024). *A longitudinal cross-country dataset on agricultural productivity and welfare in sub-saharan Africa* (Policy Research Working Papers No. 10976). World Bank.

Blackorby, C., & Donaldson, D. (1988). Money metric utility: A harmless normalization? *Journal of Economic Theory*, *46*(1), 120–129.

Blundell, R. (1998). Consumer demand and intertemporal allocations: Engel, Slutsky, and Frisch. *Econometric Society Monographs*, *31*, 147–168.

Blundell, R., Pistaferri, L., & Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review*, *98*(5), 1887–1921.

Browning, M., Bourguignon, F., Chiappori, P.-A., & Lechene, V. (1994). Income and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy*, *102*(6), 1067–1096.

Browning, M., Deaton, A., & Irish, M. (1985). A profitable approach to labor supply and commodity demands over the life cycle. *Econometrica*, *53*, 503–543.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). *Geographic displacement procedure and georeferenced data release policy for the demographic and health surveys* (DHS Spatial Analysis Reports No. 7). ICF International. Calverton, MD.

Byakatonda, J., Openy, G., Sempewo, J. I., & Mucunguzi, D. B. (2021). Over a century evidence of historical and recent dryness/wetness in sub-humid areas: A Uganda, East African case. *Meteorological Applications*, *28*(5), e2028.

Caron, J., Fally, T., & Markusen, J. R. (2014). International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics*, *129*(3).

Cashdan, E. (1985). Coping with risk: Reciprocity among the Basarwa of northern Botswana. *Man*, *20*, 454–474.

Chiappori, P.-A., Samphantharak, K., Schulhofer-Wohl, S., & Townsend, R. M. (2014). Heterogeneity and risk sharing in village economies. *Quantitative Economics*, *5*(1), 1–27. http://onlinelibrary.wiley.com/doi/10.3982/QE131/epdf

Cochrane, J. H. (1991). A simple test of consumption insurance. *Journal of Political Economy*, *99*, 957–976.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*, 1–45.

Conley, T. G., & Ligon, E. (2002). Economic distance and cross-country spillovers. *Journal of Economic Growth*, *7*(2), 157–187. http://journals.kluweronline.com/issn/1381-4338/contents

Conley, T. G., & Molinari, F. (2007). Spatial correlation robust inference with errors in location or distance. *Journal of Econometrics*, *140*(1), 76–96.

Connor, G., & Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of financial economics*, *15*(3), 373–394.

De Giorgi, G., Frederiksen, A., & Pistaferri, L. (2020). Consumption network effects. *The Review of Economic Studies*, *87*(1), 130–163.

Deaton, A. (1990). *On risk, insurance, and intra-village consumption smoothing* [Manuscript].

Deaton, A. (1997). *The analysis of household surveys*. The John Hopkins University Press.

Dercon, S. (2005). Risk, insurance, and poverty: A review. *Insurance against poverty*, 9–37.

Dercon, S., & Krishnan, P. (2000). In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy*, *108*(4), 688–727.

Engel, E. (1857). Die productions-und consumtionsverhaltnisse des königreichs sachsen. *Zeitschrift des Statistischen Biireaus des Königlich Sachsischen Ministeriums des Innern*, (8 & 9), 1–54.

Fafchamps, M. (1992). Solidarity networks in preindustrial societies: Rational peasants with a moral economy. *Economic Development and Cultural Change*, *41*(1), 147–174.

Frisch, R. (1959). A complete scheme for computing all direct and cross demand elasticities in a model with many sectors. *Econometrica*, *27*(2), 177–196.

Hansen, B. (2022). *Econometrics*. Princeton University Press.

Haushofer, J., & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *The Quarterly Journal of Economics*, *131*(4), 1973–2042.

Hayashi, F., Altonji, J., & Kotlikoff, L. J. (1996). Risk-sharing between and within families. *Econometrica*, *64*(2), 261–294.

Heltberg, R., Oviedo, A. M., & Talukdar, F. (2015). What do household surveys really tell us about risk, shocks, and risk management in the developing world? *The Journal of Development Studies*, *51*(3), 209–225.

Hoddinott, J. F., & Yohannes, Y. (2002). *Dietary diversity as a food security indicator* (Food Consumption and Nutrition Division Discussion Paper No. 136). International Food Policy Research Institute. Washington, DC.

Houthakker, H. S. (1960). Additive preferences. *Econometrica*, 244–257. https://www.jstor.org/stable/pdf/1907719.pdf

Jack, W., & Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, *104*(1), 183–223.

Jensen, R. T., & Miller, N. H. (2008). Giffen behavior and subsistence consumption. *American Economic Review*, *98*(4), 1553–77.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American statistical Association*, *70*(351a), 631–639.

Kaplan, G., & Violante, G. L. (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics*, *2*(4), 53–87.

Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *Quarterly Journal of Economics*, *129*(2).

Kinnan, C. (2022). Distinguishing barriers to insurance in Thai villages. *Journal of Human Resources*, *57*(1), 44–78.

Lee, J. W., & Lee, S. (2025). Monetary non-neutrality in a multisector economy: The role of risk-sharing. *Review of Economic Dynamics*, *55*, 101241.

Levine, D. K., & Zame, W. R. (2002). Does market incompleteness matter? *Econometrica*, *70*(5), 1805–1839.

Lewbel, A. (2010). Shape-invariant demand functions. *The Review of Economics and Statistics*, *92*(3), 549–556.

Ligon, E. (2016a). All λ-separable demands and rationalizing utility functions. *Economics Letters*, *147*, 16–18. http://dx.doi.org/10.1016/j.econlet.2016.08.010

Ligon, E. (2016b). Some λ-separable frisch demands with utility functions. *Economics Bulletin*, *36*(1), A8. http://escholarship.org/uc/item/1s06c2zp

Ligon, E. (2020). Estimating household welfare from disaggregate expenditures. https://escholarship.org/uc/item/2cg4q0b9

Ligon, E. (2023). `Cfedemands`: Tools for estimating and computing Constant Frisch Elasticity (CFE) demands. https://pypi.python.org/pypi/CFEDemands/

Ligon, E., Thomas, J. P., & Worrall, T. (2002). Informal insurance arrangements with limited commitment: Theory and evidence from village economies. *Review of Economic Studies*, *69*(1), 209–244. http://restud.oxfordjournals.org/content/69/1/209.full.pdf+html

Lise, J., & Yamada, K. (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *The Review of Economic Studies*, *86*(5), 2184–2219.

Little, R. J. (2021). Missing data assumptions. *Annual Review of Statistics and Its Application*, *8*(1), 89–107.

Mace, B. J. (1991). Full insurance in the presence of aggregate uncertainty. *Journal of Political Economy*, *99*, 928–956. http://www.jstor.org/stable/2937653

Mazzocco, M., & Saini, S. (2012). Testing efficient risk sharing with heterogeneous risk preferences. *American Economic Review*, *102*(1), 428–468. http://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.102.1.428

McKenzie, D. J. (2003). How do households cope with aggregate shocks? Evidence from the Mexican peso crisis. *World Development*, *31*(7), 1179–1199.

Morduch, J. (1995). Income smoothing and consumption smoothing. *Journal of Economic Perspectives*, *9*(3), 103–114.

Munshi, K., & Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, *106*(01), 46–98.

Obstfeld, M. (1993). *Are industrial-country consumption risks globally diversified?* (Working paper No. 4308) [Manuscript], National Bureau of Economic Research, Inc.

Ogaki, M., & Zhang, Q. (2001). Decreasing relative risk aversion and tests of risk sharing. *Econometrica*, *69*(2), 515–526.

Platteau, J.-P., & Abraham, A. (1987). An inquiry into quasi-credit contracts: The role of reciprocal credit and interlinked deals in small-scale fishing communities. *Journal of Development Studies*, *23*(4), 461–490.

Ray, D. (1998). *Development economics*. Princeton University Press.

Samphantharak, K., & Townsend, R. M. (2018). Risk and return in village economies. *American Economic Journal: Microeconomics*, *10*(1), 1–40.

Santàeulalia-Llopis, R., & Zheng, Y. (2018). The price of growth: Consumption insurance in China 1989–2009. *American Economic Journal: Macroeconomics*, *10*(4), 1–35.

Theloudis, A., Velilla, J., Chiappori, P.-A., Giménez-Nadal, J. I., & Molina, J. A. (2025). Commitment and the dynamics of household labour supply. *The Economic Journal*, *135*(665), 354–386.

Thomas, D., & Frankenberg, E. (2007). Household responses to the financial crisis in Indonesia: Longitudinal evidence on poverty, resources, and well-being. In *Globalization and poverty* (pp. 517–560). University of Chicago Press.

Townsend, R. M. (1978). On the optimality of forward markets. *The American Economic Review*, *68*(1), 54–66. https://www.jstor.org/stable/pdf/1809687.pdf

Townsend, R. M. (1994). Risk and insurance in village India. *Econometrica*, *62*(3), 539–591.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *The Journal of Economic Perspectives*, *9*(3), 83–102.

Uganda Bureau of Statistics. (2012). National panel survey 2010-2011. https://microdata.worldbank.org/index.php/catalog/study/UGA_2010_UNPS_v02_M

Uganda Bureau of Statistics. (2013). National panel survey 2011-2012. https://microdata.worldbank.org/index.php/catalog/study/UGA_2011_UNPS_v02_M

Uganda Bureau of Statistics. (2015). National panel survey 2013-2014. https://microdata.worldbank.org/index.php/catalog/study/UGA_2013_UNPS_v02_M

Uganda Bureau of Statistics. (2017). National panel survey 2015-2016. https://microdata.worldbank.org/index.php/catalog/study/UGA_2015_UNPS_v02_M

Uganda Bureau of Statistics. (2020). National panel survey 2018-2019. https://microdata.worldbank.org/index.php/catalog/study/UGA_2018_UNPS_v02_M

Uganda Bureau of Statistics. (2021). National panel survey 2019-2020. https://doi.org/10.48529/NQZX-F196

Uganda Bureau of Statistics (UBOS). (2012). National panel survey 2005-2009. https://doi.org/10.48529/PND4-7D68

Wanjala, W., Onyango, A., Makayoto, M., & Onyango, C. (2016). Indigenous technical knowledge and formulations of thick (ugali) and thin (uji) porridges consumed in Kenya. *African Journal of Food Science*, *10*(12), 385–396.

Williams, B. (2020). Identification of the linear factor model. *Econometric Reviews*, *39*(1), 92–109.

Zhang, Q., & Ogaki, M. (2004). Decreasing relative risk aversion, risk sharing, and the permanent income hypothesis. *Journal of Business and Economic Statistics*, *22*(4), 421–430.

Appendix A. Proof of Theorem 1

In this appendix we provide a proof of Theorem 1. We first supply a lemma pertaining to the homogeneity of λ-separable expenditure systems, and then provide solutions to the generalized Pexider equation. With these preliminary results in hand we establish that any rationalizable system of expenditures which is λ-separable takes the form of the generalized Pexider equation, and map the general solutions of the equation into corresponding demands and utilities.

A.1. **A Lemma Pertaining to Homogeneity.**

Lemma A.1. *If demands are rationalizable (Definition 1) and expenditures on good j are λ-separable (Definition 2), then the functions ϕ_j, a_j and b_j are either all logarithmic or ϕ_j and a_j are both positive homogeneous of some degree β_j, while b_j is positive homogeneous of degree $-\beta_j$.*

Proof. Any rationalizable expenditures x_j must be homogeneous of degree one in $(p, 1/\lambda)$. The definition of λ-separability implies that

$$x_j(p,\lambda) = \phi_j^{-1}\left(a_j(p) + b_j(\lambda)\right)$$

is similarly homogeneous of degree one. The function ϕ_j must then either be homogeneous of degree β_j, with $\phi_j(x_j) = x_j^{\beta_j}$, or else $\phi_j(x_j) = \log(x_j)$. In either case Frisch quantities can be written as

$$c_j(p,\lambda) = f_j(p\lambda) = \frac{1}{p_j}\phi_j^{-1}\left(a_j(p) + b_j(\lambda)\right) - d_j(p)$$

for some function d_j homogeneous of degree zero.

We consider the power and logarithmic cases in turn.

First suppose that $\phi_j(x) = x^{\beta_j}$. Then the sum $a_j + b_j$ must also be homogeneous of degree β_j in $(p, 1/\lambda)$, and the individual functions a_j and b_j respectively either homogeneous of degree β_j and $-\beta_j$ or else the zero function. It follows that $f_j(p,\lambda) = \phi_j^{-1}\left(a_j(p)/p_j^{\beta_j} + b_j(\lambda)/p_j^{\beta_j}\right) - d_j(p)$, and that $a_j(p)/p_j^{\beta_j}$ and $b_j(\lambda)/p_j^{\beta_j}$ are either zero or positive homogeneous of degree zero, so that

$$\left(a_j(p\theta) + b_j(\lambda/\theta)\right) = \theta^{\beta_j}\left(a_j(p) + b_j(\lambda)\right) = \theta^{\beta_j}a_j(p) + \theta^{\beta_j}b_j(\lambda)$$

for any positive scalar θ. Differentiating this with respect to $1/\lambda$ establishes that b_j' is homogeneous of degree $\beta_j - 1$, so that b_j is homogeneous of degree β_j (by Euler's theorem of positive homogeneous functions). A similar argument involving the gradient with respect to p establishes the same for a_j.

For the logarithmic case, $\phi(x_j) = \log(x_j) = \log(p_j) + \log(c_j)$ implies that

$$f_j(p, \lambda) + d_j(p) = \exp\left(a_j(p) + b_j(\lambda) - \log(p_j)\right)$$

which must be positive homogeneous of degree zero in $(p, 1/\lambda)$. This implies that for any $\theta > 0$

$$a_j(\theta p) + b_j(\lambda/\theta) - \log(\theta p_j) = a_j(p) + b_j(\lambda) - \log(p_j),$$

which in turn implies that

$$a_j(\theta p) + b_j(\lambda/\theta) = a_j(p) + b_j(\lambda) + \log(\theta),$$

implying that both a_j and b_j are linear in logs of (p,λ). □

A.2. **Generalized Pexider Equation Applied to Vector Spaces.** We now introduce our main tool for solving the functional equations implied by separability and rationalizability; this tool is an application of what is called the generalized Pexider equation, when the domain of application is limited to real vector spaces.

Consider the generalized Pexider equation

$$k(x + y) = g(x)l(y) + h(y) \tag{11}$$

where

$$g(x) = \frac{k(x) - h(0)}{l(0)} \tag{12}$$

$$\varphi(y) = \frac{l(y)}{l(0)} \tag{13}$$

$$\psi(y) = h(y) - h(0)\frac{l(y)}{l(0)} \tag{14}$$

$$k(x + y) = k(x)\varphi(y) + \psi(y) \tag{15}$$

$$\kappa(x) = k(x) - k(0) \tag{16}$$

$$\kappa(x + y) = \kappa(x)\varphi(y) + \kappa(y). \tag{17}$$

Next we give statements of two related lemmata. The first is just a statement of the solution of the well-known functional equation of Cauchy applied to real vector spaces; the second is a statement of the solution to what is sometimes called Cauchy's exponential equation, again for real vector spaces.

Lemma A.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, with f continuous at a point. Then if*

$$f(x + y) = f(x) + f(y) \tag{18}$$

then $f(x) = Cx$ *for some constant* $m \times n$ *matrix* C*.*

Also

Lemma A.3. *Let* $h : \mathbb{R}^n \to \mathbb{R}^m$*. If*

$$h(x + y) = h(x)h(y)$$

then either $h(x) = 0$ *or* $h(x)=e^{f(x)}$*, where* f *is an arbitrary solution to Cauchy's equation (18).*

Corollary 1. *Any solution to the functional equation of Lemma A.3 which is continuous and non-constant is of the form*

$$h(x) = \exp(Cx),$$

where C *is a constant matrix and the* exp *operator is element by element.*

The following is just a restatement of Theorem 15.1 of Aczél and Dhombres, 1989, and describes all solutions to the generalized Pexider equation (11) over the general domain of Abelian groupoids.

Theorem A.2. *For any* x, y *in an Abelian groupoid, solutions to (11) will satisfy one of:*

(1) *If* $\varphi(x) = 1$ *for all* x*, then* $\kappa(x)$ *is an arbitrary function;* $\psi(x) = \kappa(x)$*; and* $k(x) = \kappa(x) + B$*. Or;*
(2) *if* $\varphi(x_0) \neq 1$ *for some* x_0*, then we have* $C = \frac{\kappa(x_0)}{\varphi(x_0)-1}$*; and* $\kappa(x) = C[\varphi(x) - 1]$*; and two sub-cases:*
   (a) $C = 0$*;* $\kappa(x) = 0$*;* $\varphi(x)$ *arbitrary;* $k(x) = B$*;* $\psi(y) = B(1 - \varphi(y))$*; or*
   (b) $C \neq 0$*;* $k(x) = C\varphi(x) + B$*;* $\psi(x) = B(1 - \varphi(x))$*; where* $\varphi(x)$ *satisfies* $\varphi(x + y) = \varphi(x)\varphi(y)$ *(Cauchy's exponential equation).*

If we restrict the domain under consideration to a real vector space, then we can give explicit solutions to (11), as follows:

Proposition 1. *For any* $x, y \in \mathbb{R}^n$*, solutions to (11) will satisfy one of:*

(1) *If* $\varphi(x) = 1$ *for all* x*, then* $\kappa(x) = \psi(x) = Cx$ *and* $k(x) = Cx + B$*, where* $B \in \mathbb{R}^m$*. Or;*
(2) *if* $\varphi(x_0) \neq 1$ *for some* x_0*, then we have* $C = \frac{\kappa(x_0)}{\varphi(x_0)-1}$*; and* $\kappa(x) = C[\varphi(x) - 1]$*; and two sub-cases:*
   (a) $C = 0$*;* $\kappa(x) = 0$*;* $\varphi(x)$ *arbitrary;* $k(x) = B$*;* $\psi(y) = B(1 - \varphi(y))$*; or*

(b) $C \neq 0$; $k(x) = C\varphi(x) + B$; $\psi(x) = B(1 - \varphi(x))$; $\kappa(x) = Cx$; *and one of:*
  (i) $\varphi(x) = 0$;
  (ii) $\varphi(x) = \exp(Ax)$; *or*
  (iii) $\varphi(x) = \exp(f(x))$, f *nowhere continuous.*

Proof. Just a specialization of Theorem A.2 to the case in which domain is a real vector space, which then allows subsequent application of Lemma A.2 and Lemma A.3. □

A.3. **Proof of Theorem.** We're now in a position to prove our main result:

Proof. First, Lemma A.1 establishes that (ϕ_j, a_j, b_j) in (8) are all either logarithmic or positive homogeneous of some degree $(-)\beta_j$.

In the logarithmic case the logarithm of demand for good j can be written $\log f_j(\lambda p) = [-\log p_j + a_j(p)] + b_j(\lambda)$, which not only has expenditures λ-separable, but also quantities λ-separable. Then the main result of Ligon (2016a) applies, with $\phi_j = \log$, yielding the result that $\log(f_j(\lambda p)) = \tilde{\alpha}_j - \beta_j \log(p_j\lambda)$. Let $c_j = f_j(\lambda p)$ and solve for $p_j\lambda$, obtaining $p_j\lambda = \alpha_j c_j^{-1/\beta_j}$, where $\alpha_j = e^{\tilde{\alpha}_j}$ must be positive.

In the homogeneous case, we have

$$(p_j f_j(\lambda p))^{\beta_j} = a_j(p) + b_j(\lambda),$$

or

$$f_j(\lambda p)^{\beta_j} = p_j^{-\beta_j} a_j(p) + p_j^{-\beta_j} b_j(\lambda).$$

This takes the form of the generalized Pexider equation (11), with $x = \log\lambda$ and y the vector $\log p$, when the vector-valued function $k(x + y) = [f_j(\exp(x + y))^{\beta_j}]$, $h(y) = [a_j(\exp(y))e^{-\beta_j y_j}]$, $g(x) = [b_j(\exp(x))]$, and $\ell(y) = [e^{-\beta_j y_j}]$. Now, we seek to apply Proposition 1, which gives solutions to the system of functional equations 11–16. Part of this system is the function $\varphi(y)$. Using our knowledge that $\ell_j(y) = e^{-\beta_j y_j}$ and (13), it follows that in this equation the function $\varphi(y) = \ell(y) = [e^{-\beta_j y_j}]$. Now, consulting the different possible cases of Proposition 1 we see that with this solution of φ the only cases that can apply are the cases indicated by 2a and 2bii. The former implies that $f_j(\lambda p)^{\beta_j}$ is a constant, so that (to be consistent with the properties of Frischian demands) $\beta_j = 0$. But then the function ϕ_j isn't increasing, and the only solutions that are relevant to our problem are the solutions 2bii. These imply that $k(z) = A\exp(Cz) + B$, with A, B and C constant matrices, and $\varphi(y) = \exp(Cy)$. Thus C is a diagonal matrix, with diagonal elements $-\beta_j$. Then we have

$$k(z) = f_j(e^z)^{\beta_j} = \alpha_j e^{-\beta_j z} + \sigma_j,$$

the last substituting $A = \alpha_j$ and $B = \sigma_j$ in this scalar equation. With a change of variables $z = \log p_j + \log \lambda$ we obtain $f_j(\lambda p)^{\beta_j} = \alpha_j/(p_j\lambda)^{\beta_j} + \sigma_j$. Noting that $u_j(c) = p_j\lambda$ and solving for this gives us the solution for marginal utilities. □

Appendix B. Details of Estimation

Here we describe the steps involved in our estimation of the demands given by (9). We assume that observed household characteristics z_{it} affect log expenditures linearly, and that all households within a given market-year in Uganda face the same relative prices.[18] The latter assumption allows us to estimate the influence of prices by using market-year dummy variables, while the former allows us to express the influence of household characteristics on demand for good j as $\gamma_j z_{it}$, where γ_j is a vector of parameters to be estimated, and where t is understood to index market-years. Substitute $w_{it} = -\log \lambda_{it}$, and let $y^j_{it} = \log x^j_{it}$. Then (9) becomes

$$y^j_{it} = \gamma_j' z_{it} + a_{jt} + \beta_j w_{it} + \epsilon^j_{it}. \tag{19}$$

If we observed either β_j or w_{it} this would be an entirely standard linear regression problem. We do not, making this what Hansen (2022) calls a factor model with additional regressors.[19]

We make four standard assumptions. First, that (z_{it}, w_{it}) are orthogonal to the error ϵ^j_{it}, the usual identifying assumption in a linear regression context. Second, that w_{it} is orthogonal to z_{it}. While not necessary for identification (see Williams, 2020), this assumption simplifies both the interpretation and estimation of w_{it}. Third, that errors ϵ^j_{it} are homoskedastic across goods, though we allow clustering at the market-year level, so that $\mathrm{Var}(\epsilon^j_{it}) = \mathrm{Var}(\epsilon^{j'}_{it})$. Both the second and third assumptions are testable, and can be relaxed. Finally, as with any factor analysis some normalization is required for identification; here it's convenient to take $\mathrm{Var}(w_{it}) = 1$.

Estimation is straight-forward, and proceeds in three steps:

(1) Regress log expenditures on observed characteristics and time-market dummies

$$y^j_{it} = a_{jt} + \gamma_j' z_{it} + e^j_{it},$$

yielding estimates of the vector of the parameter vector γ and residuals $\hat{e}^j_{it}$.

[18] We take a market to correspond to one of four regions. There are eight years, resulting in a total of 32 different relative price vectors. If *levels* of prices differ within a region this will be reflected in marginal utilities of expenditure.

[19] Our estimation approach follows Hansen's procedure, though by assuming that w_{it} is orthogonal to z_{it} we avoid the need for iteration.

(2) The residuals $\hat{e}^j_{it}$ provide an unbiased estimate of $\beta_j w_{it} + \epsilon^j_{it}$. Let $\hat{e}_{it}$ denote the J-vector of residuals for household i at t, and let $\Sigma = \mathrm{E}\, e_{it} e_{it}^\top$ be its covariance matrix. From (19) and our assumptions above this implies

$$\Sigma = \beta\beta^\top + \sigma I.$$

This is a classical factor analysis problem, and the (quasi-) maximum likelihood estimate of β is proportional to the first principal component of the sample covariance matrix $\hat{\Sigma}$ (Anderson & Rubin, 1956). Our normalization $\mathrm{Var}(w_{it}) = 1$ supplies the missing factor of proportionality.

(3) Finally, we use a standard "regression" method (Bartholomew et al., 2011) to estimate the w_{it}. We return to (19), but now treat the w_{it} as unknown parameters to be estimated in the regression

$$y^j_{it} = a_{jt} + \gamma_j' z_{it} + \hat{\beta}_j w_{it} + \tilde{\epsilon}^j_{it},$$

where the difference from the original (19) is that now we can include the estimated $\hat{\beta}_j$ as (generated) regressors.

B.1. **Observations.**

- Of the three steps outlined above, two are least squares regressions, the properties of which will be familiar to economists. The second step employs factor analysis—a widely used procedure in other fields but less common in economics. A key consideration in factor analysis concerns the assumed structure of the disturbance covariance matrix across goods, say $\Psi = \frac{1}{NT}\,\mathrm{E} \sum_{i,t} \epsilon_{it}\epsilon_{it}^\top$. Here we assume a form of homoskedasticity, with $\Psi = \sigma I$, in which case the first principal component of $\hat{\Sigma}$ provides an unbiased and efficient estimate of β. Some alternative assumptions:
  - If Ψ is diagonal, the standard approach (Anderson & Rubin, 1956) uses maximum likelihood estimation assuming normally distributed disturbances. Even when this distributional assumption isn't satisfied, the (quasi-) maximum likelihood estimator retains good properties and can be implemented through a simple iterative procedure.
  - If Ψ is unrestricted, Bai (2003) provides conditions under which our principal components estimator still possesses desirable properties. The caveat is that the argument is asymptotic, requiring both J and TN to approach infinity. In our application, while TN is reasonably large, J is measured only in dozens.

  - Connor and Korajczyk (1986) establish consistency of our principal components estimator under a weaker sufficient and (Bai, 2003) necessary condition than our assumption that $\Psi = \sigma I$. Instead, they require only an asymptotic version of this, requiring that errors across goods be uncorrelated and homoskedastic as NT grows large, but holding the number of goods J fixed.
  - Williams (2020) establishes identification results for linear factor models for fixed J and very modest restrictions on Ψ. However, to our knowledge, no practical estimators have yet been developed for this case.
- With a consistent estimator of Σ, we can estimate β. Notably, it is possible to estimate Σ consistently even when some food expenditure data is missing, as is the case in our application (Little, 2021).

Appendix C. Direct Evidence on the Effects of Covariate Shocks

C.1. **Introduction.** In this appendix we draw on sources of data not needed in the main analysis to try to directly explore the effects of covariate shocks on a variety of outcomes, with the aim of tracing whether and how such shocks affect production, income, and consumption; what mechanisms households use to cope with these shocks; the effects of covariate shocks on local prices (both levels and relative); and finally on food consumption for the household. We provide direct evidence that: (i) Preferences are non-homothetic (expenditure shares vary with income and prices); (ii) Covariate shocks are indeed spatially correlated, in the sense that variation between clusters is greater than would be expected by random chance; (iii) People who reported having experienced a shock also tended to report declines in agricultural production and income; (iv) People reported coping with covariate shocks by reducing their consumption, using savings, or increasing labor, while they reported coping with idiosyncratic shocks by relying on help from relatives and friends; (v) Farmgate prices of locally produced goods increase significantly when a covariate shock is realized; (vi) Covariate shocks are associated with differences in the relative market prices consumers face; and (v) Covariate shocks of drought, floods, or pests are all associated with statistically significant changes in diet, and in particular a decreased demand for food obtained from animal sources such as dairy products and different kinds of meat.

C.2. **Preferences are non-homothetic.** With homothetic preferences the composition of the consumption portfolio is irrelevant, but for our approach the composition

is central. Can we see direct evidence in the Ugandan expenditure data for or against homotheticity? Figure C.1 uses the fact that homothetic preference structures imply that *aggregate* expenditure shares (of the sort that appear in national income and product accounts) will be the same as *average* expenditure shares. The figure takes logs of each of these, and then differences. It is apparent that these shares are not the same at all, with goods to the right (e.g., beer) playing a larger role in the consumption portfolios of households with larger budgets, and goods to the left (e.g., sorghum) playing a more important role for poorer households.

C.3. **Incidence and covariance of shocks.**

C.3.1. *Incidence of Shocks.* The data we have from the LSMS surveys includes self-reported data on a variety of "shocks" the household may have experienced (Heltberg et al., 2015, discuss the collection of this sort of data in a variety of different household surveys), generally elicited using the prompt "Did you experience [SHOCK] in the last twelve months?"[20] Where the answer is "Yes", the respondent is asked about the timing of the shock (in what month did the shock occur, and for how long did it last). There is some modest variation across rounds in the language used to describe different sorts of shocks (see Appendix D), but one can distinguish two classes. The first is idiosyncratic, shocks which directly involve a particular household. Frequently reported idiosyncratic shocks include health issues (serious illness or accident), thefts of property, and death (death of "income earners" is reported separately from the death of other household members). The second is covariate, shocks which seem likely to affect many households within a local area, though not necessarily equally. Frequently reported covariate shocks are drought (or "irregular rains"), agricultural pests, floods, and adverse agricultural prices (unusually expensive inputs or unusually low prices for output). Table C.1 reports the incidence of these shocks across different rounds.

To get a notion of frequency from Table C.1, there are about 2800 households observed per year. Drought is by far the most frequently reported shock. Meteorologists regard 2005–08 as a period of major drought for Uganda, with 2010–11 and 2014–15 periods of minor drought (Byakatonda et al., 2021), consistent with the household reports in Table C.1. Health shocks (both illness and accidents) are the second most

[20]There is some variation in the elicitation in different rounds, and the first 2005-06 round in particular uses a longer reporting period.

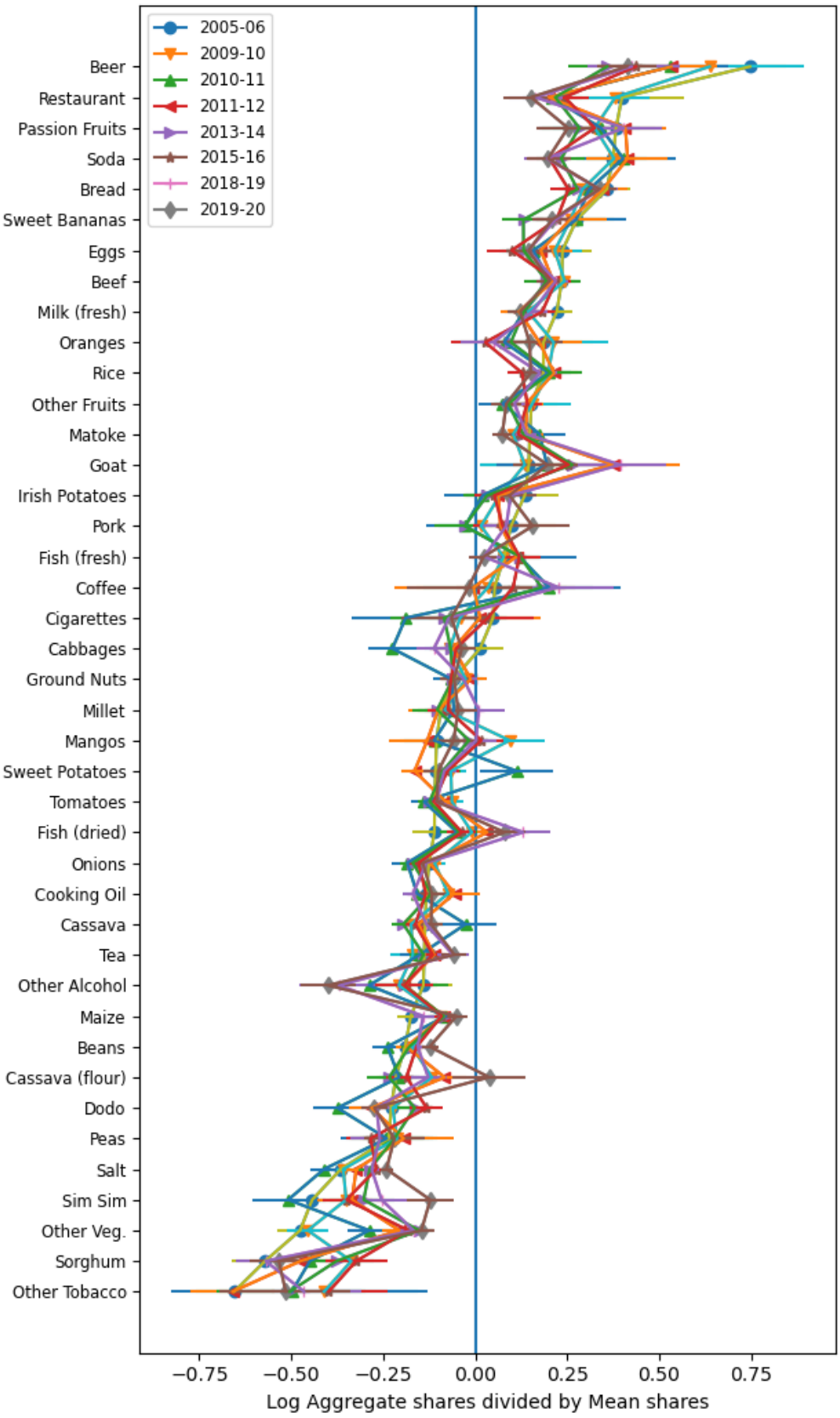


FIGURE C.1. Log of aggregate shares minus log of mean shares for different years (ordered by aggregate share averaged across years), with 95% confidence intervals.

TABLE C.1. Reported incidence of different kinds of shocks by year.

Shock	2005	2009	2010	2011	2013	2015	2018	2019	Total
Health	82	377	301	156	133	88	190	197	1524
Theft	349	233	96	55	76	62	75	83	1029
Death (non-earner)	423	74	58	35	66	35	44	49	784
Death of earner	99	27	17	19	30	19	20	19	250
Drought	1234	1344	710	560	914	598	736	529	6625
Floods	426	61	102	148	98	62	74	117	1088
Pests	475	219	77	92	71	53	130	84	1201
Prices	71	113	54	65	67	12	78	29	489

frequently reported shocks, followed by pests (which includes both crop pests and livestock disease), floods, death of non-earner, theft, and adverse changes in agricultural prices (i.e., increases in the costs of inputs; decreases in the value of outputs).

C.3.2. *Covariate shocks are spatially correlated.* We use several distinct sorts of evidence to establish that the "covariate" shocks we have data on do indeed covary. First, we test for geographical "clustering" of reported shocks within the Ugandan LSMS data. Second, we consider similar data from other LSMS surveys in Africa (eight countries in all), and construct maps of the reported incidence of different shocks, which we think provides informal but compelling visual evidence of spatial covariation. Third, we use these same data to estimate spatial autocorrelation functions, and establish that below a certain distance these estimated autocorrelations functions would be very unlikely to be unobserved under the null of no spatial correlation. The sampling scheme for the LSMS surveys involve randomly selecting geographical clusters (e.g., villages or parishes), and then randomly selecting households from within those clusters. One would then suppose that the realizations of "covariate" shocks ought to involve significant variation "between" clusters. Here we test that supposition by decomposing the variance of reports of different shocks in each year into variation attributable to variation in reporting "within" clusters and "between" clusters.

Table C.2 provides estimates of the proportion of variance which is "between", along with the p-values associated with the one-sided null hypothesis that this variance is no greater than would be expected if there was no clustering. Several features of this table are worth mentioning. First, there is statistically significant clustering (between variation) for drought in every year, for floods in most years, for pests in

TABLE C.2. Proportion of total variance in shock reports attributable to variation between clusters, by year. Figures in parentheses are p-values† associated with the null hypothesis that the "between" variancc is no greater than the pooled variance.

Round	Drought	Floods	Pests	Prices	Health	Death (non-earner)	Death (earner)	Theft
2005-06	0.11***	0.09***	0.08**	0.06	0.06	0.08**	0.07	0.07
	(0.00)	(0.00)	(0.01)	(0.84)	(0.70)	(0.01)	(0.27)	(0.15)
2009-10	0.18***	0.12***	0.11***	0.11***	0.08	0.09*	0.07	0.09
	(0.00)	(0.00)	(0.00)	(0.00)	(0.45)	(0.09)	(0.75)	(0.12)
2010-11	0.21***	0.17***	0.13***	0.09	0.11**	0.09	0.10	0.10
	(0.00)	(0.00)	(0.00)	(0.67)	(0.02)	(0.75)	(0.28)	(0.27)
2011-12	0.19***	0.15***	0.12***	0.12***	0.09	0.10	0.12**	0.11**
	(0.00)	(0.00)	(0.01)	(0.01)	(0.32)	(0.14)	(0.02)	(0.05)
2013-14	0.34***	0.27**	0.19	0.17	0.23	0.22	0.16	0.19
	(0.00)	(0.02)	(0.58)	(0.81)	(0.13)	(0.34)	(0.77)	(0.62)
2015-16	0.46***	0.29	0.19	0.25	0.23	0.23	0.17	0.29
	(0.00)	(0.16)	(0.90)	(0.48)	(0.67)	(0.60)	(0.81)	(0.22)
2018-19	0.32***	0.27*	0.20	0.20	0.24	0.23	0.10	0.23
	(0.00)	(0.07)	(0.85)	(0.72)	(0.18)	(0.42)	(0.95)	(0.42)
2019-20	0.29***	0.28	0.19	0.21	0.26	0.30	0.19	0.29
	(0.00)	(0.10)	(0.93)	(0.73)	(0.23)	(0.13)	(0.66)	(0.13)

† All p-values computed using randomized permutation tests, with 1000 draws.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.
N Total observations are 28,588.
T All regressions involve 8 rounds, 4 markets, and an average of 550 clusters per round.
H A total of 5771 distinct households (an average of 3045 households per round).

four of eight years, and for prices in two of eight years. Some of these failures to reject the null hypothesis may simply be due to the shock being reported infrequently; for example, there were only 12 reports of adverse agricultural prices in total in 2015-16, so the power of the test in that year is extremely low. Second, while for covariate shocks there's most often unanimity within a cluster (either no one says there's been a covariate shock or everyone says there has been), so that within-variance is zero for these clusters, there are also many clusters in which some but not all households report experiencing the shock. Third, statistically significant clustering is observed less frequently for the shocks we've categorized as "idiosyncratic," however, we still reject no clustering somewhat more frequently than we would expect under the null. Three of the four reported "idiosyncratic" shocks involve sickness or death, and it's

reasonable to suppose that a spatially correlated spread of disease may sometimes be implicated in these reports. While there's important variation across years in the proportion of households that report drought, we would like to know that the data are consistent with those shocks being covariate, in the sense that if a person in one cluster reports a particular sort of shock, then another nearby person is likely to report the same. Figure C.2 documents all the drought reports across the eight different LSMS-ISA countries over all the years (2005–20) for which we have data, while Figure C.3 and Figure C.4 do the same for floods and pests, respectively.[21] Dots correspond to sample clusters, and the "redder" the dot, the higher the proportion of households which reported drought.

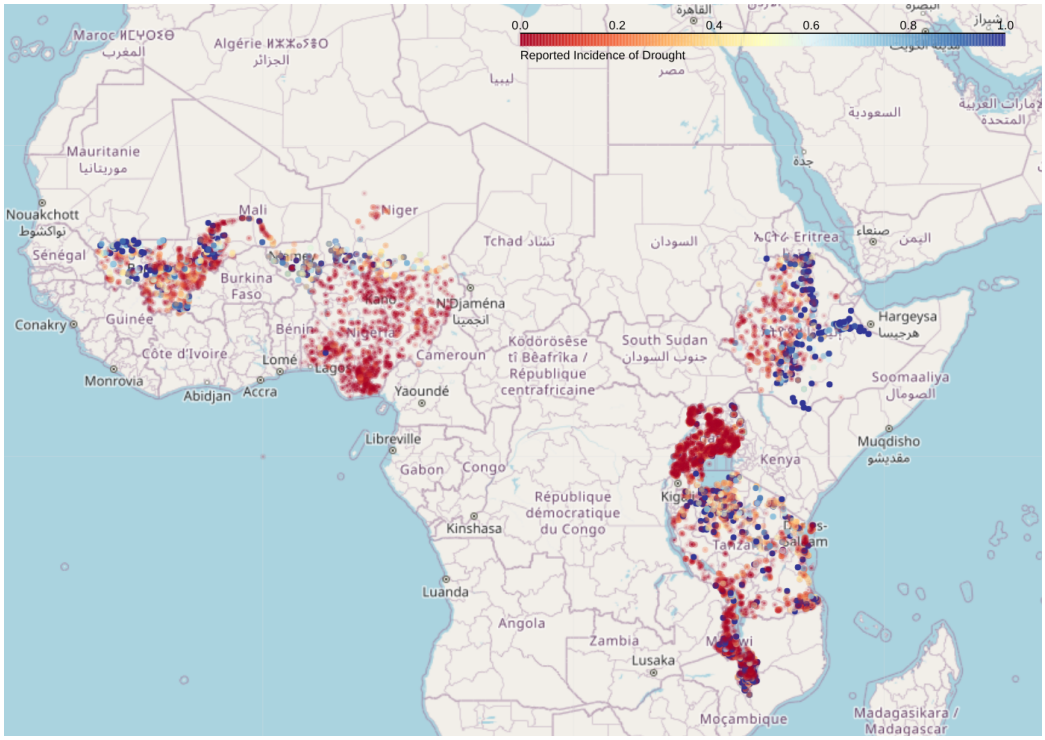


FIGURE C.2. Drought reports in the LSMS-ISA data across Africa. Each dot corresponds to a sample cluster; the "bluer" the dot the higher the proportion of households in the cluster which reported drought. Note that observations from later sample years often partly obscure data from earlier years.

[21]For our examination of data across African countries we rely on the dataset described by Bentze and Wollburg (2024). Note that this dataset categorizes shocks somewhat differently than we do in our analysis of data from Uganda, and evidently uses a narrower definition of "floods" than we do.

Casual observation of Figure C.2 suggests not only that there are many dark dots (high proportions of households within a cluster agreeing that there's drought), but also that darker dots tend to be close to each other, consistent with the notion that these reports really are reflecting covariate shocks.

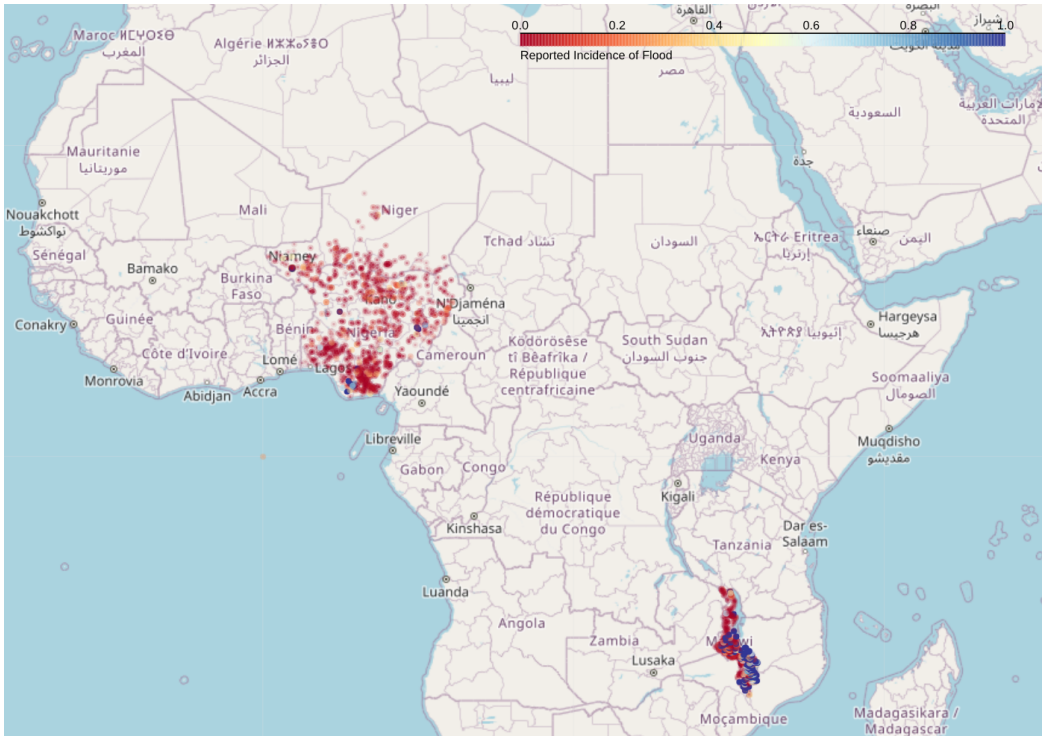


FIGURE C.3. Flood reports in the LSMS-ISA data across Africa. Each dot corresponds to a sample cluster; the "bluer" the dot the higher the proportion of households in the cluster which reported floods. Note that observations from later sample years often partly obscure data from earlier years.

- Spatial Autocorrelations of Shocks Across Clusters We've seen evidence of correlation in shock reports within clusters (villages or parishes). Is there also correlation across nearby clusters? We have data on cluster location, which we use to we take a more formal approach to testing the hypothesis of no spatial dependence across clusters in shock reports. We use the data from shock reports across Africa described above, and then follow Conley and Ligon (2002) in computing non-parametric spatial autocorrelation functions for reports of drought, floods, or pests (adverse price shocks are not reported in this cross-country dataset). This yields pointwise estimates of the correlation of shock reports at different distances; we're interested in whether or not the

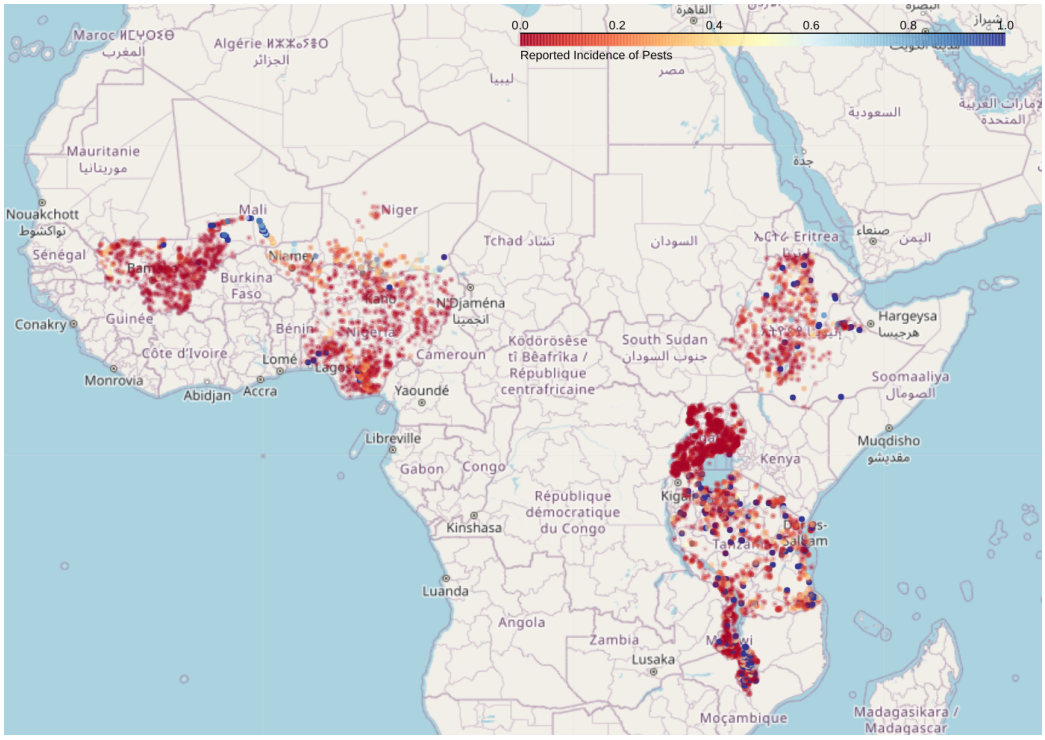


FIGURE C.4. Pest reports in the LSMS-ISA data across Africa. Each dot corresponds to a sample cluster; the "bluer" the dot the higher the proportion of households in the cluster which reported pests. Note that observations from later sample years often partly obscure data from earlier years.

observed spatial autocorrelations could have been generated by chance under the null hypothesis of no autocorrelation.

There are two important sources of error one should be aware of. First, the public LSMS data intentionally obscures location in order to protect respondent identity by mapping the location of each household to a single cluster "point", and then randomly shifting the location across space. However, the rules for this random displacement are clear and well-specified (Burgert et al., 2013), so we are able to take the effects of this displacement on estimation into account in constructing our tests. The second source of error is more profound—the underlying statistical model assumes that while *distance* matters for correlation, *direction* does not (i.e., an assumption of isotropy). At very large scales this is obviously wrong in our data—going in the direction of the Sahara matters for drought. However, the distances we consider are on the order of hundreds of kilometers, so very large-scale features such as

the Sahara are in some sense far away from most of the clusters in our data. But isotropy may be a bad assumption even at smaller scales for some shocks such as floods, where the probability of nearby flooding is much greater if one follows the direction of a watercourse than otherwise. Fortunately, Conley and Molinari (2007) provide evidence that hypothesis tests using our non-parametric estimator are quite robust to these sorts measurement error in location.

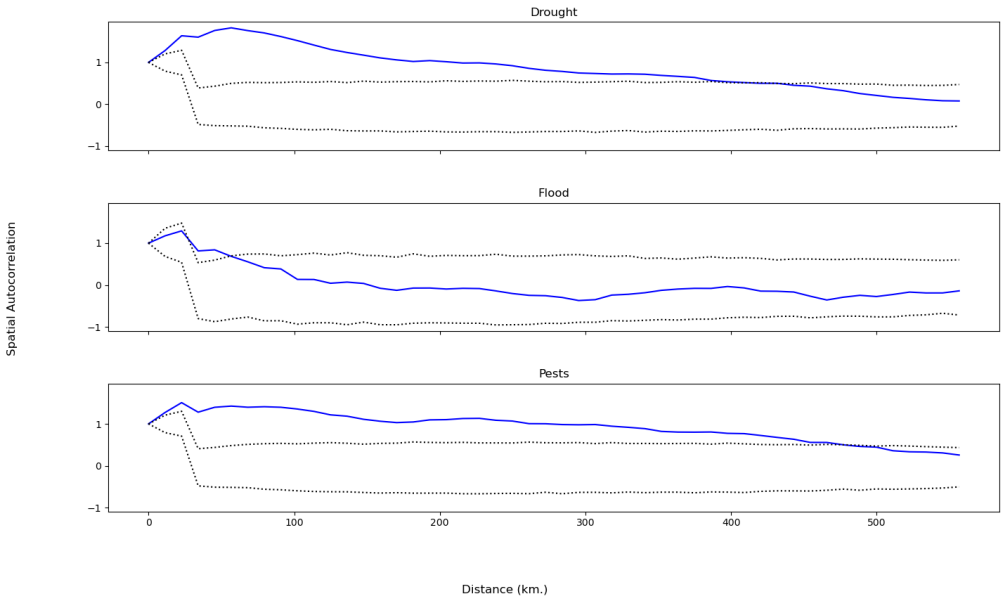


FIGURE C.5. Estimated spatial autocorrelation functions for shock reports. Dotted area indicates a point-wise 95% "acceptance" region of the null hypothesis of no correlation at a given distance.

We proceed by using provided data on cluster locations, then computing the great-circle distance between each pair of clusters in each year. We then use the non-parametric spatial autocovariance estimator of Conley (1999), using a rectangular kernel with a bandwidth distance of roughly 30 kilometers. For each year we obtain an autocorrelation function by dividing the autocovariance function at each distance d by the value of the function at a distance of zero (i.e., by the variance of cluster incidence per year). We use a sort of augmented randomized permutation inference (with 1000 draws) to construct 95% acceptance regions for the estimated functions, and construct an average across years, weighted by the number of observations per year.

The augmentation of our randomized permutation deserves some remark—this is intended to take into account the intentional random displacement of geographical coordinates imposed on the public data (Burgert et al., 2013). There are two key features of this displacement for us: first, one percent of all locations are moved some distance in a random direction, where that distance is drawn from a uniform distribution on $[0, 100]$ kilometers. Then *all* locations are similarly displaced, but only up to five kilometers. Under the null of no autocorrelation location doesn't matter, so we can treat the observed locations as though they were measured without error, and apply the displacement algorithm to each permutation of locations to discover the effect on inference.

Results are shown in Figure C.5. In our randomized permutation setting distances at which the estimated spatial autocorrelation function lies above the acceptance region are distances at which the p-value associated with the null hypothesis is less 2.5% (for a one-tailed test). Here we see evidence of positive spatial autocorrelation for all three shocks. This evidence is least pronounced for floods, where significant autocorrelation is observed at distances of about 30–60 km. Here the assumption of isotropy may be playing an important role, as two nearby clusters may well be at different elevations. There is evidence of positive autocorrelation at a much larger range of distances for both drought (0–400 km) and pests (0–470 km).

C.4. **Covariate shocks & reported declines in production, income, consumption, and assets.** The LSMS survey module on shocks elicits subjective reports on whether the shock resulted in a decline in the outcomes production, income, assets, or consumption. Table C.3 reports the proportion of people who, having reported that they'd experienced a particular sort of shock (such as drought) in the last twelve months, *also* responded affirmatively to the question: "As a result of the [SHOCK] was there a decline in your household's [OUTCOME]?"

Thus, of households who reported drought in the last twelve months, 92% asserted that the drought resulted in a decline in production, and 80% reported a decline in income. The logic of these responses seems fairly clear: a drought will directly affect agricultural production, and so *ceteris paribus* a farmer will have less income from marketing his smaller output.

The remaining two outcomes seem less direct. For a shock seems likely to affect consumption only indirectly, by affecting available resources (such as income or assets)

TABLE C.3. Subjective reports as to whether shock resulted in a decline in particular outcomes. These reports are conditional on the respondent reporting that they'd experienced the shock in the last 12 months.

Shock Affected	Drought	Floods	Pests	Prices
Production	0.92	0.89	0.73	0.47
Income	0.80	0.67	0.77	0.84
Consumption	0.50	0.41	0.37	0.44
Assets	0.25	0.25	0.35	0.27

which might be used to finance consumption. And assets could be affected in two different ways: via reduced resources (leading to lower rates of investment or sale of assets), or by directly damaging physical assets.

C.5. **Covariate shocks increase local farmgate prices.** Respondents to the LSMS surveys in Uganda report farmgate prices for goods they produce, and so we can see how these prices respond to covariate shocks. One concern one might have is that respondents who report shocks might also be inclined to report higher prices. We address this by calculating mean prices for every year-market. For each good we then regress changes (over time) in log prices on the proportion of sample households in every market-year that report different aggregate shocks and on a complete set of year effects.

We observe changes in log mean prices[22] for 20 goods over four markets and over seven periods, so the sample is small (just 28 observations of price changes for each good). Nevertheless, results from this exercise are shown in Figure C.6 for the four covariate shocks "drought", "floods", "pests", and adverse "prices". We have farmgate prices in every market-year for 20 different goods. Point estimates on "drought", "pests", and "prices" are positive for all goods; for "floods" they are positive in 9 of 20 cases. Though a t-test would reject the null of zero effect in only a handful of cases (12,1,7, and 3, respectively), this is unsurprising given our small sample. A more appropriate test of whether shocks positively affect farmgate prices in general is the non-parametric sign test. Under the null of no effect each point estimate is equally likely to be positive or negative; the p values associated with this null hypothesis are 0.41 in the case of floods, and less than 0.000001 for drought, pests, or prices.

[22]Using log median prices delivers broadly similar results.

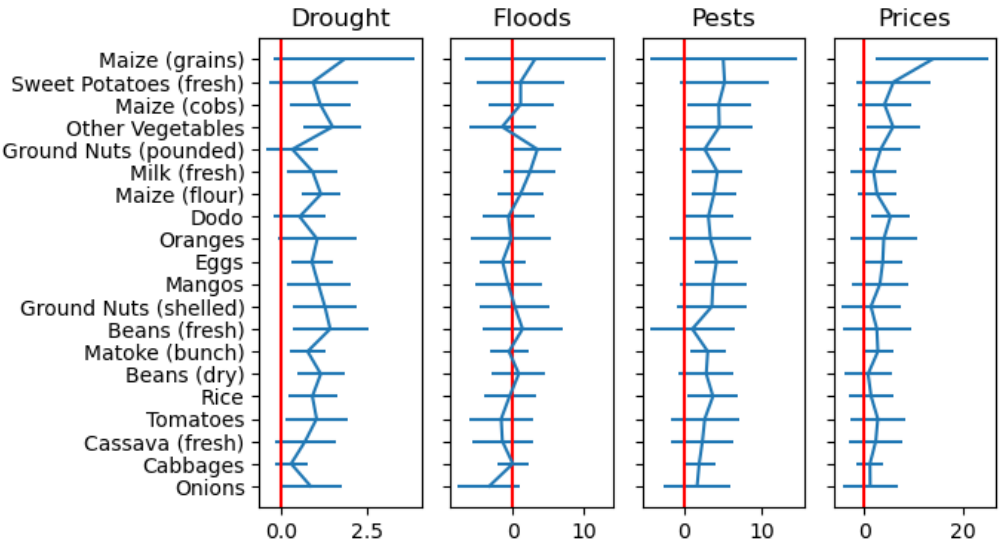


FIGURE C.6. Shocks and Farmgate Prices. Point estimates from a regression of changes in mean year-market (farmgate) log prices on the proportion of households reporting particular covariate shocks in a given year-market; includes year effects. Bars indicate 95% confidence intervals. Ordered by average magnitude of coefficient across shocks.

C.6. **Mechanisms used to cope with shocks.** Related, for each shock they reported, respondents were asked how they coped with the shock, and offered up to three responses. Table C.4[23] indicates the frequency with which different coping strategies were reported. For example, 51% of respondents who reported experiencing drought said that they coped by reducing consumption, 45% said that they used savings, and 30% said they reduced leisure. All of these mechanisms can be thought of as forms of self-insurance. "Help from relatives/friends" is also commonly invoked (20% of households who experienced one of these covariate shocks reported such help). It should not be surprising that this kind of aid from friends and family the dominant mechanism for idiosyncratic shocks, reported by 52% of those who report experiencing any of the idiosyncratic shocks found in Table C.1—a feature of covariate shocks is that others are likely to also experience them at the same time. And this risk-sharing seems to be fairly effective in addressing idiosyncratic shocks, as fewer than 14% of households who report idiosyncratic shocks say they reduced consumption in response.

C.7. **Covariate shocks and relative prices.** If a covariate shock has a proportionate effect on all prices, then this would be equivalent to a reduction in income, with

[23]See Table D.7 for details of how responses are coded.

TABLE C.4. How did households cope with different shocks? These reports are conditional on the respondent reporting that they'd experienced the shock in the last 12 months. Respondents could report up to three different coping mechanisms, so columns do not sum to one.

How Coped	Drought	Floods	Pests	Prices	Any
Reduced consumption	0.51	0.35	0.24	0.21	0.44
Used savings	0.45	0.47	0.38	0.37	0.44
Reduced leisure	0.30	0.25	0.14	0.16	0.27
Help from relatives/friends	0.21	0.20	0.14	0.10	0.20
Other	0.10	0.12	0.24	0.35	0.13
Changed cropping practices	0.10	0.08	0.21	0.16	0.11
Sold assets	0.08	0.07	0.10	0.06	0.08
Used credit	0.08	0.07	0.05	0.07	0.07
Help from local government	0.03	0.03	0.06	0.01	0.03
Reduced investments in human capital	0.03	0.01	0.02	0.02	0.03
Sent children elsewhere	0.01	0.02	0.01	0.00	0.01
Migrated	0.01	0.01	0.01	0.00	0.01
Rented out assets	0.00	0.00	0.01	0.00	0.00
Average # Coping Methods	1.91	1.68	1.58	1.53	1.83

no change in relative prices. With non-homothetic preferences this would still result in changes in the composition of expenditures, but these changes would be entirely due to changes in consumers' MUEs.

However, the kinds of shocks we consider seem likely to change *relative* prices. The exact nature of these changes will be complicated, depending not only on the effects of the shock on the supply of locally produced goods, but on income and price elasticities for all goods. For example, a drought in an region where cultivation of maize and beans is important might increase bean prices more than maize, as maize tends to be traded over longer distances. be expected to increase local bean prices (via a shift in local supply), but might also reduce local demand for other goods via an income effect. In fact, if local production is dominated by the production of only a few goods, then we would expect a shock which affected the local supply of those goods to be associated with higher prices for those goods, while for other goods we would tend to expect income effects to cause prices to be lower than otherwise.[24]

[24]However, this simple story is complicated by the possibilities of substitution and inferior goods—for example, an increase in local bean prices might lead to an increase in demand for a substitute such as matoke, or the income effect might result in an increase in demand in an apparently inferior good such as millet.

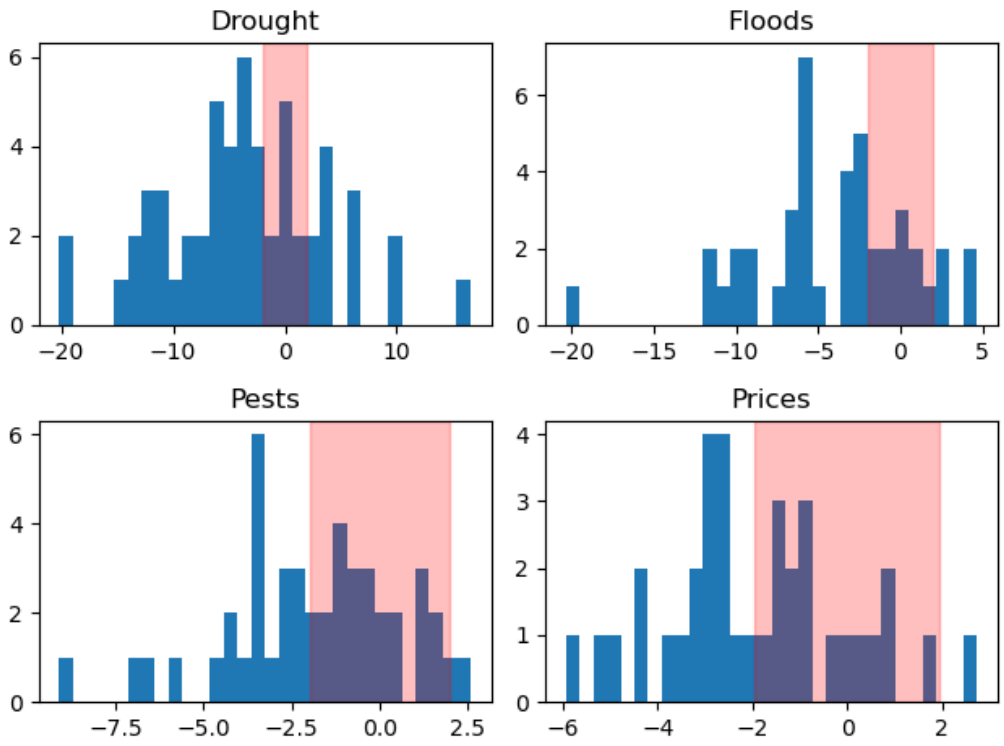


FIGURE C.7. Histograms of t-statistics testing hypothesis that covariate shocks have no effect on prices, by good. Shaded areas are 95% acceptance regions.

Here we wish to test whether the realization of a market level (regional) covariate shock leads to changes in relative prices. We earlier looked at the effect of covariate shocks on changes in farmgate prices of goods which are locally produced. Because here our focus is in part on income effects here we look instead at reported market prices, which encompasses the larger set of goods we use for estimating MUEs, and which are the relevant prices for consumers.

Our approach here is to compute the deviation of mean log price of each good in each market-year from the average country-wide log price. We then divide the sample of prices according to whether households report a particular shock versus households that don't report the shock. We then simply construct a t-test of the null hypothesis that average log prices are the same across these two samples.

Figure C.7 shows the result of these tests for the four covariate shocks we've considered to this point. Under the null hypothesis of the shock not affecting relative prices we'd expect five percent of computed t statistics to fall outside the acceptance region $(-1.96, 1.96)$. In fact, 82%, 74%, 44%, and 53% of the tests (respectively)

reject equality. Further, as suggested above, a negative shock tends to lead to *lower* prices for most goods.

C.8. **Quantities of food demanded.** How does diet respond to shocks? We consider the effects of different covariate shocks on the consumption of different food items by regressing different outcomes on the proportion of households in each market-year who reported a particular shock, along with a complete set of time effects, so that the identifying variation is across market regions within a year.

We consider three related outcomes, all related to what we'll call "food acquired": the sum of all food purchased, consumed out of own production, received as a gift, or consumed outside the household by any household member during the last week. The first outcome is simply quantities of food acquired, including "zeros" for food items that the household didn't report at all. To handle differences in units and scale quantities are standardized. Second, we consider changes in diet on the 'extensive' margin, using an indicator for whether the household reported acquiring any positive amount of food on reports of shocks; one can thus interpret estimated shock coefficients as the probability of consuming a given good conditional on the shock. This set of regressions is of some particular interest because it can be regarded as a measure of how dietary diversity responds to covariate shocks. Third, where amounts are positive we use the logarithm of quantities; the resulting coefficients are interpretable as elasticities related to the 'intensive' margin.

Results are presented in Figures C.8, C.9, and C.10. These three figures report estimated coefficients associated with different shocks, along with horizontal bars that indicate 95% confidence intervals. Goods are ordered according to magnitude of coefficients averaged across different shock regressions.

We first consider the quantity regression; coefficients are reported in Figure C.8. Several points seem worth mentioning. First, one might entertain the hypothesis that a covariate shock might reduce consumption of all kinds of food. This idea is consistent with the logic of homothetic preferences such as the CRRA, where consumption scales linearly with budget and all income elasticities are unity. But we've already seen evidence that there are large differences in the relative income elasticities of different kinds of food. Figure C.8 suggests that the principal effect of a covariate shock is to change the composition of the food bundle, rather than reducing quantities of all foods–just looking at food items, average consumption of roughly half increases in response to shock. And as with prices, patterns of substitution seem to be very

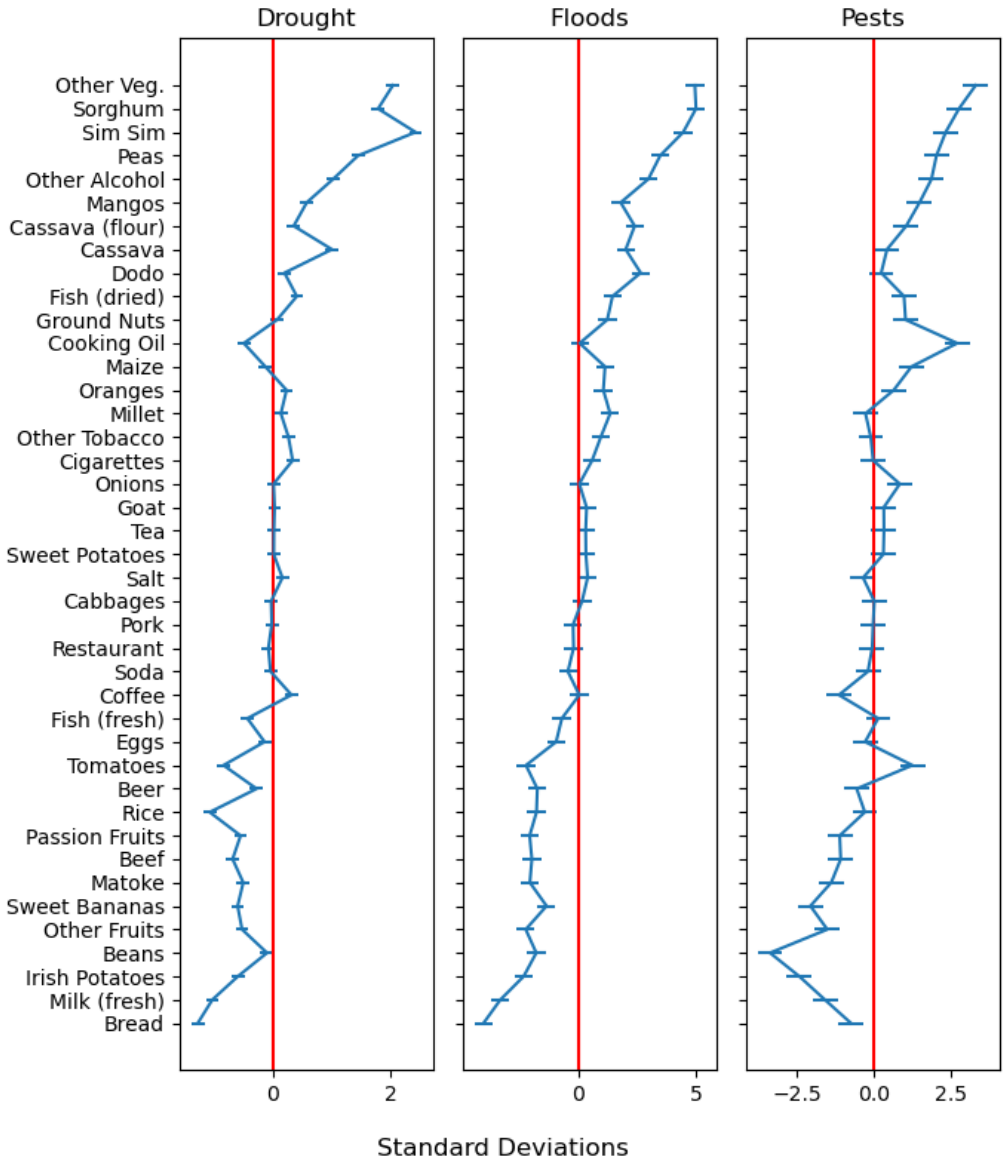


FIGURE C.8. Food Quantities and Shocks. Estimated coefficients from regression of standardized quantities of different food items on the share of households in a given market-year reporting given covariate shock (year effects included). Bars indicate 95% confidence intervals. Goods are ordered according to magnitude of coefficients averaged across different shock regressions.

important. For example, average sim sim (sesame) consumption increases dramatically for all three covariate shocks reported in the figure. From our estimates of the demand system (see Figure 1) sim sim is among the less income elastic goods, so we

might expect its role in the diet to expand when income is low. But other goods are even less income elastic, such as cassava. It appears that sim sim provides a protein-rich complement to cassava and sorghum in some staple dishes, substituting for meat (Wanjala et al., 2016), or being processed to substitute for other cooking oils (one of the more income elastic goods); collectively these patterns of complementarity and substitution may account for the large increase in sim sim consumption observed when there's a covariate shock.

We next try to decompose the average dietary changes observed in Figure C.8 into changes on the extensive (Figure C.9) and intensive (Figure C.10) margins. Figure C.9 can be thought of providing estimates of the change in the share of households acquiring a particular food conditional on the proportion of households within its market region reporting a given shock. So, for example, shocks are associated with a larger share of households consuming "other vegetables", sorghum, or sim sim, and with a dramatically smaller share consuming matoke (cooking bananas, a favored staple), fresh milk, or "other fruit."

We saw in Figure C.8 that the number of foods that increased in quantity in response to a covariate shock was roughly equal to the number that decreased. That is not the case here: any of the covariate shocks we consider lead to decreases in the expected number of different kinds of food acquired by the household. This tells us that on average these shocks lead to a decrease in dietary diversity, a widely used measure of food security (Hoddinott & Yohannes, 2002).

We next consider the intensive margin. Figure C.10 reports results from regressions of log quantities of different kinds of food on our measure of shocks. Here we regress log quantities (discarding zeros) on a set of time dummies and on the proportion of households in a given market-year reporting a particular shock. A regression of a log quantity on a proportion yields a coefficient which we can interpret as an elasticity. Foods with the largest elasticities include mangoes, different forms of cassava, and oranges. The fact that consumption of mangoes and oranges increases in response to covariate shocks may seem surprising, but both are tree crops widely grown in Uganda and well known for their drought tolerance. Note from C.9 that demand on the extensive margin responds very little to shocks, so we might guess that given that these are available in the market they may save as substitutes for other foods such as sweet bananas that are not at all tolerant of drought. Another interesting case is "other tobacco"; the prevalence of use increases when there's a covariate shock, but at the same time the log quantity falls. This suggests that shocks may induce widespread

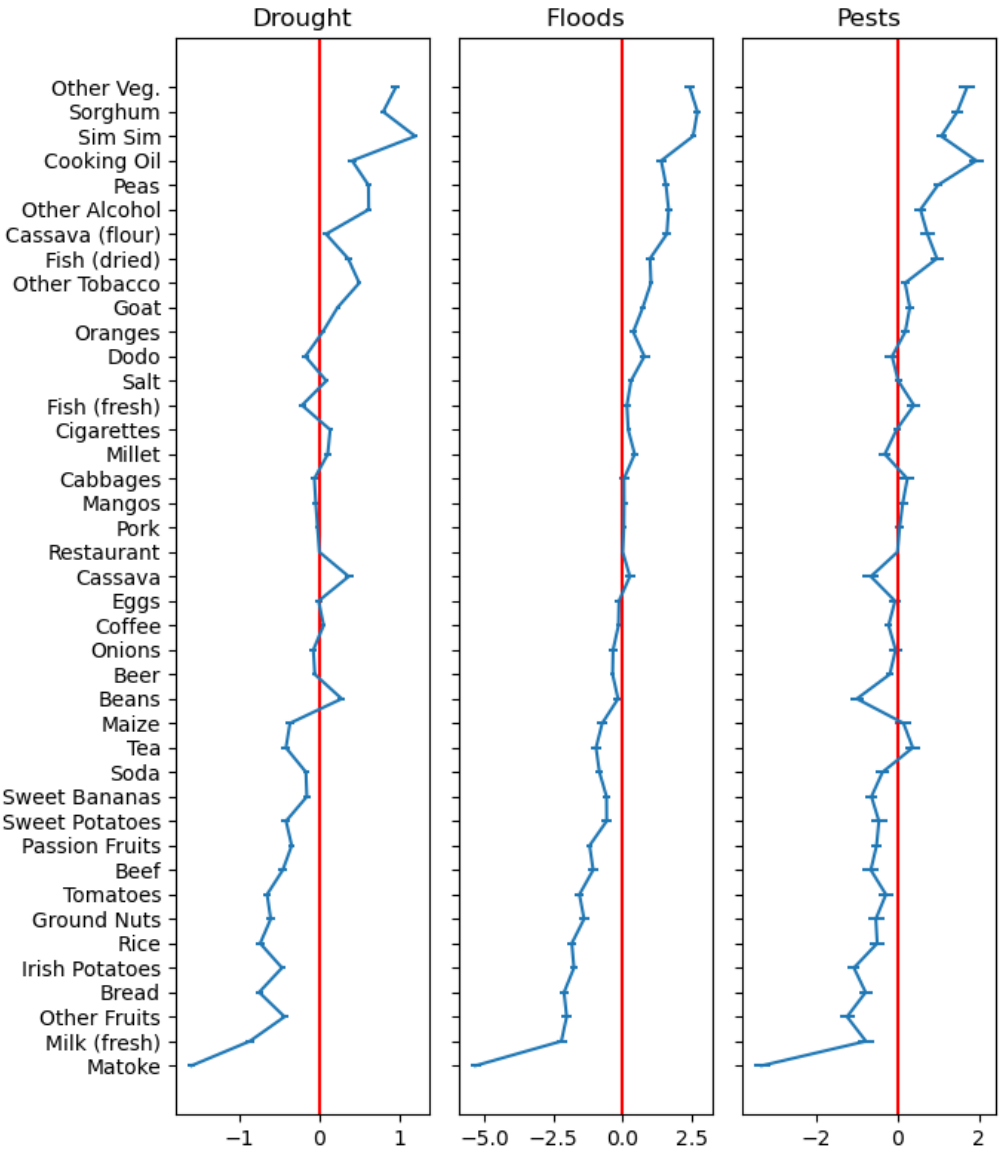


FIGURE C.9. Positive Quantities and Shocks. Estimated coefficients from regression of an indicator of positive quantities of different food items on the share of households in a given market-year reporting given covariate shock (year effects included). Bars indicate 95% confidence intervals. Goods are ordered according to magnitude of coefficients averaged across different shock regressions.

adoption but modest use of this stimulant. We've already noted the large extensive response of sim sim consumption to shocks. However, on the intensive margin there's

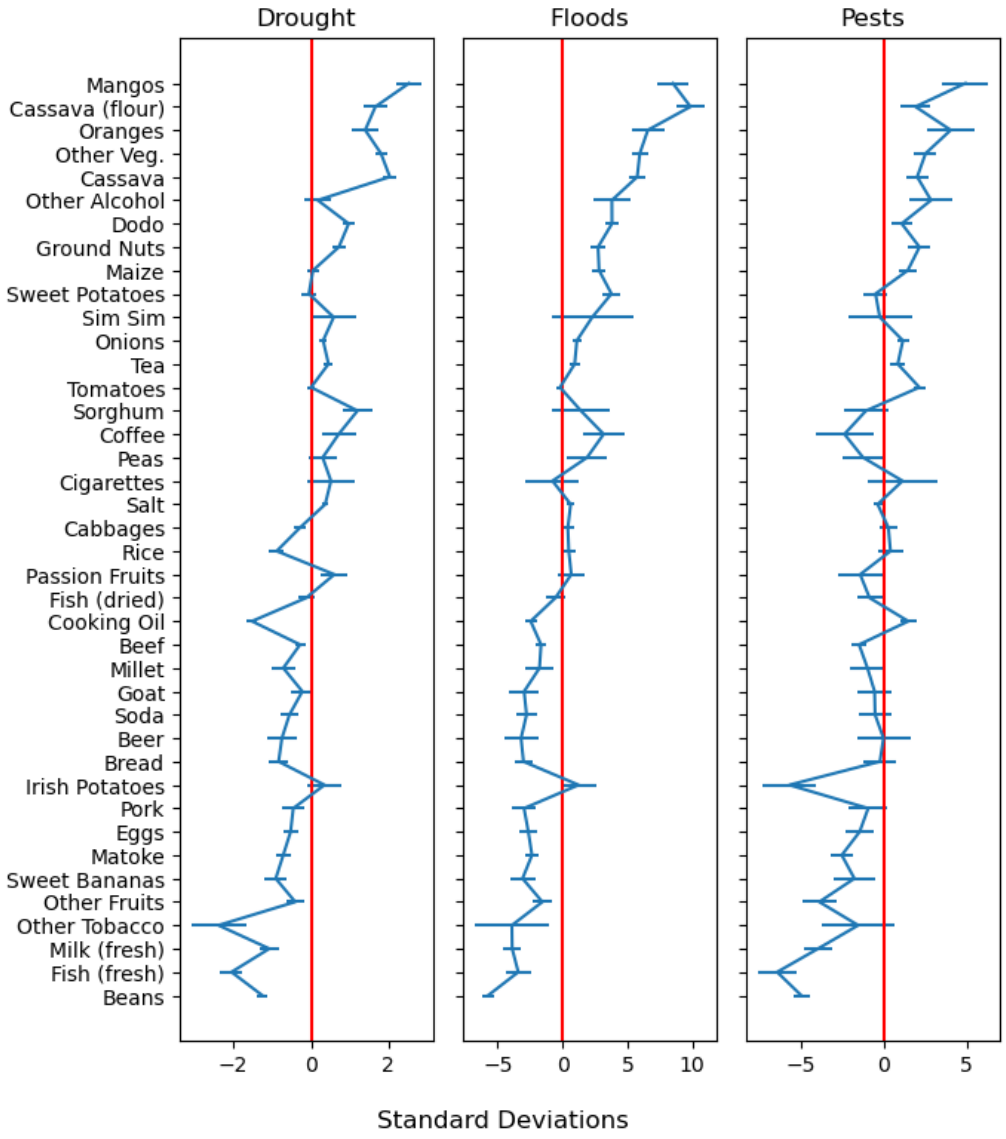


FIGURE C.10. Log Quantities and Shocks. Estimated coefficients from regression of log quantities of different food items on the share of households in a given market-year reporting given covariate shock (year effects included; zeros dropped). Bars indicate 95% confidence intervals. Goods are ordered according to magnitude of coefficients averaged across different shock regressions.

very little response to shocks, bolstering our interpretation that demand for sim sim is importantly related to substitution and complementarities with other foods.

One general observation about each of these figures is the fact that changes in demand for different goods tend to be highly correlated across the different covariate

shocks. Coefficients from both the level and log quantity regressions have a Spearman rank correlation coefficient of 0.92 between drought and flood shocks, for example, while the same correlation in the "positive quantity" regression is 0.91. The correlation between flood and pest coefficients is systematically higher than the correlation between drought and pest, but even the latter ranges from 0.54–0.68. We might read this as evidence that each of these shocks is operating more through its effect on a shared income shock, rather than through prices, since it seems likely that different shocks should have different effects on the local supply of different goods (this is borne out by relatively low correlations between coefficients across shocks in Figure C.6, which range from 0.23–0.61).

Finally, we emphasize that while the results in this section have found evidence that reported covariate shocks affect production, prices, and consumption in ways that seem consistent, that same evidence is all consistent with the hypothesis that the costs of these shocks is *shared* in a fashion consistent with full-insurance. The only evidence we've really produced for or against the full insurance hypothesis as it pertains to covariate shocks has come from the tests in Section 7.

Appendix D. For Online Publication: Data on Shocks online

Data on self-reported shocks comes from a "shocks" module in different rounds of the Ugandan LSMS surveys. There are minor differences in elicitation and in the list of shocks reported in certain years. Table D.5 reports the slight aggregation of shocks we use, while Table D.7 similarly indicates how we've harmonized responses to the question "How did your household cope with this [SHOCK]?".

Another idiosyncratic variable we're interested in is income, if only because of how often it's used in risk-sharing tests. There are three main sources of income for the households in these data: earnings from employment; profits from non-farm enterprises, and income from farming activities. The last is particularly difficult to construct,[25] and so we restrict our attention to non-farm income, constructed as the sum of earnings and profits from non-farm enterprises. In only 9724 of 22,791 total usable observations do households report any non-farm income; in the remaining cases

[25]Income from farming activities would in principle include profits from crop and livestock production, but both revenues and costs are poorly measured in these surveys. Revenues are difficult to compute because for many households a considerable amount of production is directly consumed. Though the surveys elicit respondents' assessment of the value of this own-consumption, it would be an obvious mistake to add this to income and then test whether income is correlated with the value of consumption. Costs are also poorly measured, and for a related reason—the on-farm labor of household members is poorly measured and difficult to value.

TABLE D.5. Different shock labels across eight rounds of Ugandan data, with harmonized labels.

Existing Label	Label
Conflict/Violence	Conflict
Death of Income Earner(s)	Death of earner
Death of Other Household Member(s)	Death (non-earner)
Drought	Drought
Drought/Irregular Rains	Drought
Erosion	Erosion
Fire	Fire
Floods	Floods
Irregular Rains	Drought
Landslides	Erosion
Landslides/Erosion	Erosion
Loss of Employment of Previously Employed Household Member(s) (Not Due to Illness or Accident)	Lost Earnings
Other (Specify)	Other
Reduction in the Earnings of Currently (Off-Farm) Employed Household Member(s)	Lost Earnings
Serious Illness or Accident of Income Earner(s)	Health
Serious Illness or Accident of Other Household Member(s)	Health
Theft	Theft
Theft of Agricultural Assets/Output (Crop or Livestock)	Theft
Theft of Money/Valuables/Non-Agricultural Assets	Theft
Unusually High Costs of Agricultural Inputs	Prices
Unusually High Level of Crop Pests & Disease	Pests
Unusually High Level of Crop Pests & Disease	Pests
Unusually High Level of Livestock Disease	Pests
Unusually Low Prices for Agricultural Output	Prices

we treat our measure of income as missing. We would not argue that this is a very good measure of income, but would argue that the full risk-sharing hypothesis implies that variation in this poor measure of income should not affect households MUEs.

TABLE D.6. Incidence of shocks by round

Shock	2005	2009	2010	2011	2013	2015	2018	2019
Conflict/Violence	268	34	27	46	11	12	22	25
Death of Income Earner(s)	99	27	17	19	30	19	20	19
Death of Other Household Member(s)	423	74	58	35	66	35	44	49
Drought	1234	—	—	—	735	526	556	360
Drought/Irregular Rains	—	1344	710	560	—	—	—	—
Erosion	—	—	—	—	15	3	19	13
Fire	105	26	21	20	17	18	12	9
Floods	426	61	102	148	98	62	74	117
Irregular Rains	—	—	—	—	179	72	180	169
Landslides	—	—	—	—	1	—	2	2
Landslides/Erosion	—	21	5	17	—	—	—	—
Loss of Employment of Previously Employed Household Member(s) (Not Due to Illness or Accident)	—	9	10	8	5	10	5	6
Other (Specify)	111	101	60	57	62	45	50	75
Reduction in the Earnings of Currently (Off-Farm) Employed Household Member(s)	—	28	3	10	6	4	15	15
Serious Illness or Accident of Income Earner(s)	82	189	152	91	86	56	107	119
Serious Illness or Accident of Other Household Member(s)	—	188	149	65	47	32	83	78
Theft	349	—	—	—	—	—	—	—
Theft of Agricultural Assets/Output (Crop or Livestock)	—	127	48	21	35	27	41	40
Theft of Money/Valuables/Non-Agricultural Assets	—	106	48	34	41	35	34	43
Unusually High Costs of Agricultural Inputs	71	60	19	27	49	7	10	12
Unusually High Level of Crop Pests & Disease	—	137	40	61	54	35	—	—
Unusually High Level of Crop Pests & Disease	292	—	—	—	—	—	120	79
Unusually High Level of Livestock Disease	183	82	37	31	17	18	10	5
Unusually Low Prices for Agricultural Output	—	53	35	38	18	5	68	17

TABLE D.7. Different coping labels across eight rounds of Ugandan data, with harmonized labels.

Existing Label	Label
Unconditional help provided by relatives/friends	Help from relatives/friends
Unconditional help provided by local government	Help from local government
Sold land/building	Sold assets
Sold durable household assets (agricultural or nonagricultural)	Sold assets
Sold durable household assets (agricultural or non-agricultural)	Sold assets
Sold assets	Sold assets
Sent children to live elsewhere	Sent children elsewhere
Rented out land/building	Rented out assets
Relied on savings	Used savings
Reduced expenditures on health and education	Reduced investments in human capital
Reduced consumption	Reduced consumption
Other(specify)	Other
Other (specify)	Other
Obtained credit	Used credit
Nothing else	Nothing
More wage employment	Reduced leisure
Increased agriculture labour supply	Reduced leisure
Household member(s) took on more non-farm (wage- or self-) employment	Reduced leisure
Household member(s) took on more farm wage employment	Reduced leisure
Household member(s) migrated	Migrated
HH member took on more non-farm (wage- or self-) employment	Reduced leisure
HH member took on more farm wage employment	Reduced leisure
Distress sales of animal stock	Sold assets
Did nothing	Nothing
Changed dietary patterns involuntarily (Relied on less preferred food options, reduced the proportion or number of meals per day, skipped days without eating, etc)	Reduced consumption
Changed dietary patterns involuntarily	Reduced consumption
Changed cropping practices (crop choices or technology)	Changed cropping practices
Changed cropping practices	Changed cropping practices
Rented out land/building	Rented out assets
99.0	Nothing

University of California, Berkeley