Livelihood strategies in the rural Kenyan highlands

DOUGLAS R BROWN
World Vision Canada, Mississauga, Ontario, Canada

EMMA C STEPHENS
Department of Economics, Cornell University, Ithaca, NY, USA

JAMES OKURO OUMA
Kenya Agricultural Research Institute, Embu, Kenya

FESTUS M MURITHI
Kenya Agricultural Research Institute, Nairobi, Kenya

CHRISTOPHER B BARRETT*
Department of Applied Economics and Management, Cornell University, Ithaca, NY, USA

Abstract

The concept of a livelihood strategy has become central to development practice in recent years. Nonetheless, precise identification of livelihoods in quantitative data has remained methodologically elusive. This paper uses cluster analysis methods to operationalize the concept of livelihood strategies in household data and then uses the resulting strategy-specific income distributions to test whether the hypothesized outcome differences between livelihoods indeed exist. Using data from Kenya's central and western highlands, we identify five distinct livelihood strategies that exhibit statistically significant differences in mean per capita incomes and stochastic dominance orderings that establish clear welfare rankings among livelihood strategies. Multinomial regression analysis identifies geographic, demographic and financial determinants of livelihood choice. The results should facilitate targeting of interventions designed to improve household livelihoods.

Keywords: Livelihood strategy; Kenya; Smallholder agriculture; Cluster analysis

Depuis quelques années, le concept d'une stratégie de moyens d'existence est devenu un élément central dans la pratique du développement. Malgré tout, l'identification des moyens d'existence sous forme de données quantitatives, demeure méthodologiquement vague. Cet article utilise des méthodes analytiques groupées pour rendre opérationnel le concept de stratégies de moyens d'existence en matière de données concernant les ménages. Il utilise ensuite les distributions de revenu résultantes et spécifiques aux stratégies afin de vérifier si les différences hypothétiques des résultats en matière de moyens d'existence existent bel et bien. Grâce aux données provenant des régions des hautes terres du centre et de l'ouest du Kenya, nous pouvons identifier cinq stratégies distinctes de moyens d'existence. Celles-ci exposent, du point de vue statistique, les différences importantes du revenu moyen par personne et les ordres hiérarchiques stochastiques qui établissent un classement précis de l'assistance parmi les stratégies de moyens d'existence. Une analyse de régression multinomiale identifie les déterminants géographiques, démographiques et financiers du choix de moyen d'existence. Les résultats devraient faciliter la prise pour cible des interventions destinées à améliorer les moyens d'existence des ménages.

Mots clés : Stratégie de moyens d'existence; Kenya; Agriculture des petits fermiers; Méthodes analytiques groupées

* Corresponding author: cbb2@cornell.edu

1. Introduction

Rural households earn income from diverse allocations of their natural, physical and human capital assets among various income generating activities. The literature offers many reasons why such diversification occurs (see Ellis 1998 and Barrett et al. 2001 for helpful surveys). Among these might be diminishing returns on increasing investment in certain activities, synergies (economies of scope) among distinct activities, or missing markets that compel self-provision of goods or services the household desires for own consumption. Similarly, households may wish to diversify as a strategy for coping with an unexpected shock, or to minimize risk ex ante by participating in activities that generate imperfectly correlated returns. The presumption throughout the literature is that households choose such patterns of diversification so as to achieve the best possible standard of living, broadly defined. The chosen combination of assets and activities is often referred to as the household's 'livelihood strategy'. A livelihood strategy encompasses not only activities that generate income but many other kinds of choices, including cultural and social choices, that come together to make up the primary occupation of a household (Ellis 1998).

The concept of a livelihood strategy has become central to development practice in recent years. Nonetheless, given the uncountable possible proportional mixes of activities undertaken by a household, it is not always clear what constitutes a distinct livelihood rather than just a slightly different mix of activities within the same general livelihood. A precise operational definition of *livelihood* remains elusive, as does an associated method for identifying livelihoods in quantitative data. This probably helps explain why the more quantitative development scholars (e.g. economists) have been slower to adopt the concept than have the more qualitative ones (e.g. anthropologists and sociologists).

The ability to operationalize the concept of a livelihood strategy becomes especially important when one speaks of 'improving' livelihoods, to paraphrase much current development discourse. Implicit in the concept of 'improvement' is the suggestion that certain strategies offer households a higher return on their assets, not least of which is household labor. But if we cannot pin down the boundaries between distinct livelihood strategies, how can we distinguish graduation to an improved livelihood (i.e. a better outcome from a different choice) from improvement in the performance of a given livelihood (i.e. higher productivity from the same basic choice, perhaps due to improved technical or allocative efficiency of practice or technological progress)? And within the latter class of improved outcomes, how can we distinguish permanent from transitory gains? These are not merely esoteric, intellectual questions about labeling and measurement. Much development programming today hinges, if only implicitly, on the assumption that there exist discernible orderings of distinct livelihood strategies and that carefully tailored and implemented interventions can effectively facilitate graduation to more desirable livelihood strategies – ones that are associated with improved well-being of household members.

This paper tackles the problem of how to operationalize the concept of livelihood strategies in quantitative household data and how then to use the identified strategies to test whether the hypothesized outcome differences between livelihoods indeed exist. We introduce an approach based on statistical cluster analysis to identify and order distinct livelihood strategies and to identify the correlates of access to the most desirable of those livelihood strategies. Such an analysis seems a prerequisite to credible targeting of interventions aimed at improving choices and outcomes for poor rural peoples.

Several different methods of characterizing household livelihood strategies can be found in the literature. Most commonly, economists group households by shares of income earned in different sectors of the rural economy. For example, Barrett et al. (2005) analyzed the relationship between overall household income and the proportion of income earned in on-farm and off-farm activities in several African countries, noting how these proportions changed across income quartiles and that different income sources became dominant as one moved up the income distribution. Dercon and Krishnan (1996) used income share composition to examine the relationship between income, household characteristics and barriers to entry into higher return activities. Others have examined the potential determinants of diversified income portfolios for rural smallholders (Reardon et al. 1992). The common denominator of this literature is that data on realized incomes underpin most classifications.

We have two concerns with this approach. First, sustainable livelihoods approaches typically focus on asset allocations across distinct activities, i.e. on agents' behaviors rather than on income outcomes that are heavily impacted by stochastic, exogenous factors (Ellis 1998; Scoones 1998; Bebbington 1999). Consequently, categorizations based on realizations of stochastic incomes involve a subtle but important disconnect from the underlying concept. Second, the categorizations used are based on inherently arbitrary decisions by external observers – e.g. why is the sector of employment of primary importance? – and may thereby exaggerate some differences while understating others.

An alternative method of analyzing livelihood strategies involves direct examination of the individual household's asset endowment. The amount of income earned and even the type of activity undertaken by a household is a stochastic function of the assets it controls. Certain activities may be beyond the reach of households without access to the required financial, natural, physical, human or social capital.[1] This asset-based approach makes it possible to map a household's asset endowment into its chosen livelihood strategy and then into its (logically subsequent) stochastic income realization (Carter & Barrett 2006). Households with similar bundles of assets might be limited to similar livelihood strategies, but in any given period realize quite different incomes, although they are structurally identical.

This paper presents an asset-based approach to identifying livelihood strategies, letting the data direct us as to how best to group household assets, and then testing for prospective welfare orderings among the stochastic income distributions associated with each identified livelihood strategy. We apply this method to data from two similar agro-ecological regions in the Kenyan highlands. We find that there exist five distinct livelihood strategies discernible in these data and that one livelihood strategy second degree stochastically dominates the other four. A multinomial logit regression of household characteristics on three slightly broader classifications of livelihood strategy choice (low, medium and high return) highlights key household attributes – geographic location, family size, farming experience, access to credit and remittances – that statistically differentiate households pursuing low, medium and high return livelihood strategies. Asset endowments indeed appear quite important not only in choosing among empirically distinct livelihood strategies but also in the returns earned from these strategies. Categorization of different strategies by asset allocation offers a meaningful and tractable way to operationalize the livelihood strategies concept and to discern broader patterns in livelihood choices for rural households.

2. Assets, income and portfolio choice

Assume that a household seeks to maximize its utility defined over stochastic income[2] by allocating its given asset endowment across a set of feasible activities, i=1,…,N. Then, a simple revealed preference argument suggests that, where different asset allocation strategies yield different income distributions that can be ordered in welfare terms (e.g. via stochastic dominance criteria), any households observed to have adopted a lower return livelihood strategy must have faced a constraint that limited its choice set relative to those of its neighbors, as no one would freely choose to draw from a stochastically dominated distribution when they had access to a dominating alternative. Identifying low return livelihood strategies and the households that choose them is thus critical to targeting interventions intended to advance poverty reduction policy objectives.

Slightly more formally, assume that:

$$(1) \qquad y_i = f_i(A_i) + \varepsilon_I$$

[1] The sustainable livelihoods framework divides livelihood assets into five groups: human, natural, financial, social and physical capital (DFID 2001).

[2] So long as utility is monotonically increasing in income, this is a very mild simplifying assumption.

where f_i is an increasing function relating household assets in activity i, A_i, to the stochastic return from an activity y_i, with ε_i an error term that represents unexpected shocks to activity income and any measurement error. Household total income $Y = \Sigma_i\ y_i$. If the household maximizes its utility defined over income realizations, then the household's choice is simply an optimal allocation of its asset endowment, A_0:

$$\max_{A_i} U(\Sigma_i\ y_i = \Sigma_i\ f_i(A_i) + \varepsilon_i) \quad s.t. \Sigma_i A_i \leq A_0 \tag{2}$$

The resulting choice – an allocation of the household's assets across various activities – represents its chosen livelihood strategy.

Given this conceptual structure for livelihood strategy choice, at any given point in time, the household compares the marginal utility associated with alternative asset allocations, taking into consideration the expected income from a given allocation – described by $f_i(A_i)$ – as well as the full underlying distribution of each ε_i, which generates a welfare ordering among alternatives generating exactly the same expected income.[3] Under the simplest formulation of the above problem, assuming constant returns to scale for all activities and complete, competitive markets that equalize returns across activities, $f_i(A_i) = R'A_i$, because each asset earns a fixed rate of return, R, no matter the volume of the asset or the activity to which it is allocated. More generally, however, returns on assets might vary across activities and by scale or scope if there exist barriers that restrict households' ability to allocate assets freely across all activities observed in an economy – for example, borrowing constraints that limit households' ability to supplement their endowments temporarily, market access differences that cause cross-sectional variation in returns across households, or asset specificity that matches particular assets (e.g. cows) with certain activities (e.g. dairy production). As a result, differences in standards of living may persist across households (Dercon & Krishnan 1996; Carter & Barrett 2006).

Given the difficulty of observing risk preferences, it is typically easier to observe sample information about the probability distribution associated with each activity i. Economists therefore commonly use stochastic dominance analysis to rank alternative risky choices given only weak information on preferences. In particular, first- and second-order stochastic dominance generates clear welfare orderings among alternative choices under the mild assumptions that utility is increasing and weakly concave in income (Whitmore & Findlay 1978).[4] By the optimum principle, asset allocations under the problem defined in equation (2) – the observed livelihood strategies describing household allocation of assets across a portfolio of activities – should not be second-order stochastically dominated by any other feasible choice. In any given period, a particular asset allocation may yield a low income relative to some feasible alternative(s), i.e. a poor draw from the relevant distribution. But that should result purely from the inherent stochasticity of the income-generating process, not from errors of household livelihood choice, which one would expect would vanish over time in reasonably stable systems as people experiment with different portfolio mixes and observe their neighbors' experiments as well.

3. Data

The households under investigation in this study live in two highland agricultural areas of central (Embu District) and western (Vihiga District) Kenya that exhibit several crucial similarities (Place et al. 2005).

[3] In the interests of simplicity of exposition, we treat this as a static problem and ignore the possibility of investment in assets and the laws of motion governing assets and productivity.

[4] First-order dominance only requires monotonicity; second-order dominance requires concavity, reflecting income risk aversion, both mild assumptions regarding preferences.

Both consist largely of high potential agricultural land above 1200 meters elevation above sea level, and receive around 1500 mm of rainfall per year on average, distributed bimodally. While the soils in both areas are of similar texture, those in central Kenya have a slightly better natural fertility due to volcanic parent materials. Population densities are high, but more so in western Kenya, with Vihiga District averaging 886 people/km^2 (IEA 2002). As a result, average farm sizes are slightly smaller in western Kenya than in the central highlands.

The data used in this paper come from detailed household surveys fielded in 2002–2003 in four villages in Embu District and in one village, Madzuu, in Vihiga District. A village is the smallest unit in the Kenyan administrative structure. The geographical sizes and population of villages vary from district to district. In Embu, 113 randomly sampled households were interviewed, while 127 were interviewed in Vihiga.[5] The surveys collected a wide variety of information on agricultural production at the plot level, and household level data on livestock holdings, off-farm income activities, earnings and household characteristics.

In striking contrast to the broadly similar geographic conditions, economic conditions between the two Districts vary widely. Embu's closer proximity to the capital city of Nairobi gives rural households in this region several key advantages, including access to larger agricultural markets, much lower transport costs and better and more remunerative off-farm income opportunities (Place et al. 2005). In general, Embu households possess more productive assets than do their Vihiga counterparts and receive higher prices for both their agricultural output and off-farm labor. The combination of superior access, somewhat better soils and greater asset holdings results in significant regional differences in income distributions. The per capita daily income distribution for farm households in central Kenya first order stochastically dominates that of the western highlands households in our sample (Figure 1).

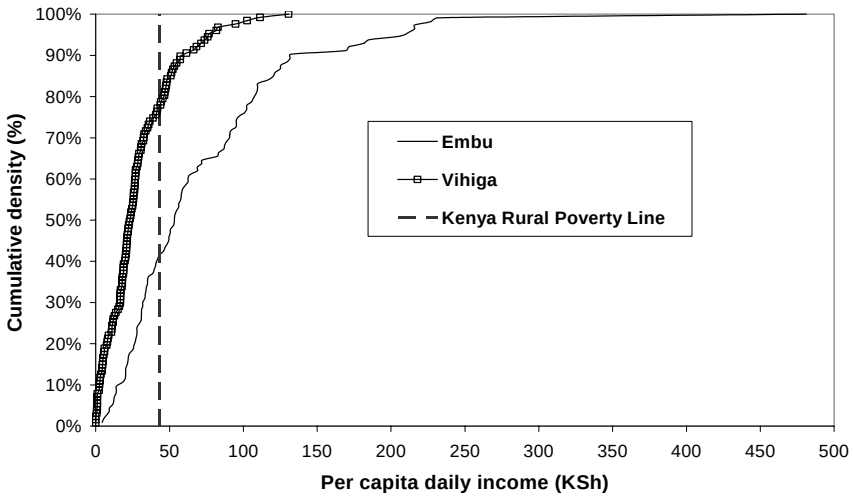

Figure 1: Cumulative density of per capita income distribution in Embu and Vihiga samples. Per capita daily poverty line shown = KSh43 (CBS 2004)

But what is equally obvious from Figure 1 is that within each area there exists a considerable dispersion of household incomes. This should serve as a caution against strict geographic determinism when making inferences about income differentials, and helps to motivate our interest in identifying distinct livelihood strategies and disaggregating the analysis among such livelihoods.

[5] Of the 127 households in Vihiga, 89 were part of a randomly sampled group from a 1989 study in the area, and the data from both Embu and Vihiga are part of a larger panel study. More detail on the Vihiga sample and the broader study from which these data are derived are provided in Barrett et al. (2006). More detail on the Embu data can be found in Wangila et al. (2005).

4. Common livelihood activities

Households in the sample engage in a variety of farm and non-farm activities. We attempted to identify common livelihood activities by summarizing plot-level crop data, types of animal species held and skilled or unskilled off-farm income. As is typical in household-level data, at this level of detail, subtle variations in each household's portfolio holdings made it appear as if each farmer in the sample followed a slightly different livelihood strategy from all the others. Such extreme disaggregation of course prevents necessary simplification for the purposes of policy analysis. We therefore sought a theoretically and statistically defensible means of aggregating the data into distinct livelihood categories.

The categories were chosen with many criteria in mind, including clearly distinct input requirements and different roles within the biophysical system from which Kenyan farming households generate income. We grouped crop production activities into three distinct groups, with two sub-groups: annual food crops (cereals, legumes, roots and tubers, and vegetables), perennial fodder crops (i.e. Napier grass), and perennial cash crops – with separate sub-group classifications for coffee and tea, the two primary cash crops at the Embu and Vihiga sites. We similarly created four livestock groups: improved dairy cattle based on cross- or pure-bred exotic breeds, local breed dairy cattle, non-dairy (i.e. beef) cattle, and small ruminants (e.g. goats, sheep) and pigs. Each of these livestock groups requires different management strategies and produces different outputs, hence their identification as distinct activities. Finally, off-farm income-generating activities were grouped into two categories according to the required skills and associated average earnings: low return, unskilled and high return, skilled.[6] Unskilled employment and self-employed opportunities are universally available, while the more remunerative (on average) options are only open to those possessing the relevant skills, education, or both. We thus have N=11 distinct activities in the empirical analysis that follows: annual food crop production, perennial cash crop production, coffee, tea, perennial forage crop production, improved dairy cattle, local (unimproved) dairy cattle, non-dairy cattle, small ruminants and pigs, unskilled off-farm employment and skilled employment.

Table 1 summarizes the frequencies and means for each activity type in Embu, Vihiga and the sample as a whole. Although most of the activity types are found at both sites, Vihiga farmers are primarily engaged in food crop production, with fewer households observed cultivating forage or cash crops (and none grow coffee). They keep fewer improved dairy and non-dairy cattle, and more household members were engaged in unskilled off-farm employment than in Embu. The observed means of the relevant activities, as well as the proportion of each population engaged in each activity type, differed significantly between the Embu and Vihiga sites at either the 5% or 1% level, as shown in Table 1, except in the case of skilled employment and small ruminant holdings. This serves to further emphasize that the livelihood strategies chosen by households in Embu and Vihiga are quite different, with Vihiga households clearly more often involved in low return activities. As well, Vihiga households appear to possess fewer agricultural assets than their Embu counterparts, with smaller average areas for agricultural crops and smaller herd sizes (with the exception of local dairy cattle).

[6] A household is listed as having low return, unskilled off-farm income if at least one member of the household is engaged in agricultural wage work, non-agricultural wage work, ox-cart business, driver, goods transport and/or petty trade. High return, skilled off-farm employment includes teaching, artisan/blacksmith, anyone receiving a pension (which is associated with continued income from a high return wage position), craft sales, shop keeping/trader, salaried employment, bar business, butcher, medical clinic, civil servant or military/police.

Table 1: Frequencies and means for different activities in Embu and Vihiga

Activities	Percentage of Embu households	Percentage of Vihiga households		Means			
				Embu (N=113)	Vihiga (N=127)		Full sample (N=240)
Food crops[a] (acres)	98.2	100.0	-	1.61	0.94	**	1.25
Cash crop perennials[b] (acres)	54.9	8.7	***	0.31	0.08	***	0.19
Napier grass (acres)	20.4	1.6	***	0.07	0.003	***	0.03
Coffee (acres)	6.2	0.0	***	0.04	0.00	**	0.02
Tea (acres)	45.1	8.7	***	0.23	0.08	**	0.15
Local dairy cattle (head)	4.4	62.2	***	0.06	1.17	***	0.65
Improved dairy cattle (head)	62.0	25.2	***	1.26	0.41	***	0.81
Non-dairy cattle (head)	27.4	3.2	***	0.36	0.04	***	0.19
Small ruminants (head)	37.2	27.6	-	1.46	0.71	*	1.06
Low return off-farm income (number of individuals)[c]	8.9	49.6	***	0.10	0.73	***	0.43
High return off-farm income (number of individuals)[d]	30.1	31.5	-	0.38	0.40	-	0.39

*=significant at 10%, **=significant at 5%, ***=significant at 1%

[a] Households engaged in 'food crop' activity have at least one plot containing maize, beans, Irish potatoes, sweet potatoes, cassava, kale/leafy greens, cabbage, tomato, arrowroot, cowpeas, carrots, climbing beans, sunflower, groundnuts and/or millet.

[b] Cash crop perennials include coffee, tea, yams, bananas, pawpaw (papaya), passion fruit, sugar cane, mango, macadamia and/or avocado.

[c] Households have low return off-farm income if at least one member of the household is engaged in this activity. Activities include agricultural wage work, non-agricultural wage work, ox-cart business, driver, transporting goods, and petty trade.

[d] Households have high return off-farm employment if at least one member of the household is engaged in this activity. Activities include teaching, artisan/blacksmith, anyone receiving a pension, craft sales, shopkeeping/trader, salaried employment, bar business, butcher, medical clinic, civil servant or military/police.

5. Identifying livelihood strategies via cluster analysis

In the livelihoods literature, strategies are commonly identified by broad but inherently arbitrary rules that partition the sample among exogenously defined groupings of activities, for example by looking at activity choices in different income quartiles (Barrett et al. 2005). In this paper we opt instead to let the data tell us how best to partition the sample into statistically distinct livelihood strategies reflecting how households allocate assets across the 11 different activities identified in the preceding section. We do this via cluster analysis techniques.

Cluster analysis is a statistical data reduction method for summarizing a large number of sample observations by assigning them to a smaller, tractable number of distinct groups – or 'clusters' – of observations. The core idea is that there are some latent common features that enable one to agglomerate individual observations into a small number of groups based on similarity along particular, pre-determined dimensions.[7] Similarity is measured with reference to a particular statistic (e.g. mean or median) of the cluster and does not require any assumptions about the distributions describing observations within or between clusters. Cluster analysis is therefore a highly flexible and intuitive, albeit computationally

[7] Stata 9.0 was used for the cluster analysis. See Everitt et al. (2001) for a good introduction to cluster analysis techniques.

intensive, method for letting the data speak for themselves in defining focal asset allocations that might usefully define distinct livelihood strategies.

In this paper, we use asset allocation across the 11 activities we have identified as the relevant dimensions and an algorithm that groups observations around central points by minimizing distances of observations in a group from the central point while simultaneously maximizing the distance between the points. More precisely, we performed a k-means cluster analysis (Jansen et al. 2003) to assign each household to a distinct group based on the following variables: land area cultivated in (i) food crops, (ii) pure stand perennials, (iii) pure stand Napier grass, (iv) pure stand coffee and (v) tea; and number of (vi) improved dairy cattle, (vii) local (unimproved) dairy cattle, (viii) non-dairy cattle and (ix) small ruminants owned; and number of household members engaged in (x) unskilled and (xi) skilled off-farm employment and self-employment activities.

K-means cluster analysis is a non-hierarchical method of partitioning data into a predetermined number of groups. Observations are initially randomly assigned to each of the k clusters, and then reassigned using an iterative method so as to minimize within-cluster variance and maximize between-cluster variance. In polythetic (i.e. multivariate) cluster analysis, the between-cluster variance is measured with respect to the Euclidean norm of the cluster means across the vector of variables used as defining characteristics.[8] Convergence is achieved when any further reassignment of observations across groups would increase within-cluster variance.

Based on statistical results[9] and common sense checks of the resulting groupings, we identified five distinct livelihood strategy clusters in these data (i.e. k=5). The mean values for each activity in each livelihood strategy are summarized in Table 2. Daily per capita household income was computed as the sum of the value of annual crop agricultural output, evaluated at average local prices, plus any earnings from livestock or livestock product sales, off-farm income and remittances; that sum was then divided by the headcount of individuals in the household and converted to a daily measure.

[8] K-medians cluster analysis yielded qualitatively identical results in these data. Those results are omitted in the interest of brevity.
[9] Available from corresponding author on request.

Table 2: Livelihood strategies estimated via K-means cluster analysis (Mean value of livelihood activity by cluster)

Clustering variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Sample means
Food crops (acres)	0.50	0.98	1.44	1.03	0.70	1.25
Cash crop perennials (acres)	0.00	0.09	0.15	0.11	3.08	0.19
Napier grass (acres)	0.00	0.02	0.01	0.04	0.70	0.03
Coffee (acres)	0.00	0.00	0.01	0.01	0.50	0.02
Tea (acres)	0.00	0.05	0.12	0.09	2.35	0.15
Local dairy cattle (head)	1.00	0.84	0.55	0.85	0.00	0.65
Improved dairy cattle (head)	0.00	0.39	0.92	0.88	2.00	0.81
Non-dairy cattle (head)	0.00	0.16	0.20	0.21	0.20	0.19
Small ruminants (head)	1.00	0.76	0.78	2.59	2.20	1.06
Low return off-farm income (number of individuals)	3.40	1.22	0.00	0.79	0.00	0.43
High return off-farm income (number of individuals)	0.00	0.00	0.27	1.50	0.80	0.39
Other variables						
Fraction of households that are in Vihiga (%)	100.0	83.7	38.8	70.6	0.0	52.9
Daily per capita income (KSh)	11.78	31.46	50.06	65.48	136.30	49.44
(standard deviation)	(10.02)	(29.36)	(42.74)	(92.94)	(71.64)	(53.48)
Number of households	5	49	147	34	5	240
Fraction of households (%)	2.1	20.4	61.3	14.2	2.1	100
Strategy name	**Part-time subsistence smallholder/unskilled worker**	**Mixed smallholder**	**Staples producer**	**Off-farm high return**	**Diversified commercial**	**Whole sample**

The first strategy (cluster 1), **part-time subsistence smallholders/unskilled workers**, is employed by a small number of farm households. They have the lowest mean income of the five strategies and the smallest land holdings (their mean land area is just 0.5 acres). On average, they have one local dairy cow and one small ruminant (sheep or goat). They allocate all of their land to annual food crops. But because they are not self-sufficient in food crop production, they supplement their income with low-wage, unskilled off-farm work. No household members obtain skilled employment or self-employment and, on average, three to four of them are engaged in unskilled off-farm employment, underscoring how important local labor markets are to the well-being of these poorest households. In our sample, cluster 1 households appear only in Vihiga.

The households in cluster 2, **mixed smallholders**, represent 20.4% of the total sample. Most of them (83.7%) are from the western Kenyan (Vihiga) sub-sample. They farm a little more than twice as much land (1.09 acres) as those in cluster 1, but with nearly all of it (89.9%) in food crops, and just a small fraction in perennial forage and cash crops, mostly tea. Cluster 2 households' total livestock holdings are similar to those of the part-time subsistence smallholders/unskilled workers, but they are more likely to have an improved dairy cow and some non-dairy cattle. Average employment in unskilled off-farm work is just over one third that of cluster 1 households, but still without any skilled employment or self-employment. Cluster 2 households' mean per capita daily income is one and two thirds times greater than that of cluster 1 households, but still below the rural Kenyan poverty line of KSh43, which is equivalent to US$0.53/day per capita, and far below the dollar-a-day extreme poverty line frequently used in international comparisons.[10]

[10] In 2002, US$1 = 78.749 Kenya Shillings (KSh) (https://www.cia.gov/cia/publications/factbook/geos/ke.html)

Cluster 3, **staples producers**, the largest of the five, represents just over 60% of the total sample. On average, they farm 1.6 acres of land, i.e. more than three times as much as cluster 1 and half again as much as cluster 2. While about 90% of their land holdings are in annual food crops, as with the cluster 2 households, the staples producers grow more of both annual food and perennial cash crops by virtue of their greater land holdings. Indeed, cluster 3 households have the largest area in food crops of all the five groups, hence the 'staples producers' label. They have more livestock on average than households in the first two clusters, with improved dairy cattle far more common. None of these households engage in unskilled off-farm employment, although they do enjoy off-farm earnings from skilled employment, roughly one quarter of a person per household, on average. The key features which distinguish cluster 3 from cluster 2 households are the greater total stock of land and livestock – especially improved dairy cattle – and access to skilled employment.

The distinguishing feature of cluster 4, **off-farm skilled employment**, representing a little more than 14% of the sample households, is their greater reliance on skilled off-farm employment as a source of income. They also keep more small ruminants than any of the other clusters. They keep improved and local breeds of dairy cattle in roughly equal proportions (nearly one head of each, on average, per household), and farm just over one acre on average, again about 90% in annual food crops. Unlike the staples producers, many cluster 4 households supplement their farm and skilled off-farm employment with unskilled off-farm employment. They are thus the most diversified in terms of varied activities, earning returns across semi-subsistence and commercial farming along with unskilled and skilled off-farm employment.

Cluster 5 exhibits characteristics best described as **diversified commercial**. Their average per capita income is over twice that of the next highest strategy (cluster 4). Households in this cluster put far less emphasis on food crop production – devoting as much land area to perennial fodder production as they do to annual food crop production – and putting nearly 70% of their land into perennial cash crops, mainly tea and coffee. Their farming operations integrate relatively large improved dairy herds – more than twice as large as any other cluster's – and sizable small ruminant herds – more than twice as large as any but cluster 4's – to constitute highly diversified commercial farms. They have no local dairy cattle and no household members engaged in unskilled employment, but do supplement their on-farm income with some skilled, off-farm employment. Cluster 5's livelihood strategy is the only one that generates an average per capita income greater than a dollar a day.

5.1. Are some livelihood strategies superior to others?

Cluster analysis provides a convenient, intuitive way to let the sample data speak for themselves in identifying distinct groupings that seem meaningful in interpreting the different strategies in play in the rural Kenyan highlands. Perhaps the most striking result is the mean income differences between the distinct livelihood strategies, ranging from KSh12 per day per capita for cluster 1 to more than KSh136 per day per capita for cluster 5. One-way analysis of variance confirms that the variation in daily per capita income is statistically significantly different between several clusters. Pairwise comparison of income levels in the five clusters found differences between cluster 5 and each of the other clusters and between clusters 4 and 2 that are statistically significant at the 5% level.[11]

We further explore these apparent differences in returns on different livelihood strategies using stochastic dominance analysis. The cumulative per capita daily income densities for each livelihood strategy group are plotted in Figure 2, assuming these densities consistently approximate the income distribution facing households engaged in each strategy.[12] We can then test for the stochastic dominance between each pair of

[11] A Bonferroni correction was applied to the one-way analysis of variance results to correct for possible spurious inference due to making multiple comparisons between group means.

[12] Given random sampling from the population, this should be true asymptotically, although the assumption is somewhat stronger in these small samples.

livelihood-specific income distributions. A particular livelihood strategy first-order stochastically dominates another strategy if and only if, for every possible income level, the strategy has a lower cumulative density, reflecting a greater likelihood of drawing higher incomes (Whitmore & Findlay 1978). Using this criterion, the diversified commercial strategy appears to first-order stochastically dominate the three lowest return strategies (clusters 1, 2 and 3), and the staples producer and off-farm skilled employment strategies dominate the two smallholder strategies.

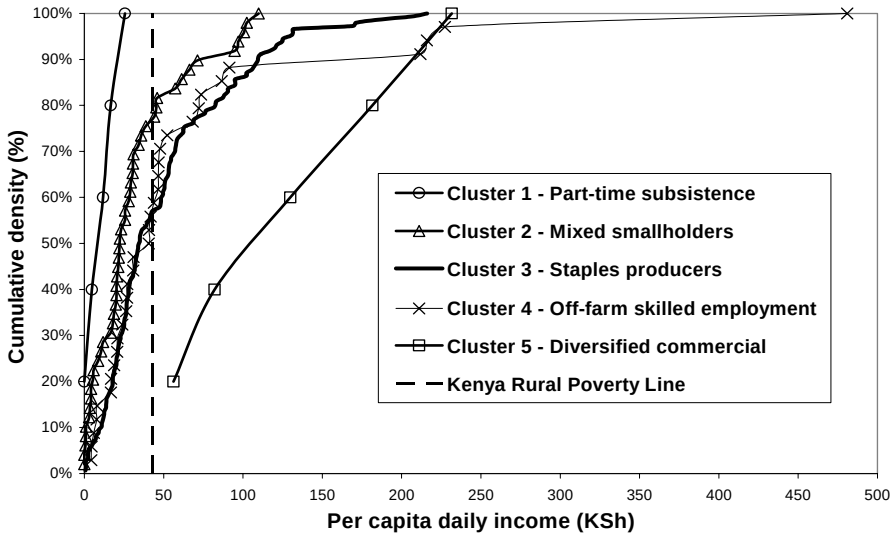

Figure 2: Cumulative density of income distribution for clusters 1 to 5 using K-means cluster analysis

The associated income distributions are measured with error, however. Therefore we estimated confidence bounds around each distribution, following Davidson & Duclos (1997), and tested for first, second and third order dominance between the distributions.[13] The resulting confidence bounds are large, given the small sample sizes. Nonetheless, second-order dominance can be inferred for the diversified commercial strategy over the other four livelihood strategies. Provided that the relatively mild assumption that households are risk averse and prefer more income to less holds true, a second-order dominance ranking implies that the diversified commercial strategy is preferred over all others.

5.2 Are there barriers to adoption of the dominant livelihood strategy?

Very few households in the sample chose the dominant, diversified commercial agriculture livelihood strategy, however. This suggests the possibility of significant barriers to adoption of the most remunerative livelihood strategy. In order to test for patterns in the adoption of distinct livelihood strategies, we performed a multinomial logit regression on livelihood strategy choice as a function of household characteristics.[14] Owing to the small numbers of households observed in the part-time smallholder/unskilled worker and the diversified commercial strategies, the households pursuing these strategies were combined with those in the neighboring strategies to generate a trinomial dependent variable. Key household level covariates are summarized in Table 3.

[13] Approximately 50 test points were considered, up to a maximum per capita income level of KSh200.
[14] Multinomial probit estimation results in almost exactly the same parameter estimates and statistical significance. We omit those results.

Table 3: Summary of household characteristics

Variable	Mean	S. D.	Min	Max	Missing
Smallholders (strategies 1 and 2)					
Household in Vihiga (y/n)	0.85	0.36	0	1	0
Household size[a] (numbers)	5.06	1.89	1	9	0
Farm size (acres)	1.20	2.73	0.20	20.06	0
Total livestock (TLU)[b]	1.43	1.20	0	5	0
Age of household head (years)	54.81	15.12	29	88	1
Head's years of farming experience (years)	22.53	14.53	1	67	3
Self-reported access to credit (y/n)	0.22	0.42	0	1	5
Receives remittances (y/n)	0.20	0.41	0	1	0
Household head has no education (y/n)	0.09	0.29	0	1	0
Household head has secondary education or above (y/n)	0.26	0.44	0	1	0
Staples producers (strategy 3)					
Household in Vihiga	0.39	0.49	0	1	0
Household size (numbers)	5.01	2.35	1	13	0
Farm size (acres)	1.82	3.21	0.13	26.00	0
Total livestock (TLU)	1.75	1.43	0	7	0
Age of household head (years)	57.70	15.17	23	88	5
Head's years of farming experience (years)	30.16	15.89	0	67	17
Self-reported access to credit	0.60	0.49	0	1	3
Receives remittances	0.35	0.48	0	1	0
Household head has no education	0.22	0.41	0	1	0
Household head has secondary education or above	0.21	0.41	0	1	0
Off-farm high return and commercial (strategies 4 and 5)					
Household in Vihiga	0.62	0.49	0	1	0
Household size (numbers)	5.74	1.97	2	10	0
Farm size (acres)	1.32	1.44	0.00	5.50	0
Total livestock (TLU)	2.23	1.69	0	8	0
Age of household head (years)	48.74	12.83	28	77	0
Head's years of farming experience (years)	19.77	13.72	1	54	4
Self-reported access to credit	0.69	0.47	0	1	4
Receives remittances	0.51	0.51	0	1	0
Household head has no education	0.03	0.16	0	1	0
Household head has secondary education or above	0.51	0.51	0	1	0

[a] In this study, a household referred to members of the immediate family who use a common resource such a piece of land and make joint livelihood decisions as a unit. Household size refers to number of individuals in the immediate family.
[b] 1 Tropical Livestock Unit = 1 head of cattle = 10 sheep = 10 goats = 10 pigs.

The parameter estimates and marginal effects are presented in Table 4. The likelihood of choosing the staples producer livelihood strategy over the mixed smallholder strategy (the comparison case) is substantially lower in the more degraded western Kenyan soils. This may be partly due to higher population density, as the probability of being a staples producer modestly, but significantly, decreases with household size. The only other covariate that proved statistically significant in explaining choice between the two lesser performing livelihood strategies was farmer experience, which increases the probability of choosing the staples producer strategy, consistent with the idea that this reflects superior productivity that induces full engagement of all unskilled household labor on-farm. The significance of the village dummy accords with the distribution of households in the sample and reinforces the point that the lowest return strategies are most often found in Vihiga, where both high population density, which leads to extremely small farm sizes, and lack of access to a large market such as Nairobi must surely play a part in this phenomenon. The years of farm experience variable may represent the ability over time to adopt somewhat higher return strategies, but also the inability of asset accumulation alone to move households into the highest return livelihoods.

Table 4: Multinomial logit regression of livelihood strategy choice

Variable[a]	Coefficient		Std. error	Marginal effects
Staples producers vs. smallholders				
Household in Vihiga	-1.598	***	0.608	-0.288
Household size (numbers)	-0.169	*	0.104	-0.035
Farm size (acres)	0.414		0.272	0.050
Total livestock (TLU)	-0.078		0.161	-0.037
Age of household head (years)	-0.023		0.022	-0.002
Head's years of farming experience (years)	0.047	**	0.020	0.009
Self-reported access to credit	0.636		0.517	0.007
Receives remittances	0.731		0.476	0.012
Household head has no education	0.134		0.796	0.082
Household head has secondary education or above	0.437		0.535	0.018
Constant	2.075	*	1.164	
Off-farm high return and diversified commercial vs smallholders				
Household in Vihiga	-0.303		0.716	0.108
Household size (numbers)	0.011		0.131	0.017
Farm size (acres)	0.386		0.286	0.005
Total livestock (TLU)	0.229		0.197	0.033
Age of household head (years)	-0.035		0.029	-0.002
Head's years of farming experience (years)	0.004		0.028	-0.004
Self-reported access to credit	1.353	**	0.626	0.092
Receives remittances	1.349	**	0.586	0.089
Household head has no education	-0.768		1.329	-0.078
Household head has secondary education or above	0.712		0.618	0.041
Constant	-0.524		1.433	
Number of observations	203			
Percent correctly predicted	63.8			
Log-likelihood value	-148.58			
Pseudo R-squared	0.211			

*=significant at 10%, **=significant at 5%, ***=significant at 1%

[a] Other demographic characteristics, such as gender of the household head, were included in earlier versions of this analysis but were found not to be significant and were eliminated from these regressions to preserve degrees of freedom.

Credit access and receipt of remittances are the only statistically significant determinants of household self-selection into the combined off-farm high return/ diversified commercial strategy. Recall that households in these two clusters have more land invested in cash crops, own more improved dairy cattle and have more household members engaged in skilled off-farm income than the rest of the sample, on average. The ability to diversify into these higher return activities appears to be a function of their relatively greater financial liquidity, facilitated by access to credit and remittances, in comparison to households pursuing the lower return livelihood strategies, consistent with previous studies on the importance of financial liquidity to livelihood choice and household welfare (Dercon & Krishnan 1996; Ellis 1998; Mosley 2001; Barrett et al. 2005).

6. Conclusions

This paper introduces a novel approach to identifying distinct livelihood strategies in household survey data using cluster analysis. The resulting data-driven partitioning of the data enables us to test the hypothesis that some livelihood strategies demonstrably offer households higher returns on investment of their assets yet are unattainable for some households given their endowments, including their geographic location.

The results from a sample of 240 farming households in Kenya's central and western highlands yield an intuitive partitioning of the complex set of activities – we work with 11 distinct, aggregated activity categories – into five distinct livelihoods. Households with more land and more individuals working off-

farm in skilled employment are able to achieve higher average per capita incomes than their neighbors. The two most remunerative livelihood strategies, on average, first- or second-order stochastically dominate each of the other strategies' income distributions. Yet, despite the differences between the strategies in the sample, only the highest return livelihood generated an average income above the one US dollar-a-day benchmark and only the three highest earning strategies yielded mean per capita daily incomes in excess of the Kenyan rural poverty line of KSh43 per person per day. So variations in livelihood strategies exist, but the overall picture is still one of considerable and broad-based poverty in the rural Kenyan highlands. Further use of the methodology described here among other, more representative, population samples may suggest focal points for targeting interventions to help households adopt higher return livelihood strategies.

Acknowledgements

We acknowledge with gratitude support from the United States Agency for International Development (USAID), through grant LAG-A-00-96-90016-00 to the BASIS CRSP, and through the Strategies and Analyses for Growth and Access (SAGA) cooperative agreement, number HFM-A-00-01-00132-00, from the Rockefeller Foundation, and from the Coupled Natural and Human Systems Program of the Biocomplexity Initiative of the National Science Foundation, through grant BCS-0215890. We are grateful for very helpful comments received from seminar audiences at Cornell University and two anonymous reviewers. The views expressed here are those of the authors and do not represent any official agency. Any remaining errors are ours alone.

References

Barrett, CB, Reardon, T & Webb, P, 2001. Non-farm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics and policy implications. Food Policy 26 (4), 315–31.

Barrett, CB, Bezuneh, M, Clay, D & Reardon, T, 2005. Heterogeneous constraints, incentives and income diversification strategies in rural Africa. Quarterly Journal of International Agriculture 44 (1), 37–60.

Barrett, CB, Marenya, PP, McPeak, JG, Minten, B, Murithi, FM, Oluoch-Kosura, W, Place, F, Randrianarisoa, JC, Rasambainarivo, J & Wangila, J, 2006. Welfare dynamics in rural Kenya and Madagascar. Journal of Development Studies 42 (2), 248–77.

Bebbington, A, 1999. Capitals and capabilities: A framework for analyzing peasant viability, rural livelihoods and poverty. World Development 27 (12), 2021–44.

Carter, MR & Barrett, CB, 2006. The economics of poverty traps and persistent poverty: An asset based approach. Journal of Development Studies 42 (2), 178–99.

CBS (Central Bureau of Statistics), 2004. Geographic Dimensions of Well-being in Kenya. Government of Kenya, Nairobi.

Davidson, R & Duclos, J-Y, 1997. Statistical inference for the measurement of the incidence of taxes and transfers. Econometrica 65 (6), 1453–65.

Dercon, S & Krishnan, P, 1996. Income portfolios in rural Ethiopia and Tanzania: Choices and constraints. Journal of Development Studies 32 (6), 850–75.

DFID (Department for International Development), 2001. Sustainable Livelihoods Guidance Sheets. DFID, London.

Ellis, F, 1998. Household strategies and rural livelihood diversification. Journal of Development Studies 35 (1), 1–38.

Everitt, BS, Landau, S & Leese, M, 2001. Cluster Analysis. Fourth edition. Oxford University Press, New York.

IEA (Institute for Economic Affairs), 2002. The Little Fact Book: The Socio-Economic and Political Profiles of Kenya's Districts. IEA, Kenya.

Jansen, HGP, Damon, A, Pender, J, Wielemaker, W & Schipper, R, 2003. Policies for sustainable development in the hillsides of Honduras: A quantitative livelihoods approach. Discussion paper,

Environment and Production Technology Division, International Food Policy Research Institute (IFPRI), Washington, DC.

Mosley, P, 2001. Microfinance and poverty in Bolivia. Journal of Development Studies 37 (4), 101–33.

Place, F, Murithi, FM & Barrett, CB, 2005. A comparison and analysis of rural poverty between the western and central Kenyan highlands. BASIS (Broadening Access and Strengthening Input Market Systems) Policy Brief No. 4.

Reardon, T, Delgado, C & Matlon, P, 1992. Determinants and effects of income diversification amongst farm households in Burkina Faso. Journal of Development Studies 28 (2), 264–96.

Scoones, I, 1998. Sustainable rural livelihoods: A framework for analysis. University of Sussex. IDS (Institute of Development Studies) Working Paper No. 72.

Wangila, J, Nambiro, E, Ouma, J, Simiyu, D & Murithi, F, 2005. USAID BASIS CRSP project on rural markets, natural capital and dynamic poverty traps in East Africa: Embu site baseline study report. Mimeo.

Whitmore, GA & Findlay, MC, 1978. Stochastic Dominance: An approach to decision-making under risk. Lexington Books, Toronto.