TESTS OF THE ARBITRAGE PRICING THEORY USING MACROECONOMIC VARIABLES IN THE NIGERIAN STOCK MARKET

Ahmed E. UWUBANMWEN[1] and Joel OBAYAGBONA[2]

Abstract

The paper examines Arbitrage Pricing Theory (APT), using four macroeconomic variables; namely, exchange rate, an index of industrial production, nominal money supply and price of oil within the context of the Nigerian stock market. Two methods were used in the analysis along with descriptive statistics. First, factor analysis was used to identify underlying factors that contribute to stock pricing, secondly, the VECM technique was used to estimate the causal relationships between stock returns and the macroeconomic factors in Nigeria in line with the APT. Quarterly data covering the period 1985 to 2009 was employed in the empirical analysis, providing enough degrees of freedom for the estimation. The short run dynamic model revealed that money supply and oil prices are important factors in stimulating stock returns and that rising stock returns, on the other hand, tends to improve industrial production. The long run results show that sustained increases in both oil price and industrial production could cause stock returns to rise over time. However, money supply has a perverse negative effect on stock returns in the short run while exchange rate fails the significance test both in the short run and long run.

Key Words: Arbitrage Pricing, Macroeconomic variables, Stock Market
JEL Code: G1, G12

[1] Ph.D, Department of Banking and Finance, University of Benin, Benin City. Nigeria.
GSM: 08133370822
[2] Department of Banking and Finance, University of Benin, Benin City. Nigeria.
GSM: 08038874566
Email address: joere2010@yahoo.co.uk

1. Introduction

The theory of Arbitrage Pricing Model (APT) is a finance tool that has become very relevant in the pricing of stocks returns. This theory holds that the expected return of a financial asset can be modelled as a linear function of several macro-economic factors or theoretical market indices, where sensitivity to changes in each factor is represented by a factor-specific beta coefficient. The model-derived rate of return is used to price the asset correctly – as the asset price should equal the expected end of period price discounted at the rate implied by the model. If the price diverges, arbitrage would return it back to line.

The theory was first propounded by a renowned economist, Ross (1976), as a result of much criticisms occasioned by the inherent problems, shortcomings or weaknesses embedded in the Capital Asset Pricing Model (CAPM) on both theoretical and empirical grounds as evidenced by its unrealistic assumptions, difficulty of its empirical testing. Besides, a security's expected rate of return is a function of only one factor (the general stock market), whereas other multiple factors such as macroeconomic factors (relative sensitivity to inflation, Gross National Product (GNP), interest rates, oil prices, exchange rates, a diversified stock index, tax rate, etc. could also influence the security's returns relative to those of other securities.

Several tests of Arbitrage pricing theory have been empirically carried out by several scholars in the field of finance and economics in the more developed stock markets. For instance, Antoniou et al. (1998) applied it to the London stock market, Dhankar and Esq (2005) to Indian stock market, Berry *et al.* (1988) to S & P 500, and Chen *et al.* (1986) to New York stock market, Azeez and Yonezawa (2003) to Japanese stock exchange and Anatolyev (2005) to Russian stock exchange. However, much of this work has not been done within the context of the Nigerian capital market. In view of this therefore, the study seeks to examine empirically, the validity of the Arbitrage Pricing theory in the Nigerian stock market, using four macroeconomic variables, namely, the Exchange rate (EXRT), an index of industrial production (INDQ), the nominal money supply (MS) and the price of oil (POIL). More importantly, the study seeks to provide answers to the following research question: do macroeconomic variables such as Exchange rate (EXRT), an index of industrial production (INDQ), the nominal money supply (MS), and price of oil (POIL) affect securities returns relative to those of other securities in the Nigerian stock market? Answers to these research questions will be very useful in making an informed judgment about securities'

investment and their attendant returns, within the context of the Nigerian capital market.

The study is structured such that section one contains the introduction, section two deals with a brief review of APT and review of related literature, section three addresses the methodologies of the study, section four focuses on the empirical analysis and findings, and while the conclusion is in section five.

2. A Brief Review of Arbitrage Pricing Theory

In the literature, the APT states that there is a linear relationship between a security's return and some specified factors. That is to say, it relates the expected rate of return on a sequence of primitive securities to their factor sensitivities, suggesting that factor risk is of crucial importance in asset pricing (Gilles and Leroy, 1990). In equilibrium, according to the APT, the expected return on a security E(Ri) is given by:

$$E(R_i) = R_{FR} + \beta_{i1}(F_1 - R_{FR}) + \beta_{i2}(F_2 - R_{FR}) + \text{L} + \beta_{in}(F_n - R_{FR})$$

Where;

$E(R_i)$ = Expected rate of return on security i

B_{i1} = Sensitivity of security i to economic factor 1

β_{i2} = Sensitivity of security i to economic factor 2

F_1 = Expected value of factor 1

F_n = Expected value of factor n

β = Beta coefficient

R_{FR} = Risk free rate of return (Olowe, 2008: 233-234).

The above theory is however subject to the following assumptions: (i) Investors prefer more returns (ii) Investors are risk averse (iii) Investors have homogeneous risk expectations (iv) the Capital market does not have any transaction cost and there are no taxes.

Although the APT is concerned with a multi-factor model (that is, multiple Beta (β) model), it does not itself reveal the identity of its priced factors – the number and nature of these factors is likely to change over time and between economies, which essentially made it to be empirical in nature.

2.1 Review of Related Literature

Gibbons (1981) empirically examined if the number of factors affecting portfolio returns remains the same across three different portfolio groups. To determine the relevant number of factors required to describe the governance structure of 41 stock portfolios, and 9 bond portfolio, he utilized the appropriate likelihood ratio technique and concluded that, when one analysed stock and bond portfolios together, additional factors common to both groups had an influence on returns. These results however, were not evident when one group of portfolios was analyzed. Kryzanowski and To (1982) examined the number of factors that determine security returns and the sample size in terms of time periods and secondly, the security returns and the size of the groups being factored. He used Raw and alpha factor analysis for the first test to determine in each of the six time intervals the relevant number of factors that is related to the security returns. The results showed that, on average, the number of factors associated with returns remained the same across various samples of the same size and across different time intervals. But with the second text, a security group of size 50, four (overlapping) subgroups containing 10, 20, 30 and 40 respectively were randomly drawn with the use of Raw and alpha factor analysis. The result also showed that the number of relevant factors increased with the group size. However, further analysis of Gibbons and Kryzanowski and to revealed that their results would have been more statistically significant if they had utilized more groups of portfolio (securities).

Diacogiannis (1984) in his study of London stock exchange as it relates to APT and using time series data had similar results like that of Gibbons and Kryzanowski and To. However he concluded with a note of caution that though the findings indicate that the number of factors changes as the group size changes and that such results only highlight the fact that the methodology used for testing the APM is not the appropriate one, and previous tests of the APM are not necessarily tests of the model. That APM may be true, but the existing statistical methodology does not provide an unambiguous test of the model for the London stock exchange.

Recently, Cagnetti (2010) in his empirical study of the Italian stock market (ISM) from January 1990-June 2001 using factor analysis showed that over 40% of the sample size of 30 shares together with the Mibtel market index, are normally distributed, and the relationship between β and return in the Italian stock market over the above period was weak and the capital asset pricing model (CAPM) has poor overall explanatory power. The Arbitrage pricing theory (APT) on the other hand, which allows multiple sources of systematic risks to be taken into account, performs better than the CAPM, in all the tests considered. Securities and portfolios in the Italian stock market seem to be significantly influenced by a number of systematic forces and their behaviour can be explained only through the combined explanatory power of several factors or macro economic variables. Connor and Korajezyk (1992) citing Roll and Ross (1980) affirm that when they estimated factor risk Premia and test the APT restrictions with a sample of daily returns on 1260 firms over the period from July 1962 to December, 1972, and dividing the cross sectional sample into 42 groups of thirty firms each and performing an analysis on each group with a cross sectional regression of asset excess returns on β, the results show that as many as four factors have significant risk Premiums.

Furthermore, Roll and Ross (1980) also test the APT by including the same sample standard deviation of the asset an instrument in cross –sectional regression, but the estimate of the standard deviation wasn't predetermined. The results however show three of the forty-two (42) groups of δ (standard deviation).

Lehmann and Modest (1988) performed time series based tests of APT restriction $<=$ 0. They divide the period from 1963 to 1982 into four five year sub periods. Firms traded on the NYSE and AMEX that do not have missing daily data over a sub period comprise the sample. Several other adjustments and application were made, and in the final analysis, they concluded that while the APT is rejected on the basis of the regressions with size based portfolios, its apparent ability to explain the dividend yield and variance effects that are unexplained by the CAPM (with standard proxies for the market portfolio) make it good alternative model of asset pricing.

Huberman and Kandel (1987), Fama and French (1993) find that the Multifactor model do a much better job in explaining asset returns (i.e, values of α close to zero) than do standing single index models. MCElroy and Burmeister (1988) postulated macroeconomic variables as observable factors and use non linear time series regression to estimate the parameters of the factor model-using monthly returns on 70 individuals stocks from January 1972 to December 1982 as the set of test assets and five

pre specified factors that are similar to the factors used by Chen Roll, and Ross (1986). They concluded that the multi-factor model is a useful empirical framework or linking macroeconomic innovations to expected asset returns, (Connor and Korajezyk (1992).

Burmeister and Wall (1986), Berry (1988), Connor and Uhlaner (1988), Ferson and Harvey (1991), Wei, Lee, and Chen (1991), and Cragg and Donald (1992), all had similar findings like those of Chan, Chen, and Hsieh, (1985).

3. Methodology

Two methods are used in the analysis of this study. Factor analysis was used in an attempt to identify underlying factors or variables. This method has been used in many studies on the APT such as Cagnetti (2010) and Paavola (2006). Generally, factor analysis is used in data reduction to identify a small number of factors that explain most of the variance observed in a much larger set of variables. This method is used to screen variables for subsequent analysis.

Four variables were selected for the factor analysis based on these factors: the exchange rate (EXRT), an index of industrial production (INDP), the nominal money supply (MS) and price of oil (POIL). The main factors that help to explain variables in stock returns (R) in the Nigerian stock market are highlighted using this methodology.

Subsequent to the factor analysis, an empirical model is specified and estimated in order to empirically determine the main factors that determine share prices within the context of APT. It is hypothesized that the stock returns depends on the variables selected as macroeconomic factors. The model is specified as:

$$R = (MS, INDQ, POIL, EXRT) \qquad (1)$$

Where:

R = all share index (used to proxy share prices)
MS = money supply
INDQ = index of industrial output
EXRT = exchange rate of the naira
POIL = price of oil

Stock return is computed as:

$$R = \left(\frac{ASI - ASI_{t-1}}{ASI_{t-1}}\right) X\ 100$$

The macroeconomic outlook of the variables used in the analysis may imply simultaneity among them. Hence, the Granger causality test, which is a preliminary aspect of an autoregressive-based analysis, is used to provide the background for estimating dynamic relationships. Based on the Granger Causality the Vector Error Correction Modelling (VECM) technique is employed in the estimation of the relationships.

The study estimates the following VECM to determine the long and short run dynamics between stock prices and the macroeconomic variables:

$$Y_{it} = \beta_0 + \sum_{k=1}^{m} \beta_k Y_{it-k} + \sum_{l=1}^{n} \alpha_j X_{it-j} + ECM_{it-1} + u_t \quad (2)$$

$$X_{it} = \gamma_0 + \sum_{k=1}^{m} \delta_k X_{it-k} + \sum_{l=1}^{n} \varphi_j Y_{it-j} + ECM_{it-1} + \vartheta_t \quad (3)$$

where:

Y_i = stock price returns

X_i = macroeconomic factors

Thus five equations are to be estimated based on the formulations in equations 2 and 3.

u_t and ϑ_t = mutually uncorrelated error terms (i.e. zero mean white noise error terms)

k' and 'j' = the number of lags

ECM = error correction term that is included in the short run VECM.

α_j, β_k, δ_j, and φ_k are all parameters to be estimated.

The real exchange rate is expected to have a negative input on the price of equities since it is a cost of funds. Thus, a negative effect is expected between exchange rate and stock returns. Funds become scare as the exchange rate depreciation, this will lead to rise in domestic prices including share prices. All the other variables positively impact on share prices since positive movements in them tend to promote economic activities

and drive up investors' participation in the stock market. Quarterly time series data covering the period 1985 to 2009 are used for the analysis.

4. Empirical Analysis

4.1 Descriptive Statistics

Table1 presents the descriptive statistics for quarterly data consisting the period 1985 – 2009. The table shows that all the macroeconomic variables have positive mean return values. The average all share index value is N11711.9 for the period which is relatively high. The standard deviation for all share index is quite close to the mean value, thereby indicating that the price index of share prices have been relatively steady over the years. Stock returns has a mean value of 5.68 and a high standard deviation of 8.46, suggesting that returns have been very unstable in the market over time. The standard deviation, of INDQ and RINT are dispersed from their respective mean values and this implies that these are most volatile variables in terms of annual movements over time. Surprisingly POIL variable has a relatively steady quarterly covariation as can be seen from the standard deviation value of 23.55 which is close to the mean value. Relative to the mean, money supply has the highest standard deviation followed by stock returns. Apparently, these values are the most variable in the series. Money supply (along with oil price) also has a large kurtosis value, suggesting that the data may not be normally distributed.

Table 1: Descriptive statistics for variables

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
ASI	11711.93	5672.700	57990.20	127.3000	14817.75	1.541288	4.874693
R	5.68	6.10	25.88	-21.30	8.46	-0.46	4.07
MS	1655018	425473.7	10730793	21882.2	2581280	2.037814	6.237462
INDP	16486.66	14601.69	29990.92	11568.64	4667.729	1.46	4.06
EXRT	61.82	21.89	150.92	0.89	55.60	0.28	1.24
POIL	31.63	21.15	127.35	11.26	23.55	2.01	6.91

Source: Author's computations

4.2 Factor Analysis

We run the factor analysis to find out how well we have interpreted the theory of the APT in selection of our variables and especially the number of factors. The results from factor analysis can be seen in Table 2 below. Originally, this factor analysis gives guidance to how many factors should be used. As it turns out, the result shows that two of the variables account for over 74 percent of the variance. This is the output of the SPSS calculations. However, this number is much smaller than the number of variables used in the tests of the APT in the study.

Table 2: The Results from Factor Analysis

Component	Total	% of Variables	Cumulative
1	2.831	47.18	47.08
2	1.662	27.70	74.89
3	0.906	15.10	
4	0.338	5.63	
5	0.185	3.04	
6	0.008	1.35	

Source: Author's computations

The number of variables used in the APT was extended to four based on the high level of significance of Bartlett's test of sphericity as shown in the table below. This test detects the partial correlations of the examined variables. The chi-square value of the Bartlett's sphericity test is 81.51 which is significant at the 1 percent level. The conclusion from the result is therefore indicative of the fact that more than just two of the variables are principal components.

Table 3: Bartlett's Test Result

Approx. chi square	81.51
Degree of freedom	15
Significance	0.0000

Source: Author's computations

4.3 Empirical Analysis

According to Gordon (1995), most economic time series are non-stationary and only achieved stationary at the first difference level or at a higher level. The Augmented dickey Fuller (ADF) test is employed in order to analyze unit roots. The results are presented in levels and first difference. This enables us determine in, comparative terms, the unit root among the time series and also to obtain more robust results. Table 4 presents results of ADF test in levels without taking into consideration the trend in variables. The reason for this is that an explicit test of the trending pattern of the time series has not been carried out. Given the critical ADF value of -2.891, none of the series is stationary in levels but each of them attains stationarity after first differencing. This shows that the variables are all integrated of order one and possess unit roots.

Table 4: Unit Root Test for Variables

Series	Level	Difference	Integration
R	-1.679	-5.747**	I[1]
MS	0.565	-9.851**	I[1]
INDP	-2.242	-3.502**	I[1]
EXRT	-2.521	-8.612**	I[1]
POIL	-0.637	-8.499**	I[1]

** Indicates the rejection of the null hypothesis of a unit root at the 5% significance level. The critical value at the 5% significance level is -2.891, with constant and trend.

Having established that the series in the analysis are not stationary in their levels, we move on to determine if they are cointegrated. The results from the multivariate cointegration test are presented in Table 5 below. As can be seen from Table 4, both the λ-max and the trace test statistics indicate that there are two significant cointegrating vectors between stock returns, money supply, industrial production, exchange rate and price of oil. This implies that a long run relationship exists among these variables.

Table 5: Johansen Multivariate Cointegration Tests Results.

Trace Test			*Maximum Eigenvalue Test*		
Null Hypothesis	Test Statistic	Critical Value	Null Hypothesis	Test Statistic	Critical Value
r = 0**	102.02	69.81	r = 0**	44.49	33.88
r ≤ 1 **	57.52	47.85	r ≤ 1*	29.47	27.58
r ≤ 2	28.05	29.79	r ≤ 2	16.25	21.13
r ≤ 3	11.81	15.49	r ≤ 3	10.25	14.26
r ≤ 4	1.554	3.841	r ≤ 4	1.554	3.84

**(**) denotes rejection of the hypothesis at 5% (1%) significance level.*

The macroeconomic outlook of the variables used in the analysis may imply simultaneity among them. Hence, the Granger causality test, which is a preliminary aspect of an autoregressive-based analysis, is used to provide the background for estimating dynamic relationships. The results of the Granger causality tests are reported in Table 6 below. As is generally the case, the F-test is conducted on the null hypotheses in order to determine the direction of causality between each pair of variables. The rejection of each of the null hypothesis is based on the significance of the F-value for the particular relationship.

The test result shows clearly that there is a feedback relationship between stock returns and money supply, suggesting that while money supply Granger causes stock returns, money supply also responds to movements in stock returns over time. Unidirectional relationships exist between stock returns and the other variables in the analysis. It is also seen that causality both industrial production and oil prices Granger cause stock returns without a reverse relationship. However, the pattern of Causality actually runs from stock returns to exchange rate and not the other way round. These directions of causality indicate that simultaneity issues a germane among these variables. We thus adopt an appropriate estimation technique in the study to investigate the empirical relationships.

Table 6: Granger Causality Test results

Null Hypothesis:	F-statistic	Probability	Decision	Causality
MS does not Granger Cause R	3.553	0.033	Reject	Feedback
R does not Granger Cause MS	3.099	0.049	Reject	
INDP does not Granger Cause R	2.601	0.079	Reject	Unidirectional
R does not Granger Cause INDP	0.689	0.505	Accept	
EXRT does not Granger Cause R	1.384	0.256	Accept	Uni-directional
R does not Granger Cause EXRT	3.093	0.050	Reject	
POIL does not Granger Cause R	3.143	0.048	Reject	Uni-directional
R does not Granger Cause POIL	0.384	0.683	Accept	

The existence of cointegration among the variables allows us to implement the Vector Error Correction Modelling (VECM) technique, which describes the systematic disequilibrium adjustment process and the short-run transmission mechanism. The endogenous variables in the system include lagged variables of the GDP growth, stock prices, and interest rate and the error correction term from the cointegrating equation. The use of lags is expected to internalize the implications of expectations among the variables. The result of the VECM is presented in Table 7 below.

Several interesting transmission patterns emerge from the examination of Table 7. We observe that the estimated lagged error-correction term (ECM_{t-1}) emerges as an important channel of influence. The statistically significant error-correction term (apart from that of the exchange rate equation), confirms the existence of long run relationships between stock returns and all the macroeconomic variables. In other words, the series quickly adjusts to eliminate any deviations from the long-run equilibrium relationships that they may share with each other.

Table 7: Dynamics between Stock Returns and Macroeconomic Variables

	ECM_{t1}	*ΔR*	*ΔMS*	*ΔINDP*	*ΔEXRT*	*ΔPOIL*
Panel A: Stock Returns (ΔR) Equation in the VECM	-0.06**	1.01**	2.95E-06**	0.003	-0.003	0.063*
Panel B: Money supply (ΔMS) Equation in the VECM	7613.63**	-4889.11	-0.451**	-1.982	626.1	6901.74*
Panel : Industrial Production (ΔINDP) Equation in the VECM	2.67**	13.39**	-7.23E-05	1.084**	0.402	-0.61
Panel D: Exchange Rate (ΔEXRT) Equation in the VECM	-0.08	-0.06	3.81E-06	0.0001	0.045	-0.23*
Panel E: Oil Price (ΔPOIL) Equation in the VECM	-0.25**	-0.36	1.79E-05**	0.008	-0.065	0.47**

As can be seen in Panel A of Table 6, money supply and oil prices have positive influence on stock returns through the significant error correction term. Neither industrial production nor exchange rate seems to exert any significant influence on stock returns in the market. This result suggests that in the short run, money supply and oil prices are the strongest factors that push stock market activities. As domestic money supply rises, aggregate spending also rise which includes participation in the stock market. Moreover, Nigeria being highly dependent on oil revenue is like to experience improvement in economic activities when oil prices rise. There are quite likely spillover effects from the increased oil revenue into the capital market. Thus, in the short run, money supply, oil prices and stock market activities are linked.

Oil prices are still significant in the Panel B of the result table. This indicates that oil prices tend stimulate money supply in the short run, implying that oil prices may actually exert an indirect effect on stock returns by changing the movements in money supply. In the result in Panel C, stock return is highly significant in the industrial production equation. Returns in the stock market positively stimulate industrial

production; a booming stock market tends to cause improvement in the industrial sector in the short run.

For the exchange rate Panel, neither the lagged coefficients (apart from oil price), nor the error correction term is statistically significant, indicating that disequilibrium in exchange rate may not be effectively restored by short term changes in the other variables. Finally, stock return is not a significant factor in oil price determination. However, the significant ECM term indicates that long run equilibrium in oil price is ensured.

The long run position of the estimated relationship with respect to the effects of the macroeconomic variables on stock returns is reported in Table 8 below. The long run result is taken as the cointegrating equation from the VECM. In the result, MS, INDP and POIL are all significant. However, INDP has a negative coefficient, suggesting that industrial production seems to reduce stock returns in the long run. This may reflect the fact that new or emerging firms with prospects may be more active in the market than older companies that have been in the market over time. Both money supply and oil prices have positive coefficients indicating that they both promote stock returns in the long run. Exchange rate once again fails the significance test, showing that it is an unimportant factor in stock returns both in the short run and in the long run. The exchange rate may not be a factor of interest to policy makers when devising means of impacting stock market activities.

Table 8: Long Run VECM Result

Variable	Coefficient
MS(-1)	8.38E-06**
INDP(-1)	-0.006**
EXRT(-1)	-0.022
POIL(-1)	1.559**
C	33.22076

5. CONCLUSIONS

Stock markets in African countries are usually characterized by inefficiencies and are hence unsuitable for the application of certain financial theories (Appiah-Kusi and Menyah, 2003; Mlambo and Biekpe, 2007). This study has set out to examine the relevance and the efficacy of the arbitrage pricing theory within the context of the

Nigerian stock market, using four selected macroeconomic variables as determinants of equity market returns and their impact on assets pricing in Nigeria. In order to present a more holistic dynamic relationship, the VECM was adopted in the analysis. Oil prices showed up as the most critical factor, both in the short run and long run, in affecting stock returns. The use of oil money in Nigeria must therefore be well managed so as to improve the stock market rather than allowing the vagaries in oil revenue to strangulate the market. In the overall evaluation, these results give us guidance on how we may determine the equity prices in the Nigerian stock exchange by examining some prevalent macroeconomic conditions. Moreover, it has justified the APT with respect to the Nigerian case by showing that expectations in macroeconomic variables may act well in determining the behaviour of stock returns.

References

Anatolyev, S. (2005). A ten-year retrospection of the behaviour of Russian stock returns, *BOFIT* Discussion paper (9) Helsinki: Bank of Finland, Institute for Economies in Transition.

Antoniou, A., Garrette, I. & Priestley, R. (1998). Macroeconomic variables as common pervasive risk factors and the empirical content of the Arbitrage pricing Theory: *Journal of Empirical Finance,* 5,(3), 221-240.

Appiah-Kusi, J. & Menyah, K. (2003). Return predictability in African stock markets. *Review of Financial Economics*, 12: 247-270

Azeez, A. A. & Yonezawa, Y. (2003). Macroeconomic factors and the Empirical content of the Arbitrage pricing Theory in the Japanese stock market. *Japan and the world Economy, 1-24.*

Berry, M. A., Burmeister, E. & McElroy, M. (1988). Sorting out risks using known APT factors. *Financial Analyst Journal,* 44, (2), 29-41.

Burmeister, E. & Wall, K. D. (1986). The arbitrage pricing theory and macroeconomic factor measures, *The Financial Review,* 21:1-20.

Cagnetti A. (2010). CAPM and APT in the Italian stock market: an empirical study. Unpublished article.

Chan, K. C., Chen, N. & Hsieh, D. (1985). An exploratory investigation of the firm size Effect, *Journal of Financial Economics,* 14, 451-471.

Chen, F. (1981). Arbitrage Asset Pricing: Theory and Evidence. Ph.D. Dissertation, University of California, Los Angeles

Chen, N. (1983). Some empirical tests of the theory of arbitrage pricing. *The Journal of Finance,* 38, (5), 1393-1414.

Chen, N. F, & Ingersoll, E. (1983). Exact pricing in linear factor models with finitely many assets: A note, *Journal of Finance.*

Connor, G. & Korajezyk, K. (1986). Performance measurement with the Arbitrage Pricing Theory: A new framework for analysis. *Journal of Financial Economics*, 15,(3), 373-394.

Connor, G. & Korajezyk, R. (1992). The attributes behaviour and performance of U.5. mutual funds. *Review of Quantitative Finance and Accounting* 1: 5-26.

Connor, G. & Uhlaner, R. (1988). New cross-sectional regression tests of beta pricing models, *Working paper*, (School of Business Administration, University of California, Berkeley, CA).

Cragg, J. G. & Stephen, G. D. (1992). Testing and determining the arbitrage pricing structure from regression on macro-variables, Working paper, (University of British Columbia, Vancouver, BC)

Dhankar, S. & Esq, R. S. (2005). Arbitrage pricing theory and the capital assets pricing model- evidence from the Indian stock market, *Journal of Financial Management and Analysis*, 18, (1), 14-28.

Diacogiannis, G. (1984). "Some Empirical Evidence on the Intertemporal Stationarity of the Security Returns Distributions for the London Stock Exchange". Forthcoming in the *Journal of Accounting and Business Research.*

Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Ferson, W. E. & Harvey, C. R. (1991). The variation of economic risk premiums, *Journal of Political Economy,* 99, 385-415.

Gehr, A. (1978). "Some Tests of the Arbitrage Pricing Theory". *Journal of the Midwest Finance Association,* 91 -105.

Gibbons, M. R. (1981). "Empirical Examination of the return Generating Process of the Arbitrage Pricing Theory", Working Paper, University of Chicago.

Gilles, R. & Leroy, S. F. (1990). On the Arbitrage pricing Theory, *Economic Theory*, 1,(3), 213-229.

http://dictionary.bnet.com/definition/arbitrage+pricing+theory.html

http://www.faqs.org/abstracts/Business/An-asset-pricing-theory-unifying-the-CAPM-and-APT-Temporal-aggregate-and-the-continuous-time-capital.html

http://www.faqs.org/abstracts/business/New-tests-of-the-APT-arbitrage-pricing-theory-and-their-implications.html

http://www.mywikibiz.com/Arbitrage_pricing_theory

http://www-personal.umich.edu/~kathrynd/JEP.FamaandFrench.pdf

Huberman, G. and Kandel, S. (1987). Mimicking portfolios and exact arbitrage pricing: The 42, (1), 1-9.

Hughes, P. (1982). "A Test of the Arbitrage Pricing Theory". Working Paper, University of British Columbia.

Kryzanowski, L. & To, M. C. (1983). "General Factor Models and the Structure of Security Returns". *Journal of Finance and Quantitative Analysis,* 31 - 32.

Lehmann, B. & Modest, D. (1988). The empirical foundations of the Arbitrage pricing Theory 1: The empirical tests, *Journal of Financial Economics,* 21, (2), 213-254.

McElroy, M. B. & Burmeister, E. (1988). Arbitrage pricing theory as restricted nonlinear multivariate regression model, *Journal of Business & Economics Statistics,* 6, 29-42.

Mlambo, C. and Biekpe, N. (2007). "The efficient market hypothesis: Evidence from ten African stock markets" *Investment Analysts Journal* – No. 66.

Olowe, R. A. (2008). "Financial Management, Concepts, Financial System and Business Finance," Brierly Jones Nigeria Limited, Lagos, Nigeria.

Paavola, M. (2006). Test of arbitrage pricing theory using macroeconomic variables in The Russian Equity Market. Bachelor's Thesis.

Reinganum, M. R. (1981). "The Arbitrage Pricing Theory: Some Empirical Results". *Journal of Finance,* 313 - 321.

Roll, R. & S. A. Ross. (1980). An Empirical Investigation of the Arbitrage Pricing Theory, *Journal of Finance:* 35, (5), 1073-1103.

Roll, R. (1977). "A Critique of the Asset Pricing Theory's Tests: Part I.: On the Past and Potential Testability of the Theory". *Journal of Financial Economics,* 129-176.

Ross, S. A. (1976). The arbitrage theory of asset pricing. *Journal of Economic Theory,* 13, 1, 341–360.

Wei, K. C. J., Lee, C. & Chen, A. H. (1991). Multivariate regression tests of the arbitrage pricing theory: The instrumental variables approach, *Review of Quantitative Finance and Accounting* 1, 191-208.

LAND REGISTRATION AND LAND INVESTMENT: THE CASE OF TIGRAY REGION, NORTHERN ETHIOPIA

Fitsum Hagos[1]

Abstract

The study explored whether land titling fostered tenure security and, thereby, increased investment on land improvement. We assessed the determinants of the probability and intensity of investment by using random effects and modified random effects probit and truncated regression models on 437 randomly selected households operating 1696 plots from eighteen selected communities (tabias) located in the five zones of the Tigray Region. Findings indicated that registration enhanced holders' feeling of tenure security, there was significant increase in probability and composition of investments, and increased private initiatives. The likelihood and the intensity of conservation were low on land lost in the last redistribution or taken away by the public for different purposes. Length of tenure, initial investment, and access to food-for-work positively influenced the likelihood and intensity of conservation. Households with more livestock, land holding and adult male labor (although significant only in the random effects probit and at 10 percent level) were found to be more likely to make investments on land. Moreover, the intensity of investment was significantly influenced by the year of registration. Finally, households operating rented-in land were found to be less likely to and invested less indicating that tenants commit fewer resources to long-term investments because they strive to maximize immediate benefits. There were various time invariant household and plot level characteristics that influenced the probability and intensity of conservation. This calls for policy makers to minimize the potential sources of insecurity such as threats of future land redistribution and land taking without proper land compensation. Moreover, land registration/certification is vital for creating tenure security; this has to out scaled throughout the country.

Key words: titling, tenure security, conservation investment, random effects model; Ethiopia, Africa.

JEL Classification: C23, C.24, D13, D23.

---

[1] Ph.D; Researcher in International Water Management Institute
E-mail: F.Hagos@CGIAR.ORG

1. Introduction

Land is one of the most critical assets to the livelihood of farming households in developing countries. In Ethiopia, until recently, rural households' access to land was met through regular government sponsored land redistribution and informal land transactions. Formal land sales have been prohibited during the last three decades or more. This is because land is declared the property of the state; it may not be sold or mortgaged. Peasants and pastoralists have only user rights (Rahmato, 1992; FDRE, 1995; Joireman, 2001).

The land issue has a strong bearing on a wide range of issues and policy concerns, including agricultural development, food security, natural resource management, poverty reduction and even human rights (Feder and Nishio, 1999; Rahmato, 2003; Deininger and Chamorro, 2004). In Ethiopia, while there is still policy debate on the choice of ownership type, there is a move in the policy realm to give rural land users a title of 'ownership', in short title, by issuing user certificates, to the land they received during the last land redistribution or through inheritance from close kins. One such case is the process that unfolded in Tigray Regional National State, northern Ethiopia after 1997. The regional Government of the Tigray undertook land inventory and registered all rural lands before issuing title certificates. This process continued later on in areas where land registration took place (e.g. in the recent resettlement areas) and to other people who obtained land recently through partial community redistribution processes.

Titling through land registration is widely believed to improve efficiency of land use and agricultural production by increasing farmers' incentives to adopt new technology, on-farm investment and soil conservation practices (Feder and Nishio, 1999; Rahmato, 2004). The government of Ethiopia (GoE) aims to boost farmers' sense of security, which, in turn, is expected to encourage investment on erosion-reducing and land quality enhancing technologies (FDRE, 2005a).It is also possible that the GoE's policy measures might have triggered other outcomes, intended or otherwise. The expected changes associated with land registration could be the development of land and credit markets. This case provides, hence, an important study for drawing important policy lessons on the impacts of land registration/certification on on-farm investment.

There is an on-going discussion about the differentiated effects of land registration/certification may have on tenure security and labor-intensive land

conservation investment. The evidence so far is mixed. There are evidences from Asia and elsewhere that indicates that such government sponsored titling enhances tenure security (Feder et al., 1988; Binswanger, 1996; Firmin-Sellers and Sellers, 1999). On the other hand, there are evidences, especially from Africa, that call against land registration and tilting as the cost of land registration is quite high, and the effect of land registration is contrary to expectations (Atwood, 1990; Place and Mighot-Adholla, 1998; Jacoby and Minten, 2007).

The study by Holden et al., (2007) assessed the impact of registration/titling on the functioning of land markets and their results revealed that the land reform contributed to increased land rental market participation. However, there exist no previous studies that have attempted to quantify the impacts of registration on tenure security and land investment or crop productivity.

This study, which is the first of its kind in the region explores whether exogenous registration and certification improved land users' feeling of security and whether land investment increased after certification. The study utilizes data from 437 randomly selected households operating 1696 plots collected from eighteen selected communities (*tabias*) located in the five zones of the Tigray Regional National State in 2004/05. In the survey, we tried to capture investments made before and after registration to assess whether there was significant increase in long-term land investment after the registration. Once this was established, we explained the probability and intensity of investment after registration by controlling for covariates using probit and truncated regression random effects models. The dependent variable of intensity of investment is measured by the length (in meters) of conservation investments, which may lead to measurement error[2], especially in the light of changes in type of conservation measures. We used a modified random effects model framework proposed by Mundlak (1978), whereby we included on the right-hand side of each equation the mean value of plot varying explanatory variables. The effect of certification is estimated by the inclusion of the year of registration/certification in the model and relies on sufficient variation in this variable (Deininger and Chamorro, 2004).

The paper is outlined as follows: in part 2 we present a short description of the process of land registration and certification followed by presentation of a simple conceptual

[2]A measurement error on the dependent variable does not destroy the unbiasedness property of the estimators but the standard errors (variances) of the estimated coefficients could be higher (Verbeek, 2000).

framework to understand the effect of land registration on tenure security and land investment. In section three we present the empirical model and the methods of estimation. Sections 4 and 5 present the study sites and sampling strategy employed and the empirical findings respectively. The final part concludes by summarizing the findings and drawing relevant policy conclusions.

2. Background and Process of Land Certification in Tigray

The land tenure system in Ethiopia has been substantially affected by past and current government policies (Rahmato, 1992; Joireman, 2001; EEA/EEPRI, 2002; Teklu, 2003). Land is declared the property of the state; hence, it may not be sold or mortgaged. Peasants and pastoralists have only user rights. Holding rights are defined in the Federal Constitution of Ethiopia as "the right any peasant shall have to use rura11and for agricultura1 purposes as well as to lease and, while the right remains in effect, bequeath it to his family member; and includes the right to acquire property thereon, by his labor or capita1, and to sell, exchange and bequeath [the] same" (FDRE, 1995 Art 2 Sub Art. 3). Art. 51 of the constitution states that the Federal Government shall enact laws for the utilization and conservation of land and other natura1 resources (FDRE, 1995). Art. 52 also states, that Regional Governments have the duty to administer land and other natura1 resources according to Federa1 laws and on behalf of the same (FDRE, 1995). A law was enacted in July 1997 on "Rura1 Land Administration Proclamation, No. 89/1997" that vested Regiona1 Governments with the power of land administration (defined as "the assignment of holding rights and the execution of distribution of holdings"[3] (FDRE, 1997 Art. 2.6). The Federal land policy states that farmers have a perpetual use right on their agricultural holdings, and that this right will be strengthened by issuing certificates and keeping registers (FDRE, 1997a; Deininger *et al.*, 2006).

A new land policy enacted in the Tigray Regional State in 1997[4], and subsequently revised three times, prohibited further redistribution, except in few cases, as indicated below. By prohibiting further land redistribution, the policy hopes to end further land fragmentation. The 1997 and subsequent laws also formalized land-lease practices

[3] The latest Federal legislation on land administration and use, proclamation No. 456/2005, also calls for regional states to come up with proclamations and regulations to implement the federal proclamation.

[4] This first law follows the spirit of the 1997 Federal Rural Land Administration Proclamation while the latest one Proclamation No. 136, 2007 follows the spirit of Federal legislations No 455/2005 and 456/2005.

among farmers and between farmers and investors with contracts up to three years if the leaser uses traditional technology and up to 20 years if he/she uses modern technology. However, the definition of use of "modern technology" was left to be decided by the regional guideline. The policies also triggered the process of registration and certification of holdings of users. In line with the provisions of the constitution, the land polices granted holders use rights, rights to bequeath, and rights to rent land. In the spirit of this legal provision, the Tigray Regional State undertook land inventory surveys and issued use-right certificates to current users (Hagos *et al.*, 1999) starting 1998. By doing so the regional government hoped to boost farmers' sense of ownership or land security and encourage investment on erosion reducing and landing quality enhancing technologies without the state losing its right of ownership to land. Lately, different regional governments in Ethiopia, namely, Amhara and Oromia, have also initiated a process of land registration and certification (Deininger *et al.*, 2006: p. 5).

The latest land policy also outlines obligations of the land users: not to cut trees on farm, protect plot boundaries, undertake soil and water conservation measures and plant trees, among others. Failing to do so could lead upto withdrawal of the household's holding rights (FDRE, 2005a; RNST, 2007: p. 10), which is contrary to tenure security per se. The policy also provides guarantees to land users against expropriation without proper compensation (FDRE, 2005b) and is expected to increase farmers' incentives to make long-term land investments.

The aim of this study is to assess the impact of land registration and certification undertaken in Tigray Regional State after 1997.

The process of land registration and certification in Tigray Regional State, the problems faced, people's attitude on the registration and institutions involved is well documented by Haile *et al.*, (2005) and Deininger *et al.*, (2007). We present a brief description of the last land redistribution, registration and certification in Tigray Regional State so that we understand the context under which this took place. Most of the cultivated land had been redistributed for the last time in Tigray Regional State between 1989 and 1991 although exact dates vary a bit from place to place. By the time of the last distribution, most people, even many of those who had lost land, said they thought a fair distribution had been achieved (Haile *et al.*, 2005). Land distributions have (with a few exceptions) only occurred – to date – when 1) a new micro-dam created newly irrigated land and the existing land holdings were reduced and distributed; 2) when people abandoned

land (e.g., left their village for more than 2 years) or died without "legal heirs"[5] and the land returned to the *tabiabaito* (local village government) for distribution; or 3) when some government improvement or infrastructure development (e.g., roads) cause land to be taken away rendering compensation necessary. Otherwise, the vast majority of land holdings have not changed hands since about 1991 in Tigray Regional. Nonetheless, farm households still expect future land redistribution even if it is ruled out in the current land policy.

It is roughly twenty years since the last redistribution, most of the children under eighteen at that time have come of age but have not received land. This probably accounts for something like half to two-thirds of adults today. Some have received land through inheritance but due to high population growth rate and prohibitions against further fragmentation of land, most have not obtained land of their own. The problem of landless youth is today one of the growing problems in Ethiopia (EEA, 2007) in general and Tigray Regional National State in particular and is becoming a cause of growing tenure insecurity.

Registration and certification of most cultivated land was completed in Tigray Regional State between 1996 and 1998 and appears not to have changed the size of holdings. The registration of cultivated land had been preceded for seven or eight years by fairly strong and clear policy statements that there would not be any further land distributions in the foreseeable future. So, existing landholdings were simply registered. For each household the plots of cultivated land held by the household (whether brought to it through marriage or given in the last distribution or a more recent distribution) were recorded in *Tigrigna*, the local language, by hand in ink on a pre-printed page in a record book at the *tabia* office, with each page listing each parcel of land held by the household. The approximate size of each plot (in *tsimdi*, a local unit of measure that is the amount of land that can be cultivated in one day using oxen plough and averages about 0.25 hectares), the type of land of each plot (poor, medium, and fertile) is registered. The record book describes each plot only by the local area name or a geographical marker and the names of the neighbouring landholders on the north, south, east and west sides are registered. In addition, the family size is registered. The certificate was typically issued in the name of the household head, who in most cases is male. The certificate is nearly identical in form and content to the registration book

[5] The Tigray land law excludes some heirs allowed by the Civil Code, e.g., those with another source of income.

page and is also written in ink on a pre-printed form in the local language; that is Tigrigna in all of our sites.

In most cases "technicians" worked together with the local Agricultural Development Agent and community members (usually men who had been involved in the last land distribution), performed a study of land ownership and recorded the land details on a pre-printed form. In almost all cases these findings of the study were then reviewed publicly with the landholders. In some cases landholders were involved in the study of their land.

This procedure is commonly described as low cost registration.

3. Conceptual Framework

The notion that the greater tenure security accorded by possession of registered land title will be associated with higher levels of investment is a key element in the literature (Feder *et al.*, 1986; Feder *et al.*, 1988) and the relationship between possession of title and higher levels of land-attached investments has been confirmed (Binswanger, 1996; Deininger and Chamorro, 2004) although there are also evidences of weak or non-existent land-attached investments in spite of land titling (Atwood, 1990; Mighot-Adholla *et al.*, 1991; Carter *et al.*, 1991; Roth *et al.*, 1994; and Place and Mighot-Adholla,1998).

From the literature (Feder *et al.*, 1986; Feder *et al.*, 1988; Besley, 1995; Feder and Nishio, 1999; Deininger and Chamorro, 2004) secure property rights in land are generally considered to be a precondition for economic growth and development for three reasons: (i) land titles have positive effect on land tenure security and provide investment incentives for owners to undertake land-related investments; (ii) land titling reduces the transaction costs in land markets thus helping decrease cost and increase allocative efficiency; and (iii) formal land titles improve access to institutional credit by creating collateral value for land. We briefly outline each of these linkages and put it together into a schematized graph in Figure 1 below.

Land tenure security, which accrues from land registration/certification, removes uncertainties on whether or not land owners can reap the benefits from any long-term investment they make such as on-farm soil and water conservation, water harvesting structures and trees. With positive expectations about the exclusive enjoyment of any

returns earned from investment, landowners develop interest in investing in land improvements as well as making land-based investments in agriculture. This boosts demand for investment which in turn increases demand for complementary inputs including labour and agricultural inputs (including capital). There are empirical evidences in support of the positive impact of land registration on investment (see Feder et al., 1986; Feder *et al.*, 1988; Binswanger, 1996; Feder and Nishio 1998; Firmin-Sellers and Sellers, 1999; Deininger and Chamorro, 2004).

Land registration provides the necessary information to overcome the asymmetries in information available between two contracting parties to a land transaction. Consequently, land registration plays an important role in reducing land transaction costs and thereby raising the efficiency of any planned investment. This may enhance efficiency in land use by creating a market in land and/or increases market efficiency, thereby enabling property rights in land to move from less efficient to more efficient users of land. There is emerging evidence that land registration contributed to increased land rental market participation in Ethiopia (Holden *et al.*, 2007).

Land titles are associated with collateral in the following way: when borrowers apply for loans, land titles are often pledged as collateral. The pledging of titles as collateral, accompanied by registration of mortgage transactions, helps to overcome the problem of asymmetric information and the related incentive problems of moral hazard and adverse selection. These collateral arrangements are crucial for lending institutions and the credit market because they partly or fully shift the risk of loan loss from the lenders to the borrowers since a default on the loan would trigger the loss of collateral to the borrower. A combination of an increase in investment demand and credit supply associated with land registration leads to more investment, greater use of variable inputs, higher output per unit land area, greater income and higher land values.

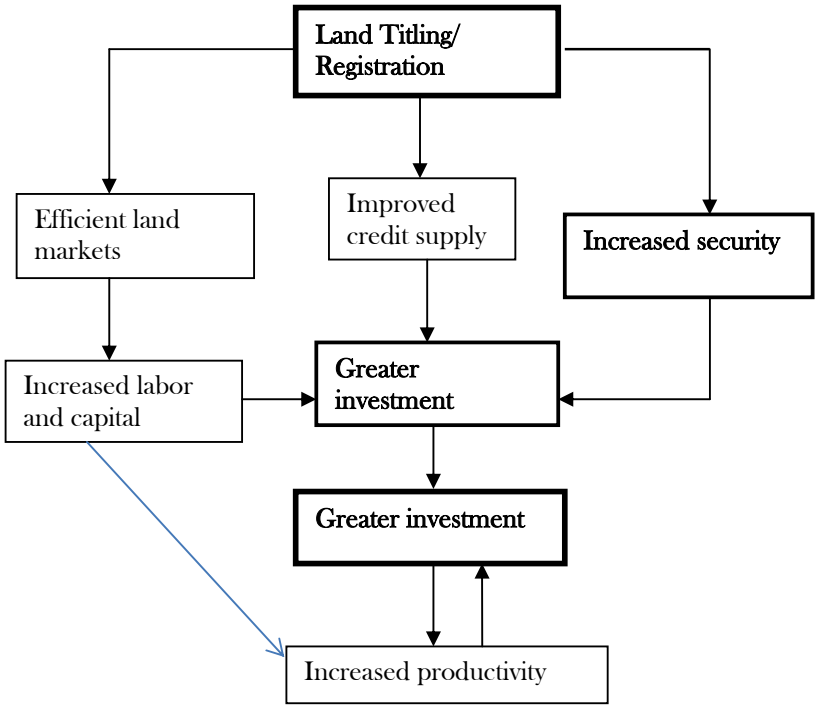
Evidence from Asia and Latin America suggests that formalizing land ownership, through registration and titling, can deliver large productivity gains. Feder and Nishio (1998) found that land registration led to higher land values in Thailand, Philippines, Indonesia, Honduras, Brazil and Peru. Carter and Olinto (2003) found that farmers in Paraguay have a positive investment demand effect.

On the other hand, evidences from Africa have found hardly any linkage between land titling and better credit market performance(Roth *et al.*, 1994).Cross-sectional data from Ghana, Kenya and Rwanda on the incidence of land improvements and on land

yields provide little support for the view that limitations under indigenous law on the right to transfer land are a constraint on productivity (Mighot-Adholla *et al.*, 1991) underlying that indigenous tenure systems may be preferable to improvements in tenure regimes through land registration. Recently Jacoby and Minten (2007) reported that, in Madagascar, having a title has no significant effect on plot-specific investment and correspondingly little effect on land productivity and land values. Results from Kenya also indicate that land registration and titling had weak impacts on perceived land rights of farmers, credit use and terms, crop yields, or concentration of land holdings (Place and Mighot-Adholla,1998).

Based on the reviews made on both theoretical and empirical literature, the study develops the following conceptual framework, which explicates the interaction(s) among variables of interest.

Figure 1: Conceptual framework on the effect of registration and land tilting on land related investment:



Source: Modified version of Feder et al., 1988

In this paper, we focused on the land tenure security and investment incentive effects instead of the land title, collateral and credit linkages because foreclosure of land is not within the law in Ethiopia because land is not used as collateral to access formal credit markets by smallholder farmers, but commercial farmers do so, while livestock or other property may be confiscated for indebted people. The land markets, transactions and efficiency linkages of land registration/titling in Ethiopia has been assessed by Holden *et al.*, (2007) by exploring on the functioning of land markets and market participation before and after registration. This paper intends to explore if land registration/certification has led to improved tenure security and whether this is translated to increased investment in land conservation, the so-called land tenure security and investment incentives effect, by focusing on the case study from Ethiopia. Furthermore, we identify the most important determinants of land conservation investment by smallholder farmers in northern Ethiopia.

4. Empirical Model and Econometric Approaches

Availability of data on investment before and after titling allowed estimating if there is a significant difference in the likelihood and level of investment which in turn enables us to capture the impact of land registration and certification on land-related investments before and after titling. We could have explained differences in investment before and after (see Deininger and Chamorro, 2004), if (1) the same households have been observed before and after registration, (2) it is possible to run a household fixed effects model. The later required adequate variation in plot level variables and year of registration/certification within the same household, which was hardly the case here. We, hence, used a random effects probit and truncated regression models to identify determinants of probability and intensity of investment. We also tried a modified random effects model as suggested by Mundalk (1987).The tobit model imposes a structure that is often too restrictive: exactly the same variables affecting the probability of adoption determine the level (intensity) of adoption and with the same sign (Verbeek, 2000). This is not a realistic assumption. We, thus, used the truncated regression model by taking only the positive values of levels of adoption to relax this restriction and identify the determinants of intensity of investment.

The approach we took is as follows. We established whether there was significant difference in the likelihood and level of investment (by type) using simple mean separation tests (t-test and X^2/ Z-tests). Next we explained the determinants of

probability of investment and level of investment using maximum likelihood estimation techniques.

Let the amount of conservation made on farm plot by a household i be given by:

$$y_i = x_1\beta_1 + \mu_1 \quad (1)$$

where y_i is the length of conservation structures per land area that is expected to depend on the vector x_1 regressors. y_i is non-continuous variable which is censored from below. We assumed that the level of investment made by each household on each plot is a result of rational decision. We used the physical measures of land-related investment that were made after registration. The measures include stone and soil bunds, terraces, tree planting, gully stabilization measures and water harvesting structures, among others. We regress the level of investment on set of covariates x_1 such as initial status of investment, year of registration[6], whether all plots were certified, a vector of household characteristics and resource endowments,, whether the households lost land because of past land redistribution and/or land taking, and a vector of time invariant plot characteristics such as tenure regime and observable plot characteristics such as soil type, slope, topography, soil fertility and susceptibility to erosion (using a subjective measure) and plot distance from homestead and other household characteristics such as age, sex of the household head and education of the head to obtain an estimate of the impact of an exogenous change in land rights on investment. We also controlled for household level asset holdings (land area, labor, livestock holding, and household's cash situation) and village level factors such as distance to market and agro ecology.

The participation equation, whether the household decides to invest or not, is given by:

$$y_2 = 1[x\delta_2 + v_2 > 0] \quad (2)$$

[6] Households do not get holding certificates automatically once their land is registered; there is a need to control both for registration and certification, even if they could be correlated.

where (x, y_2) are always observed whereas y_i (in eq. 1)is observed only when $y_2 = 1$. We assume that (u_1, v_2) is independent of x with mean zero implying that x is exogenous, and $v_2 \sim N(0,1)$. We used the same set of explanatory variables in equation (2) and in equation (1). Equations (1) and (2) are usually estimated using random effects probit and tobit regression models[7].

Random effects model estimators are consistent only under the assumption that the unobserved heterogeneity is uncorrelated with the observable explanatory variables. To obtain consistent estimates of the effects of registration/certification, we need to control for unobserved heterogeneity that may be correlated with observed explanatory variables. One way to do that is to exploit the panel nature of our data set (repeated cross-sectional plot observations per household). We used a modified random effects model framework proposed by Mundlak (1978), whereby we included on the right-hand side of each equation the mean values of plot varying explanatory variables. The rationale for including the mean values of plot varying explanatory variables as right hand side variables is to check if the unobserved heterogeneity that influence decisions are somewhat related to the observables (For a similar approach see Kassie *et al.*, 2008). In this case let, the model is given as:

$$y_i = x_1\beta_1 + \upsilon_h + \mu_1, \qquad (3)$$

where υ_h captures the unobserved household characteristics that affect household's decision to invest in land conservation. Assuming that the unobserved effects υ_h are linearly correlated with some of the observed explanatory variables,

$$\upsilon_h = \bar{x}\gamma + \eta_h, \eta_h \sim iid(0, {}_{\eta}^{2}) \qquad (4)$$

where $\bar{x}$ is the mean of plot-varying explanatory variables within each household, γ is the corresponding vector of coefficients, and η is the random error term uncorrelated with the $\bar{x}$'s. In our case, it is important to include average plot characteristics such as average soil quality, plot size, depth, slope, and the average size of initial investments,

[7]Heckman's selection model was not used because the presence of selection bias was not accepted (p-value 0.397). The results are not reported here.

which may have greater impact on adoption decisions. The vector γ will be equal to zero if the unobserved explanatory variables are uncorrelated with the random effects. For sake of comparison, we report results from the random and modified random effects models in this paper.

5. Study Site and Data Description

The study is based on a cross sectional data covering 437 randomly selected households, operating about 1695[8] plots gathered in June-Sept. 2005[9] in Tigray, northern Ethiopia. The survey covered 18 villages (tabias), four tabias strategically selected from each of the three zones (central, eastern, and southern) and 5 from the North Western and 1 from Western zones[10]. The last two tabias (1 each from Western and North Western) were purposively selected from the low land areas recently affected by the on-going human resettlement program to explore the effects of resettlement on tenure security and land investment. The selection of the 16 communities was based on stratification of differences in distance to market, agricultural potential conditions (due to variations in altitude and rainfall variability), population density and presence of irrigation projects (Hagos and Holden, 2006). The study assessed farmers' perception of land registration/certification, and its impacts, and the magnitude and quality of conservation investments made on farm through private initiative and/or public-led programs between before and after registration. The data gathered a host of household related variables as well as plot level data on the plots' biophysical features, including production and input use data, which are used for statistical analysis as reported in the subsequent sections.

6. Results and Discussion

Although land registration started in 1998, land registration and certification was still going on during the survey period (2004-2005) in some areas. This is especially the case in the lowland resettlement areas of the region. About 80 percent of the plots were registered during 1998-99. In the registration process it was found that about 96 percent

[8] There were some missing values in some observations of the data. We used only 1439 observations in the probit model.

[9] Although the data is relatively old, there is no radical land policy shift so far. The current issue is whether registration can be high cost (e.g. geo-information and cadastral survey based) instead of low cost.

[10] There have been recent changes in the boundary and number of zones in the region. The former Western zone is subdivided into North Western and Western zones.

of the plots were registered (only 4 percent indicated as not registered) and the demarcation process was also made both on paper and on the ground. In the registration process, it was found that in about 14 percent of the plots there was a change of boundary and about 8 percent of the plots were registered to somebody other than the owner. In about 13 percent of the plots there was dispute with a neighbour during the registration and demarcation of the boundaries. It was also found that about 14 percent of the plots are not certified. Title certificates were given to the household head (99.3 percent).

From the results of the survey, we found that there were households who received land through land redistribution up until 2005. The bulk of the plots (about 80 percent), however, were allotted to the current holders before 1991/92. During these redistributions about 31 percent of the households claim to have lost land while 69 percent did not. About 39 percent of the households believed that there will be future land redistribution even if land redistribution is ruled out by law (against 61 percent that did not expect land redistribution). As a consequence households expect that they will lose land due to redistribution. About 44 percent of the respondents who expect land redistribution believed that they will lose land against 56 percent who believed that they will not lose land. The major reasons for fearing future land redistributions included increasing population size, landlessness and too small land holding. Those who do not fear to lose land in a future land redistribution expect to get more land (7.2 percent), or at least would get their share (6.4 percent), or have landless member (3 percent) and they hoped will get an additional land from a redistribution. Of those who do not expect future land redistribution, only about 5.7 percent of the households indicated that there will not be further redistribution as it is prohibited by law and titles have already been issued and about 1 percent believed that resettlement is reducing pressure on further redistribution. Furthermore, 18 percent of the households also reported land takings for various reasons by the government (against 82percent who reported no land takings) and about 43 percent of those who lost land reported to have received compensation. In nutshell, non-negligible proportions of the households are apprehensive about or keen to see of future land redistributions perhaps indicating that the feeling of tenure insecurity is pervasive even after the land registration/certification although the latest law, as indicated, does not absolutely promote tenure security.

Table 1: Description and summary of important variables (n= 1695)

Variable name	Description	Mean
Latest redistribution	Proportion of land redistributed until 1991	0.79
Registration	Whether land is register or not (dummy)	0.96
Certification	Whether land is certified or not (dummy)	0.86
Change of boundary	Change of boundary during registration (dummy)	0.14
Registered to somebody	Plots were registered to somebody other than the owner (dummy)	0.08
Dispute	dispute with a neighbour during the registration (dummy)	0.13
Lost land	lost land during the latest land redistribution (dummy)	0.31
Future land redistribution	Expect future land redistribution (dummy)	0.39
Future land taking	Expect future land taking (dummy)	0.26
Will lose land	will lose land in future redistribution (dummy)	0.44
Land takings	Land takings by government after registration (dummy)	0.18
Compensation	Received compensation after land taking (dummy)	0.43
Access to formal credit market	Increased access to formal credit market because of land registration (dummy)	0.18
Access to informal credit market	Increased access to informal credit market because of land registration (dummy)	0.21
Temporary sell	Temporary transfer through sell (dummy)	0.10
Permanent sell	Permanent transfer through sell (dummy)	0.01
Tenure security	More secure because of land registration (dummy)	0.77
Tenure regime: owner operated	Proportion of owner-operated plots	0.75
Tenure regime: rented-in	Proportion of rented in plots	0.14
Tenure regime: rented-out	Proportion of rented-out plots	0.11
Conserved before	Proportion of plots conserved before registration (dummy)	0.34
Conserved after	Proportion of plots conserved after registration (dummy)	0.42
Conservation by holders before	Conservation by holders before registration (dummy)	0.52
Conservation through public programs before	Conservation through public programs before registration (dummy)	0.48
Conservation by holders after	Conservation by holders after registration (dummy)	0.66
Conservation through public programs after	Conservation through public programs after registration (dummy)	0.44

Comparison of important variables between those who invest and did not invest

Variable name	Description	Invest		Not invest	
		Mean	SD	Mean	SD
Labour endowment (adult male)	Average number of male adults	1.74	1.28	1.56	1.23
Land holding	Average parcel area	1.26	1.39	1.75	2.80
Livestock (oxen)	Average oxen holding	1.16	1.19	1.16	1.13
Credit access	Average farm loan size (in ETB)	4763.3	43107.0	3222.1	36280.6
Distance	Average distance (in min)	19.4	27.2	35.01	61.53

Source: 2004/05 survey.

When asked about the benefits of land registration, about 18 percent of the households perceived that land after certification can be used to access formal capital markets and slightly higher (21percent) to access informal credit markets. About 10 percent also believe that they can temporarily sell it against about 1 percent of the households who perceived that they can permanently sell their land, although the law stipulates differently.

Increased tenure security seems to have been also one of the most important benefits of land registration/certification. About 77 percent of the households believed that they feel more secure about their holding rights after registration. Nonetheless, a non-negligible portion of the households feel insecurity of tenure even after the land registration/certification.

Furthermore, be that as it may, whether this increased feeling of tenure security is translated into actual changes in behaviour is another main interest of this paper. The results show that about 34 percent of the plots were conserved before the registration as compared to 42 percent of the plots after the registration. Hence, there is a slight increase in the number of plots conserved after the registration. The dominant types of land investment, before registration, include establishment of stone terraces followed by soil bunds, gully control, tree planting and other conservation investments on 24, 11, 2, 2 and 5 percent of the plots respectively. Soil bunds and terraces remained dominant conservation structures established by farmers on their plots after registration both accounting for 26 and 14 percent of the plots. Gully stabilization (3 percent) and tree planting (2.4 percent) remain important as well. However, there are new introduction to land investment, most notably construction of water harvesting structures (3.3 percent) such as ponds and wells, since recently.

Before registration, about 52 percent of the investments were made by owners themselves and tenants, and 48 percent by public programs both mass mobilization and Cash/Food-For-Work programs (C/FFW). After registration, about 66 percent of the investments were made by owners themselves and tenants (i.e. who rent-in land), and 44 percent by public programs, both mass mobilization and C/FFW programs indicating an increase in private investment while public investments have also showed a decline.

In the resettlement areas, new settlers occupy about 4 percent of the plots of which nearly 80 percent are registered and certified. About 20 percent of the households expect that there will be further settlers in the future triggering fear of losing more land.

About 7 percent of the households indicated that there is land related conflict in the new resettlement areas. About 27 percent believed that those conflicts are between old and new settlers. One of the major sources of land related conflict in the resettlement areas are illegal expansion of farm holdings and unequal distribution of land. About 97 percent of the respondents from the resettlement areas, however, believed that there are institutions involved in solving land related disputes.

We explored if there is a significant difference in the proportion of plots conserved by households before and after registration. The simple mean separation result reported in Table 1 indicates that there is a significant change in conservation after registration (p-value = 0.000) Likewise, as to whether there is a significant difference in private investment before and after registration, the results indicated that there is a significant increased in private investment after registration(p-value = 0.000).

It is interesting to look at the changes in composition of investment before and after registration. From the mean separation tests, we gathered that while the proportion of stone terraces remains the same, soil bunds and other conservation measures have changed significantly (see Table 2).

We also conducted a mean separation test on the changes in levels of investment made both in terms of the level of physical measures and labour man days used for conservation. Both tests show that there are no statistically significant differences in levels of investment before and after registration, although the mean level is higher after registration. The measurement problem, as indicated earlier and the change in composition could be a reason for lack of difference in the level of investment. We can, hence, conclude that although there is significant increase in likelihood of making investment on land, the intensity remains the same. But this is unconditional mean; we did not control for the effect of other covariates.

Table 2: Mean proportion/separation tests (n= 1695)

Variables	Mean proportion		Difference (p-value)	X^2/Z/t-test (p-value)
	Before	After		
Proportion of conserved plots	0.335(0.011)	0.415 (0.01)	-0.076 (0.016)	-4.62 (0.000)
Proportion of private investment	0.18 (0.01)	0.27 (0.011)	-0.093 (0.01)	-6.49 (0.000)
Proportion of stone terraces	0.23 (0 .010)	0.25 (0 .010)	-0.021 (0 .015)	-1.45(0.145)
Proportion of soil bunds	0.105 (0.01)	0.138 (0.01)	-0.033 (0 .011)	-2.95 (0.003)
Proportion of other conservation investments	0.046 (0.005)	0.08 (0.006)	-0.034 (0.008)	-4.14 (0.000)
Level of investment (length of conservation measures in meters)	22.9 (2.08)	25.6 (3.73)	-2.66 (3.57)	-0.7446 (0.456)
Level of investment (amount of labour man days)	18.83 (2.96)	16.59 (1.89)	2.23 (2.12)	1.0523(0.293)

Source: 2004/05 survey

The results of the regression analyses, probit and truncated regression random effect and modified random effects models, for the investments after registration while controlling for the level of initial investment, household and plot level covariates, and changes in policy (i.e. year of registration and whether all plots are certified) are reported below. The year of registration and whether plots are certified or not had no significant effect on the probability of investment indicating there is no difference between those who got their land registered and secured their certificate earlier on and those who didn't obtain certificates. This may imply that registration and certification may not have created the required security incentive for increased land investment. Year of last redistribution, which could be a good proxy for the duration of holding, is found to have a negative effect on land investment implying that those who obtained land during the recent land redistributions were less likely to invest on their land. In other words, households that kept their holdings longer were found to be likely to make investment compared to households that obtained land recently. Households that reported to have lost land during the last land redistribution were also found less likely to make investment on their land pointing to the disincentive effect of recurrent redistribution and the associated loss of holdings. Households operating rented in land were found to be less likely to invest on land indicating that they may want to maximize immediate benefits without committing more resources to long-term investments. This may point to incentive problems as renting duration is usually for one year or two but rarely longer. Plots located far away from the homestead are less likely to be conserved

mainly because the cost (e.g. transport costs) of making and maintaining those investments is very high. These results are also confirmed earlier results on investment in Ethiopia (Hagos and Holden, 2006).

Most interestingly, initial investment had a very significant effect on the likelihood of investment after registration. What this result indicates is that those who have been making investment earlier on are also making the investments after registration/certification. This may imply that household specific characteristics influence decision rather than registration. In other words, regardless of whether land registration or certification has taken place or not, there are households that are likely to make investments on their land. A host of factors related to asset holdings and access to labour and capital markets had significant effect on probability of investment. Households with more livestock, land holding and adult male labour (although only in the random effects probit and at 10 percent level of significance) were found to be more likely to make investments on land. Households whose average land holding is relatively larger were found to be more likely to make investments on land perhaps pointing to space requirement of the conservation technologies. But this result was significant in the modified random effects model. Furthermore, households with access to food-for-work markets were found to be more likely to invest. However, those who have obtained higher loan from formal credit organizations were less likely to invest perhaps indicating that those who access for more loan invest it somewhere else than in land investment.

Table 3: Survey probit regression

Dept. variable: Investment after registration (0/1)						
Variables	Random effects			Modified random effects model		
	Coef.	Std. Err.	Marginal effect	Coef.	Std. Err.	Marginal effect
Household characteristics						
Age of household head	-0.001	0.004	-0.001	-0.0004	0.005	-0.0001
Education of household head (literate)	0.026	0.049	0.008	0.044	0.052	0.013
Female-headed household (Reference male-headed)	0.234	0.162	0.075**	0.283	0.167*	0.092*
Asset holding						
Number of Male adults	0.060	0.042	0.018*	0.041	0.043	0.013
Number of female adults	-0.002	0.065	-0.001	-0.001	0.065	-0.0003
Livestock holding	0.047	0.031	0.014**	0.051	0.032	0.015**
Land holding	0.001	0.001	0.001	-0.014	0.004**	-0.003**
Access to FFW income	0.0004	0.0002**	0.0001***	0.0005	0.0002**	0.0001***
Amount of Farm loan taken	-0.0001	0.0001	-0.000	-0.0001	0.0001**	-0.00004**

Land policy related variables						
Last year of land redistribution	-0.069	0.017***	-0.021***	-0.071	0.018***	-0.22***
Year of registration	0.015	0.048	0.005	0.025	0.051	0.007
Lost land during last redistribution (yes)	-0.147	0.104	-0.045*	-0.150	0.106	-0.045*
Land taken after registration (yes)	-0.094	0.132	-0.028	-0.005	0.131	-0.001
Whether plot is registered (yes)	-0.257	0.345	-0.086	-0.034	0.339	-0.011
Whether plot is certified (yes)	0.199	0.193	0.058	0.250	0.203	0.072
Duration of land holding (since year)	0.005	0.007	0.001	-0.017	0.011	-0.005
Tenure form: rented-in (reference owner operated)	-0.302	0.141**	-0.086**	-0.235	0.146*	-0.068*
Tenure form: rented-out(reference owner operated)	-0.197	0.171	-0.057	0.218	0.210	0.071
Whether plot conserved before registration (yes)	1.092	0.111***	0.390***	0.761	0.123***	0.265**
Plot characteristics						
Plot area (tsimad)	0.028	0.018	0.008	-0.001	0.001	-0.0002
Distance (in minutes)	-0.005	0.002***	-0.001***	-0.007	0.002***	-0.002***
Soil type: Walka (reference Baekel)	-0.037	0.123	-0.011	-0.084	0.120	-0.025
Soil type: Sandy (reference Baekel)	-0.067	0.117	-0.020	-0.128	0.113	-0.038
Soil type: Red soil (reference Baekel)	0.069	0.1222	0.022	-0.085	0.129	-0.026
Soil type: Red soil: other (reference Baekel)	0.532	0.248**	0.189*	0.136	0.244	0.044
Soil depth: medium (reference shallow)	0.026	0.104	0.008	0.008	0.106	0.002
Soil depth: deep (reference shallow)	-0.006	0.128	-0.002	0.029	0.141	0.009
Soil slope: foothill (reference flat)	0.220	0.122**	0.071**	0.088	0.119	0.027
Soil slope: mid hill (reference flat)	0.358	0.130**	0.121***	0.147	0.157	0.047
Soil slope: steep hill (reference flat)	-0.041	0.279-	-0.012	-0.333	0.298	-0.090
Soil quality: medium (reference poor)	0.078	0.115	-0.024	-0.093	0.124	-0.028
Soil quality: good (reference poor)	0.011	0.147	0.003	0.085	0.180	-0.025
Susceptibility to erosion: medium (reference high)	0.250	0.159	0.081*	0.381	0.179**	0.126**
Susceptibility to erosion: low (reference high)	0.271	0.135**	0.087**	0.355	0.157**	0.0115**
Susceptibility to erosion: none (reference high)	0.216	0.135	0.068*	0.274	0.142**	0.087**
Village level characteristics						
Distance to woreda market	-0.0001	0.0005	-0.000	-0.0002	0.001	-0.0001
Distance to DA office	0.002	0.0002	0.000	0.002	0.003	0.001
Altitude: Midland (reference highland)	0.040	0.120	0.012	0.049	0.126	0.015
Altitude: lowland (reference highland)	-0.094	0.143	-0.028	-0.125	0.152	-0.037
Average time-invariant plot characteristics						
Average soil type	-	-	-	0.157	0.088*	0.048**
Average soil depth	-	-	-	0.015	0.127*	0.004
Average soil slope	-	-	-	0.239	0.129*	0.074**
Average soil quality	-	-	-	0.099	0.159	0.031
Average susceptibility to erosion	–	-	-	0.064	0.106	0.019
Average form of tenure	-	-	-	-0.373	0.159**	-0.115**
Average plot area	-			0.118	0.044**	0.036**
Average distance	-	-	-	0.002	0.001	0.001
Average duration of holding (in yrs)	-	-	-	0.031	0.015**	0.009**
Average probability of investment	-	-	-	0.891	0.243***	0.275***
Intercept	94.409	84.07	-	58.25	92.29	
	Number of obs = 1439			Number of obs = 1439		
	F(39, 354) = 6.56			F(49, 354) = 5.73		
	Prob> F = 0.0000			Prob> F = 0.0000		

Note: *, **, *** significant at 10, 5 and 1% respectively.
Source: 2004/05 survey

Some plot level characteristics were also found to be significant. Investment on land was significantly higher on plots that are moderately sloppy and moderately susceptible to erosion. This results show that households avoid making investments on highly susceptible and steep slopes perhaps the cost of investment are prohibitively high. The likelihood of investment increased with average soil type. Plots with soil type that deviate from Baekel (mixed sandy and loam soil) are likely to be conserved. The likelihood of investment decrease with average form of tenure: plots that are not owner-operated are less likely to be conserved. The likelihood of investment increase with average duration of holding strengthening duration of ownership increases the likelihood of land investment. Finally, the likelihood of investment increase with average probability of investment. The inclusion of time-invariant plot characteristics does not change much the results, although not a negligible number of the included variables themselves turned out to be significant, individually and jointly. We could conclude that the major results are quite robust.

Many of the results from the probit model were also confirmed in the truncated model random effects and modified random effects models. We also get new insights as many variables turned out significant in explaining levels of investment made by households. In this case, duration of holding significantly determines the level of investment made by households. Households that received land very recently were found to have made significantly lower investment than households that kept their holdings for a long time. Unlike in the probit models, intensity of investment seems to have significantly been influenced by the year of registration, i.e. level of investments has decreased as the year of registration was delayed. Households that reported to have lost land during the last land redistribution were found to have higher land investment. But those households whose land was taken away because of infrastructural development were found to have made lower investments strengthening the disincentive effect of land takings. Unlike the results in probit model, rented-in plots have higher investment. This implies that they are unlikely to be conserved but once they are selected for conservation they have higher intensity, although only in the random effects model. The influence of initial investment on the level of investment was also found to be significant and positive confirming that households to make higher investments on their land in the presence or absence of land registration. The results here strengthen that there are household levels characteristics that predispose households to carry on making land investments in the presence or absence of land policy changes. Consistent with the results in the probit model, households with access to food-for-work markets were found to be have made higher investment on land underscoring the significance of access to food-for-work in

relieving households' cash constraints and enhancing long-term investment (Holden et al, 2006; Bezu and Holden, 2004). Furthermore, household factors such as age of household head and education of head have a significant influence on the intensity of investment. In this case, households with relatively older heads, although not significant in the modified random effect model, invested less while households with better educated heads made higher investment on land underlying the importance of better human capital endowment for land investment.

From environmental and plot level variables, the level of investment made on land depends on altitude indicating agro-ecological variations in land degradation and the need for SWC. The level of investment is significantly higher in midland communities than high altitude. The level of investment also varies by soil and plot level characteristics. Accordingly, the level of investment varied by soil type, quality, depth and slope of the soil. Investments were found to be higher on relatively deeper soils than shallow soils; on relatively more fertile soils than poor soils pointing to the economic considerations of investment by farmers. Distance has also negative effect on the level of investment as we found in the probability of investment, although not significant in the intensity model. This could be understood in the light of the fact that making investments and undertaking proper maintenance on far away plots is costly. Farm lands located away from homesteads usually turn to common grazing lands during the dry season increasing the chance of destruction of conservation structures by livestock.

Table 4: Truncated random effects model

Dept. variable: Intensity of investment

Variables	Random effects			Modified random effects model		
	Coef.	Std. Err.	Marginal effect	Coef.	Std. Err.	Marginal effect
Household characteristics						
Age of household head	-44.68	16.97***	-0.320***	-4.36	16.41	-0.033
Education of household head (literate)	968.77	221.03***	6.94***	787.43	216.15***	6.10***
Female-headed household (Reference male-headed)	191.95	573.89	1.37	862.89	584.36	6.68
Asset holding						
Number of Male adults	201.38	148.49	1.44	236.08	161.5	1.83
Number of female adults	109.06	246.05	0.78	68.35	259.4	0.529
Livestock holding	-278.78	126.38**	-1.99**	-29.42	115.85	-0.228
Land holding	87.67	72.77	0.62	78.17	0.835	0.606
Access to FFW income	3.632	0.789***	0.026***	2.518	0.762***	0.019***
Amount of Farm loan taken			-0.001	-0.093	0.278	-0.0007

Land policy related variables						
Last year of land redistribution	-239.14	112.69**	-1.71**	-191.06	114.05*	-1.481
Year of registration	-829.21	304.05***	-5.93***	-818.13	111.5***	-6.341***
Lost land during last redistribution (yes)	1648.92	466.55***	11.81***	2063.8	513.5***	15.998***
Land taken after registration (yes)	-2213.40	668.24***	-15.85***	-1704.97	652. 85***	-13.21***
Whether plot is registered (yes)	-878.50	2025.5	-6.29	-967.56	1919.06	-7.50
Whether plot is certified (yes)	1139.19	828.11	8.15	1463.27	32.75	11.34
Duration of land holding (since year)	6.476	589.17	0.046	-84.16	32.75***	-0.652***
Tenure form: rented-in (reference owner operated)	1266.03	589.17**	9.06**	349.8	600.31	2.71
Tenure form: rented-out (reference owner operated)	297.16	709.17	2.13	-625.70	832.7	-4.84
Whether plot conserved before registration (yes)	3.522	0.734***	0.025***	1614.1	490.7***	12.51***
Plot characteristics						
Plot area (tsimad)	-341.34	236.11	-2.44	-206.82	207.8	-1.60
Distance (in minutes)	-8.824	9.639	-0.06	-5.329	8.469	-0.041
Soil type: Walka (reference Baekel)	-1878.0	576.53***	-13.45***	-2572.25	634.08***	-19.94***
Soil type: Sandy (reference Baekel)	-3055.77	649.82***	-21.88***	-1950.05	617.23***	-15.12***
Soil type: Red soil (reference Baekel)	-1278.58	579.83**	-9.16**	-1458.8	654.65**	-11.31**
Soil type: other (reference Baekel)	520.8	935.39	3.73	11.43	951.94	0.088
Soil depth: medium (reference shallow)	678.16	446.82	4.85	396.83	501.29	3.076
Soil depth: deep (reference shallow)	1813.03	586.13***	12.98***	1122.0	649.03*	8.69*
Soil slope: foothill (reference flat)	977.94	542.05*	7.00*	1019.04	503.38**	7.89**
Soil slope: mid hill (reference flat)	3225.9	652.53***	23.10***	2888.9	783.1**	22.39**
Soil slope: steep hill (reference flat)	790.34	1464.01	5.66	456.22	1390.56	3.53
Soil quality: medium (reference poor)	1333.36	581.41**	9.55**	673.08	671.46	5.217
Soil quality: good (reference poor)	2650.7	768.18***	18.98***	1712.90	920.37*	13.27*
Susceptibility to erosion: medium (reference high)	979.88	911.77	7.01	-822.43	897.98	-6.37
Susceptibility to erosion: low (reference high)	1125.28	848.54	8.06	-816.42	769.71	-6.32
Susceptibility to erosion: none (reference high)	147.37	904.90	1.05	-1873.69	917.42**	-14.52**
Village level variables						
Distance to woreda market	-0.182	2.128	-0.001	1.90	2.06	0.014
Distance to DA office	27.37	10.15***	0.0196***	23.09	8.72***	0.179***
Altitude: Midland (reference highland)	3627.9	666.89***	25.98***	2216.48	557.80***	17.181***
Altitude: lowland (reference highland)	207.01	915.33	1.48	-469.64	853.03	-3.640
Average time-invariant plot characteristics						
Average soil type	-	-	-	121.49	356.65	0.941
Average soil depth	-	-	-	331.91	512.40	2.572
Average soil slope	-	-	-	-173.56	583.94	-1.345
Average soil quality	-	-	-	715.31	744.1	5.544
Average susceptibility to erosion	-	-	-	-281.78	744.07	-2.18
Average form of tenure	-	-	-	1010.76	547.34*	7.835*
Average plot area				29.93	358.52	0.232
Average distance	-	-	-	-0.117	6.243	-0.001
Average duration of holding (in yrs)	-	-	-	65.59	61.87	0.508
Average probability of investment	-	-	-	-3845.7	1182.02***	-29.81***
Intercept	2100080	648603***	15042.2***	2031827	-	-
	Number of obs = 420					
	Wald chi2(39) = 99.48					
	Prob> chi2 = 0.0000					

Note: *, **, *** significant at 10, 5 and 1% respectively.
Source: 2004/05 survey

Contrary to our expectations, we found that investments were significantly higher by households that are located far away from the Development Agent Office, which we used as a proxy for access to extension service. Whether this is capturing the effects of access to extension service or not is difficult to tell. The intensity of investment increases with average form of tenure (not owner-operated) and decrease with average probability of investment.

7. Conclusions and recommendations

The evidence that comes out of this cross sectional study is that land registration and certification has enhanced household's feeling of security but it is not a single dominant factor that affects farmers' decision on investment. Land registration/certification is, thus, a necessary policy measure to induce positive security effects on holdings with positive effect on land investments. Yet high population growth, increasing landlessness and land takings in peri-urban settings do pose serious implication on insecurity to users.

This study also indicated that there was significant increase in probability of land investment after registration although the mean level of investment statistically remained about the same. There is also a change in the composition of investments: increased investment on trees, gully stabilization and water harvesting structures than the usual soil and water conservation measures, and increase proportion of plots conserved through private investment.

Households that kept their land holdings relatively longer were found more likely to make and have made long (more) investment compared to households that obtained land recently. Likewise, households that reported to have lost land during the last land redistribution were also found to have made lower investments strengthening the disincentive effect of insecurity caused by recurrent redistribution and the associated loss of holdings. The intensity of investment are significantly influenced by the year of registration, i.e. level of investments has decreased as the year of registration was delayed indicating the security effect induced incentive on land conservation. This underscores the need for having the feeling of tenure security for increased long-term land investment. This calls for policy makers the need to minimize the potential sources of insecurity such as threats of future land redistribution and land taking without proper land compensation. The prohibition of further distribution and the proclamation 455/97on compensation (FDRE, 2005b) is in the right direction but it

requires meticulous enforcement. Households who rented in land want to maximize immediate benefits without committing more resources to long-term investments. Accordingly, they were found to be less likely to invest on land conservation. This points to incentive problems related to renting as the duration is usually for one year or two and rarely longer. Hence, there is a need to reduce the incentive problems of land rental markets perhaps by formalizing and extending the duration of land rental rights among farmers. Initial investment had a very significant effect on the likelihood and intensity of investment after registration, the results here reinforces that there are household level characteristics that predispose households to carry on making land investments in the presence or absence of land policy changes. But there is no doubt that land registration/certification strengthens those household predispositions. This calls for the expansion of land certification to other parts of Ethiopia; but the increased effect of high-cost registration including cadastral survey on security and land investments was not the focus of this study. Furthermore, households with access to food-for-work markets were found to be more likely and have made higher investment on land underscoring the significance of access to food-for-work in relieving households' cash constraints and enhancing long-term investment. Such programs not only reduce households' vulnerability to food insecurity but also generate required resources to make long-term investments. There is, hence, a need for continued support to households in the form of food-for-work.

Finally, some time invariant environmental plot level factors such as altitude, soil type, quality and depth, distance and plot area are found to have significant effect on investment. Long-term land investment, therefore, is determined by hosts of physical, socio-economic and institutional factors that need to be taken into account in planning any such event.

References

Atwood, D. (1990). Land registration in Africa: The impact on agricultural production. *World Development,* 18(5): 659-671.

Besley, T. (1995). Property rights and investment incentives: theory and evidence from Ghana. *Journal of Political Economy,* 103, 903-937.

Binswanger, H. (1996). The Political Implications of Alternative Land Policies. Chapter 6. In: van Zyl, J. and Hans Binswanger. (Eds.), *Market-assisted Land Reform: How will it work?* Oxford University Press, Oxford.

Bezu, S. and Stein Holden. (2004). Impact of Food-For-Work in Tigray, Ethiopia. Norwegian University of Life Sciences, Department of Economics and Resource Management. Unpublished.

Carter, M. R, and Olinto, P. (2003). Getting institutions "right" for whom? Credit constraints and the impact of property rights on the quantity and composition of investment. *American Journal of Agricultural Economics,* 85(1): 173-186.

Carter, M., Wiebe, K. and Blarel, B. (1991). Tenure security for whom? Differential impacts of land policy in Kenya, Land Tenure Center Research Paper No. 106, University of Wisconsin-Madison.

Deininger, K., Chamorro, J. S. (2004). Investment and equity effects of land regularization: the case of Nicaragua. *Agricultural Economics,* 30, 101-116.

Deininger, K., Zevenbergen, J. Ali, D. A. (2006). Assessing the certification process of Ethiopia's rural lands. Paper presented in Colloque international "Les frontières de la question foncière – At the frontier of land issues", Montpellier, 2006.

Deininger, K., Ali, D. A., Holden, S., and Zevenbergen, J. (2007). Rural land certification in Ethiopia: Process, initial impact, and implications for other African countries. World Bank Policy Research Working Paper 4218. The World Bank, Washington, D.C.

Ethiopian Economic Association (EEA). (2007). Report on the Ethiopian Economy 2005/06. Volume V. Addis Ababa.

EEA/EEPRI. (2002). Land tenure and agricultural development in Ethiopia. Research report. Ethiopian Economic Association/ Ethiopian Economic Policy Research Institute.

FDRE. (1995). Constitution of the Federal Republic of Ethiopia. Federal Negarit Gazeta, Proclamation No. 1/1995. Addis Ababa.

FDRE. (1997). Rura1 land administration proclamation, No. 89/1997. Addis Ababa.

FDRE. (2005a). Federal rural land administration and use proclamation. Proclamation No. 456/2005". Federal Negarit Gazeta. 11th year no. 44. 15th July 2005. Addis Ababa.

FDRE. (2005b). Expropriation of landholdings for public purposes and payment of compensation. Proclamation No. 455/2005". Federal Negarit Gazeta. 11th year no. 43. 15th July 2005. Addis Ababa.

Feder, G., T. Onchan and T. Raparla. (1986). Land Ownership Security and Access to Credit in Rural Thailand, Report No. ARU 53.The World Bank, Washington, D.C.

Feder, G. and Feeny, D. (1993). The theory of land tenure and property rights. In: The *Economics of rural organization: Theory, Practice and Policy*, (eds.) Hoff, K., A. Braverman and J. E. Stiglitz. John Hopkins, Baltimore, MD, 240-258pp.

Feder, G. and Nishio, A. (1999). The benefits of land registration and titling: Economic and social perspective. *Land Use Policy,* 15, 941-951.

Feder, G., Onchan, T., Chamlamwong, Y. and Hongladarom, C. (1988). Land policies and farm productivity in Thailand, Johns Hopkins University Press, Baltimore, MD.

Firmin-Sellers, K., Sellers, P. (1999). Expected failures and unexpected successes of land titling in Africa. World Development, 27 (7): 1115-1128.

Gebremedhin, B. (1998). The economics of soil conservation investments in Tigray region of Ethiopia, Ph.D dissertation, Department of Agricultural Economics, Michigan State University, East Lansing, MI, 258 pp.

Hagos, F., J. Pender and N. Gebreselassie. (1999). Land degradation and strategies for sustainable land management in the Ethiopian highlands: Tigray region, Socio-economic and Policy Research Working Paper, 25, International Livestock Research Institute, Nairobi, Kenya.

Hagos, F., Holden, S. (2002). Incentives for conservation in Tigray, Ethiopia: Findings from a household survey. Department of Economics and Social Sciences, Agricultural University of Norway, unpublished.

Hagos, F., Holden, S. (2006). Tenure security, resource poverty, public programs and household plot level conservation investment in the highlands of northern Ethiopia. *Agricultural Economics*, 34, 183-196.

Haile, M., Witten, W., Kinfe Abreha, Sintayehu Fisseha, Adane Kebede, Getahun Kassa, Getachew Reda. (2005). Land registration in Tigray, Northern Ethiopia. Research Report 2, IIED.

Holden S., Yohannes, H. (2002). Land redistribution, tenure insecurity and intensity of production: A study of farm households in Southern Ethiopia. *Land Economics*, 78 (4), 573-590.

Holden, S., Barrett, C., and Hagos, F. (2006). Food-for-Work for poverty reduction and the promotion of sustainable land use: Can it work? *Journal of Environment and Development*, 11:1-26.

Holden, S., Deininger, K., and Ghebru, H. (2007). Impact of land certification on land rental market participation in Tigray Region, Northern Ethiopia. Paper submitted for the Nordic Development Economics Conference in Copenhagen June 18-19, 2007

Jacoby, H. G. Minten, B. (2007). Is Land Titling in Sub-Saharan Africa Cost-Effective? Evidence from Madagascar. *The World Bank Economic Review*, 21(3): 461-485.

Joireman, S. F. (2001). Property rights and the role of the state: Evidence from the Horn of Africa. *Journal of Development studies*, 38 (1), 1-28.

Kassie, M., Pender, J., Yesuf, M., Kohlin, G., Bluffstone, R. and Mulugeta, E. (2008). Estimating returns to soil conservation adoption in the Northern Ethiopian Highlands. *Agricultural Economics*, 38: 213-232.

Migot-Adholla, S. E., Hazell, P., Blarel, B. and Place, F. (1991). Indigenous land rights systems in Sub-Saharan Africa: A constraint on productivity? *The World Bank Economic Review, 5* (1): 155-175.

Mundlak, Y. (1978). On the pooling of time series and cross sectional data. *Economertica* 64(1): 69-85.

Place, F., Migot-Adholla, S. (1998). The economic effects of land registration on smallholder farms in Kenya: evidence from Nyeri and Kakamega districts. *Land Economics*, 74, 360-373.

Rahmato, D. (1992). *The Land question and reform policy: Issues for debate. Dialogue*, Addis Ababa University Printing Press.

Rahmato, D. (2003). A brief note on research priorities on land matters. Excerpt from a paper. Attp://www.worldbank.org.

Rahmato, D. (2004). Searching for tenure Security? The Land System and New Policy Initiatives in Ethiopia. FSS Discussion paper no. 12. Forum for Social Studies. Addis Ababa.

Regional National State of Tigray (RNST). (1997). Regional rural land utilization proclamation (23/1989 E.C.), Mekelle.

Regional National State of Tigray (RNST). (2007). Revised rural land administration and use proclamation. (Proclamation 136/2000 E.C). Mekelle.

Roth, M., Unruh, J and Barrows, R. (1994). Land registration, tenure security, credit use, and investment in the Shebelle Region of Somalia. In: J. Bruce and S. E. Mighot-Adholla (eds.), Searching for Security of Tenure in Africa. Dubuque, IA: Kendall/Hunt.

Teklu, T. (2003). Rural lands and evolving tenure arrangements in Ethiopia: Issues, evidence and policies. FSS discussion paper No. 10. Forum for Social Studies. Addis Ababa.

Verbeek, M. (2000). *A guide to modern econometrics*. John Wiley & Sons, Ltd.

THE SUSTAINABILITY AND OUTREACH PERFORMANCE OF ETHIOPIAN MICROFINANCE INSTITUTIONS: DOES SCALE OF OUTREACH MATTER?

Letenah Ejigu[1]

Abstract

This study analyzes sustainability and outreach performance of Ethiopian MFIs using scale of outreach as a classification variable. Performance is compared against local and international benchmarks and across time. Secondary data of 13 MFIs collected for 6 years (2003-2008) is used for analysis. The result reveals scale of outreach matters when performance is compared against local benchmarks. It further indicates the presence of commercial and social oriented MFI clusters. Comparison with international benchmarks shows mixed evidence across scale of outreach, but the whole industry is weak in terms of depth of outreach. The trend analysis indicates the industry suffers from lack of clear women targeting policy and erratic profitability as measured by ROA.

Key terms: Microfinance, Sustainability, Outreach
JEL Code: G21

[1] PhD, Postgraduate Research and Community Service Coordinator, College of Business & Economics, Bahir Dar University; P.O. Box 79, Bahir Dar Ethiopia
E-mail:etalem2000@yahoo.com, Cell-phone: +251 918 761517

1. Introduction

Microfinance is defined as financial services targeted to low income clients (Ledgerwood, 1999). It includes saving, credit, micro insurance, payment service & money transfer. Although the industry has a long history, the practice of formal microfinance can be traced back prior to the works of the Grameen Bank in Bangladesh (Christen, 1997). Over the years, microfinance has registered impressive growth on product offerings, repayment rate, financial sustainability and loan portfolio (Ledgerwood, 1999; Woller and Schreiner, 2006).

MFIs are developed in response to big market failures in the financial systems of many countries. Such market failures are created due to the reluctance of traditional banks to lend to the poor for various reasons (Aghion & Morduch, 2005). First, disbursing small loans to a large number of borrowers is a costly business proposition. Second, lending to the poor is risky due to information asymmetry. Poor people have no physical collateral in the event of default and getting information on their credit history is difficult. Third, the poor lack skills to use borrowed money productively. MFIs have circumvented all these assumptions and attained a huge success (e.g. they secured high repayment rates). They did so by using innovative group lending methods, accessing subsidy from donors and providing to the poor non-financial services like training and assistance in marketing.

With respect to its setup, microfinance is a hybrid organization that combines traditional banking with the social goal of poverty reduction at the grass roots level. Like banks, it aspires to live using its own means by following a commercial approach of doing business. Such a goal is associated with the concept of sustainability[2]. On the other hand the social goal of poverty reduction calls for making services available to many poor clients at affordable rates, despite high transaction costs of doing such business. This issue is associated with the concept of outreach[3]. The hybrid nature of

[2]Sustainability is conceived as "full cost recovery or profit making and the building of microfinance institutions that can last into the future without continued reliance on government subsidies or donor funds." (Conning, 1999, p.52)

[3]Outreach refers to "the efforts by microfinance organizations to extend loans and financial services to an ever wider audience (breadth of outreach) and especially toward the poorest of the poor (depth of outreach)." (Conning, 1999, p.52). There are other broader conceptualization of outreach. For example Schreiner (1999) classified outreach in to six dimensions: worth to clients, cost to clients, depth, breadth, length and scope. But such broad approaches are not pursued due to data limitations and unclear operationalizations.

such organizations urges them to use innovative methods of fighting poverty. This attracts the attention of many stakeholders like academicians, policy makers and practitioners.

Despite the huge success of the industry as a whole[4], the twin goals of sustainability and outreach are not yet met. For instance, many MFIs are not financially sustainable, with 41% seeking donor support to keep afloat (Mersland & Storm, 2010). MFIs reach only 3% of the world's 500 million poor lacking access to financial services, indicating less breadth of outreach (Marzys, 2006). On the depth of outreach, studies indicate that MFIs fail to reach the poorest of the poor (Hashemi & Rosenberg, 2006) and do not have clear targeting rules (Hishigsuren, 2004). The microfinance model itself is increasingly questioned. Some authorities on the subject say reaching the poorest clients and being financially self-sufficient are competing objectives (Mersland & Storm, 2007; Okumu, 2007) whereas others like Schreiner, (1999), Christen, (1997) and Cull, DemigÄuc-Kunt, & Morduch (2007) are of the view that such contradictions do not exist.

Various studies have explored sustainability and outreach of MFIs in Ethiopia and documented weak performance of the sector (Pfister et al, 2008; Kereta, 2007; Kidane, 2007, Amha, 2007, Ejigu, 2009). They reach around 6% of the poor, only 38.4% of targeted women and depend on subsidies. Most of these studies, with the exception of Ejigu (2009), did not apply rigorous statistical analysis. Ejigu's study, however, uses one year (2007) data to classify MFIs across different scales of outreach[5], which could not capture the dynamics within the sector. This study, using data from 13 MFIs collected over 6 years (2003-2008), tries to address such gaps in a more rigorous way. The sustainability and outreach performance of MFIs is compared with global benchmarks of Microfinance Information eXchange (MIX)[6] and local benchmarks (i.e. across the different scales of outreach). Using graphs, the trend of such performance is also depicted for the industry as a whole.

[4]Muhammad Yunus and the Grameen Bank are awarded the Nobel Peace Prize in 2006 "for their efforts to create economic and social development from below." (Sengupta & Aubuchon, 2008).

[5]Scale of outreach is measured by Gross Loan Portfolio (GLP) of MFIs. GLP itself is a product of the two measures of outreach, number of borrowers (measuring breadth of outreach) and average loan size (measuring depth of outreach).

[6]MIX is a non-profit initiative dedicated to the dissemination of quality MFIs data globally. See www.mixmarket.org

The findings of this paper are important to MFIs management and regulatory bodies such as the National Bank of Ethiopia. Comparison with MIX averages enables to know lessons to be learned from abroad. Comparing MFIs locally helps benchmarking among the MFIs themselves, instituting policy changes on entry barriers and facilitators, and changing the regulatory frameworks for MFIs with different missions. The trend analysis helps to check up whether MFIs are moving in line with their missions. The paper also informs future researchers on a scientific way of conducting studies of performance analyses.

The rest of the paper is organized as follows. Section two reviews the existing literature on performance measurement in the context of microfinance and some empirical studies conducted in this area. Data and methodology issues are discussed in section three. Section four discusses results of the study and section five presents concluding remarks.

2. Review of Literature

2.1 The Development of Performance Measures in Microfinance Institutions

When MFIs were initially constituted, they were having a mission of reaching many poor households offering them low interest rates (Bateman, 2010). Low interest rates were favored because it was presumed that the cost of credit is a significant bottleneck for the poor to access credit and cheap credit, by liberating the poor from exorbitant money lenders interest rates, helps them overcome poverty. To make credit cheap, MFIs were heavily subsidized by governments & donors. This lending approach is called 'welfarist' since microcredit is designed as part of an integrated program of poverty alleviation and welfare improvement.

Starting in the 1990s, rural finance experts from Ohio State University began criticizing the 'welfarist' approach claiming that subsidized credit based on faulty conceptions leads to worst practices of high unpaid rates and transaction costs resulting in the failure of many microcredit programs (Congo, 2002). This leads to another paradigm in the development of microfinance. Considered as the 'institutionalist' approach, it seeks to establish institutions which can offer saving and credit services on sustainable and commercial basis and free from subsidy. To achieve this, the MFIs charge high interest rates due to high costs of doing business (Gurgand, Pederson, & Yaron, 1996).

The debate between 'welfarists' and 'institutionalists' is termed as the 'microfinance schism' and there seems little agreement between the two schools of thought regarding the goal of MFIs and how to achieve it. The disagreement between the two schools also leads to the use of various metrics to measure the performance of MFIs. 'Welfarists' are interested on demand or client side outcome and focus on the impact of MFIs on the lives of the poor and outreach to the poor. 'Institutionalists', on the other hand, focus on the supply side issue of building sustainable institutions which are free from subsidy. They are interested in tracking repayment rate, transaction cost and financial self reliance (Congo, 2002). Nevertheless, there is no unanimous agreement among stakeholders until now on the metrics used to gauge the performance of MFIs. Many authorities use sustainability and outreach (Gurgand et al.1996) and this paper follows the same approach.

2.2 Empirical Evidence on the Sustainability and Outreach Performance of MFIs

Several studies on sustainability and outreach performance of MFIs have been conducted at global, regional as well as country levels. At global level, Cull et al (2007) found that the microfinance industry is financially and operationally sustainable with FSS, 1.03 and OSS, 1.16, but ROA[7] still negative, -0.027. Depth of outreach seems somewhat good with average loan size per GNI per capita of 67.6% and percentage of targeted women borrowers served accounting for 64.9%. Mersland and Storm (2007) also confirm the sustainable but low ROA of the global microfinance industry. This indicates that due to the social goal of poverty reduction, MFIs are not very much profitable like banking or any other commercial businesses. They also found average number of borrowers served as 12,805. The figure is low and indicates a huge financial exclusion of the poor.

The regional studies show mixed results. Lafourcade et al., (2005) found that African MFIs are lagging behind other parts of the world in terms of both sustainability and outreach. Hartarska's (2005)study in Central and Eastern European (CEE) region showed MFIs to be profitable with ROA, 3.04% but serve few borrowers,7268.

[7]FSS, OSS and ROA are commonly used measures of sustainability and see the methodology section later to know what they stand for.

On country-level studies, Okumu (2007) based on Ugandan MFIs showed that MFIs are operationally self sustainable with OSS, 1.21; the breadth of outreach is small, 6729 borrowers being served; their depth of outreach is small with average loan size per GNI per capita being 1.43. Congo (2002) examines Burkina Faso MFIs and obtained lower sustainability and outreach performance.

Studies on sustainability and outreach performances of MFIs in Ethiopia also revealed mixed results. For instance, Kereta (2007) shows that the industry has registered a high growth rate of 22.9% in terms of borrowers over the period 2003-2007; depth of outreach is small with the percentage of women borrowers reached standing at 38.4%; financial sustainability is improving although still donor dependent. The author also documented no clear tradeoff between sustainability and depth of outreach. Kidane (2007), examined Amhara Credit & Saving Institution (ACSI), one of the largest MFIs in Ethiopia, and found that it has served more than half a million clients; it is operationally and financially self sufficient with 119.9% and 115.3%, respectively. Pfister et al (2008) documented that Ethiopian MFIs have limited outreach, reaching almost two million clients only in a country of 77 million people although dependent on government and mother NGOs, indicating weak financial sustainability. Ejigu (2009) showed that Ethiopian MFIs have less depth of outreach, extend large loan sizes and the number of women clients reached is low. They have good breadth of outreach as a large number of borrowers are reached. Their profitability seems to depend on their scale of outreach, i.e. the larger MFIs being more sustainable than the small ones. There is also a tradeoff between serving the poorest borrowers and being financially sustainable. Almost no studies have conducted performance analysis by comparing with benchmarks like MIX or local averages using proper statistical tools. This paper building on the work of Ejigu (2009), tries to address the empirical research gaps.

3. Data and Methodology

3.1 Data

As of 2010, there are 30 MFIs in Ethiopia registered and licensed by the National Bank of Ethiopia (NBE). For sustainability and outreach performance indicators, however, adequate time series data is not available from many of the MFIs. Only from 13 MFIs which have a six year data (from 2003-2008), 78 MFI-year observations are used as a sample for the study. All the data is taken from the MIX database. For conformity with the study period of 2003-2008, benchmark figures from the MIX are taken for the years 2004-2008 as data available for 2003.

The data collected is rather unbalanced because while many mature MFIs reported the full six year data, a few new MFIs have only a two year data. But the average number of years for the whole sample is around 5.3, making it a reasonably balanced data with less missing observations. The final data used for analysis, after ignoring missing values, is 53 MFI-year observations with the average number of years per MFI being 5.3 while the number of MFIs is 10.

3.2 Methodology

At the outset, MFIs are classified based on scale of outreach as the sample shows a huge variation on such variable as opposed to other metrics. According to MIX (2008) MFIs outside the Latin American and Caribbean region, to which Ethiopia is a part, are classified as follows. Small MFIs are those below $2 million Gross Loan Portfolio (GLP), medium MFIs are those having a GLP between$2 and $8 million and large MFIs are those with a GLP of above $8 million. The average GLP figure over the sample period of 2003-2008 is used for classification purpose and this is different from Ejigu (2009) who used only a single year (2007). Using average rather than single year GLP helps to capture the dynamics in the industry.

Three different approaches are used to fulfill the objective of the study. One sample tests are used to check the presence of any significant difference between the performances of MFIs with the MIX benchmark. For local comparison, ANOVA (simple or robust)[8] and Kruskal Wallistest[9], as appropriate, with post hoc tests (Scheffe or Games-Howell)[10] are used. Lastly, graphs in level and growth forms are used for trend analysis.

Measures

There are many measures of sustainability. These include Subsidy Dependence Index (SDI), self sufficiency measures, adjusted profitability ratios, arrears rate and efficiency and productivity. But from these the paper used self sufficiency measures of OSS and

[8]One sample t test and ANOVA are parametric test of difference used when the critical assumptions of normality and homogeneity of variance are fulfilled. If such assumptions don't hold, it is better to transform the data, use tests that are robust to the violation of such assumptions (e.g. robust ANOVA) or resort to non-parametric tests that don't require the fulfillment of these assumptions (Field, 2005).

[9]Kruskal-Wallis test is a non-parametric test used when the ANOVA can't be implemented.

[10]Scheffe post-hoc test is used if the homogeneity of variance assumption is fulfilled whereas the Games-Howell procedure is used in cased whereas such assumption is violated.

adjusted profitability measure of ROA as they are heavily used in prior literature (Barres, 2006). Financial Self Sufficiency (FSS) is another widely used measure, but such data is not available from MIX database.[11]

For breadth of outreach, the number of borrowers is used as a measure. For depth of outreach two measures are used: average loan size (measured as the ratio of GLP to number of borrowers) and the percentage of women borrowers (measured as the ratio of women borrowers to total borrowers). Less average loan size and more percentage of women borrowers served as an indicator of good depth of outreach.

Disbursing small loans is perceived as reaching the poorest as this group of people, due to their small business activity, are interested in small loans. For international comparison, average loan size adjusted for Gross National Income (GNI) per capita is used. The measure of average loan size is subject to criticism. For instance, Christen (1997) argues that due to the heterogeneity of loan products in terms of maturity and purpose, its uses can be misleading. Thus, it may not reflect the target market and the clients' poverty level.

Brau & Woller (2004) argue that women are preferred in the business of microfinance due to three reasons. Firstly, women invest the loans in productive activities or in improving family welfare more often than men, who are assumed to consume rather than invest loan funds. Secondly, access to financial services empowers women, both financially and socially. Thirdly, women are considered poorer than men. The often expressed fear in lending to women is that they bring about less economic growth than men. However, a study by Kevane and Wydick (2001), cited in Brau & Woller, (2004) reveals that there is no significant difference between men and women in generating business sales. Finally, it has to be underscored that all the measures used are consistent with prior literature such as Cull et al, (2007), Mersland & Storm, (2007), Okumu, (2007).

[11]OSS is calculated as [(financial revenue)/ (financial expense + net loan loss provision + operating expense)]. ROA is measured as the ratio of adjusted net operating income net of taxes divided by adjusted average total assets. OSS is not adjusted for subsidy whereas FSS and ROA are adjusted for subsidy. (Ledgerwood, 1999).

3.3 Descriptive Statistics

Table 1 presents descriptive statistics and mean growth rate of sustainability and outreach indicators for the industry as a whole. The industry served an average of 123,034 borrowers which is considered as a fairly good outreach. The coefficient of variation figure is very high indicating a huge difference between the MFIs in terms of number of borrowers reached. The industry also registered the highest growth rate of 22.24% in number of borrowers which is expected from a young industry. The average loan size is $128 (equivalent to Br. 2048 assuming an average exchange rate of Br 16 to the dollar over the six year period of 2003-2008). Half of the MFIs clients are women. This is poor performance as MFIs are presumed to target mostly women. The coefficient of variation of the percentage of women borrowers is low indicating similar women targeting policy between MFIs. The industry is operationally self-sustainable (OSS, 1.3) whereas ROA is low (1%) indicating that it is not profitable like mainstream banks and other commercial businesses. The high coefficient of variation of ROA also indicates a huge difference between the various MFIs profitability rates. The descriptive statistics results indicate at a glance that the MFIs are better in sustainability than in outreach. This is consistent with the empirical findings reported earlier and the overall commercialization trend observed in the global microfinance industry.

4. Results

4.1 Correlation Results

The correlation coefficients between sustainability and outreach indicators are depicted in Table 2. The results show that as MFIs serve a large number of borrowers, they ignore the poorest of the poor. As average loan size increases the percentage of women borrowers served decreases. As expected, serving more borrowers leads to achieve sustainability due to economies of scale. Sustainability and depth of outreach are found to be negatively related, i.e. the former has positive relationships with average loan size and negative relationships with the percentage of women borrowers served. All these results indicate there is a tradeoff between serving the poorest and being financially self-sufficient. Consistent with Ejigu (2009), Mersland and Storm (2007) and Okumu (2007), reaffirm the fear of 'welfarist' approach that the commercialization of microfinance will bring death to the industry's original mission of serving the poorest of the poor.

Surprisingly the two measures of depth of outreach, average loan size and the percentage of women borrowers are not significantly negatively correlated. This partially reaffirms Christen's (1997) fears that the use of average loan size as depth of outreach measure may be misleading due to the heterogeneity of loan products in terms of maturity and purpose. Hence more studies are called for using clear theoretical frameworks. As expected, the two measures of sustainability (OSS and ROA) are positively related implying they are measures of the same construct.

In sum, scale of outreach seems to matter as the large MFIs are better in breadth of outreach and sustainability and the small MFIs in depth of outreach indicators. Two big clusters are observed from the analysis. Small MFIs have a social mission of reaching the poorest of the poor, but they are unsustainable whereas the large MFIs have a commercial mission of reaching large number of possibly marginally poor borrowers in a sustainable way. A natural question to ask is which MFI should be chosen? This depends on the ethical school of thought one subscribes to and an assessment of the country's economic conditions and poverty level. Using utilitarian theory of benefits to the mass, the large MFIs are preferred whereas if one believes in moral relativism, the small MFIs may be chosen[12]. Scanning the Ethiopian economy roughly indicates that it is a country of a large number of marginally poor people and the poorest are judged to be small. Using such line of reasoning, opting for the large MFIs makes more sense.

a. Comparison with MIX Benchmarks

Table 3 presents the one sample test results that compare the sustainability and outreach[13] performance of each category of MFIs with their respective MIX benchmarks. The result indicates that small MFIs are not good in either sustainability or depth of outreach compared to MIX. They are only good in breadth of outreach. Medium MFIs are at par in terms of sustainability and depth of outreach and below MIX in breadth of outreach. The large MFIs are better in sustainability and breadth of outreach and below MIX in depth of outreach. The overall MFIs industry is at par in sustainability, better in breadth of outreach and below MIX in terms of depth of outreach.

[12]For the various ethical theories, see a book by Jennings (2009).

[13]At some places the two measures of depth of outreach (average loan size and the percentage of women borrowers) show contradictory result due to the no correlation between them. The result based on the percentage of women borrowers is heavily relied upon in many cases due to criticism in using average loan size as poverty measure.

In general the international comparison shows mixed results. Scale of outreach seems to matter in sustainability as the large MFIs have better sustainability as compared to their MIX benchmark than small MFIs. But in outreach, it does not seem to matter. Both the large and small MFIs beat their MIX benchmarks in breadth of outreach while they are beaten in depth of outreach.

In the earlier section, the correlation result shows that the large MFIs are preferred in the local context for sustainability and breadth of outreach and the small MFIs for depth of outreach. There was a clear demarcation between the two. In the international comparison, the large MFIs are preferred again for sustainability and breadth of outreach. But the small MFIs failed to be chosen for excellence in depth of outreach which weakens the tradeoff result found earlier. Thus, the overall result shows that large MFIs are deemed better and hence using utilitarian theory, preference should be given to them.

The fact that Ethiopian MFIs surpass the global average in breadth of outreach should not give a wrong signal. If breadth of outreach performance was compared with the potential demand for microfinance service (poverty level) in each respective region, possibly, Ethiopian MFIs could lose as the poverty level is high. So, more research using a variant of relative benchmark is called for in the future. The low performance in depth of outreach is consistent with the findings of Kereta (2007), Ejigu (2009) and Lafourcade et al (2005) who showed that African MFIs fall behind in terms of women borrowers as compared to other regions of the world.

4.3 Local Comparison across the Scale of Outreach

Table 4 presents the results of the local comparison of MFIs (classified by scale of outreach).The results show natural ordering of MFIs in terms of number of borrowers, average loan size and sustainability indicators but reverse ordering in the percentage of women borrowers. This means that the larger MFIs have a large number of borrowers, average loan size and OSS while they serve a small percentage of women borrowers as compared to small MFIs. Such a result is the same as the correlation result which shows that the small MFIs are better in depth of outreach indicators and the large MFIs in sustainability and breadth of outreach. This again supports the tradeoff literature that documented serving the poorest and being financially self-sufficient are contradictory goals.

4.4 Trend Analysis for the Microfinance Industry

The trend in sustainability and outreach performance are shown in level and growth forms in Graphs 4.1 up to 4.10. Beginning with the borrowers' graph, the linear trend shows an increasing pattern indicating a large number of borrowers are served over the years. This indicates heavy growth emphasis given to such variable which is expected in a young industry. Looking at the growth rate, it shows a declining trend in recent periods. The major constraint to push for large number of borrowers may be lack of finance. This might have been created by, among others factors, the global financial crises that hit many industries recently which reduced donor support to humanitarian causes. Secondly, the demand for microfinance credit may decrease if the estimated poverty impact is low. Bateman (2010) cites some recent global studies that reveal such hard facts.

On average loan size, both the linear and growth rate graph show increasing trend. Increase in average loan size may or may not be a worrisome signal. It will be a mission drift only if loan size to new clients has grown over the years whereas increase in loan size to old clients is considered as progressive, lending after successful repayment (Cull et al, 2007). No detailed data is available to confirm which proposition is at work in Ethiopia and hence further inquiry is needed. The other complicated issue is that the real mission[14] of different MFIs may not be the same. Some may follow a purely commercial strategy, others purely social goals and still others (perhaps the majority of MFIs) a mix of these two. Increase in average loan size may be a mission drift for the purely social oriented MFIs, an optimal strategy for MFIs pursuing both social and commercial goals and a best strategy for the purely commercial oriented MFIs. Hence more detailed and micro level analysis should be done on this issue by future researchers.

For the percentage of women borrowers, both the linear and growth rate graphs show erratic patterns. This implies there is no clear woman targeting policy and targeting seems to be done haphazardly. The OSS graph shows it is stable and erratic in a few periods which may not be worrisome. ROA shows erratic trends which are worrisome.

[14]Real mission is conceived as the one at the heart of owners and managers. Nominal missions may be those that are posted in brochures, banners and found at public relation offices. Sometimes, these two missions may be divergent for some MFIs.

This may be due to erratic income of clients (agricultural income) and unplanned investment decisions in assets.

5. Concluding Remarks

This study attempted to assess the sustainability and outreach performance of Ethiopian MFIs by comparing them with benchmarks like MIX averages, local benchmarks and through time. The study extended the earlier studies by Kereta (2007) and Ejigu (2009) in a more advanced way.

From the correlation analysis and local comparison, the small MFIs are better in depth of outreach indicators and the large MFIs in sustainability and breadth of outreach. The result indicates many things. First, it shows that scale of outreach matters for performance of MFIs. Second, the market shows a cluster of social and commercial oriented MFIs which calls for different strategies and modes of operation. Third, due to such clustering, the MFIs can not attain the double objective of reaching the poorest and being sustainable simultaneously. It seems they have to make a choice. Using utilitarian ethical school of thought, the large MFIs that serve the larger marginally poor people have to be promoted as opposed to small MFIs that strive to reach the poorest of the poor. The promotion of large MFIs includes inter alia increasing the minimum capital required to establish microfinance business, mobilization of saving and commercial debt source of finance, merger with small MFIs and other measures. If the small MFIs can not be scrapped altogether, separate regulatory frameworks should be introduced for such MFIs. Surprisingly, the two measures of depth of outreach, average loans size and the percentage of women borrowers, are not significantly correlated which needs further inquiry. As expected the two measures of sustainability, OSS and ROA, are significantly and positively correlated.

With the international MIX comparison, the result is somewhat mixed and contradicts with the correlation and local comparison results. Scale of outreach seems to matter for sustainability only and not for outreach. This also further invalidates the tradeoff that is expected between serving the poorest and being financially self-sufficient. More research is also required to clear such contradiction. Ethiopian MFIs surpass the global average in breadth of outreach and slack in depth of outreach. Surpassing in the absolute number of borrowers, however, should not give a wrong signal. If performance in breadth of outreach is measured against the potential demand for MFIs, they fall behind. The weakness in depth of outreach indicates the commercial

oriented nature of the sector which constantly challenges the viability and continuity of small social oriented MFIs.

The trend analysis indicates the industry has shown impressive growth in number of borrowers. Average loan size increased over the years, women targeting and ROA are erratic while, OSS is stable. From this result, more micro-level studies are called for to explain the following:

- The growth in average loan size, specifying whether the increase refers to new or old clients,
- The mission of MFIs as some have pure social mission, others pure commercial and the majority a mix of the two,
- The erratic nature of women targeting at least by the social oriented MFIs.

The commercial MFIs have also to tackle the erratic ROA by diversifying income sources and making planned investment decisions.

References

Aghion, A. B. and Morduch, J. (2005). *The Economics of Microfinance*. Cambridge, MA: MIT Press.

Amha Wolday. (2007). Managing growth of microfinance institutions: Balancing Sustainability and reaching large number of clients in Ethiopia. AEMFI Occasional paper No 18, Addis Ababa, Ethiopia.

Barres, I. (2006). Financial Self Sufficiency (FSS).The Micro banking Bulletin.13:21-22, Microfinance Information Exchange Market Inc: Washington DC

Bateman M. (2010). *Why Doesn't Microfinance Work? The destructive rise of local neoliberalism*. Zed Books Ltd, London.

Brau J. C., Woller G. (2004). Microfinance a comprehensive review of existing literature. *Journal of Entrepreneurial Finance and Business Ventures* 9(1):1-26.

Christen, R. P. (1997). *Banking Services for the poor: Managing for financial success*. ACCION International, Washington DC.

Congo, Y. (2002). Performance of Microfinance Institutions in Burkina Faso. Discussion Paper No. 2002/01, United Nations University, WIDER

Conning, J. (1999). Outreach, Sustainability and Leverage in monitored and peer-monitored lending. *Journal of Development Economics*, 60, 51-77

Cull, R.; DemigÄuc-Kunt, A., and Morduch, J. (2007). Financial performance and outreach: A global analysis of leading microbanks. *Economic Journal.* 117(517): 107-133.

Ejigu, L. (2009). Performance analysis of sample microfinance institutions of Ethiopia. *International NGO Journal*, 4 (5): 287-298.

Field, A. (2005). *Discovering Statistics using SPSS*. second edition, Sage Publication

Gurgand, M.; G. Pederson, and Yaron, J. (1996). Rural finance institutions in Sub-Saharan Africa: Their outreach and sustainability. *Savings and Development*, 20 (2): 135–69.

Hartarska, V. (2005). Governance and performance of microfinance institutions in Central and Eastern Europe and the newly independent states. *World Development*. 33: 1627-1648.

Hashemi, S. and Rosenberg, R. (2006). Graduating the poorest in to microfinance: Linking safety nets and financial services. CGAP Focus Note 34.

Hishigsuren, G. (2004). Scaling up and mission drift: Can microfinance institutions maintain a poverty alleviation mission while scaling up? (Unpublished PhD thesis) Southern New Hampshire University.

Jennings, M. M. (2009). *Business Ethics Case Studies and Selected Readings*. 6th edition, South Western Cengage Learning.

Kereta, B. (2007). Outreach and financial performance analysis of microfinance institutions in Ethiopia. African Economic Conference, Addis Ababa.

Kevane, M. & Wydick B. (2001). Microenterprise Lending to Female Entrepreneurs: Sacrificing Economic Growth for Poverty Alleviation? *World Development* 29, 1225-1236.

Kidane, A. (2007). Outreach and Sustainability of the Amhara Credit and Saving Institutions (ACSI) Ethiopia. (Unpublished MA thesis), Norwegian University of Life Sciences.

Lafourcade, A. L.; Isern J.; Mwangi P., and Brown, M. (2005). Overview of the outreach and Financial Sustainability and outreach of Microfinance Institutions in Africa. The MIX Market Inc.

Ledgerwood, J. (1999). *Sustainable Banking with the Poor Microfinance Handbook: An Institutional and Financial Perspective*. The World Bank, Washington, DC.

Marzys, O. R. (2006). Microfinance Institutions: Profitability at the Service of Outreach? A Study of the Microfinance Industry in the ECA Region. (Unpublished MA thesis).

Mersland, R. and Strom, O. (2010). Microfinance Mission Drift. *World Development*, 38 (1), 617-629

_____. (2007). *Sustainability and outreach and corporate governance in Microfinance Institutions*. Agder University, Norway

Microfinance Information eXchange (MIX). (2008). Benchmark reports. Washington DC.

Okumu, L. J. (2007).The Microfinance industry in Uganda: Sustainability, Outreach and Regulation. (Unpublished PhD Dissertation) University of Stellenbosch, South Africa

Pfister, M. W., Gesesse D., Amha Wolday, Mommartz, R., Duflos W., Steel E. (2008). *Access to finance in Ethiopia: Sector Assessment study*.2, GTZ.

Sengupta, R. and Aubuchon, C. P. (2008). The Microfinance Revolution: An overview. *Federal Reserve Bank of St. Louis Review*, 90(1), 9-30

Schreiner, M. (1999). Aspects of outreach: a framework for the discussion of the social benefit of Microfinance. Microfinance Risk Management, St. Louis

Woller, G. M. and Schreiner M. (2006).Poverty lending, financial self sufficiency and the six aspects of outreach. SEEP Network Publication, New York

Appendix

Table 1: Descriptive Statistics

Variables	N	Mean	SD	Min	Max	Coefficient of Variation	Mean growth rates (2003-2008) (%)
Borrowers	60	123034	18374	434	710576	14.9	22.24
AvLnSz	60	128	64.46	32	314	0.5	16.79
Women	57	0.53	0.2	0.14	0.93	0.37	3.35
OSS	58	1.3	0.54	0.15	2.3	0.41	5.07
ROA	54	0.01	0.08	-0.4	0.1	8.8	-15.1

Table 2: Correlations among sustainability and outreach indicators

Variables	Borrowers	Women	AvLnSz	ROA	OSS
Borrowers	1				
Women	-0.48*	1			
AvLnSz	0.5*	-0.17	1		
ROA	0.34*	-0.29*	0.11	1	
OSS	0.66*	-0.38*	0.36*	0.8*	1

*values significant at 5%

Table 3: Comparison with MIX benchmark: One sample t test

Scale of outreach	Variable	N	Mean	SE	MIX Benchmark (2004-2008 average)	T	P
Small	LnBorrower	26	11904	0.22	5332	2.5	0.01*
	Women	25	0.6	0.03	0.69	-2.01	0.05*
	LnAvLnSzGNI	26	0.47	0.07	0.47	0.203	0.84
	OSS	26	0.95	0.08	1.11	-2.14	0.04*
	ROA	Left					
Medium	Borrower	12	16734	1934	18132	-1.1	0.29
	Women	11	0.66	0.03	0.67	0.14	0.88
	AvLnSzGNI	12	0.72	0.03	0.72	0.3	0.76
	OSS	12	1.27	0.1	1.18	1.09	0.29
	ROA	Left					
Large	Borrower	22	311068	39639	185503	3.2	0.004*
	Women	21	0.37	0.03	0.58	-5.9	0.000*
	LnAvLnSzGNI	22	0.93	0.05	1.11	-2.6	0.01*
	OSS	20	1.8	0.07	1.25	7.07	0.000*
	ROA	18	0.05	0.005	0.01	7.31	0.000*
Whole MFIs	LnBorrower	60	123,034	183,746	67,908	15	0.000*
	LnWomen	57	0.53	0.2	0.65	-4.9	0.000*
	LnAvLnSzGNI	60	0.7	0.3	0.76	-2.97	0.004*
	OSS	58	1.3	0.54	1.18	1.77	0.08
	ROA	Left					

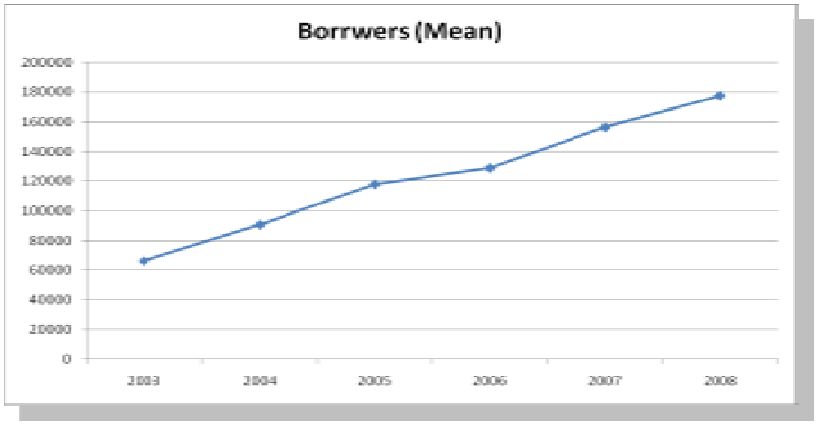
*values significant at 5%

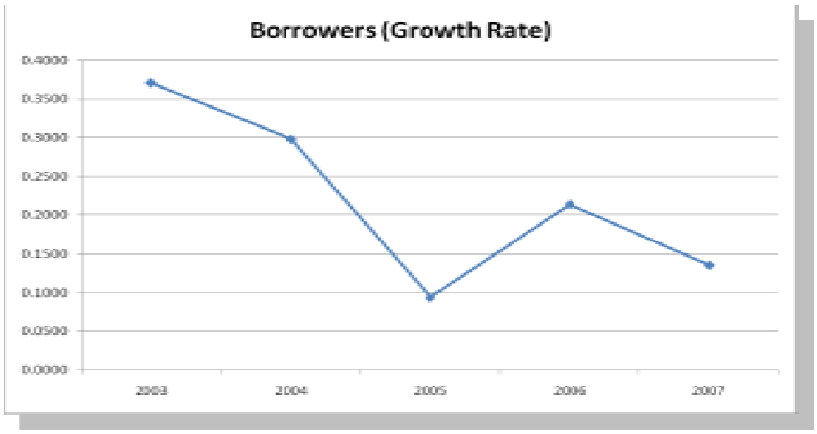
Table 4: Comparison across each category of the MFIs: ANOVA

Variable	Scale of outreach	Mean	SD	F/	P	Post Hoc tests Games-Howell/Scheffe/	P
LnBorrower	Small	11904	1.13	94.01	0.000*	Small vs. medium	0.057*
	Medium	15987	0.48			Small vs. large	0.000*
	Large	313644	0.83			Medium vs. large	0.000*
LnAvLnSz	Small	91	0.49	19.87	0.000*	Small vs. medium	0.006*
	Medium	131	0.23			Small vs. large	0.000*
	Large	173	0.35			Medium vs. large	0.219*
LnWomen	Small	0.6	0.4	14.78	0.000*	Small vs. medium	0.604
	Medium	0.66	0.16			Small vs. large	0.000*
	Large	0.37	0.44			Medium vs. large	0.000*
OSS	Small	0.95	0.42	28.14	0.000*	Small vs. medium	0.03*
	Medium	1.27	0.36			Small vs. large	0.000*
	Large	1.8	0.34			Medium vs. large	0.003*
ROA	Small	-0.032		5.918	0.005*	Small vs. medium	0.04*
	Medium	0.034				Small vs. large	0.002*
	Large	0.05				Medium vs. large	0.15

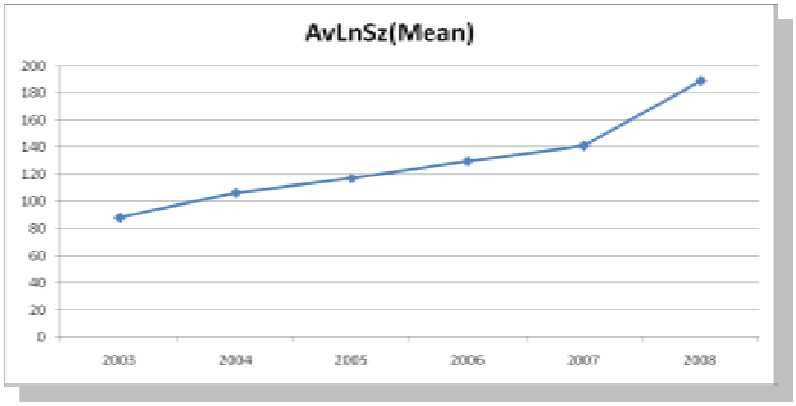
*values significant at 5%

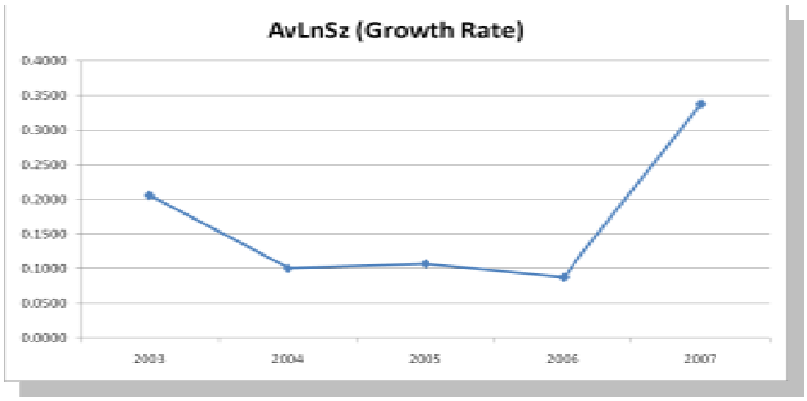
Graph 4.1 & 4.2: The mean number of borrowers and the growth trend.



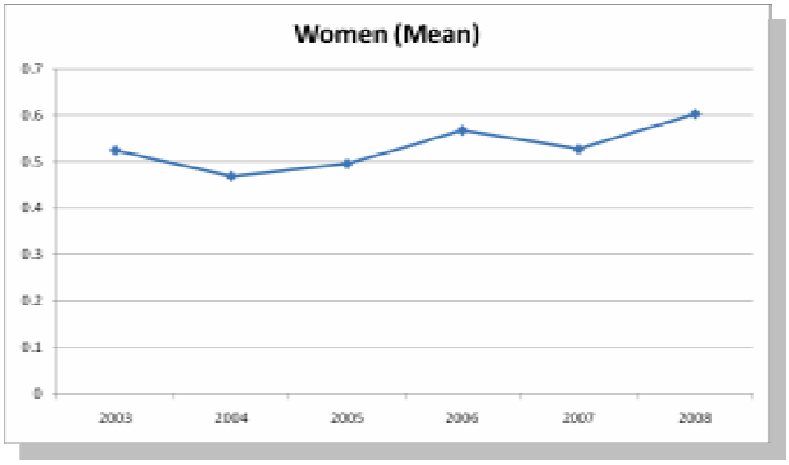


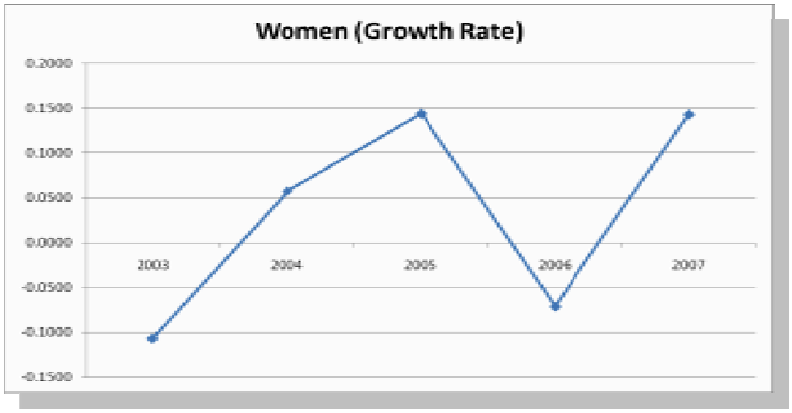
Graph 4.3 & 4.4: The mean average loan size and its growth trend.



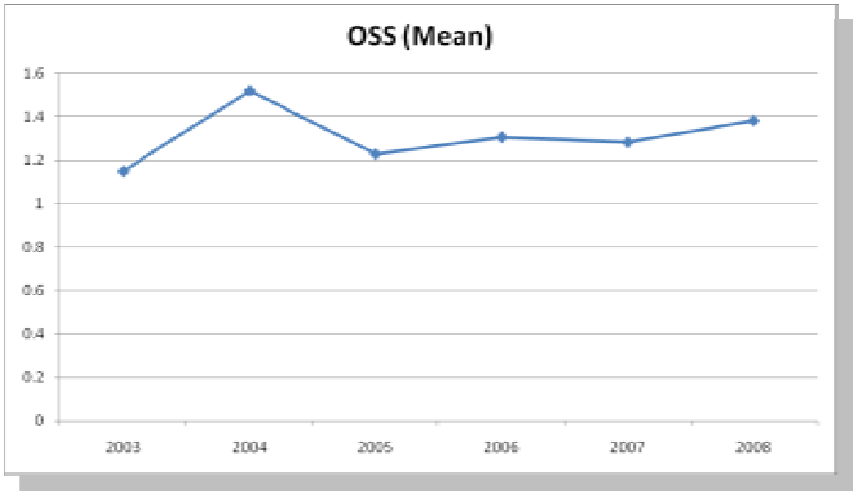


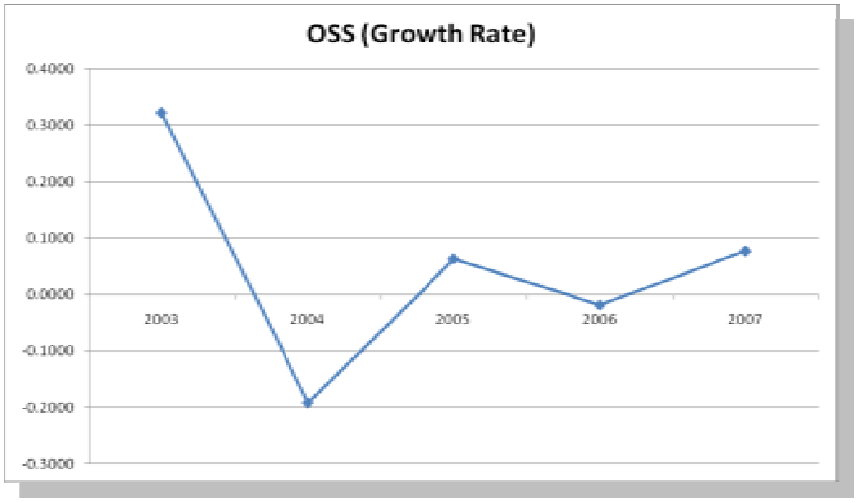
Graph 4.5 & 4.6: The mean percentage of women borrowers and its growth trend.



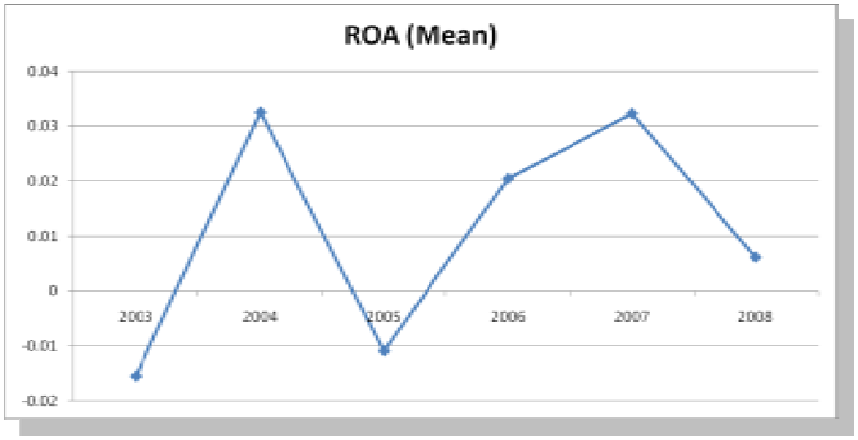


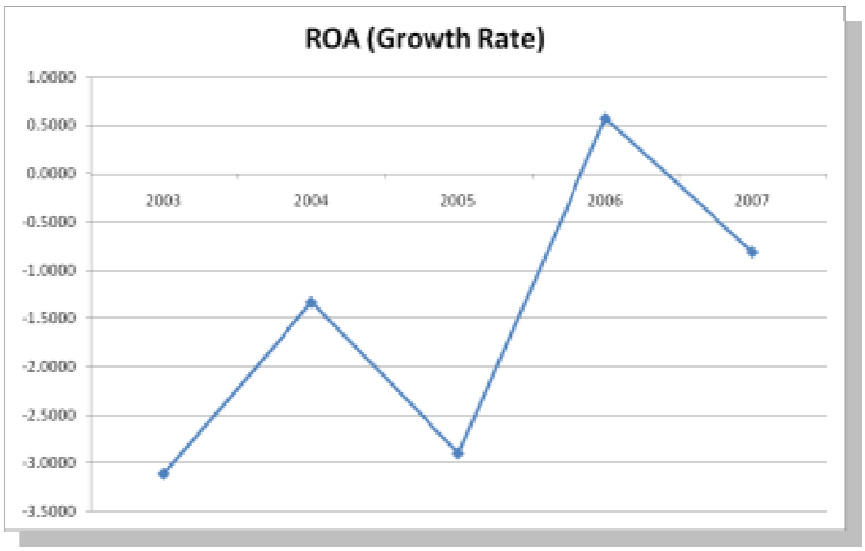
Graph 4.7 & 4.8: The mean OSS and its growth trend.





Graph 4.9 & 4.10: The mean ROA and its growth trend.





ESTIMATING CONSUMPTION-BASED POVERTY IN THE ETHIOPIA DEMOGRAPHIC AND HEALTH SURVEY[1]

Mark Schreiner[2]

Abstract

Inequalities in health are linked with poverty, but quantifying the health/poverty nexus is hampered by data constraints. In particular, the most common measure of poverty compares consumption with poverty lines, but consumption surveys often do not collect detailed health data. Conversely, the large repository of internationally-comparable Demographic and Health Surveys has detailed health data but no consumption data. This has led DHS researchers who want to control for socio-economic status use an asset index defined in terms of housing characteristics and the ownership of durable goods. While this is a valid conception of poverty, it is difficult to compare it with the more-common consumption-based measure. This paper presents a simple poverty scorecard for Ethiopia based on the poverty-mapping approach of Elbers et al. (2003). It allows researchers to estimate the likelihood that consumption is below a given poverty line using nine verifiable, inexpensive-to-collect indicators found in both Ethiopia's 2005 DHS and in the 2004/5 Household Income, Consumption, and Expenditure Survey. It turns out that the poverty scorecard and the DHS asset index do not generally rank people the same, so estimates of consumption-based poverty in the DHS should use the poverty scorecard, not the DHS asset index. The bias and precision of scorecard estimates compare well with that of other tools, suggesting that government could use it to track poverty in years between national household expenditure surveys.

Key words: Poverty measurement, asset index, poverty mapping, Africa, Ethiopia, health equity

---

[1] The final version of this article was submitted in February 07, 2013.

[2] Mark Schreiner directs Microfinance Risk Management, L.L.C., microfinance.com. He is also a Senior Scholar at the Center for Social Development in Washington University in Saint Louis.

Acknowledgements

HICE data come from Ethiopia's Central Statistical Agency; DHS data come from IFC Macro. Thanks go to Karen Foreit, Kiersten Johnson, Rachel Sanders, and Bridgette Wellington. This work was supported by MEASURE Evaluation PRH, which is funded by the U.S. Agency for International Development (USAID) through cooperative agreement associate award number GPO-A-00-09-00003-00 and is implemented by the Carolina Population Center at the University of North Carolina at Chapel Hill, in partnership with Futures Group, Management Sciences for Health, and Tulane University. The opinions expressed are those of the author and do not necessarily reflect the views of neither USAID nor the U.S. government.

1. Introduction

This paper applies a simple poverty scorecard (Table 1) to Ethiopia's 2005 Demographic and Health Survey (DHS) to estimate the likelihood that a given person has consumption below a given poverty line. This allows researchers to look at how health outcomes vary with socio-economic status as seen through a consumption lens.

Table 1: Simple scorecard for estimating consumption-based poverty in the Ethiopia DHS

Entity	Name	ID	Date (DD/MM/YY)
Client:			Joined:
Region:			Today:
Service point:			HH size:

Indicator	Value	Points
1. How many people usually live with the household?	A. Nine or more	0
	B. Eight	2
	C. Seven	6
	D. Six	9
	E. Five	14
	F. Four	21
	G. Three	27
	H. One or two	42
2. What is the highest grade the female head/spouse has completed?	A. Four or less	0
	B. No female head/spouse	0
	C. Five or six	4
	D. Seven to nine	10
	E. Ten or higher	13
3. What is the main material of the walls of the residence?	A. Stone with mud, stone with lime/cement, or cane/trunks/bamboo/reed	0
	B. Bamboo/wood, uncovered adobe, plywood, carton, no walls, or other	5
	C. Cement, bricks, cement blocks, covered adobe, or wood planks/shingles	11
4. What type of toilet facility do members of your household usually use?	A. Non-flush or none	0
	B. Flush	5
5. What type of fuel does your household mainly use for cooking?	A. Wood or straw/shrubs/grass, or animal dung	0
	B. All others	5
6. Does the household have a bed?		0
		5
7. Does the household have a radio?		0
		8
8. Does any member of this household own any land that can be used for agriculture?		0
		6
9. Does the household own any cattle, sheep or goats?	A. No	0
	B. Yes	5

Total score:

Such a consumption-based poverty measure is useful because while the DHS surveys are the largest repository of nationally representative data on population, health, HIV, and nutrition (covering more than 75 countries, often for multiple rounds), they do not collect data on consumption. DHS researchers seeking to relate health outcomes to socio-economic status have had to rely on an asset index that in recent years comes pre-packaged with DHS data (Rutstein and Johnson, 2004).

The DHS asset index is widely used. Constructed with Principal Components Analysis (PCA), it defines socio-economic status in terms of housing characteristics, asset ownership, and agricultural employment. PCA does not explicitly model any particular conception of poverty; rather, it finds the linear combination that maximizes the explained variation among a given set of indicators. Nevertheless, the resulting DHS asset indexes seem to be related to socio-economic status, especially when this is conceived as "permanent income" or "expected long-term control over resources". The indexes are correlated in intuitive ways with outcomes such as fertility (Bollen, Glanville, and Stecklov, 2007), use of emergency obstetric care (Pitchforth et al., 2007), maternal and child mortality (Knowles et al., 2008), food security (Dekker, 2006, for Ethiopia), child health and nutrition (Sahn and Stifel, 2003), and education (Filmer and Pritchett, 2001).

Table 2 is Ethiopia's 2000 DHS asset index (Gwatkin et al., 2007). It has 23 indicators and 113-point values. The DHS index ranks people on a relative scale; a higher value of the index implies higher socio-economic status or lower poverty.[3]

The contribution of this paper is to allow DHS analysis in terms of consumption-based poverty: a person is poor if the monetized value of his/her per-capita household consumption is below a poverty line such as the Millennium Development Goals' $1.25/day at 2005 purchase power parity (PPP). Consumption-based poverty lines are commonly used by governments, the World Bank, policymakers, and others.

[3] A PCA asset index may be seen as a measure of absolute poverty as defined by its indicators and points, and as such, it can be used to measure poverty over time and across countries (Booysen et al., 2008; Sahn and Stifel, 2000). Most applications, however, treat the index as a relative measure.

Table 2: DHS asset index for Ethiopia (2000)

Question	Score if "Yes"	Score if "No"	Item score
1. In your household is/are there . . .?			
Electricity	0.16178	-0.04632	
One or more radios	0.10087	-0.03931	
One or more televisions	0.23657	-0.01390	
One or more bicycles	0.17138	-0.00198	
One or more motorcycles. Scooters	0.18746	-0.00005	
One or more cars, trucks	0.26830	-0.00248	
One or more telephones	0.24991	-0.00911	
One or more electric mitads	0.23539	-0.01491	
One or more kerosene lamps, pressure lamps	0.00624	-0.00124	
One or more beds, tables	0.04349	-0.04492	
One or more cattles, camels	-0.05021	0.05846	
One or more horses, mules, donkeys	-0.05443	0.01551	
One or more sheep, goats	-0.04996	0.02424	
2. Do any of the members of your household own . . .?			
A house	-0.03046	0.12970	
Crop land	-0.05011	0.10455	
Cash crops	-0.05072	0.01277	
3. Do the members of your household work their own or family's agricultural land?	-0.05908	0.02255	
4. What is the principal source of drinking water for your household?			
Piped water in dwelling	0.20251	-0.00039	
Piped water outside dwelling	0.20125	-0.02419	
Piped water in public faucet	0.07932	-0.01497	
Well in dwelling	-0.04956	0.00298	
Covered well	-0.02492	0.00144	
River, canal, surface water	-0.04904	0.01974	
Open spring	-0.05232	0.02165	
Covered spring	-0.02578	0.00098	
Rainwater	-0.06333	0.00012	
Other	0.01952	-0.00005	
5. What is the principal type of fuel for cooking use by your household?			
Electricity	0.23445	-0.00037	
Gas	0.26576	-0.00040	
Biogas	0.15508	-0.00020	
Kerosene	0.19738	-0.02609	
Charcoal	0.13642	-0.00459	
Wood	-0.03490	0.09703	
Dung, manure	-0.02591	0.00318	
Other	0.02881	-0.00001	
6. What is the principal type of toilet facility used by your household?			
Private flush toilet	0.25570	-0.00235	
Private latrine	0.11641	-0.03987	
VIP latrine	0.19368	-0.00174	
Bush, field as latrine	-0.04642	0.12343	
7. What is the principal material used for the floors in your household?			
Dirt, sand, dung	-0.03345	0.17974	
Cement	0.18049	-0.01260	
Wood plank	0.20963	-0.00078	
Parquet, polished wood	0.24016	-0.00250	
Bamboo, reed	0.12413	-0.00067	
Vinyl, asphalt	0.20579	-0.00944	
Tiles	0.22425	-0.00156	
Carpet	0.06956	-0.00121	
Other	0.18014	-0.00076	
8. What is the principal material used for the roof of your household?			
Corrugated iron	-0.03323	0.00290	
Cement, concrete	-0.05660	0.06001	
Wood and mud	-0.04760	0.00024	
Thatch	0.04694	-0.00004	
Bamboo, reed	-0.05413	0.00437	
Plastic sheet	0.11469	-0.05285	
Mobile roof of nomads	0.11469	-0.05285	
Other	-0.05212	0.00037	
9. How many people are there for each sleeping room in your household?	[(# people-3.89)/ 2.14] x -0.029		
Total household asset score (sum of individual score items):			

Source: Gwatkin et al., 2007

While the consumption-based conception of poverty is not more valid than the asset-based conception, it is more commonly used and better understood.[4] Thus, DHS research might be more relevant for policy if it compared health outcomes not only with an asset index but also with consumption-based poverty. The rub is that collecting consumption data is complex and costly (Sahn and Stifel, 2003; Deaton and Zaidi, 2002). For the example of Ethiopia's 2004/5 Household Income, Consumption, and Expenditure Survey (HICE), enumerators visited each household 16 times, asking each time about a lengthy list of consumption items. The cost of doing this explains why the DHS does not collect data on consumption.

The scorecard in this paper allows researchers to estimate consumption-based poverty in Ethiopia's 2005 DHS. The estimates come from a three-step method that follows that of poverty mapping (Elbers et al., 2003). First, potential poverty indicators are matched between a survey that collects consumption data (Ethiopia's 2004/5 HICE) and another survey that does not collect consumption data (Ethiopia's 2005 DHS). Second, a poverty scorecard is constructed based on the 2004/5 HICE, using only indicators that appear in both the 2004/5 HICE and in the 2005 DHS. Third, the scorecard is applied to the 2005 DHS to estimate consumption-based poverty.

This poverty-mapping approach rests on three difficult-to-test assumptions. The first is that scorecard indicators are well-matched across the two surveys, so that all else constant, households are as likely—for example—to report owning a radio in the 2004/5 HICE as in the 2005 DHS. Unfortunately, the strength of the match is never certain, as questions may be worded differently, offer different response options, or appear in a different context.

The second strong assumption is that the relationships between indicators and poverty are constant over time (Christiaensen et al., 2010). This is plausible for the Ethiopia data analyzed here, as the 2004/5 HICE covered July 2004 and February 2005, and the 2005 DHS covered April to August 2005. The assumption is less plausible for longer time gaps and in periods of sharper socio-economic change.

[4] Arguments in favor of the asset-based view include Carter and Barrett (2006), Schreiner and Sherraden (2006), and Sherraden (1991). In practice, the two views, though distinct, are tightly linked, as income/consumption are flows of resources received/consumed from the use of stocks of assets (wealth). The two views are low-dimensional simplifications—due to practical limits on definitions and measurement—of a higher-dimensional and more complete conception of the production of human well-being. Section 6 below discusses the correlation between health outcomes, asset indexes, and consumption-based poverty likelihoods.

The third strong assumption is that the scorecard is applied to nationally representative groups (Elbers, Lanjouw, and Leite, 2008; Tarozzi and Deaton, 2007). This holds for Ethiopia's 2005 DHS.

If all three assumptions hold, then scorecard estimates of consumption-poverty rates are unbiased, that is, their average in repeated samples matches the true rate.

The scorecard can be used to estimate the poverty likelihood of a person responding to a survey. It can also be used to estimate the poverty rate of a group of people, such as those who use public-sector health services or women whose most recent childbirth was not attended by trained personnel.

The results here suggest that the approach here has three main possible uses. First, researchers can the poverty scorecard use as a substitute for (or as a complement to) the DHS asset index when relating socio-economic status to health outcomes. Second, local government and pro-poor organizations can use the scorecard to inform their social-performance management by estimating head-count poverty rates for small regions in which the DHS has a larger sample than does a national expenditure survey. Third, national policy-makers can use the scorecard to monitor poverty in years between the less-frequent and more-costly national household expenditure surveys.[5]

The next section documents data, poverty lines, and indicator-matching. Section 3 describes scorecard construction. Section 4 defines the concept of poverty likelihood and details the estimation of consumption-based poverty rates. Section 5 compares the poverty scorecard with the pre-packaged DHS asset index in terms ranking people in Ethiopia's 2005 DHS. Section 6 places the scorecard in the context of related work. The final section is a summary.

[5] When applied to representative cross-sections at two points in time, then the scorecard measures the net change in poverty, but it does not reveal who moved up or down. When applied repeatedly to the same set of households, then the scorecard can reveal poverty dynamics, an approach with a long history in Ethiopia (Alem, 2011; Bigsten and Shimeles, 2008; Dercon, 2006; Kedir and McKay, 2005; Bigsten et al., 2003).

2. Data, poverty lines, and indicator-matching

This section briefly discusses the building blocks of the analysis. It documents the data and poverty lines used to construct and test the poverty scorecard, and it also discusses the process and results of matching indicators across surveys.

2.1 Data

The poverty scorecard is constructed from a random sample of half the people clustered in the 21,297 households in the 2004/5 HICE. The other half of the HICE is used for testing accuracy.

2004/5 HICE refers to the combination of two Ethiopian surveys, the 2004 Household Income, Consumption and Expenditure Survey (consumption data) and the February 2005 Welfare Monitoring Survey (scorecard indicators). All households in the 2005 WMS are also in the 2004 HICE, and both surveys are nationally representative and were conducted by Ethiopia's Central Statistical Agency (CSA).

2.2 Poverty lines

Ethiopia has no official, published poverty lines. Dercon (1997) uses the 1995/6 HICE and information from the CSA to derive a food line and a food-plus-non-food line in 1995/6 prices. Unfortunately, there are no regional and temporal price indexes available for converting these lines to 2004/5 prices. Other major poverty documents for Ethiopia (Woldehanna, Hoddinott, and Dercon, 2008; Ministry of Finance and Economic Development, 2006 and 2002) apparently use Dercon's lines and convert consumption to 1995/6 prices, but none of these documents provide enough information for replication. This paper uses the international "extreme" poverty line of $1.25 per person per day at 2005 purchase-power parity (Ravallion, Chen, and Sangraula, 2009), unadjusted for regional differences in cost-of-living.[6]

[6] In Ethiopia, prices vary widely by region (Gebremedhin and Whelan, 2008). Not adjusting for this overstates poverty in low-price regions and understates it elsewhere; weakening scoring's ability to reflect the real links between indicators and poverty. The World Bank normally does not adjust the $1.25/day line for intra-country price differences, so doing it here would harm inter-country comparability.

Ethiopia's $1.25/day 2005 PPP line is derived from the 2005 PPP exchange rate for "individual consumption expenditure by households" of ETB2.75 per $1.00 (World Bank, 2008) and consumer price indexes (Loening, Durevall, and Birru, 2008) of 77.43 for July 2004 and 82.88 on average for 2005, using the formula in Sillers (2006):

$$(2005\text{PPP exchange rate})\cdot \$1.25\cdot\left(\frac{\text{CPI}_{\text{July 2004}}}{\text{CPI}_{\text{Ave.2005}}}\right)=\left(\frac{\text{ETB2.75}}{\$1.00}\right)\cdot \$1.25\cdot\left(\frac{77.43}{82.88}\right)=\text{ETB3.21}.$$

The scorecard is built with the $1.25/day line and person-level weights. Because policy-makers may want to use different or various poverty lines, scores from the single scorecard are also calibrated to poverty likelihoods for $1.75/day and $2.50/day.

Table 3 reports poverty rates in the 2004/5 HICE. For the $1.25/day line, 34.8 percent of Ethiopians are poor. For $1.75/day, 61.8 percent are poor, and for $2.50/day, 85.8 percent are poor. Oromiya Region has the lowest poverty rate by $1.25/day (29.3 percent), and Tigray and SNNPR have the highest (44.6 percent).

Table 3: Sample sizes, poverty lines, and poverty rates for all-Ethiopia and by construction/validation sub-samples, regions, and poverty lines

		International 2005 PPP		
	Households	$1.25/day	$1.75/day	$2.50/day
Poverty line (ETB/person/day)	N/A	3.21	4.49	6.42
Poverty rate (% people)				
All Ethiopia	21,297	34.8	61.8	85.8
Construction sub-sample	10,675	34.9	61.8	85.8
Validation sub-sample	10,622	34.8	61.8	85.8
Poverty rate (% people) by region				
Tigray	1,716	44.6	67.2	84.0
Afar	951	34.9	55.9	78.8
Amhara	3,985	37.5	66.4	89.7
Oromiya	4,622	29.3	57.9	84.8
Somali	1,152	42.1	65.3	85.7
Benshangul-Gumuz	1,064	31.9	55.7	83.4
SNNPR	3,084	44.6	67.2	84.0
Harari	635	34.9	55.9	78.8
Addis Ababa	3,417	37.5	66.4	89.7
Dire Diwa	671	35.4	56.7	76.2

Source: 2004/5 HICE. SNNPR is Southern Nations, Nationalities, and People's Region.

2.3 Matching indicators across the HICE and DHS

The validity of the estimation of consumption-based poverty in this paper rests on the assumption that indicators in the 2004/5 HICE mean the same as in the 2005 DHS. Ideally, both surveys would be fielded at the same time and their indicators would have identical wording, offer identical response options, appear in identical contexts, and elicit identical distributions of responses. Furthermore, ideal indicators would have a balanced distribution of responses (for example, half own an asset, and half do not, rather than 95 percent are non-owners) with variation linked with the poverty of people close to a given poverty line.

The Appendix documents the matching of scorecard indicators for Ethiopia. In the poverty-mapping approach, indicators are considered as "matched" only if the distributions of responses across the two surveys are not statistically different at conventional levels such as $p < 0.10$ or $p < 0.05$. Based on a Chi-squared test for the independence of the response distributions, this standard would reject eight of the nine indicators in the scorecard here, even when most of the ideal conditions appear to be met. Of course, this may be particular to these Ethiopia surveys, and it need not reflect on the usefulness or validity of the general approach.

To be able to proceed, this paper accepts as "well-matched" indicators for which the difference in response percentages across surveys does not much exceed 10 percent of the most common response.

In broad terms, the indicators are of four types. The first is household size. This is the most powerful indicator. As might be expected, larger households more likely to be poor (Lanjouw and Ravallion, 1995).

The second type of indicator is well-matched but has highly unbalanced responses and serves mostly to distinguish the least-poor five or ten percent of people from the more-poor masses. These are the education of the female head, the type of toilet facility, and the main type of cooking fuel.

The third type of matched indicator is the type of wall. Here, responses are more evenly distributed, so power is better for distinguishing among people close to the relevant poverty lines. The wording of some response options, however, differs greatly

between the surveys, and although the responses can be grouped to give similar frequencies, it is not clear how such different wordings can lead to such close matches.[7]

The fourth and final type of matched indicators concern asset ownership: beds, radios, agricultural land, and cattle, sheep, and goats. These indicators appear to fulfill many of the ideal requirements, yet reported ownership rates differ a lot across surveys.[8] Nevertheless, these indicators are included because, apart from household size and wall type, they are the only way to distinguish among the more and less poor in the non-elite masses, particularly in rural areas.

To sum up, indicators for Ethiopia do not match up as well as might be hoped. While this does not affect the construction of the scorecard from the 2004/5 HICE, it does matter for the application of the scorecard to the 2005 DHS. The weaker the match, the weaker the link between measured accuracy in the 2004/5 HICE and assumed accuracy in the 2005 DHS. Unfortunately, the extent of this inaccuracy is untestable.

3. Scorecard construction

The first step in scorecard construction is to identify potential matched indicators in the 2004/5 HICE and the 2005 DHS. About 46 potential indicators were identified in the areas of:

- Family composition (such as number of household members in an age range)
- Education (such as school attendance by children in an age range)
- Housing (such as the main type of cooking fuel)
- Ownership of durable goods (such as beds and radios)
- Ownership of agricultural assets (such as land and livestock)

Initial screening eliminated 22 potential indicators due to very weak matching, highly unbalanced response distributions, or similarity with other indicators that are simpler or more intuitive.

[7] It is possible that the wordings are more similar in local languages.

[8] See the Appendix for details. The causes of the differences is unknown, but they may hinge on differences in enumeration. For example, many households have non-functioning assets (such as broken radios, or radios with no batteries), and DHS enumerators may have handled such cases differently than HICE enumerators.

The scorecard is built using the $1.25/day 2005 PPP line and Logit regression on the construction sub-sample from the 2004/5 HICE. Indicator selection uses both judgment and statistics (forward stepwise). The first step is to use Logit to build one scorecard for each candidate indicator. Each scorecard's accuracy is taken as "c", a measure of ability to rank by poverty status (SAS Institute Inc., 2004).

One of these one-indicator scorecards is then selected based on several factors (Schreiner et al., 2004; Zeller, 2004), including improvement in accuracy, likelihood of acceptance by users (determined by simplicity, cost of collection, and "face validity" in terms of experience, theory, and common sense), sensitivity to changes in poverty, variety among indicators, applicability across geographic regions, and verifiability.

A series of two-indicator scorecards are then built, each based on the one-indicator scorecard selected from the first step, with a second candidate indicator added. The best two-indicator scorecard is then selected, again based on "c" and judgment. These steps are repeated until additional indicators do not improve power.

The final step is to transform the estimated Logit coefficients into non-negative integers such that total scores range from 0 (most likely below a poverty line) to 100 (least likely below a poverty line). This linear transformation makes the scorecard's points simple for users, and it does not affect ranks.

This algorithm is the Logit analogue to the familiar R^2-based stepwise with least-squares regression. It differs from naïve stepwise in that the criteria for selecting indicators include not only statistical accuracy but also judgment and non-statistical factors. The use of non-statistical criteria can improve robustness through time and helps ensure that indicators are simple and intuitive.

The single poverty scorecard here applies to all of Ethiopia. Tests for Indonesia (World Bank, 2012), India and Mexico (Schreiner, 2006 and 2005), Sri Lanka (Narayan and Yoshida, 2005), and Jamaica (Grosh and Baker, 1995) suggest that segmenting scorecards by urban/rural does not improve ranking accuracy much, although it may improve the accuracy of estimated poverty rates (Tarozzi and Deaton, 2007).

4. Estimates of poverty likelihoods for individuals, and estimates of poverty rates for groups

This section describes how scores are converted to poverty likelihoods, that is, the probability that an individual person has consumption below a given poverty line. It also explains how the poverty likelihoods of individuals in a group are aggregated to estimate the group's consumption-based poverty rate. The accuracy of estimates of poverty rates is measured with the validation sample of the 2004/5 HICE, providing the best-available guess of accuracy when the scorecard is applied with the 2005 DHS.

4.1 Poverty likelihoods and their calibration with scores

The sum of scorecard points for a person is called the score. As described above, scores range from 0 to 100. While higher scores indicate less likelihood of being below a poverty line, the scores themselves are only ordinal and do not have equal-interval or ratio units. For example, doubling the score does not double the likelihood of being above a poverty line.

To get equal-interval and ratio units, scores are converted to poverty likelihoods, that is, probabilities of being below a poverty line. This is done via simple look-up tables. For the example of the $1.25/day 2005 PPP line, scores of 10–14 have a poverty likelihood of 25.9 percent, scores of 15–19 have a poverty likelihood of 20.1 percent, and so on (Table 4).

The poverty likelihood associated with a score varies by poverty line. With the $1.75/day 2005 PPP line, for example, scores of 10–14 are associated with a poverty likelihood of 85.9 percent (Table 4).

Table 4: Poverty likelihoods by score and poverty line

Score	Poverty likelihood		
	$1.25/day	$1.75/day	$2.50/day
0–4	45.3	71.9	91.3
5–9	32.6	80.1	95.5
10–14	25.9	62.7	85.9
15–19	20.1	57.9	77.4
20–24	16.6	41.6	65.8
25–29	9.3	30.2	53.6
30–34	4.8	22.2	43.5
35–39	3.8	9.4	27.9
40–44	1.7	8.0	25.4
45–49	1.1	5.1	14.1
50–54	1.0	4.2	9.6
55–59	0.0	0.0	3.7
60–64	0.3	0.3	1.3
65–69	0.0	1.9	5.4
70–74	0.0	0.0	4.7
75–79	0.0	0.0	2.5
80–84	0.0	0.0	0.0
85–89	0.0	0.0	0.0
90–94	0.0	0.0	0.0
95–100	0.0	0.0	0.0

A given score is non-parametrically associated ("calibrated") with a poverty likelihood by defining the poverty likelihood as the share of households in the 2004/5 HICE construction sub-sample with the score who are below a given poverty line.

For the example of the $1.25/day 2005 PPP line (Table 5), there are 3,108 (normalized) people in the construction sub-sample with a score of 10–14, of whom 806 (normalized) are below the poverty line. The estimated poverty likelihood associated with a score of 10–14 is then 25.9 percent, because 806 ÷ 3,108 = 25.9 percent. The same method is used to calibrate scores with estimated poverty likelihoods for the other poverty lines.

Although the points in the scorecard are transformed Logit coefficients, scores are not converted to poverty likelihoods via the Logit formula of 2.718281828^{score} x $(1+2.718281828^{score})^{-1}$. This is because the Logit formula is esoteric and difficult to compute by hand. It is more intuitive to define the poverty likelihood as the share of households with a given score in the construction sample who are below a poverty line. Thus, the (transformed) Logit coefficients are used to order people by relative ranks, and the ranks are then calibrated with absolute poverty likelihoods.

Table 5: Derivation of poverty likelihoods linked with scores, example poverty line of $1.25/day 2005 PPP

Score	People below poverty line		All people at score		Poverty likelihood (%)
0–4	696	÷	1,536	=	45.3
5–9	484	÷	1,485	=	32.6
10–14	806	÷	3,108	=	25.9
15–19	956	÷	4,763	=	20.1
20–24	1,036	÷	6,250	=	16.6
25–29	712	÷	7,645	=	9.3
30–34	477	÷	9,923	=	4.8
35–39	474	÷	12,349	=	3.8
40–44	179	÷	10,595	=	1.7
45–49	117	÷	10,991	=	1.1
50–54	95	÷	9,562	=	1.0
55–59	0	÷	6,484	=	0.0
60–64	16	÷	4,840	=	0.3
65–69	0	÷	2,971	=	0.0
70–74	0	÷	2,698	=	0.0
75–79	0	÷	2,184	=	0.0
80–84	0	÷	1,736	=	0.0
85–89	0	÷	398	=	0.0
90–94	0	÷	483	=	0.0
95–100	0	÷	0	=	0.0

Number of people normalized to sum to 100,000.

4.2 Estimates of a group's poverty rate

A group's estimated poverty rate is the average of the estimated poverty likelihoods of the individuals in the group. To illustrate, suppose a program samples three people on Jan. 1, 2010 and that they have scores of 20, 30, and 40, corresponding to poverty likelihoods of 16.6, 4.8, and 1.7 percent ($1.25/day, Table 4). The group's estimated poverty rate is the people's average poverty likelihood of (0.166 + 0.048 + 0.017) ÷ 3 = 7.7 percent.[9]

4.3 Accuracy of estimates of poverty rates

As long as the relationship between indicators and poverty does not change and as long as the scorecard is applied to a representative sample of people from the same population from which it was constructed, the scorecard produces unbiased estimates of poverty rates. Unbiased means that in repeated samples, the average estimate matches the true value.

Of course, the relationship between indicators and poverty does change over time and across sub-groups within Ethiopia's population, so the scorecard will generally be biased to some unknown extent when applied—as it must be in practice—after the end of the HICE fieldwork in February 2005 or when applied to non-nationally representative groups. To the extent that indicators are mismatched, it will also be biased when applied to the 2005 DHS. Unfortunately, this bias cannot be measured, and accuracy as measured for the 2004/5 HICE validation sample is the best available approximation of accuracy for the 2005 DHS.

How accurate are scorecard estimates of poverty rates for nationally representative samples in the period when the 2004/5 HICE was in the field? Table 6 reports estimates of bias (average differences between estimated and true poverty rates) as well as precision (confidence intervals for the differences) for the scorecard applied to 1,000 bootstrap samples of size n = 16,384 from the 2004/5 HICE validation sample. For $1.25/day, the scorecard is too low by 0.8 percentage points; on average, it estimates a poverty rate of 34.0 percent for the validation sample, but the true value is 34.8 percent

[9] The group's poverty rate is not the poverty likelihood associated with the average score. Here, the average score is (20 + 30 + 40) ÷ 3 = 30, and the poverty likelihood associated with the average score is 4.8 percent. This is not the 7.7 percent found as the average of the three poverty likelihoods associated with each of the three scores.

(Table 3). For $1.75/day, bias is –1.4 percentage points, and for $2.50/day, bias is –0.3 percentage points.[10]

In terms of precision, the 90-percent confidence interval for a group's estimated poverty rate at a point in time with n = 16,384 is 0.5 percentage points or less (Table 6). This means that in 900 of 1,000 bootstraps of this size, the difference between the estimate and the true value is within ±0.5 percentage points of the average difference. In the specific case of $1.25/day and the validation sample, 90 percent of all samples of n = 16,384 produce estimates that differ from the true value in the range of –0.8 – 0.3 = –1.1 to –0.8 + 0.3 = –0.5 percentage points, as –0.8 is the average difference and ±0.3 is its 90-percent confidence interval.

As shown in Schreiner (2013), the standard error[11] of the estimated poverty rate is

$$\alpha \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$, where:

p is the proportion of sampled households below the poverty line,

n is the sample size, and
α is a factor specific to the country, scorecard, and poverty line.
α factors below 1.0 (such as 0.44 and 0.62 for the $1.25/day and $1.75/day lines, Table 6) imply that the scorecard is more precise that direct measurement, while factors above 1.0 (such as 1.04 for the $2.50/day line) imply the converse.

[10] There are differences, in spite of the estimator's unbiasedness, because the scorecard comes from a single sample. The average difference would be zero if samples were repeatedly drawn from the same population and split into sub-samples before repeating the entire scorecard-building process.
[11] This does not correct for sampling from a finite population.

Table 6: Bias, precision, and sample-size α for bootstrapped estimates of poverty rates for groups of people at a point in time for the scorecard applied to the 2004/5 HICE validation sample

	International 2005 PPP Poverty Line		
	$1.25/day	$1.75/day	$2.50/day
Estimate minus true value			
Scorecard applied to 2004/5 HICE validation sample	–0.8	–1.4	–0.3
Precision of difference			
Scorecard applied to 2004/5 HICE validation sample	0.3	0.4	0.5
α factor for standard errors			
Scorecard applied to 2004/5 HICE validation sample	0.44	0.62	1.04

Precision is measured as 90-percent confidence intervals in units of +/– percentage points.
Differences and precision estimated from 1,000 bootstraps of size n = 16,384.
α is estimated as described in Schreiner (2013).

5. Value-added by the poverty scorecard

This section asks whether the poverty scorecard and the DHS asset index produce similar rankings. If not, then the poverty scorecard may have something new and useful to offer.

Large differences in rankings are possible—at least in principle—because the poverty scorecard and the asset index define poverty differently (consumption versus assets). Also, the two tools are constructed differently; subject to usability constraints, the scorecard uses Logit to choose indicators and points to maximize the accuracy of ranking based on consumption poverty, while the asset index uses PCA to maximize the explained variance among a pre-selected set of indicators.

Nevertheless, differences might be small. After all, the two tools use many similar indicators, and the correlation between assets and consumption may be strong.

The extent of differences in rankings matters because it would be convenient if the DHS index ranked people about the same as the poverty scorecard. In that case, the DHS asset index could be calibrated to poverty likelihoods and researchers could estimate consumption-based poverty rates using the familiar asset index.

Table 7 shows the quintile correspondences for people in the 2005 Ethiopia DHS when ranked by the two tools. The sample is sorted and divided into equal-sized quintiles twice, once by the scorecard's score and once by the asset index, so each row total and each column total should be 20 percent. There are 25 cells in the matrix (5 quintiles in the rows, by 5 quintiles in the columns). Each cell contains the percentage of all people who rank in a given row quintile by the poverty scorecard and who also rank in a given column quintile by the asset index. The five cells on the diagonal show the percentage of people who fall in the same quintile by both tools (first quintile on both the scorecard and the asset index, second quintile on both the scorecard and the asset index, etc.) If the correspondence across the two rankings were perfect, all diagonal cells would be 20 percent and all off-diagonal cells would be zero. At the other extreme of no correspondence, all the cells would be 4 percent.

Table 7: Correspondences of quintile ranks for the poverty scorecard and DHS asset index applied to the 2005 Ethiopia DHS

Poverty scorecard	DHS asset index 1	2	3	4	5
1	6.7	4.9	3.8	2.9	1.8
2	4.6	4.3	4.3	4.0	2.8
3	3.5	4.0	4.5	4.4	3.7
4	3.1	3.7	3.9	4.9	4.4
5	2.2	3.2	3.6	3.8	7.3

The actual correspondence in Table 7 is better than random,[12] as diagonal cells always exceed 4 percent and most off-diagonal cells are less than 4 percent. Except for the first/most-poor quintile (6.7 percent) and the fifth/least-poor quintile (7.3 percent) where ranking is easiest, however, the correspondence is not much better than random. About 27.7 percent of people fall in the same quintile in both rankings (versus 20 percent if random). Large differences are common; 38 percent differ across the two rankings by two quintiles or more, and 16 percent differ by three quintiles or more. In sum, there are large differences in ranking by the two tools.

A simpler test is to replace the quintiles with a single cut-off. Table 8 shows two examples, one with a 35th-percentile cut-off (corresponding to Ethiopia's poverty rate for $1.25/day)

[12] $p < 0.01$ for a Chi-square test.

and a second example with a 85th-percentile cut-off (corresponding to Ethiopia's poverty rate for $2.50/day). Having a single cut-off increases the share of people in cells on the diagonal who are classified the same by both tools, as some cells that are off-diagonal with quintiles are now part of the (larger) diagonal cells (Friedman, 1997).

Table 8: Correspondences of ranks with cut-offs at the 35th and 85th percentiles, poverty scorecard and DHS asset index applied to the 2005 Ethiopia DHS

Cut-off at 35th percentile (poverty line of $1.25/day 2005 PPP)

		DHS asset index	
		< 35th	> = 35th
Poverty scorecard	< 35th	16.2	18.8
	> = 35th	18.8	46.2

Cut-off at 85th percentile (poverty line of $2.50/day 2005 PPP)

		DHS asset index	
		< 85th	> = 85th
Poverty scorecard	< 85th	75.2	9.8
	> = 85th	9.8	5.2

In the 2005 Ethiopia DHS, less than half of the people (46 percent = 16.2 ÷ (16.2 + 18.8)) with poverty scores below the 35th percentile also have asset scores below the 35th percentile. Thus, on the most-poor end of the scale, there are large differences between ranks between the two tools. On the least-poor end, however, agreement is greater, at 71 percent (= 46.2 ÷ (46.2+ 18.8)).

Increasing the cut-off improves accuracy for the most-poor but worsens accuracy for the least-poor. With a cut-off at the 85th percentile, there is agreement between the two

tools for 88 percent of the most-poor (75.2 ÷ (75.2 + 9.8)) but only 35 percent of the least-poor (= 5.2 ÷ (9.8 + 5.2)).

To sum up, the poverty scorecard and the DHS asset index do not generally concentrate large shares of the same people among low scores (most-poor) nor among high scores (least-poor). The two tools are not good proxies for each other. If the goal is to measure consumption-based poverty, then it is better to use an estimator tailored for that purpose than to utilize an estimator designed for asset-based poverty.

6. Estimating consumption-based poverty with the poverty-mapping approach in the DHS

This paper is not the first to build a consumption-based poverty scorecard using only indicators matched to a DHS (or DHS-like) survey. This section asks two questions of previous (non-Ethiopian) work. First, how does their accuracy and precision compare with that of the scorecard here? And second, are poverty scores more strongly linked with health outcomes than asset-index scores? Of course, the answers to these questions are related to the overall usefulness of the poverty scorecard, but they are distinct from the main point of this paper, namely, that consumption-based poverty can be estimated in the DHS, even though the DHS does not collect consumption data.

6.1 How accurate is this scorecard versus others?

This sub-section describes three cases where comparisons of bias and precision (standard errors or confidence intervals) are possible. Some papers (for example, Filmer and Pritchett, 2001; Kijima and Lanjouw, 2003) are omitted because they compare health only with true (reported) consumption (not predicted consumption) or because their scorecards use only a subset of the indicators used here.

6.1.1 Stifel and Christiaensen

Stifel and Christiaensen (2007) seek to an intuitive and inexpensive way to track changes in poverty. They build three scorecards (Nairobi, other urban, and rural) using consumption data from Kenya's 1997 Welfare Monitoring Survey and indicators matched to Kenya's DHS. The scorecards are applied to the 1993, 1998, and 2003 DHS to estimate changes in poverty in years without consumption surveys. Like most poverty scorecards—but unlike the one in this paper—Stifel and Christiaensen regress

the logarithm of per-capita household consumption against a set of indicators, many of which are similar to those in this paper.

When Stifel and Christianensen's scorecards constructed with Kenya's 1997 WMS is applied to the same data used to construct the scorecards in the first place (that is, in-sample), bias ranges from –1 to –2 percentage points. Such in-sample tests overstate the accuracy that can be expected when the estimator is applied out-of-sample to new data that was not used to construct the estimator.[13] When the poverty scorecard here is applied out-of-sample—that is, to data not used to construct the scorecard—bias ranges from –0.3 to –1.4 percentage points (Table 6). Thus, the scorecard here is not more biased than that of Stifel and Christianensen (and they understate bias).

For precision, Stifel and Christianensen report a standard error of 1.7 percentage points for an in-sample poverty-rate estimate (n = 10,639). Ignoring again the in-sample overstatement of precision, the implied α factor is about 3.5, suggesting that the scorecard here (α of 0.44 to 1.04, Table 6) is more precise; its confidence intervals are at least three times narrower.

6.1.2 Simler, Harrower, and Massingarela

Simler, Harrower, and Massingarela (2003) use poverty mapping as a simple, inexpensive way to track changes in poverty rates without complex, costly consumption surveys. They build 11 scorecards (one per province) using Mozambique's 1996/7 National Household Survey of Living Standards, using only indicators matched to Mozambique's DHS-like 2000/1 Core Welfare Indicator Survey. The scorecards predict the logarithm of consumption using indicators on education, housing, asset ownership, community averages, and GIS variables.

Based on an in-sample test with the 1996/7 consumption survey, bias is –3.9 percentage points, and the α factor is 2.3. These numbers are much larger than those for the Ethiopia scorecard here.

[13] For example, if the scorecard here is applied in sample, bias is exactly zero.

6.1.3 Azzarri et al.

Azzarri et al. (2005) construct a poverty scorecard from the 2002 Albania Living Standards Measurement Survey and then apply it to a sub-sample of households who were revisited in 2003. Thus, the indicators are perfectly matched. Like the others reviewed here, the scorecard predicts the logarithm of consumption, and indicators are selected with stepwise regression. Azzarri et al. also include some subjective indicators. In an in-sample test with the 2002 data, bias is –4.8 percentage points, much greater than for the Ethiopia scorecard here. Azzarri et al. do not report standard errors, so precision cannot be compared with that of the scorecard here.

6.2 Which tool is more closely related with health outcomes?

The poverty scorecard and the DHS asset index rank people differently, but this need not imply differences in their relationship with health (Wagstaff and Watanabe, 2003). This subsection discusses two papers that compare how health relates with poverty scores and with DHS-like asset scores. Other papers that look at health vis-à-vis true (not predicted) consumption are not discussed.

6.2.1 Sahn and Stifel

In a seminal paper covering nine countries, Sahn and Stifel (2003) look at whether child health (percentage stunted, and mean height-for-age z scores) is more closely related to ranks based on consumption from a poverty scorecard or ranks from a DHS-like asset index. As usual, they predict the logarithm of per-capita consumption.

On the one hand, Spearman correlation coefficients and correspondence indexes suggest that, "in terms of predictive capabilities, it does not matter which welfare measure is used" (Sahn and Stifel, p. 480). On the other hand, they find that—in 17 of 22 cases—the gradient with child-health outcomes between the first and fifth quintiles was greater for the asset index than for predicted consumption from a scorecard.

In the end, Sahn and Stifel fail to reject the hypothesis of no differences: "In the context of estimating models of nutrition, we find no compelling reason to believe that either reported or instrumented [predicted by a scorecard] expenditures serve as a better proxy for economic welfare than the asset index" (p. 485).

6.2.2 Filmer and Scott

Filmer and Scott (2012) compare ranks for reported consumption, scorecard-predicted consumption, and DHS-like asset indexes. Several results from their tests with 11 countries are of interest here.

First, "predicted per-capita expenditure [from a scorecard] yields the most similar household rankings to per-capita expenditure" (p. 18). This is no surprise; if matching true consumption is the goal, then poverty scorecards are better than asset indexes. Still, Filmer and Scott report that asset indexes are also highly correlated with true consumption.

Second, "despite household re-rankings, conclusions about inequalities across quintiles in education outcomes, health-care-seeking behavior, fertility, and child mortality, as well as labor-market outcomes, are not very sensitive to the particular economic-status measure used to classify households" (p. 22).[14] That is, Filmer and Scott's scorecards do better than their asset indexes at estimating consumption-based poverty and about as well as asset indexes with health outcomes.

Filmer and Scott's third point is that scorecards and asset indexes "show vastly different gradients in household composition" because scorecards do not adjust for household economies of scale (for example, one bathroom can serve five people at less than five times the cost of serving one person). Thus, the two "equivalence" results just described may not apply to the poverty scorecard here; Filmer and Scott's scorecard omits household size (and education and employment as well), but household size is by far the most powerful predictor of consumption-based poverty in Ethiopia.

The scorecard here uses more types of indicators than those of Filmer and Scott, and so their results—that asset indexes and scorecards perform about the same—may not apply here. In particular, the scorecard here uses a wider variety of types of information, so it is likely to perform better than those in Filmer and Scott. Further tests, however, are beyond the scope of this paper.

[14] This fits Wagstaff and Watanabe (2003, 19 countries), but not Lindelow (2006, one country).

7. Conclusion

The poverty scorecard provides a way to estimate consumption-based poverty for people and for groups in the 2005 Ethiopia DHS. The approach follows poverty mapping in that it constructs a scorecard based on a consumption survey (the 2004/5 HICE) using only indicators that are also in the 2005 DHS (which does not collect consumption data). Researchers can then apply the scorecard to the 2005 DHS and analyze how health outcomes vary with consumption-based poverty.

The poverty scorecard ranks people differently than the DHS asset index; asset-based poverty (a longer-term concept) is not a good proxy for consumption-based poverty (a shorter-term concept). While both conceptions of poverty are legitimate, the consumption-based definition dominates discussion among both the polity and policymakers. Thus, including consumption-based estimates may give DHS research greater policy relevance.

Like poverty mapping, poverty scoring makes three assumptions about its two data sources: that they represent the same population, that they represent the same time period, and (least tenably in the case of Ethiopia) that the indicators are well-matched. The poverty-scoring approach here improves on traditional poverty mapping in that it reports out-of-sample bias and a simple summary measure of precision.

Of course, the results here hold only for Ethiopia's 2004/5 HICE and 2005 DHS; they may or may not generalize to other countries and data sources. The approach, however, can be tested in any country with both a DHS and a national household expenditure survey.

From the point of view of policy in Ethiopia, the poverty scorecard is accurate enough in terms of bias and precision to provide useful estimates of head-count poverty rates in years when there is a DHS but no HICE.[15] More-frequent quantitative feedback on the direction of overall household well-being can help keep policy on-track from the point of view of poverty alleviation.

[15] In the same way, the approach can be used to create poverty maps, based on the DHS, in years between censuses.

References

Alem, Yonas. (2011). "Poverty Dynamics and Intra-Household Heterogeniety in Occupations: Evidence from Urban Ethiopia", efdinitiative.org/news-press/workshops-etc/workshops-archive/files-1/files-ethiopia/Yonas_Alem.pdf, retrieved 6 February 2013.

Azzarri, Carlo; Carletto, Gero; Davis, Benjamin; and Alberto Zezza. (2005). "Monitoring Poverty without Consumption Data: An Application Using the Albania Panel Survey", Eastern European Economics, Vol. 44, No. 1, pp. 59–82.

Bigsten, Arne; Kebede, Bereket; Shimeles, Abebe; and Mekonnen Taddesse. (2003). "Growth and Poverty Reduction in Ethiopia: Evidence from Household Panel Surveys", World Development, Vol. 31, No. 1, pp. 87–106.

Bigsten, Arne and Abebe Shimeles. (2008). "Poverty Transition and Persistence in Ethiopia: 1994–2004", World Development, Vol. 36, No. 9, pp. 1559–1584.

Bollen, Kenneth A.; Glanville, Jennifer L.; and Guy Stecklov. (2007). "Socio-Economic Status, Permanent Income, and Fertility: A Latent-Variable Approach", Population Studies, Vol. 61, No. 1, pp. 15–34.

Booysen, Frikkie; van der Berg, Servaas; Burger, Ronelle; von Maltitz, Michael; and Gideon du Rand. (2008). "Using an Asset Index to Assess Trends in Poverty in Seven Sub-Saharan African Countries", World Development, Vol. 36, No. 6, pp. 1113–1130.

Carter, Michael R.; and Christopher B. Barrett. (2006). "The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach", *Journal of Development Studies*, Vol. 42, No. 2, pp. 178–199.

Christiaensen, Luc; Lanjouw, Peter; Luoto, Jill; and David Stifel. (2010). "The Reliability of Small-Area Estimation Prediction Methods to Track Poverty", wider.unu.edu/stc/repec/pdfs/wp2010/wp2010-99.pdf, retrieved 6 February 2013.

Deaton, Angus; and Salman Zaidi. (2002). "Guidelines for Constructing Consumption Aggregates for Welfare Analysis", World Bank Living Standards Measurement Survey Working Paper No. 135, Washington, D.C., http://go.worldbank.org/8YRCR9ERJ0, retrieved 6 February 2013.

Dekker, Marleen. (2006). "Estimating Wealth Effects without Expenditure Data: Evidence from Rural Ethiopia", *Ethiopian Journal of Economics*, Vol. 15, No. 1, pp. 35–54.

Dercon, Stefan. (2006). "Economic Reform, Growth, and the Poor: Evidence from Rural Ethiopia", *Journal of Development Economics*, Vol. 81, No. 1, pp. 1–24.

_____. (1997) "Poverty and Deprivation in Ethiopia", Center for the Study of African Economies, Oxford University, memo.

Elbers, Chris; Lanjouw, Peter; and Phillippe George Leite. (2008). "Brazil within Brazil: Testing the Poverty-Map Methodology in Minas Gerais", World Bank Policy Research Working Paper No. 4513, Washington, D.C., papers.ssrn.com/sol3/papers.cfm?abstract_id=1092691, retrieved 6 February 2013.

Elbers, Chris; Lanjouw, Jean O.; and Peter Lanjouw. (2003) "Micro-Level Estimation of Poverty and Inequality", *Econometrica*, Vol. 71, No. 1, pp. 355–364.

Filmer, Deon and Lant, Pritchett. (2001) "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrolments in States of India", Demography, Vol. 38, No. 1, pp. 115–132.

Filmer, Deon and Kinnon, Scott. (2012). "Assessing Asset Indexes", Demography, Vol. 49, pp. 359–392.

Friedman, Jerome H. (1997). "On Bias, Variance, 0–1 Loss, and the Curse-of-Dimensionality", Data Mining and Knowledge Discovery, Vol. 1, pp. 55–77.

Gebremedhin, Tesfaye Alemayehu; and Stephen Whelan. (2008). "Prices and Poverty in Urban Ethiopia", *Journal of African Economies*, Vol. 17, No. 1, pp 1–33.

Grosh, Margaret; and Judy L. Baker. (1995). "Proxy Means Tests for Targeting Social Programs: Simulations and Speculation", World Bank Living Standard Measurement Survey Working Paper No. 118, Washington, D.C., go.worldbank.org/W9OWN57PD0, accessed 6 February 2013.

Gwatkin, Davidson R.; Rutstein, Shea; Johnson, Kiersten; Suliman, Eldaw; Wagstaff, Adam; and Agbessi Amouzou. (2007). "Socio-Economic Differences in Health, Nutrition, and Population: Ethiopia", Country Reports on HNP and Poverty, Washington, D.C. World Bank, go.worldbank.org/T6LCN5A340, retrieved 6 February 2013.

Kedir, Abbi M.; and Andrew McKay. (2005). "Chronic Poverty in Urban Ethiopia: Panel-Data Evidence", International Planning Studies, Vol. 10, No. 1, pp. 49–67.

Kijima, Yoko; and Peter Lanjouw. (2003). "Poverty in India during the 1990s: A Regional Perspective", World Bank Policy Research Working Paper No. 3141, ideas.repec.org/p/wbk/wbrwps/3141.html, retrieved 6 February 2013.

Knowles, James C.; Bales, Sarah; Cuong, Le Quang; Oanh, Tran Thi Mai; and Duong Huy Luong. (2008). "Health Equity in Vietnam: A Situational Analysis Focused on Maternal and Child Mortality", background paper prepared for UNICEF Consultancy on "Equity in Access to Quality Healthcare for Women and Children" (April 8-10, Ha Long City, Vietnam), jkinusa@aol.com.

Lanjouw, Peter; and Martin Ravallion. (1995). "Poverty and Household Size", *Economic Journal*, Vol. 105, pp. 1415–1434.

Lindelow, Magnus. (2006). "Sometimes More Equal Than Others: How Health Inequalities Depend on the Choice of Welfare Indicator", *Health Economics*, Vol. 15, pp. 263–279.

Loening, Josef L.; Durevall, Dick; and Yohannes Ayalew Birru. (2009). "Inflation Dynamics and Food Prices in an Agricultural Economy: The Case of Ethiopia", Working Papers in Economics No. 347, University of Gothenburg, gupea.ub.gu.se/dspace/bitstream/2077/19464/1/gupea_2077_19464_1.pdf, accessed 6 February 2013.

Ministry of Finance and Economic Development (MoFED). (2006). Ethiopia: Building on Progress, a Plan for Accelerated and Sustained Development to End Poverty (PASDEP), 2005/06–2009/10, Volume I: Main Text, Welfare Monitoring Unit, Addis Ababa, planipolis.iiep.unesco.org/upload/Ethiopia/Ethiopia_PASDEP_2005_2010.pdf , accessed 6 February 2013.

_____. (2002). "Poverty Profile of Ethiopia: Analysis based on the 1999/00 HICE and WM Survey Results", Welfare Monitoring Unit, Addis Ababa.

Narayan, Ambar; and Nobuo Yoshida. (2005). "Proxy Means Tests for Targeting Welfare Benefits in Sri Lanka", World Bank Report No. SASPR–7, Washington, D.C., documents.worldbank.org/curated/en/2005/07/6209268/proxy-means-test-targeting-welfare-benefits-sri-lanka, accessed 6 February 2013.

Pitchforth, Emma; van Teijlingen, Edwin; Graham, Wendy; and Ann Fitzmaurice. (2007). "Development of a Proxy Wealth Index for Women Utilizing Emergency Obstetric Care in Bangladesh", *Health Policy and Planning*, Vol. 22, No. 3, pp. 311–319.

Ravallion, Martin; Chen; Shaohua; and Prem Sangraula. (2009). "Dollar-a-Day Revisited", World Bank Economic Review, Vol. 23, No. 2, pp. 163–184.

Rutstein, Shea Oscar; and Kiersten Johnson. (2004). "The DHS Wealth Index", DHS Comparative Reports No. 6, Calverton, MD: ORC Macro, measuredhs.com/pubs/pdf/CR6/CR6.pdf, accessed 6 February 2013.

Sahn, David E. and David Stifel. (2003). "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", Review of Income and Wealth, Series 49, No. 4, pp. 463–489.

_____. (2000). "Poverty Comparisons over Time and across Countries in Africa", World Development, Vol. 28, No. 12, pp. 2123–2155.

SAS Institute Inc. (2004). "The LOGISTIC Procedure: Rank Correlation of Observed Responses and Predicted Probabilities", in SAS/STAT User's Guide, Version 9, Cary, NC, support.sas.com/documentation/cdl/en/statug/63033/HTML/default/statug_logistic_sect035.htm, accessed 6 February 2013.

Schreiner, Mark. (2013). "A Simple Poverty Scorecard for Peru", microfinance.com/English/Papers/Scoring_Poverty_Peru_2010_EN.pdf, accessed 6 February 2013.

_____. (2006). "Is One Simple Poverty Scorecard Enough for India?", microfinance.com/English/Papers/Scoring_Poverty_India_Segments.pdf, accessed 6 February 2013.

_____. (2005). "Un Índice de Pobreza para México", memo for Grameen Foundation, microfinance.com/Castellano/Documentos/Scoring_Pobreza_Mexico_2002.pdf, accessed 6 February 2013.

Schreiner, Mark; Matul, Michal; Pawlak, Ewa; and Sean Kline. (2004) "Poverty Scorecards: Lessons from a Microlender in Bosnia-Herzegovina", microfinance.com/English/ Papers/Scoring_Poverty_in_BiH_Short.pdf, accessed 6 February 2013.

Schreiner, Mark and Michael Sherraden. (2006). Can the Poor Save? Saving and Asset Accumulation in Individual Development Accounts, Piscataway, NJ: Transaction.

Sherraden, Michael. (1991). Assets and the Poor: A New American Welfare Policy, Armonk, NY: M.E. Sharpe.

Sillers, Don. (2006). "National and International Poverty Lines: An Overview", Washington, D.C.: United States Agency for International Development, povertytools.org/other_documents/siller-povertylines.doc, retrieved 6 February 2013.

Simler, Kenneth R.; Harrower, Sarah; and Claudio Massingarela. (2003). "Estimating Poverty Indexes from Simple Indicator Surveys", http://www.csae.ox.ac.uk/conferences/2004-GPRaHDiA/papers/3p-SimlerMassingerela-CSAE2004.pdf, retrieved 6 February 2013.

Stifel, David; and Luc Christiaensen. (2007). "Tracking Poverty over Time in the Absence of Comparable Consumption Data", World Bank Economic Review, Vol. 21, No. 2, pp. 317–341.

Tarozzi, Alessandro; and Angus Deaton. (2007). "Using Census and Survey Data to Estimate Poverty and Inequality for Small Areas", Review of Economics and Statistics, Vol. 91, No. 4, pp. 773–792.

Wagstaff, Adam; and Naoko Watanabe. (2003). "What Difference Does the Choice of SES Make in Health Inequality Measurement?", Health Economics, Vol. 12, No. 10, pp. 885–890.

Woldehanna, Tassew; Hoddinott, John; and Stefan Dercon. (2008). "Poverty and Inequality in Ethiopia: 1995/96–2004/05", ssrn.com/abstract=1259341, accessed 6 February 2013.

World Bank. (2012). Targeting Poor and Vulnerable Households in Indonesia, www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2012/03/01/000386194_20120301004533/Rendered/PDF/672180WP00PUBL0T00English000PUBLIC0.pdf, retrieved 6 February 2013.

_____. (2008). "International Comparison Project: Tables of Results", Washington, D.C., siteresources.worldbank.org/ICPINT/Resources/icp-final-tables.pdf, retrieved 6 February 2013.

Zeller, Manfred. (2004). "Review of Poverty Assessment Tools", Accelerated Microenterprise Advancement Project, pdf.usaid.gov/pdf_docs/PNADH120.pdf, retrieved 6 February 2013.

Appendix

Matching indicators across the 2004/5 HICE and the 2005 DHS

This appendix documents, using simple cross-tabs, how responses are grouped match indicators across the 2004/5 HICE and 2005 DHS. Only indicators in the scorecard in Table 1 are included here; documentation for other indicators is available on request.

HICE: How many people live with the household six months out of the year?
DHS: How many people usually live with the household?

	% of people with a given response		
Response	HICE	DHS	Difference
1 or 2	13.2	14.4	-1.2
3	9.9	11.5	-1.6
4	15.3	15.4	-0.1
5	16.6	16.7	-0.1
6	15.9	17.7	-1.8
7	14.0	12.5	1.5
8	9.2	8.3	0.9
9 or more	5.8	4.6	1.2

Household size is the most powerful scorecard indicator, and it seems well-matched, although in general the HICE reports larger households.

HICE: What is the highest grade the female head/spouse has completed?
DHS: What is the highest grade the female head/spouse has completed?

	% of people with a given response		
Response	HICE	DHS	Difference
Four or less, or no female head/spouse	92.5	92.7	-0.2
Five or six	2.7	2.5	0.2
Seven to nine	2.3	2.2	0.1
Ten or more	2.5	2.6	-0.1

This well-matched indicator mostly distinguishes the least-poor form the poorest ninth-tenths of Ethiopians.

HICE: What is the main construction material of the walls of the dwelling unit?
DHS: What is the main material of the walls of the residence?

HICE	Item	Group	DHS	Item	Group	Diff. Group
Mud and stone	9.6		Stone with mud	7.8		
Wood and grass	8.2		Stone with lime/ cement	1.1		
Others	2.7		Cane/ trunks/ bamboo/ reed	5.6		
			Other	3.1		
		20.4	Undefined	1.7	19.3	1.1
Wood and mud	74.9		Bamboo/ wood	78.2		
Reed and bamboo	3.4		Uncovered adobe	0.0		
No data	0.0		Plywood	0.0		
			Carton	0.0		
			No walls	0.2		
		78.3	No data	0.0	78.4	-0.1
Cement and stone	0.7		Cement	0.5		
Hollow bricks	0.5		Bricks	0.1		
Bricks	0.1		Cement blocks	0.6		
			Covered adobe	0.0		
		1.3	Wood planks/ shingles	1.1	2.3	-1.0

This indicator is well-matched in terms of the percentages across grouped responses. But it is less clear why the first group (stone with mud, stone with lime/cement, etc.) implies more poverty than the second group (bamboo/wood, uncovered adobe, etc.). Also, the responses sometimes seem different across the two surveys, even though their percentages line up nicely. Nevertheless, "wood and mud" in the HICE must match up with "bamboo/wood" in the DHS, as these account for about three-fourths of all people. Likewise, "mud and stone" in the HICE must match with "Stone and mud" or "Stone with lime/cement" in the DHS, even though that (less plausibly) implies that "wood and grass" in the HICE matches with "cane/trunks/bamboo/reed" in the DHS.

HICE: What type of toilet facility does the household use?
DHS: What kind of toilet facility do members of your household usually use?

	% of people with a given response		
Response	HICE	DHS	Difference
Non-flush or none	97.6	97.8	-0.2
Flush	2.4	2.2	0.2

This is well-matched, but, like the education of the female head/spouse, only matters for a few of the elite, failing to help differentiate socio-economic status for the masses.

HICE: What is the main source of cooking fuel?
DHS: What type of fuel does your household mainly use for cooking?

HICE	Item	Group	DHS	Item	Group	Group diff.
Mainly collected firewood	72.8		Wood or straw/shrubs/grass	86.8		
Mainly purchased firewood	9.5		Animal dung	7.5		
Crop residue	11.1	93.4			94.3	-0.9
Charcoal	1		Electricity	0.2		
Kerosene	2		LPG or natural gas	0.1		
Butane gas	0.4		Biogas	0		
Electricity	0.4		Kerosene	3		
Does not use cooking fuel	0.1		Charcoal	2.2		
Other	2.7		Other	0.1		
			Other	0.1		
		6.6	No data	0	5.7	0.9

Like the education of the female head/spouse and the type of toilet arrangement, the type of cooking fuel mostly serves to identify the highest end of the Ethiopian socio-economic scale. While the percentages of people in each group match well, and while it is obvious what the "high-quality" cooking fuels are, it is nevertheless odd that among low-quality fuels, the HICE response options differ greatly from those in the DHS, even though they must both be picking up virtually the same households. For example, 7.5 percent of people in the DHS live in households that mostly cook with "animal dung", but dung is not even an option in the HICE. Conversely, 11.1 percent of people in the HICE live in households that mostly cook with crop residue, but this response does not even appear in the DHS.

HICE: Does the household currently own mattresses and/or beds?
DHS: Does the household have a bed?

	% of people with a given response		
	HICE	DHS	Difference
No	44.9	40.8	4.1
Yes	55.1	59.2	-4.1

Given a rule-of-thumb that allows differences of up to 10 percent, this indicator is well-matched. Nevertheless, the differences in questionnaire wording would seem to favor

more affirmative responses in the HICE, but the data show the opposite. Still, this indicator is well-positioned to differentiate among the very poor and the merely poor.

HICE: Does the household currently own a radio?
DHS: Does the household have a radio?

	% of people with a given response		
	HICE	DHS	Difference
No	71.9	64.3	7.6
Yes	28.1	35.7	-7.6

This seemingly simple question is perhaps the least-well-matched of all the scorecard indicators. It is not clear—beyond differences in questionnaire context—why the rate of affirmative responses would be so different.

HICE: Does the household have any agricultural holdings?
DHS: Does any member of this household own any land that can be used for agriculture?

	% of people with a given response		
	HICE	DHS	Difference
No	11.6	15.7	-4.1
Yes	88.4	84.3	4.1

The differences here may well be due to differences in wording, as the DHS wording seems less likely to be interpreted as applying only to the household head or to the respondent.

HICE: Does the household currently own cattle?
Does the household currently own sheep and goats?
DHS: How many cattle does the household own?
How many goats does the household own?
How many sheep does the household own?

	% of people with a given response		
	HICE	DHS	Difference
No	24.1	30.5	-6.4
Yes	75.9	69.5	6.4

As with the previous asset indicators, the quality of the match is disappointing. Nevertheless, the indicator is used because—at least in the HICE—it helps distinguish among the poverty levels of the great mass of rural, agricultural households.

In the general, the indicators of asset ownership are kept because, without them, the scorecard would essentially identify the 10 percent or so of "elites" in Ethiopia, and estimate the poverty of the other 90 percent almost completely based on household size.

DETERMINANTS OF INTRA-INDUSTRY TRADE BETWEEN ZAMBIA AND IT'S TRADING PARTNERS IN THE SOUTHERN AFRICAN DEVELOPMENT COMMUNITY (SADC)

Mulenga Chonzi Mulenga[1]

Abstract

Intra-Industry trade (henceforth IIT) has generally been perceived to be a feature of the industrialized countries. As the past few years have seen a rapid increase in Zambia's trade with its trading partners in the Southern African Development Community (SADC), trade statistics reveal that a substantial part of such intra-SADC trade is in fact of the IIT form. This study seeks to establish the extent of IIT between Zambia and its trading partners in the SADC region and to identify the determinants of IIT at this level.

Using a modified gravity model in a panel data framework for the 1998-2006 period, the estimation results from the Feasible Generalized Least Squares in the random effects model evaluates the existence of IIT between Zambia and its trading partners in the SADC. The empirical results reveal that gross domestic product, dissimilarities in per capita income, transportation costs (distance and common border) and colonial ties (common language) are significant factors explaining IIT between Zambia and its trading partners in the SADC. The results also reveal that IIT between Zambia and its trading partners in the SADC is positively determined by GDP, distance, and dummies for common border and common language while dissimilarities in per capita income (DPCI) depresses it.

Key words: Intra-Industry Trade, Gravity Model, Feasible Generalized Least Squares in the Random Effects Model, Panel Data, Zambia, SADC Countries.
JEL Classification: F12

[1] Lecturer and Researcher in the School of Social Sciences at Mulungushi University in Zambia.

1. Introduction

International trade involves the exchange of various commodities between countries. There are two types of trade: intra-industry and inter-industry trade. Intra-industry trade (IIT) is the simultaneous import and export of products belonging to the same group, such as the two-way exchange of differentiated textiles or vehicles while inter-industry trade refers to trade in products that belong to different industrial groups, for instance the import of textiles and the export of maize.

While there are a number of studies[2] on developing countries' IIT, previously most trade studies placed greater emphasis on a country's comparative advantage as the basis of trade rather than on economies of scale. This tendency however, ignored the IIT theories which are important in understanding and analysing trade patterns between countries which are relatively similar and produce relatively similar products. IIT arises from the fact that countries try to take advantage of economies of scale in production and because of this it has generally been regarded as a way in which countries involved in trade stand to benefit. This can be achieved through increasing trade among them, and it is in this vein that many countries in the Southern African Development Community (SADC) have realized the potential benefits and have therefore advocated for its expansion.

It has been assumed that the degree of specialization in IIT is highly correlated with the level of a country's development. Therefore, since specialization mostly characterizes manufacturing goods and not primary commodity exports on which countries in the SADC are mainly dependent for their economic survival, IIT has generally been perceived to be a feature of the industrialized countries. However, trade statistics show that substantial part of the intra-SADC trade is in fact IIT. For instance in 2004, the G-L index[3] as calculated at a four digit Harmonised System (HS) code level, revealed that Zambia's top 15 categories of products had a G-L index above 0.6 in its trade with other countries in the SADC region except South Africa (TIPS[4], 2007).[5]

[2] Studies on developing countries' IIT include Aquino (1978), Balassa (1979), Havrylyshyn and Civan (1983), Manrique (1987), Lee and Lee (1993), Stone and Lee (1995), Gonzalez and Velez (1995), Havrylyshyn and Kuznel (1997), Hu and Ma (1999), and Nilsson (1999).

[3] The G-L index estimates the proportion of trade accounted for by IIT for an industry or sector.

[4] Trade and Industrial Policy Strategy

[5] A G-L index value of 0.6 means that the proportion of IIT is high.

This study in its own right tries to establish the extent of the existence of IIT between Zambia and its trading partners in the SADC region and to identify the determinants of IIT between Zambia and its trading partners in SADC. Although IIT is considered to have potential benefits in terms of improving a country's economic prospects, the only study on Zambia's IIT is a study by the Trade and Industrial Policy Strategies (2007). The study measured the existence of IIT in certain industries using the Grubel Lloyd Index and established that the country has substantial IIT in the manufactured copper products such as insulated wire, sugar, confectionery, cement, electric energy products in which the country may exhibit some comparative advantage. No study however, has been undertaken in Zambia to measure the determinants of IIT between Zambia and its trading partners; therefore this study tries to make a modest contribution to knowledge and to Zambia's IIT.

Studies that have attempted to identify the determinants of IIT can be divided into two groups: country-specific studies and industry-specific studies. The country-specific studies explain IIT through the macroeconomic variables in each country, such as per capita income, country size, distance, and trade orientation (DeRosa and Roningen, 2003). Industry-specific studies explain an industry's IIT as a function of industry-specific variables, such as scale variables, advertising/sales ratio and firm concentration ratio (*Ibid*). Some studies[6] have attempted to combine both country and industry variables to identify determinants of IIT. This study, however, employs the country variables using the gravity model of trade which explores the trade partner composition as well as the trade commodity composition. Despite the theoretical relevance and successful empirical performance of the gravity model[7], no studies have focused on Zambia's IIT using the model although there is strong evidence for increasing IIT among developing countries.

Many studies on IIT state that IIT is prevalent among countries with almost similar economic structures. One thing to note from theoretical and empirical studies involving the determinants of IIT among developing countries is that bilateral trade depends primarily on three variables – the size of an economy, the level of development and the geographical distance between economic centres (Verdoorn, 1960, Kimura and Lee, 2004). Most studies have paid insufficient attention to the role of other country-specific factors such as adjacency, historical ties, trade intensity and exchange rate.

[6] Verdoon (1960), and Damoense and Jordaan (2010)
[7] Chidoko *et. al.,* (2006)

This study is significant in the following aspects; by evaluating the existence of IIT, the study determines whether trade in actual fact takes place among countries with similar economic structures and therefore provides policy guidelines within SADC. Furthermore, by outlining the determinants of IIT between Zambia and its trading partners in SADC, this study sheds more light on how IIT is determined by various economic factors other than the size of an economy, level of development and the geographical distance between economic centres. Therefore, this study is expected to equip trade policy makers with some insights to design strategies for improvement of overall trade in the region, and more precisely Zambia's trade balance.

The remainder of this study is organized as follows: The next section discusses the theoretical foundation of the study. Section 3 discuses the nature and extent of IIT in Zambia. In section 4 the methodological approach, data and variables used for the analysis employed will be discussed. Section 5 will then discuss the estimated results from the specified model used for this study. Finally, section 6 summarizes the study with respect to the study's contribution to the literature of IIT.

2. Theoretical Framework

Although the factor endowment theories or any other comparative advantage theory predicted that trade involves the exchange of different products and should be greater the more the countries differed in their relative production possibilities, most of the enormous growth in trade in recent years came in relatively similar goods between relatively similar countries (Neary, 2009).

Failure of the traditional trade theories to explain IIT has seen the emergence of other theories of trade. One such model is the Linder hypothesis (1961). The model argues that countries with similar levels of per capita incomes have similar preferences and in turn trade more with each other in similar but differentiated productions (Montenegro and Soto, 1996). In his argument, Linder (1961) also looked at production quality as well as tastes as the main determinants for the basis and direction of trade with the proposition that a country will produce first for home consumption and the surplus for export to countries with similar preferences. The high income countries will have low income earners and low income countries will have high income earners. Thus, the low income country will produce low quality goods and export to the markets of the rich countries for low income earners and high income countries will produce high quality goods and export them to low income countries for high income earners. This

proposition cannot be dismissed as it is evident in the increasing bilateral trade patterns in the SADC.

Another theory giving an explanation of IIT is that formulated by Krugman (1979). This marked the birth of the New Trade Theories (NTT). The NTT explain world trade based on economies of scale, imperfect competition and product differentiation which relax the strict assumptions of Traditional Trade Theories of constant returns to scale, perfect competition and homogenous goods (Do, 2006). Under these new assumptions countries can specialize in producing a narrower range of products at larger scale with higher productivity and lower cost. Then it can also increase the variety of goods available to the consumer through trade. In short trade is likely to occur even when countries do not differ in their factor endowments or technology. Krugman (1979) made two simplifying assumptions: that consumers prefer a diverse choice of brands and that production favours economies of scale. He stated that the existence of differentiated products say different versions of a car can be explained by consumer's preference for diversity but because of economies of scale, it is not profitable to spread the production of one version all over the world. Therefore production will be concentrated in a few factories and therefore in a few countries. This logic gave an explanation of how each country specialised in producing a few brands of any given type of product and in essence IIT.

3. Nature And Extent of Intra-Industry Trade in Zambia

In 1991, Zambia autonomously liberalised trade. As a policy measure to promote exports, the Zambian government liberalised the capital and current accounts and restructured its tariff structure (TIPS, 2007). These policies had a mixed impact on the overall trends of Zambian trade. Despite the negative effects that have been associated with liberalisation, such as the collapse of the manufacturing industries, the country's trade has more than doubled over the period (TIPS, 2007). In terms of direction of merchandise trade, prior to liberalization, high income countries especially Europe and Asia absorbed more than 66 percent of Zambia's exports and were the source of over 60 percent of its imports. In that period SADC absorbed only 4 percent of Zambia's exports and supplied 8 percent of its imports (TIPS, 2008). Between 1995 and 2004 the situation changed as trade with the SADC region became so dominant that it outgrew its trade with the rest of the world, as shown in the table below.

Table 1: Zambia's Export and Import Destination by Region: 1995-2004.

Region	Exports (%)			Imports (%)		
	1995-1997	2000-2002	2004	1995-1997	2000-2002	2004
COMESA	8.50	15.20	13.40	13.40	4.70	6.00
SADC	12.00	36.80	48.10	48.00	75.10	58.90
EU	19.70	16.60	26.20	22.90	10.50	14.00
USA	4.50	1.80	2.80	4.90	2.20	2.00
ASIA	50.60	28.70	7.90	9.20	5.80	15.40
OTHERS	4.70	0.90	1.60	1.60	1.70	3.70

Source: (DTIS)[8], (CSO)[9]

Historically, Asia was the largest export market for Zambian commodities, however evidence from Table 1 shows that this is no longer the case as Zambia's exports have declined from 51 percent in the 1995-1997 period to 8 percent in 2004. On the import side, the EU secured 14 percent of Zambia's imports from that region in 2004, a reduction of 8 percent from 23 percent in the 1995-1997.

Table 1 shows Zambia's import and exports to various regions between 1995 and 2004. The table shows that by 2004 the SADC region supplied 59 percent of Zambia's imports and absorbed about 48 percent of its exports. The SADC region has over the past few years experienced increased volumes of trade with Zambia and has increasingly become important to Zambia as a market for both its non-traditional and traditional exports. The reduction in trade with the rest of the world (ROW) and the increase in trade with SADC give evidence of the occurrence of IIT as countries in the SADC are assumed to have similar economic structures as well as the same levels of development.

[8] Ministry of Commerce Trade and Industry (2005). Zambia: Diagnostic Trade Integration Study.
[9] Central Satistical Office-Department of External Trade.

Table 2: Zambia's Trade with SADC: 1998-2006 (Percentage and Total Volume)

Zambia Exports	1998	2003	2006	Zambia Imports	1998	2003	2006
Angola	0.27	0.27	0.06	Angola	0.01	0.00	0.00
Botswana	2.19	0.81	0.87	Botswana	0.79	0.41	1.05
DRC	18.65	8.71	18.77	DRC	0.01	1.18	1.74
Lesotho	0.00	0.04	0.66	Lesotho	0.00	0.43	0.00
Malawi	18.09	4.80	8.41	Malawi	0.32	0.93	0.74
Mauritius	0.02	0.88	0.10	Mauritius	0.32	0.21	0.14
Mozambique	0.35	0.14	0.19	Mozambique	0.08	0.93	0.94
Namibia	1.61	0.20	2.58	Namibia	0.23	0.35	0.47
RSA	36.44	48.57	8.60	RSA	75.10	73.55	81.72
Seychelles	0.00	0.00	0.00	Seychelles	0.13	0.00	0.01
Swaziland	0.02	0.06	0.13	Swaziland	1.17	0.00	0.30
Tanzania	14.26	31.17	2.02	Tanzania	2.03	2.31	3.09
Zimbabwe	8.10	4.35	7.56	Zimbabwe	19.80	19.48	9.80
SADC (US$m)	257.2	421.1	684.3	SADC (US$m)	571.5	1081.0	1750.0
ROW (US$m)	1025.8	980.8	3694.3	ROW (US$m)	1092.8	1518.9	2916.9
SADC in Total Trade (%)	25.07	42.93	18.52	SADC in Total Trade (%)	52.30	71.21	60.02

Source: CSO, SADC Trade Database

Table 2 shows the trends in Zambia's trade with its trading partners in SADC in the period 1998 to 2006. Zambia's trade with SADC continues to increase substantially as can be seen in Table 2. The total value of Zambia's exports rose from US$ 257.22 million in 1998 to US$ 421.31 million in 2003. In 2006, the exports rose even further to US$ 684.30 million, the share of Zambian exports to SADC as a proportion of total exports rose from 25 percent in 1998 to 43 percent in 2003 but later fell to 19 percent in 2006. This could be attributed to the impact of the global economic recession on trade. On the import side, the total value of its imports rose for US$ 571.51 million in 1998 to US$ 1 081 million in 2003 and then further rose to US$ 1 750 million in 2006. In terms of import shares to SADC as a proportion of total imports, they rose from 52 percent to 71 percent in 1998 and 2003, respectively before falling to 60 percent in 2006. Most of Zambia's imports from SADC came from three countries (RSA,

Tanzania and Zimbabwe) of which RSA is the largest, representing 82 percent in 2006. Generally this was attributed to RSA's competitive advantage in production, its capacity to export a wide range of products and the increased investment undertaken by RSA companies into the Zambian economy of total imports since 2003 (TIPS, 2007). Major products imported include iron, steel, vehicles, paper and paper products, industrial equipment, petroleum products, foodstuffs and beverages (UNCTAD, 2006). Zambia's export destination within the region as of 2006 was dominated by three SADC countries, RSA (59 percent), DRC (19 percent) and Zimbabwe (8 percent). The volume of trade sent from Zambia to RSA could be explained by high industrial activity in RSA, the short distance between the two countries and the preferential market access which, via the SADC Trade Protocol, allowed Zambia to export a wide range of products on a duty and quota free basis to that partner country (TIPS, 2007). South Africa remains the country's major trade partner within the region. Other important SADC trading partners are DRC, Tanzania, Zimbabwe, Malawi, Botswana and Namibia. Zambia's major export products to SADC are cotton, stock feeds, fresh vegetables, sugar and processed foodstuffs. Other major exports are; copper, scrap metal, wood and electricity (UNCTAD, 2006). Although Zambia's trade with SADC countries outside RSA is relatively small, recent developments reveal positive trends.

Zambia has witnessed an improvement in its economic growth over the past few years and this coincides with the substantial and increasing trade taking place with its trading partners in the SADC. This has resulted in an improvement in the economic performance of Southern Africa since the mid nineties. This improved economic performance of Southern Africa results also in part from better economic policies and structural reforms that led to an improvement of macro-economic indicators (reduced inflation rates, budget deficits). Progressive trade liberalization was also an important component of the opening up of the economies and of the strengthening of export performance.

4. Model Specification and Interpretation of Results

In establishing the determinants of Intra-Industry Trade this study applies the Gravity model which is a variation of the standard gravity model used by Chidoko, *et al.*, (2006) augmented by adding an extra dummy variable for common language. The theoretical foundation of the model is the Linder hypothesis which predicts that patterns of trade will be determined by the aggregated preferences for goods within countries. Matyas and Harris (1998) observed that the gravity model has performed particularly much

better than other trade models in analysing trade flows between countries and therefore has been deemed appropriate for policy analysis by most economists.

In estimating the determinants of IIT, a log-linear function is employed so as to make the estimates less sensitive to extreme observations as well as to enable interpretation of the coefficient terms as elasticities. The logarithmic transformation of the estimated model is as follows;

$$\begin{aligned}\text{Log IIT}_{ijk} &= \beta_0 + \beta_1\text{LogGDP}_k + \beta_2\text{LogPCI}_k + \beta_3\text{LogDPCI}_k + \beta_4\text{LogTI}_{jk} \\ &+ \beta_5\text{LogEXRT}_{jk} + \beta_6\text{LogDIST}_{jk} + \beta_7\text{D}_1 + \beta_8\text{D}_2 + \varepsilon_{ik}\end{aligned} \quad (4.1)$$

where;

i represents the industry.

j is the trading country, which in this study is Zambia.

k is the partner country.

The dummies are in linear form because they assume the values of zero or one.

β_0 stands for the country effects.

Definition and Measurement of Variables in the Model

Dependent Variable

In this study the dependent variable is the IIT Index as defined by Grubel and Lloyd (1975). The IIT index measures the proportion of IIT in total trade between Zambia and country *k* as a measure of the IIT_{ijk} and can be written as;

$$\text{IIT}_{ijk} = \left[1 - \frac{\sum |X_{ijk} - M_{ijk}|}{\sum (X_{ijk} + M_{ijk})}\right] * 100 \quad (4.2)$$

where;

IIT_{ijk} is the intra-industry trade index in industry i between Zambia and country *k*.

X_{ijk} are Zambia's exports of industry *i* to country *k*.

M_{ijk} are Zambia's imports of industry *i* from country *k*.

The dependent variable lies within the range of (0, 100), depending on the importance of IIT (Musonda, 1997).

Explanatory Variables and Expected Signs

EXPLANATORY VARIABLE	EXPECTED SIGN
Real Gross Domestic Product (GDP) GDP is a basic measure of a country's economic performance and is a proxy for economic size. It is hypothesised that the greater the economic size, the higher the IIT. Therefore GDP determines the level of international trade. In agreement with this, Filippini (2003) states that just as any other economic activity, trade will generally increase with an increase in the size of the economy. In this study GDP is measured in United States Dollars (USD$) and is expected to have a positive sign.	Positive (+)
Per Capita Income (PCI) Per Capita Income is simply the GNP per capita. It is calculated by dividing the total income of a country by its population. PCI measures the level of a country's economic development and is used in comparing levels of economic development between countries. It is believed that IIT with any given trading partner may tend to be higher as PCI of the partner country is higher since IIT is a phenomenon of countries with similar economic levels of development. In this study PCI is measured in constant base year prices denominated in United States Dollars (US$).	Positive (+)
Dissimilarity in Per Capita Income (DPCI) Dissimilarity in per capita income also known as the Linder term is simply the absolute difference between the PCI of the trading countries. It is defined as follows; $DPCI_{jk} = \left\lvert PCI_j - PCI_k \right\rvert$ (4.3) Where; $DPCI_{jk}$ is dissimilarity in per capita income between Zambia and partner country k. PCI_j is the PCI for Zambia. PCI_k is the PCI of the partner country. Linder (1961) and other researchers use dissimilarities in per capita income as proxies for consumer tastes and preferences. It has been argued that countries with similar levels of PCI will have similar tastes and will produce similar but differentiated products and therefore will tend to trade more among themselves.	Negative (-)

Distance (DIST) — **Negative (-)**

Distance is the geographical distance between the economic centres of trading partners; it is a proxy for transport costs. The distance used in this study is the actual road distance between capital cities of trading countries measured in kilometres. The distance between capital cities of trading countries is likely to affect the search and transaction costs. This will in turn affect the bilateral trade as larger distances tend to be associated with greater costs. Therefore, the longer the distance, the lower the IIT between countries expected.

Trade Intensity (TI) — **Positive (+)**

Trade intensity measures the degree of trade between the two partner countries. It is hypothesised that the higher the trade intensity between trading partners, the greater the IIT. Therefore, as two countries engage in more and more trade, the level of IIT is believed to increase. It is given as follows;

$$TI_{jk} = \frac{X_{jk} + M_{jk}}{GDP_j} \qquad (4.4)$$

where;

TI_{jk} = Trade intensity between Zambia and partner country k.

X_{jk} = Zambia's exports to partner country k.

M_{jk} = Zambia's imports from partner country k.

GDP_j = Zambia's gross domestic product.

Real Exchange Rate (EXRT) — **Negative (-)**

The study uses the real exchange rate between trading partners which is calculated as follows;

$$RER_{jk} = E_{jk} \times \frac{P_k}{P_j} \qquad (4.5)$$

where;

RER_{jk} = Real exchange rate between Zambia and trading partner k.

E_{jk} = is the nominal exchange rate between Zambia and trading partner k.

P_j = Zambia's GDP deflator.

P_k = GDP deflator for the trading partner.

The real exchange rate is used because it gives a measure of an economy's competitiveness in terms of exports and imports and because it also takes into account the real as well as the nominal price changes. Empirically, it has been shown that the exchange rate in gravity type studies has been significant in explaining trade variations among countries involved in trade (Do, 2010). Appreciation of the Zambian kwacha makes exports more expensive while imports become cheaper thereby discouraging IIT.	
Common Border (D1) The dummy variable for common borders represents SADC countries with a common border with Zambia. The existence of common borders represents the possibilities of IIT in response to locational advantages (Balassa and Bauwens, 1987). Therefore, Cēterīs paribus, IIT between countries which share a common border is likely to be higher than between countries which do not share a border. $D_1 = \begin{cases} 1 & \text{if countries share a common border} \\ 0 & \text{otherwise} \end{cases}$	**Positive (+)**
Common Language (D2) The existence of a common language in both trading countries is likely to enhance a flexible flow of information and lower transaction costs, therefore increase IIT between the countries. Common language is measured as a dummy variable which is defined as follows: $D_2 = \begin{cases} 1 & \text{if countries use a common language} \\ 0 & \text{otherwise} \end{cases}$	**Positive (+)**

Estimation Technique

The model is estimated using a panel data framework in Stata. The use of panel data methodology in this study can be justified based on its advantages;

- Panel data analysis allows control of heterogeneity of cross-sectional units.
- Generates more variability, more degrees of freedom and at the same time reduces multicollinearity problems thereby improving the efficiency of the econometric estimates.

It should also be noted that panel data may lead to inconsistent estimates because it may be affected by problems of non-stationary time series, however, these problems are usually of concern when the time series is lengthy. This study uses a short time series of 9 years, therefore, panel data unit root tests and panel data cointegration tests will not be carried out.

Estimation Models

There is a distinction in the literature between static and dynamic panel data models. Static panel data models include the fixed effects and the random effects methods, while dynamic panel data models are those that include a lagged dependent variable as an explanatory variable. This study, however, considers the static panel data models as opposed to the dynamic panel data models because in the dynamic panel data models, the lagged dependent variable is correlated with the error component which complicates estimation and therefore yields biased and inconsistent estimates. Static panel data regression models can be estimated using pooled estimation, fixed effects and random effects (Asteriou, 2006).

This study uses the random effects model as opposed to pooled and the fixed effects estimation methods. The reasons for this model choice are the following: Firstly, the pooled estimation method has a tendency of giving biased results by ignoring country effects. Secondly, the fixed effects estimation method does not take time invariant variables such as distance, common border and common language into account therefore rendering the Hausman Specification test inappropriate to this study. Lastly, the use of a dummy for each cross-sectional unit in the fixed effects model creates losses in degrees of freedom.

Given the results of Appendices 2 and 3, which show that the disturbance variance of the country-specific effects varies across countries (Heteroskedastic) and the errors are serially correlated over time, it is important to control for both Heteroskedasticity and Autocorrelation. Therefore, in order to obtain consistent and efficient estimators the model is estimated by Feasible Generalized Least Squares (FGLS) in the random effects model. The assumption behind FGLS is that all aspects of the model are completely specified; here that includes that the disturbances have different variances for each panel and are constant within panel. The advantage of FGLS estimation in the random effects model is that it is able to handle both Heteroskedasticity and serial correlation.

Data Type and Sources

This study makes use of secondary data for the years 1998 - 2006 and covers 21 sectors of commodities which Zambia trades with its partners in SADC. The sample contains 11 countries in SADC which include; RSA, Zimbabwe, Malawi, Botswana, DRC, Tanzania, Namibia, Angola, Mozambique, Mauritius and Swaziland. The choice of countries was influenced by the availability of data for the variables used in the model as well as whether the commodities exhibit IIT. The data was obtained from various sources including: The Department of External Trade, the Zambian Central Statistical Office (CSO), World Development Indicators and the Penn World website: www.pennworld.com. Other sources include the International Monetary Fund, World Economic outlook database, the SADC trade database and the World Bank-World Trade Indicators (2008).

5. Results and Discussions

Diagnostic Test Results

Testing for Multicollinearity using the Correlation Matrix, the results in Appendix 1A show that PCI and DPCI were highly collinear (0.88) thus the need to correct for Multicollinearity by dropping one of the collinear variables. In this study PCI is dropped and the model is run using DPCI.

Results from the likelihood ratio test for Heteroscedasticity shown in Appendix 2 indicate the presence of Heteroskedasticity across panels. Since the presence of Heteroskedasticity across panels may lead to estimates that are consistent but not efficient, it is taken into account by the use of Feasible Generalised Least Squares (FGLS).

The study tested for Autocorrelation using the Wooldridge test for Autocorrelation in panel data and the results are presented in Appendix 3. The null hypothesis of no first order Autocorrelation was rejected at all levels of significance in favor of the alternative hypothesis of first order Autocorrelation. Since Autocorrelation is regarded as a very big problem it has to be corrected (Woodridge, 2002), in this study autocorrelation is corrected by the use of FGLS.

Regression Results and Interpretation

After dropping PCI the empirical results from the regression using Feasible Generalized Least Squares (FGLS) in the random effects model are reported in Table 3. PCI is dropped in order to control for Multicollinearity. With an overall R^2 of 0.778, it shows that the model is a good fit as approximately 78 percent of the variations in the dependent variable are explained by the independent variables. The descriptive statistics in Appendix 4 show that the dependent variable had the highest level of variability in the variables with a standard deviation of 2.2, while the dummy variables had lower standard deviations of 0.45 and 0.48 for D1 and D2 respectively, meaning that the variability was lower in the variables.

Table 3: FGLS Regression Results Table after dropping PCI

Variable	Coefficient	Standard Error	Prob. > |z|
LogGDP	0.9176383	0.2270798	0.000***
LogDPCI	-0.6029963	0.3083821	0.051*
LogEXRT	-0.0971468	0.1054887	0.387
LogDIST	1.165163	0.7008871	0.096*
LogTI	0.1633474	0.1033916	0.114
D1	3.938728	0.5812316	0.000***
D2	3.969157	1.002791	0.000***
Constant	-28.06041	8.244858	0.001***

denotes significance at 10%, ** denotes significance at 5%, * denotes significance at 1%.*
No. of observations = 99, No. of groups = 11, Time periods = 9, Prob > chi2 = 0.0000, R-sq: overall = 0.7789

Using a single equation model as specified in equation 5.1, the results show that after dropping PCI all the variables are significant with the exception of LogEXRT and LogTI. LogEXRT and LogTI have the expected signs however. The empirical result of LogEXRT suggests that fluctuation of the Zambian Kwacha has not supported IIT. Since exchange rate liberalization, the Zambian Kwacha as compared to other currencies has been unstable; this implies that the effect of the change in the exchange rate on imports and exports have been cancelling each other, thereby having no effect on IIT.

The study establishes the extent of the existence of IIT between Zambia and its trading partners in SADC and the estimation results reveal that economic size (GDP), dissimilarities in per capita income (DPCI), transportation costs (distance and common

border) and colonial ties (language) are significant factors in explaining IIT between Zambia and its trading partners in the SADC. The findings of this paper are consistent with other empirical studies[10] in explaining IIT using the gravity model.

GDP is found to be statistically significant at 1 percent and positively related to IIT, which suggests that the larger the size of the economy the larger the IIT to be conducted. The results show that an increase by 1 percent of Zambia's trading partner's GDP will increase the proportion of IIT between that trading partner and Zambia by 0.91 percent. The intuition behind this pattern is that, the larger the size of the economy, the larger the opportunities for production of differentiated goods under conditions of economies of scale and therefore the greater the demand for foreign differentiated goods in these economies. This leads to larger opportunities for trade in these goods. Zambia has shifted its trade from the European Union (EU) and Association of South East Asian Nations (ASEAN) to countries in the SADC as these countries have similar economic structures and therefore produce and trade in similar but differentiated goods (TIPS, 2008). This has led to increased production and trade in the economies for instance the increased trade flows between Zambia and RSA that have been recorded in recent years. Since RSA is a large economy, the opportunity to produce differentiated goods under economies of scale is large and therefore its demand for foreign differentiated goods from Zambia has been high leading to increased IIT between the two countries. This finding is in line with the findings of Balassa (1986) and those of Clark and Stanely (1999).

The Linder hypothesis states that countries with similar levels of PCI will have similar demand structures and will produce similar but differentiated products and therefore trade more among themselves. The Linder term in this study which is represented by Dissimilarities in Per Capita Incomes between Zambia and its trading partners is found to be consistent with the Linder theory. DPCI is found to be weakly significant and negatively related to IIT, which generally suggests that as countries become similar in their income levels, IIT becomes more pronounced. The results show that a 1 percent increase in the DPCI of trading partners will reduce the proportion of IIT by 0.60. This result shows the wider the gap in the resource endowments or demand structures of trading partners the lower the IIT. Therefore economies which share a lot in common economically will conduct more IIT as compared to those that have little or nothing in common. A study by Ekanayake (2001) shows that if PCI is interpreted as

[10] Balassa (1986), Clark and Stanely (1999), Ekanayake (2001), Chidoko, et al., (2006) and many others.

an indicator of demand structure, a greater difference in PCI implies that demand structures have become more dissimilar which indicates that the potential for IIT decreases. The explanation to this is that, for trade to exist between two countries there must be in each country a demand for differentiated products produced by the other country. Therefore, when the gap between the PCIs of the two trading partners widens, the scope of IIT tends to lessen. This finding conforms to the findings of Balassa (1986).

The estimated coefficient for DIST is found to be weakly significant and positively related to IIT. The positive sign indicates that Zambia's IIT is more pronounced with countries that are geographically further from it. This result is not in conformity with the earlier expectation that long distance discourages IIT and is in contrast to Balassa (1986) who argued that IIT will tend to be greater when trading countries are geographically close to each other. The major explanation to this could be attributed to the fact that despite the large geographical distance between Zambia and RSA, Zambia tends to conduct more trade with RSA which is further away as compared to other countries which are geographically closer. Being a landlocked country, Zambia's cheapest mode of conducting trade is through overland transportation, in particular road transport. Therefore this result could be influenced by the large trade volumes between Zambia and RSA which could be as a result of the good road infrastructure between the two countries.

In line with the findings of Grubel and Lloyd (1975) who suggested that in sharing a common border, IIT may take place in products that are functionally homogenous but differentiated by location. This study reveals that the estimated coefficient for common border is strongly significant and has the anticipated positive sign. The result shows that countries that share a common border tend to trade more than those that do not because the geographical distance between the two countries sharing a border will be relatively shorter. This in essence means that transport costs will be reduced significantly if Zambia conducts more trade with countries geographically close to her as compared to countries geographically further from her. However, for this result to have intuitive appeal there should be economic complementarity between the two trading partners involved in trade. Countries in SADC usually lack complementarity and this could be attributed to the dominance of one or two commodities in the export baskets of partner SADC countries. This finding however, shows that there exists economic complementarity between Zambia and its trading partners in SADC.

The language dummy is found to be strongly significant and has the expected positive sign. The language dummy represents the 11 SADC members used in this study with colonial ties to Zambia. The language dummy essentially indicates how colonial ties influence the magnitude of IIT. The result suggests that the seven (7) countries used in this study that have English as their official language conduct more IIT as compared to the four (4) non-English speaking countries in this study. The explanation to this could be that the existence of common language will contribute to freer information flows (Balassa and Bauwens, 1987, Stone and Lee, 1995) and therefore is expected to enhance IIT. This finding is consistent with the findings of Ekanayake (2001).

6. Summary and Conclusions

Summary of Results

The main objective of the study was to establish the extent of the existence of IIT between Zambia and its trading partners in the SADC region and to identify the determinants of IIT between Zambia and its trading partners in the SADC. In a panel data framework the study used the Feasible Generalized Least Squares in the random effects model to estimate the gravity equation covering a period of 9 years from 1998 to 2006. Although the gravity model has been criticised for being ad hoc and lacking theoretical foundation, this study reveals that it is an important empirical tool in explaining trade flows as it has been able to evaluate the existence of IIT between Zambia and its trading partners in SADC as well as to establish the determinants of this trade.

The empirical results establish the extent of the existence of IIT between Zambia and her trading partners in the SADC and reveal that apart from the common gravity equation variables (GDP, PCI and DIST), IIT between Zambia and her trading partners in SADC is also determined by other variables such as DPCI, common border and common language. The results further reveal that GDP, DIST, Common Border and Common Language have a positive impact on IIT, while DPCI depresses it. EXRT and TI, however, seem to have no effect on IIT between Zambia and its trading partners in the SADC as they are found to be statistically insignificant although with the anticipated signs.

Conclusion of the Study

Global trends reveal that IIT has gained ground in world trade and in this regard Zambia has not been an exception. Over the years, Zambia's trade with other countries in the SADC has been on the rise, trade statistics show that substantial part of the intra-SADC trade is in fact IIT. For instance in 2004, the G-L index as calculated at a four digit Harmonised System (HS) code level, revealed that Zambia's top 15 categories of products had a G-L index above 0.6 in its trade with other countries in the SADC region except South Africa (TIPS, 2007). This is surprising considering that countries in the SADC region have similar economic and productive structures (except RSA) therefore tend to produce and trade in similar but differentiated goods within the same industry.

The contributions of this study can be stated as follows; Firstly, the results suggest that IIT between Zambia and its trading partners increases, the larger the economic size (GDP) of a country. This means that economic growth will strongly affect trade relationships, that is to say IIT between Zambia and its trading partners in SADC is likely to expand as the economies become larger. Secondly, the results show that similarities in per capita income is a very important aspect in increasing IIT between Zambia and its trading partners in the SADC. Therefore, if Zambia is to increase IIT and maximize her gains from this kind of trade, she has to engage more in trade with countries with similar per capita incomes. Thirdly, in order to expand IIT, Zambia has to trade more with her neighbours and this is evident from the large and significantly positive effect of the coefficient of the common border variable. Fourthly, historical ties have been found to have a very important role to play in expanding IIT between Zambia and its trading partners in SADC. Although the results suggest that Zambia should engage more in trade with other former British colonies because of the easy information flows. Doing so, however, would limit Zambia's trade within the region and thereby affect IIT considering the fact that there has been increased trade activity in countries like; Angola, DRC, Tanzania and Mozambique which are not former British colonies.

While many studies[11] on developing countries have found the exchange rate to be a significant factor in explaining IIT, this study however finds that in the case of Zambia, the exchange rate though having the anticipated sign is insignificant. This suggests that

[11] E.g. Chidoko, *et al.* (2006), Do (2006), Simwaka (2006)

the exchange rate has not supported IIT. This finding can be explained by the fact that the Zambian Kwacha has constantly been appreciating and depreciating ever since it was liberalized. Currency appreciation causes exports to be more expensive and imports to be cheaper while currency depreciation causes imports to be more expensive while exports become cheaper. Therefore, exchange rate instability does not support IIT because the effects of the change in the exchange rate on imports and exports tend to cancel each other out. In this regard the real exchange rate cannot be used as a determinant of IIT in a country with an unstable exchange rate.

In addition, the study finds distance to be a significant factor in explaining IIT. This means that the distance between trading centres is a very important factor in explaining trade.

Lastly, for a very long time IIT has been perceived to be a feature of developed countries however, this study shows that IIT is a feature of both the industrialized countries as well as developing countries; this finding is confirmed by the significance of the dissimilarities in per capita income (DPCI) variable.

Policy Implications

Trade is considered as a very important aspect in the economic performance of a country. It is for this reason that it is important to investigate IIT, for this may be an area where substantial benefits could be reaped if properly nurtured. Therefore, there is need for policy to be aimed at expanding it in order to improve a country's economic prospects. The results reveal that IIT does in actual fact exists, therefore since this trade is beneficial to the country, there is need to direct efforts to expand this form of trade. This can be achieved through paying particular attention to the determinants of IIT as established by the gravity model in this study. Firstly, economic size (GDP) has been found to be one aspect that can increase IIT. Therefore policy must be aimed at encouraging economic growth and this can be achieved through expanding the production sectors of the economy. Expansion of the productive sectors entails an expansion in the production of goods and services and therefore leads to an increase in income (Gross Domestic Product and Per Capita Income). In order to achieve this, this paper recommends that policy makers put in place stabilization policies and an attractive business environment which will attract Investment and will therefore contribute to a higher growth rate in the economy. This study also recommends that Zambia promotes and maintains good relations with its neighbours as well as countries

with which it has historical ties with. This has potential benefits in terms of reducing transaction costs because of closeness. The other recommendation is that Zambia enters into bilateral trade agreements with her neighbours as this would result in the elimination of trade barriers and therefore enable reciprocal non-trade barrier trade between her and her neighbours. Distance is also an important determinant of IIT between Zambia and its trading partners in the SADC. As many countries in the SADC are landlocked; one of the most important features of trade in the SADC is that it is dominated by road transport. Road transport is Zambia's main link to other countries in the SADC, therefore improvement in the road infrastructure as well as reduction in the delays at border posts would be necessary steps to the expansion of IIT within the region. Improvement of the road network is particularly beneficial to the country in terms of increased export earnings to countries like DRC, Angola and Zimbabwe which have in recent years experienced growing demand for consumer goods.

A key objective of the Government is to reposition the economy with a view to take advantage of the rebound in global economic activity and trade. The promotion of trade is integral to Zambia in its efforts to find additional regional and international markets for its products. Zambia has continued to maintain a liberal trade policy regime aimed at enhancing productivity and competitiveness of Zambian products in both the domestic and international markets. The main objective of Zambia's trade policy is to contribute to economic growth and national development through the creation of viable and competitive export sectors in the economy:[12] this led to the formation of the Zambia Development Agency (ZDA). The policy seeks to achieve this objective by directing resources to the most productive areas for export production, therefore, this study can act as a guide to policy makers as they formulate National Development Plans (NDPs) in terms of ways of fostering economic growth and development in Zambia through the promotion of IIT with its trading partners in SADC.

[12] This objective has been enshrined in key national policy documents such as the Commercial Trade and Industrial Policy (CTIP), the Fifth National Development Plan (FNDP) and the Vision 2030, which articulate the country's long term development objectives (Katotoka, 2010)

References

Aquino, A. (1978). "Intra-Industry Trade and Inter-Industry Specialization as Concurrent Sources of International Trade in Manufactures," *Weltwirtschaftliches Archiv,* Vol. 114, No. 2, pp. 275–296.

Asteriou, D. (2006). *Applied Econometrics: A Modern Approach Using EViews and Microfit*, Palgrave Macmillan: Hampshire and New York.

Balassa, B. (1979). *Intra-Industry Trade and the Integration of Developing Countries in the World Economy,* in H. Glersch (ed.), *On the Economics of Intra-Industry Trade*, pp. 245-270: Tubingen.

Balassa, B. (1986). "The Determinants of Intra-Industry Specialization in United States Trade," *Oxford Economic Papers*, Paper No. 38, pp. 220-223.

Balassa, B. and Bauwens, L. (1987). "Intra-Industry Specialization in a Multi-Country and Multi-Industry Framework," *The Economic Journal*, Vol. 97, No. 388, pp. 923–939.

Chidoko, C., Zivanomoyo, J. and Sunde, T. (2006). "Determinants of Intra-Industry Trade between Zimbabwe and it's Trading Partners in the Southern African Development Community Region (1990-2006)," *Journal of Sciences*, Vol. 5, No. 1, pp. 16-21.

Clark, D. P. And Stanley, D. L. (1999). "Determinants of Intra-Industry Trade between Developing Countries and the United States," Journal of Economic Development, Vol. 24, No. 2, pp. 79-92.

Damoense, M. Y. and Jordaan A. C. (2010). "The Determinants of Bilateral Intra-Industry Trade: An Application to South Africa's Automobile Industry," *Paper presented at the 14th Annual Africa Econometrics Conference held at the Sheraton Hotel, Abuja, Nigeria, 8-10 July.*

DeRosa, D.A. and Roningen, V.O. (2003). *Zambia in Regional and Extra-Regional Free Trade Agreements: Estimates of the Trade and Welfare Impacts*, VORSIM/ Potomac Associates: Virginia.

Do, T. T. (2006). "A Gravity Model for Trade between Vietnam and Twenty-Three European Countries," Unpublished Thesis, Department of Economics and Society: Hanoi.

Ekanayake, E. M. (2001). "Determinants of Intra-Industry Trade. The case of Mexico, " *The International Trade Journal*, Vol. 15, No. 1, pp. 89-112.

Filippini, C. (2003). *The Determinants of East Asian Trade Flows: A Gravity Equation Approach*, Bocconi University: Milan.

Gonzalez, J. G. and Velez, A. (1995). "Intra-Industry Trade between the United States and the Major Latin American Countries: Measurement and Implications for Free Trade in the Americas," *The International Trade Journal,* Vol. 9, No. 4, pp. 519–536.

Grubel, H. and Lloyd, P. (1975). *Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Product,* McMillan Publishers: London.

Havrylyshyn, O. and Civan, E. (1983). *Intra-Industry Trade and the Stage of Development*, in P.K.M. Thakaran (ed.), *On Intra-Industry Trade: Empirical and Methodological Aspects*, pp. 111–140, Elsevier: Amsterdam.

Havrylyshyn, O. and Kuznel, P. (1997). "Intra-Industry Trade of Arab Countries: An Indicator of Potential Competitiveness," IMF Working Paper No. 97/ 47.

Hu, X. and Ma, Y. (1999). "International Intra-Industry Trade of China," *Weltwirtschaftliches Archiv*, Vol. 135, No. 1, pp. 82–101.

Katotoka, H. (2010). "Reflections on Evidence Based Trade Policy Formulation in Zambia," Paper Presented at the SADRN Workshop for Policy Makers: Pretoria.

Kimura, F. and Lee, H. (2004). "The Gravity Equation in International Trade in Services," *Review of World Economics Journal*, Vol. 142, No. 1, pp. 92-121.

Krugman, P. (1979). "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, Vol. 9, No. 4, pp. 469–479.

Lee, H. and Lee, Y. (1993). "Intra-Industry Trade in Manufactures: The case of Korea," *Weltwirtschaftliches Archiv,* Vol. 129, No. 1, pp. 159–169.

Linder, B. S. (1961), *An Essay on Trade and Transformation,* Almqvist & Wiksell: Stockholm.

Manrique, G. (1987). "Intra-Industry Trade between Developed and Developing Countries: The United States and the NICs," *Journal of Developing Areas*, Vol. 21, No. 4, pp. 481–493.

Matyas, L. and Harris, M. N. (1998). "The Econometrics of Gravity Models," Melbourne Institute Working Paper No. 5/98.

Ministry of Commerce (2005). "Zambia Diagnostic Trade Integration Study," World Bank Sponsored Study: Lusaka, Downloadable from http://www.integratedframework.org/countries/zambia.htm.

Montenegro, C. E. and Soto, R. (1996). "How Distorted is Cuba's Trade? Evidence and Predictions from a Gravity Model," *Journal of International Trade and Economic development,* Vol. 5, No. 1, pp. 45-68.

Musonda, F. M. (1997). "Intra-Industry Trade between Members of the PTA/COMESA Regional Trading Arrangements," *AERC Research Paper* No. 64, pp. 10-17.

Neary, J. P. (2009). *Putting the 'New' into New Trade Theory: Paul Krugman's Nobel Memorial Prize in Economics,* Oxford University Press: London.

Nilsson, L. (1999). "Two-Way Trade between Unequal Partners: The EU and the Developing Countries," *Weltwirtschaftliches Archiv*, Vol. 135, No. 1, pp. 102–127.

Simwaka, K. (2006). "Dynamics of Malawi's Trade Flows: A Gravity Model Approach," Munich Personal RePEc Archive, MPRA Paper No. 1122.

Stone, J. A. and Lee, H. (1995). Determinants of Intra-Industry Trade: A longitudinal and Cross-Country Analysis, *Weltwirtschaftliches Archiv,* Vol. 131, No. 1, pp. 67–83.

Trade and Industrial Policy Strategies (2007). Zambia: Intra–SADC-Trade Performance Review; 2000-2005: Lusaka.

Trade and Industrial Policy Strategies (2008). Zambia: Intra–SADC-Trade Performance Review; 2006-2008: Lusaka.

United Nations Conference on Trade and Development (UNCTAD). (2006). Zambia and the Multilateral Trading System: The Impact of WTO Agreements, Negotiations and Implementation, United Nations Publications: Geneva and New York.

Verdoorn, P. J. (1960). The Intra-Bloc Trade of Benelux, In E. Robinson (ed.), *Economic Consequences of the Size of Nations*, Macmillan: London.

Woodridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data,* MIT Press: London.

Appendices

Appendix 1: Correlation Matrices

Appendix 1A: Correlation Matrix with LogPCI
(obs=99)

	LogIIT	LogGDP	LogPCI	LogDPCI	LogEXRT	LogDIST	LogTI	d1	d2
LogIIT	1.0000								
LogGDP	0.2071	1.0000							
LogPCI	0.0737	0.3806	1.0000						
LogDPCI	0.0353	0.3372	0.8836	1.0000					
LogEXRT	0.0229	0.0896	0.4185	0.3502	1.0000				
LogDIST	0.3477	0.0846	0.2515	0.2894	0.3377	1.0000			
LogTI	0.2658	0.3671	0.0862	0.1580	0.0737	-0.0531	1.0000		
d1	0.3183	0.2346	0.5493	0.5338	-0.1208	-0.332	-0.4139	1.0000	
d2	.1755	0.0697	.5712	0.5841	0.4506	-0.3624	0.3568	-0.4629	1.0000

Appendix 1B: Correlation Matrix after droping LogPCI
(obs=99)

	LogIIT	LogGDP	LogDPCI	LogEXRT	LogDIST	LogTI	d1	d2
LogIIT	1.0000							
LogGDP	0.2071	1.0000						
LogDPCI	-0.0353	0.3372	1.0000					
LogEXRT	0.0229	0.0896	0.3502	1.0000				
LogDIST	-0.3477	0.0846	0.2894	-0.3377	1.0000			
LogTI	0.2658	0.3671	0.1580	-0.0737	-0.0531	1.0000		
d1	0.3183	-0.2346	-0.5338	-0.1208	-0.3321	-0.4139	1.0000	
d2	.1755	-0.0697	0.5841	0.4506	-0.3624	0.3568	-0.4629	1.0000

Appendix 2: Likelihood Ratio Test for Heteroscedasticity

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares
Panels: heteroskedastic
Correlation: no autocorrelation

Estimated covariances	=	11	Number of obs	=	99
Estimated autocorrelations	=	0	Number of groups	=	11
Estimated coefficients	=	8	Time periods	=	9

R-sq: overall = 0.7789
Wald chi2(7) = 87.79
Prob > chi2 = 0.0000

LogIIT	Coef.	Std. Err.	z	P>|z|	[95% Conf	Interval
LogGDP	1.00137	.1674127	5.98	0.000	.673247	1.329493
LogDPCI	-.5015895	.1545782	-3.24	0.001	-.8045572	-.1986219
LogEXRT	-.0707713	.0477486	-1.48	0.138	-.1643568	.0228143
LogDIST	1.152163	.3801933	3.03	0.002	.4069975	1.897328
LogTI	-.0029988	.1091277	-0.03	0.978	-.2168852	.2108876
d1	3.367093	.4506935	7.47	0.000	2.48375	4.250436
d2	3.737611	.5863582	6.37	0.000	2.58837	4.886852
_cons	-29.98961	4.666829	-6.43	0.000	-39.13643	-20.8428

Appendix 3: Wooldridge Test for Autocorrelation in Panel Data

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
F(1, 10) = 34.691
Prob > F = 0.0002

Appendix 4: Descriptive Statistics

	Obs	Mean	Std. Dev.	Min	Max
logiIt	99	1.877805	2.229945	-7.145463	4.268123
loggdp	99	22.67257	1.206293	20.89851	26.2645
logdpci	99	6.597309	1.321443	3.870002	8.521646
logexrt	99	5.132428	1.970125	.375566	11.64926
logdist	99	7.130415	.5129907	6.075786	8.049336
logti	99	-.5430343	2.295744	-6.93765	3.236551
d1	99	.7272727	.4476283	0	1
d2	99	0.6363636	.4834938	0	1