SCOPE ECONOMIES IN MICROFINANCE: EVIDENCE FROM RATED MFIS

VALENTINA HARTARSKA, CHRISTOPHER F. PARMETER, AND ROY MERSLAND

Abstract. While economies of scope of lending and mobilizing deposits in banking are justified theoretically (Diamond 1984) and found empirically (Sounders 1999), in microfinance, the existence and the magnitudes of scope economies has not been investigated. We use a recently developed semiparametric smooth coefficient model to estimate a generalized cost function to provide estimates of scope economies using an authoritative data set from rated MFIs ranging over 50 countries. We find substantial scope economies in general, 22 percent on average, as well as evidence that economies of scope vary across the type of services and areas where the MFIs operate, suggesting key insights for policy prescriptions.

1. Introduction

Microfinance emerged as an innovation in lending to the rural poor in Asia, where previous interventions in rural financial markets (directed and subsidized production credit usually disbursed by agricultural development banks) often failed. While it started as "a collection of banking practices built around providing small loans (typically without collateral) and accepting tiny savings deposits" Armendariz de Aghion & Morduch (2005, p. 1), today many microfinance institutions (MFIs) expand their services and strive to offer payment and savings facilities, insurance, housing, and longer-term loans. Moreover, MFIs today serve marginalized clientele in rural and urban settings. The goals of microfinance are to reduce poverty by promoting self-employment and entrepreneurship as well as by alleviating liquidity constraints and contributing to consumption and income smoothing.

Estimates show that there are at least 10,000 microfinance programs worldwide, with the majority still relying on traditional multilateral and bi-lateral donors for funds and most continue to focus on lending. However, two related and important trends are now emerging. The first is toward commercialization which essentially is transforming the NGO-MFIs into regulated intermediaries with the intent to lower costs by accessing deposits as well as strengthen the organizations by privatizing it so that it is better monitored by the new owners and the regulator. The second trend is a renewed interest in experimentation to mobilize savings especially rural savings. These developments should be based on observed economies of scope of extending loans and mobilizing savings but there are no studies that estimate the magnitude and sign of such economies. We present the first evidence using a large sample of MFIs.

Date: June 25, 2009: Preliminary and Incomplete – Not to be quoted without the author's permission.
Key words and phrases. Microfinance Institution, Scope Economies, Bandwidth Selection, Semiparametric Smooth Coefficient.

Understanding if the current trends of attention to savings are justified is important because funding by 'traditional' donors remains a non-trivial quantity. It is estimated that annual spending on microfinance is between US$ 800 million and 1.5 billion. Bi- and multilateral agencies account for the bulk of quasi-commercial investment in microfinance organizations through equity, guarantees, and quasi-commercial debt ((Hartarska & Holtmann 2006). These donors are becoming more coordinated in their funding and offer less funds for individual MFI experimentation with various financial products.

In addition, private foundations such as the Bill and Melinda Gates Foundation and the Dell Foundation play an increasingly active role as both donors and investors. Large private banks such as CITIBANK and HSBC have also entered what used to be a niche market (WSJ May 15, 2006). A 2008 survey by the Consultative Group to Assist the Poor (CGAP) reports that the assets under management by various Microfinance Investment Vehicles, which intermediate between foreign investors and MFIs has grown significantly to 5.4 billion at the end 2007 which was 78% increase from 2006. Most of the growth has come from private institutional investor such as TIAA-CRAFT and ABP. These (quasi-) commercial investors may fund large international microfinance networks (e.g. ACCION International, FINCA, Opportunity International), individual MFIs, or may adopt a "greenfielding" approach and create new micro and small enterprise banks. Thus, analysis of the possible efficiency gains or losses of intermediation as well under what circumstances they are realized is timely and important.

Previous studies that explore the productivity and efficiency of organizations providing microfinance are predominantly case studies describing the experience and performance of a single MFI or a cross section from group of MFIs operating in one country or in similar markets (for example Leon, 2008; Paxton, 2007). While it is useful to conduct case studies to gain insight into particular situations, it is also important to look at many institutions to make broad comparisons across the MFI population.

We are the first to estimate scope economies in microfinance. We offer two innovations in this paper. The first is that we use a non-parametric method which avoids some of the pitfalls of traditional scope economies estimation methods. It avoids the "flip- flop" problem from parametric estimation methods resulting from substitution of small values for observations with a zero output or a zero input price. Estimation results from substitution with one set of small numbers can differ significantly from estimation results from substitution with another set of small numbers, thus scope economies estimation are dubbed "flip- flop." In microfinance, the potential for such mistakes is big because most MFIs only extend loans, and so would have a zero output for savings, and many have the cost of funds completely subsidized, thus the actual input price of financial capital is zero.

In addition, the method applied here accounts for challenges typical for the microfinance industry, specifically the fact that MFIs operate in diverse environments and use diverse lending methodologies. The non-parametric method permits environmental factors to affect the existence and magnitudes of scope economies which MFIs have argued affects the cost-effectiveness of their operation ability to offer both loans and deposits. We use several environmental variables. First we use financial depth measured by the value of M3 over GDP of the country to control for the level of financial development in the country as well as population density measured as number of people per square kilometer to account for the fact that reaching borrowers and savers in remote areas is difficult.

We also control for the type of markets served such as primarily urban, rural or serving both about equally as well as the kind of lending technology - village banking, solidarity groups, or primarily individual lending. Results indicate that, on average, MFIs possess scope economies of nearly 13-19%. However, we also find that roughly 20% of our MFIs are currently operating with diseconomies of scope.

The plan of the paper is as follows. Section 2 outlines our theoretical and econometric techniques to estimate scope economies. Our rated MFI data is discussed in section 3. The results of our econometric study are presented in section 4. Conclusions and directions for future research are contained in section 5.

2. Economies of Scope and Cost Estimation

Economies of scope can emerge from two sources: (i) allocation of fixed costs over an extended product mix and (ii) cost complementarities across categories in production. Allocating fixed costs over a firm's product mix can contribute to scope economies when excess capital capacity is reduced by providing both savings and loans rather than individual provision of these services. Alternatively, cost complementarities result in scope economies when consumer information developed in the production of either savings or loans is used to reduce the monitoring requirements of the other product.

Given trustworthy estimates of scope economies, it is straightforward to asses the production consequences of narrow services provided by microfinance institutions. We describe here the general estimation of scope economies followed by a detailed discussion of a flexible, yet broad methodology using semiparametric methods that is robust to myriad functional form issues arising in empirical work.

Pulley & Humphrey (1993) define overall economies of scope as the percentage of cost savings from producing all outputs jointly as opposed to producing each output separately. Here we have only two outputs, loans and deposits so our estimates of scope economies are:

$$SCOPE = \frac{C(q_1, 0; \bar{r}) + C(0, q_2; \bar{r}) - C(q_1, q_2; \bar{r})}{C(q_1, q_2; \bar{r})}, \tag{1}$$

where $\bar{r}$ is a vector of ℓ input prices (here taken to be the relative price of labor and borrowed funds), and $C(\cdot)$ is the cost function. Given that the data used to estimate the cost function will represent a mix of firms producing loans and deposits jointly and firms specializing in the production of loans exclusively, the use of standard cost functions in production econometrics are not suitable.

For example, the transcendental logarithmic cost function (**?**):

$$\begin{aligned}\ln C =&\alpha_0 + \sum_{i=1}^{m}\alpha_i \ln q_i + 0.5\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{ij}\ln q_i \ln q_j + \sum_{i=1}^{m}\sum_{k=1}^{\ell}\delta_{ik}\ln q_i \ln r_k \\ &+ \sum_{k=1}^{\ell}\beta_k \ln r_k + 0.5\sum_{k=1}^{\ell}\sum_{s=1}^{\ell}\beta_{ks}\ln r_k \ln r_s j,\end{aligned} \tag{2}$$

cannot handle zero outputs. Given the appeal of this functional form in applied settings, various authors have dealt with the zero output (or zero input price) problem in a variety of ways. The simplest approach is to add a small number to all observations that have a zero output value (**?**) or to introduce a Box-Cox transformation parameter to all outputs, i.e. instead of using $\ln q_i$ one would replace it with $q_i^{\phi} = (q_i^{\phi} - 1)/\phi$ when ϕ is nonzero and is equal to the standard logarithmic function when it is zero. This is problematic for several reasons. First, it abstracts from the linear in parameters appeal of estimating a transcendental logarithmic function, second, if the estimate of ϕ is not statistically different from zero then further recourse is required, and lastly, all of the outputs are transformed by the same parameter. Thus, the use of a translog cost function for the study of economies of scope is in general restrictive and inappropriate for a wide array of empirical problems.

An empirical cost function for estimating scope economies was suggested by Pulley & Braunstein (1992) (PB hereafter) based off of the theoretical cost function suggested by Baumol, Panzer & Willing (1982). Their cost function is multiplicatively separable in outputs and input prices and is quadratic (as opposed to log-quadratic) in outputs, thus alleviating the empirical issue of zero valued outputs in real world data sets. More formally, the cost function deemed appropriate for estimating economies of scope by Baumol et al. (1982) is

$$C(q, r) = F(q) \cdot G(r), \tag{3}$$

where $F(q)$ is a quadratic form in outputs while $G(r)$ is a linearly homogeneous function of input prices. The empirical model suggested and estimated by PB is

$$C(q, \ln r) = F(q, \ln r) \cdot \exp\{G(\ln r)\} + u. \tag{4}$$

The reason that *both* q and $\ln r$ appear in $F(\cdot)$ is that there is no explicit reason for imposing separability between input prices and outputs. $F(q, \ln r)$ is still required to be quadratic in outputs. PB suggest that the exponential of $G(\cdot)$ is required given that one is using costs and not logarithmic costs. However, the theoretical suggestion of Baumol et al. (1982) only requires $G(r)$ to be linearly homogeneous. The composite cost model of PB can be written more succinctly as

$$C(q, \ln r) = F(q, \ln r) \cdot \tilde{G}(\ln r) + u. \tag{5}$$

While the logarithm will produce an additively separable model in $F(\cdot)$ and $G(\cdot)$ in equation 4, we feel that the generalized form of the Baumol et al. (1982) advocated initially by PB is the most appropriate way to investigate costs with the purpose of studying scope economies.

The empirical form of the composite cost function advocated by PB combines a quadratic form for outputs linked through interaction terms with a log-quadratic form for input prices. With an additive error term, the composite model is:

$$C = \Big[a_0 + \sum a_i q_i + 1/2 \sum\sum a_{ij} q_i q_j + \sum\sum g_{ik} q_i \ln r_k\Big] \cdot \exp\Big(b_0 + \sum b_k \ln r_k + \sum\sum b_{kl} \ln r_k \ln r_l\Big) + \varepsilon_i \tag{6}$$

where q_i is output i, r_i is the price of input i, a and b and g are parameters to be estimated. A variation of this model involves taking the logarithm of both sides which transforms the cost function into a composite log-quadratic structure. The following symmetry conditions need to be imposed onto the above cost function: $a_{ij} = a_{ji}$ and $b_{kl} = b_{lk}$. To ensure homogeneity, the following conditions need to bind: $\sum b_k = 1$; $\sum_k b_{kl} = 0$ for $\forall l$; $\sum_i g_{ik} = 0$ for $\forall k$.

Equation (4) can be estimated using maximum likelihood estimation routine assuming that the errors in equation are normally distributed or using a general nonlinear least squares algorithm if one was unwilling to assume the error terms belong to to the Gaussian family.

2.1. **Semiparametric Smooth Coefficient Cost Function.** While the empirical model of PB reflects a composite structure suitable for estimating scope economies, is easily restricted to impose linear homogeneity in input prices and is a nesting function of many more common empirical cost functions, Asaftei, Parmeter & Yuan (2009) recently proposed a semiparametric smooth coefficient cost function (SPSCC) that takes a similar form as that in PB but relaxes the function form restrictions on $\tilde{G}(\ln r)$. This setup with the same type of cost structure affords the researcher sufficient flexibility to model costs and investigate scope economies. Another appealing feature of this setup is that the incorporation of environmental variables thought to influence costs (such as type of market served by the MFI or load method) is taken up in a straightforward fashion that does not involve the user to provide an empirical specification.

For our purposes we estimate the model of Asaftei et al. (2009). Let the function $G(\ln r) \equiv \exp(b_0 + \sum b_k \ \ln r_k \ + \sum\sum b_{kl} \ln r_k \ln r_l)$, then equation (6) can be re-written as

$$C = \Big[\overline{a}_0 + \sum \overline{a}_i q_i + 1/2 \sum\sum \overline{a}_{ij} q_i q_j + \sum\sum \overline{g}_{ik} q_i \ln r_k\Big], \tag{7}$$

where $\overline{a}_i$, $\overline{a}_{ij}$ and $\overline{g}_{ik}$ are the coefficients a_i, a_{ij} and g_{ik} in equation (6) multiplied by $G(\ln r)$. We can therefore specify $\overline{a}_i$, $\overline{a}_{ij}$ and $\overline{g}_{ik}$ as functions of $G(\ln r)$ and an additional series of covariates related to scope economies and firm costs in general(V_i). While the theoretical properties of cost functions are well known with traditional input prices, the use of environmental variable is less understood. Thus, the ability to introduce these variables in a manner that imposes little structure on their exact specification within the cost function is pertinent.

We can write equation (7) in the following SPSCM specification,

$$y_i = \alpha(z_i) + \beta(z_i) x_i + \varepsilon_i \tag{8}$$

where $y_i \equiv C_i$, $x_i = [1 \quad q_i' \quad qq_i' \quad q \ln r_i']'$, $z_i = [\ln r_i \quad V_i]$. We do not have to introduce quadratic and interaction terms in z_i since the unknown smooth coefficient will select the appropriate higher order/interaction terms. Here, qq_i' is the $k(k-1) \times 1$ vector of squares and interactions of the outputs and $q \ln r_i'$ is the $kj \times 1$ vector of interactions between outputs and log input prices.

Another way to think of this model is that for a given level of z_i, we have a linear in parameters model where the slopes possibly differ for differing levels of z_i. Since z_i and x_i can contain the same variables, this model is more general than that of PB. One can also view the PB model as a smooth coefficient model, with $\exp(b_0 + \sum b_k \ln r_k + \sum\sum b_{kl} \ln r_k \ln r_l)$ representing the smooth coefficient on q_i in equation (4). The key difference between the semiparametric smooth coefficient model in equation (8) and the parametric smooth coefficient model in equation (4) is that the coefficients are identical, up to scale, in equation (4) while in equation (8) they can be entirely different functions altogether. At this point it is important to emphasize that imposing linear homogeneity is difficult in our semiparametric setup given that we are not imposing any structure on $\alpha(z_i)$ and $\beta(z_i)$.[1] However, this is a small price to pay since even the use of the popular translog specification, which violates global concavity, is traditionally used in production econometrics.

The semiparametric smooth coefficient model can be specified as quadratic in output, as recommended by Baumol et al. (1982), but can be more/less general in the input price structure, due to the lack of specification on the $\beta(z_i)$ and $\alpha(z_i)$. Indeed, if all the $\beta(z_i)$s turn out to be constant then we have a structure that is fully semiparametric, known in the output dimension but unknown in the input dimension. Li, Huang, Li & Fu (2002) and Li & Racine (2004) proposed an estimation procedure for the SPSCM defined in equation (8) based on local constant least squares (LCLS).

The estimation of equation (8) is as follows. Denote $\delta(z_i) = [\alpha(z_i), \beta(z_i)]$ and rewrite (8) as $y_i = \delta(z_i)X_i + \varepsilon_i$, where $X = [1 \ X_i]$. Our LCLS estimator of $\delta(z)$ becomes

$$\delta(z) = (\mathbf{X}'\mathbf{K}(z)\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}(z)\mathbf{y}, \tag{9}$$

where $\mathbf{K}(z)$ is a diagonal matrix with i^{th} element $K_i = K(\frac{z_i - z}{h})$. K_i is constructed using the generalized product kernel of Racine & Li (2004) and h is a vector of bandwidths. The reason we use a generalized kernel is that several of our additional z variables are discrete,[2] most notably our indicator for loan method and area serviced by the MFI. Smoothing these types of variables with continuous kernels is inappropriate while treating it as a dummy variable leads to a loss of efficiency in our estimates of the smooth coefficients.

Given the importance of bandwidth selection for all semiparametric methods we briefly describe our selection criterion. We use the standard least squares cross validation (LSCV) selection method which has recently been shown to have desirable properties by Hall, Li & Racine (2007). This selection method chooses bandwidths that minimize the sample counterpart of mean integrated

[1] See Racine, Parmeter & Du (2009) for an approach that could impose linear homogeneity in this setting.

[2] Let $z_i = [z_i^d, z_i^c]$, where z_i^d is a vector of discrete regressors and z_i^c is a vector assuming continuous values. One can further decompose z_i^d into subvectors of ordered and unordered discrete regressors, z_i^{do} and z_i^{du}, respectively.

square error and is defined as follows:

$$LSCV(h) = \min_h n^{-1} \sum_{i=1}^{n} (C_i - \hat{C}_{-i})^2, \tag{10}$$

where $\hat{C}_{-i}$ is the leave-one-out estimator of costs that is produced by dropping the i^{th} observation from our construction of costs in equation (8). Essentially, LSCV selects bandwidths that result in minimum average prediction error for the sample.

3. Data

The data used are collected from the rating reports of MFIs publicly available at www.ratingfund.org. In the early 2000s, microfinance rating was believed to be a market based mechanism of control which would discipline mangers by providing independent reporting for potential donors and investors. MFIs in need of raising funds or under demands by donors submitted to this independent evaluation and opened their books, signaling that they were more transparent than the average MFI. This also could reflect the fact that they were more in need of financial assistance as well. The rating fund offered financial support and required that the rating reports remain publicly available online.

The dataset used here contains all rating reports completed by June 2007. The data comprise an unbalanced panel of 244 MFIs operating in 53 countries with about 3 years of data per MFI for a total of 777 observations. Distribution of MFIs by country is presented in Table 1. Comparison with other publicly available data shows that these data have more observations from Latin America, perhaps because they needed external funds. The two outputs used - loans and savings - are measured as the dollar value of loans outstanding and the dollar value of voluntary deposits. Input price of labor is the average annual salary per employee, the cost of capital is the weighted cost of borrowed funds (deposits and loans), and the input price of physical capital is ratio of non-labor operating expense to the value of net fixed assets. Total costs (TC) are the sum of input prices and input quantities.

Summary statistics of the variables used in the scope estimation are in Table 2. Most of MFIs, 76 percent only extend loans and 24 percent offer both loans and mobilize savings. The average MFI has about \$2,2 million in loan portfolio outstanding, with a range from slightly more than \$36,000 to \$34.6 million. The volume of savings (when offered) is \$960,000 on average and the largest case is 24 million. The average value of annual salaries is \$6,707, the cost of capital (deposits and borrowed funds) is 8 percent and the ratio of non-labor operating expense to next fixed assets is 4.2 (Table 2). We see that 35% of our MFIs serve urban markets only and an additional 39% serve both urban and rural markets. Service strictly to rural markets is rare for the group of rated MFIs that compose our dataset. Additionally, we see that nearly 60% of our rated MFIs provide loans on an individual basis while another 22% use solidarity groups to provide loan services.

Table 1. Distribution of MFIs by region and country

Country	# Obs.	% Obs.	Country	# Obs.	% Obs.
Eastern Europe and Central Asia (ECA)					
Albania	10	1.29	Kazakhstan	9	1.16
Armenia	4	0.51	Kyrgyzstan	8	1.03
Azerbaijan	16	2.06	Moldova	8	1.03
Bosnia and Herzegovina	39	5.02	Romania	3	0.39
Bulgaria	4	0.51	Russia	30	3.86
Georgia	13	1.67	Tajikistan	11	1.42
Latin America (LA)					
Argentina	3	0.39	Guatemala	16	2.06
Bolivia	53	6.82	Haiti	3	0.39
Brazil	36	4.63	Honduras	22	2.83
Chile	3	0.39	Mexico	55	7.08
Colombia	22	2.83	Nicaragua	28	3.60
Dominican Republic	14	1.80	Paraguay	4	0.51
Ecuador	51	6.56	Peru	98	12.61
El Salvador	11	1.42			
Sub-Saharan Africa (SSA)					
Benin	11	1.42	Mozambique	3	0.39
Burkina Faso	5	0.64	Senegal	8	1.03
Cameroon	9	1.16	South Africa	3	0.39
Ghana	5	0.64	Tanzania	3	0.39
Kenya	13	1.67	Togo	1	0.13
Madagascar	3	0.39	Uganda	12	1.54
Mali	3	0.39	Zambia	3	0.39
Southeast Asia (SEA)					
Bangladesh	1	0.13	Indonesia	1	0.13
Cambodia	11	1.42	Mongolia	5	0.64
India	29	3.73	Philippines	10	1.29
Middle East and North Africa (MENA)					
Chad	2	0.26	Jordan	9	1.16
Egypt	14	1.80	Morocco	12	1.54
Ethiopia	24	3.09	Tunisia	3	0.39

It is the presence of MFIs with 0 savings that require us to eschew a traditional cost function approach to estimating scope economies and, in addition, the fact that some MFIs have zero input costs for funding implies that we must include input prices into our semiparametric smooth coefficient cost function in level form as well. The added generality afforded by our approach is key for analyzing these types of datasets as *ad hoc* approaches to dealing with zeros are unappealing in applied settings.

Table 2. Summary Statistics of Variables Used in Estimation

Variable	Mean	Std. Dev.	Minimum	Maximum
Loans ($)	4,233,373	5,182,054	3,586	34,600,000
Savings ($)	960,369	3,128,975	0	24,000,000
Total cost	1,115,828	1,306,629	2,689	12,200,000
Average wage ($)	6,707	4,022	18	26,573
Cost of capital (ratio)	4.25	29.75	0.04	800.00
Cost of funds (%)	0.08	0.07	0.00	0.58
Population density	65	77	2	1,050
Financial depth	0.38	0.20	0.07	1.39
Market served				
Urban	0.35	–	0	1
Rural	0.13	–	0	1
Urban and Rural	0.39	–	0	1
No information	0.13	–	0	1
Loan methods				
Village banks	0.14	–	0	1
Solidarity groups	0.22	–	0	1
Individual	0.59	–	0	1
Other	0.01	–	0	1
Unclassified	0.04	–	0	1

4. Results

For the model discussed in the previous section we estimate a SPSCM cost function with two outputs and three input prices. Input prices are scaled by physical capital wages (Price of capital ratio) to produce two relative input prices to be used in each piece of the cost function. Total costs, loans and deposits are scaled by 1,000,000 and 10,000,000, respectively to stabilize the smooth coefficients during bandwidth selection. All results were computed using the np package (Hayfield & Racine 2008) in R (R Development Core Team 2008). Bandwidths were selected via LSCV, discussed previously, using 20 multi-starts.[3] In addition to normalized input prices entering the unknown smooth coefficients, we also included the year in which the MFI was observed and its region (based off of the classification in Table 1) as well as the main type of lending methodology

[3]Multi-starts are the number of different trials used to calculated the minimum of the least squares cross validation function. Given the nonlinearity of this function with respect to multiple bandwidths, it is good practice to use numerous multi-starts to avoid obtaining bandwidths indicative of a local minimum as opposed to the global minimum.

the MFI uses (village bank, solidarity group, individual loan, etc.), the main market the MFI services (urban, rural or both), the population density of the area in which the MFI operates and the level of financial depth of the country. Given our ability to smooth discrete variables, year, loan methodology and main market served were smoothed using discrete kernels (see Li & Racine 2007).

Table 3. Bandwidths For our SPSCM Cost Function

Variable	Including Controls Bandwidth	Excluding Controls Bandwidth	Upper Bound	Cut-Off
WL	1755	695	∞	17177
WF	85.49	33.02	∞	0.45
Pop. Dens.	89.09	–	∞	173.58
Fin. Depth	13.35	–	∞	0.40
Year	0.831	0.968	1	0.80
Region	0.408	0.687	0.80	0.64
Area	0.289	–	0.75	0.60
Method	0.050	–	0.80	0.64

Before discussing our estimates of scope economies and their implications, we first highlight the bandwidths obtained via our LSCV selection procedure. Hall et al. (2007) have shown that when bandwidths are 'close' to their upper bounds that those variables are smoothed out of the regression model, i.e. they are irrelevant in terms of predicting the response variable. Table 3 presents our cross-validated bandwidths for estimation of the generalized cost function in (8) both including and excluding environmental variables (Pop. Dens, Fin. Depth, Area, Method). For continuous variables, Li & Racine (2007) suggest that bandwidths larger than 2 standard deviations of the variable are effectively smoothed away while for discrete variables, bandwidths within 80-90% of their theoretical upper bounds are irrelevant.

For ease of reference we have included our cut-offs for variables to be deemed irrelevant for both models. We see that when we do not include environmental factors directly into the cost function that the relative price of financial capital is smoothed away, suggesting that it has not effect on the smooth coefficients in our cost function. This continues to be the case when we include controls, but, of the four environmental factors that we include, three of them are not smoothed away, suggesting that the cost structure of firms depends upon the setting in which they operate. Financial depth and the relative cost of financial capital are smoothed away in our model including controls. We note that relative financial capital costs are included directly into our cost function via interactions with outputs so they still possess explanatory power for overall costs. The results here are simply to suggest that the *coefficients* of the cost function are not impacted by this variable.

This feature is an interesting one as it pertains to cost estimation. Since we include financial capital costs directly as an input and indirectly as a contextual variable which influences the cost structure, we have a model that is very general in terms of allowing for financial capital costs to influence overall cost. Thus, our estimated bandwidths point to the fact that financial capital costs do not influence the cost structure, but influence overall costs as a traditional input. This points to the fact that the level of financial capital does not sway the cost structure, suggestive that larger(richer) firms cannot tilt the playing field in their favor.

Our estimates of scope economies are presented in Table 4. We find scope economies in microfinance institutions indicating substantial cost savings from mobilizing deposits and extending loans instead of only extending loans (assuming of course that the same cost structure of the MFIs are preserved). We also find that there are substantial differences in the magnitudes of scope economies related to influence of variables outside of the cost function (that is relative input prices and outputs), namely region, year, population density, financial sector development, main market served (rural, urban or both), and main loan methodology used (village banking, solidarity groups, or individual method). The results suggest that MFIs achieve substantial reductions in costs by offering both savings and loans to their customer base. On average the scope economies when environmental factors are not included are 19 percent but fall to 13 percent when environmental factors are controlled for. Looking at our quartile measures (Median, Q1, Q3) we see that the bulk of MFIs have scope economies ranging from 4% to 29%. The standard deviation of the scope economy estimates are 30% larger when environmental factors are included, suggestive of greater dispersion when these factors are controlled for.

Additionally, focusing on scope diseconomies, we see that the MFIs who have estimated scope diseconomies have considerable worse levels of diseconomies when we account for our environmental controls. The mean estimated scope diseconomies is over 50% higher[4] and the interquartile range is nearly 33% larger. However, considering only those MFIs with scope economies we see that our summary measures are much more in unison across specifications. Here, no statistically significant differences exist across our specifications.[5]

Table 4. Scope Economy Measures

	Overall Scope Economies		Scope Diseconomies		Scope Economies	
	w/ Controls	w/o Controls	w/ Controls	w/o Controls	w/ Controls	w/o Controls
Mean	0.13	0.19	-0.28	-0.18	0.26	0.24
Median	0.09	0.13	-0.12	-0.07	0.18	0.16
Std. Dev.	0.35	0.27	0.29	0.29	0.26	0.23
Q1	-0.02	0.04	-0.28	-0.20	0.07	0.06
Q3	0.27	0.30	-0.05	-0.02	0.36	0.33

[4] (28-18)/18=5/9≈50%.

[5] We obtained six extremely large estimates of scope economies which we were unable to explain. These six scope economy estimates have been removed from the remainder of the analysis. They are D-MIRO from Ecuador in 2001, Cooperativa de Ahorro y Crèdito Erco from Ecuador in 2002, Fundacion Alternativa in Ecuador from 2004-2006 and welfare service Ernakulam from India in 2005.

We present box plots[6] and empirical distribution functions of our estimates of scope economies in Figure 1. We see that the inclusion of environmental factors changes the landscape of scope economies dramatically. The box plot suggests that the distribution of scope economies is shifted up when we omit controls as opposed to keeping them in the smooth coefficients. Additionally, it appears that a stochastic dominance relationship is present. Using a modified Kologorov-Smirnoff test for first order stochastic dominance we reject the null that the two distributions are equivalent (p-value of 0) and conclude that our estimates of scope economies including controls first order stochastically dominate those estimates when environmental variables are omitted from the smooth coefficients of the cost function. Given our conclusions regarding the bandwidths in Table 3 it appears that controlling for environmental variables is key for considering cost function estimation across a range of different MFIs.

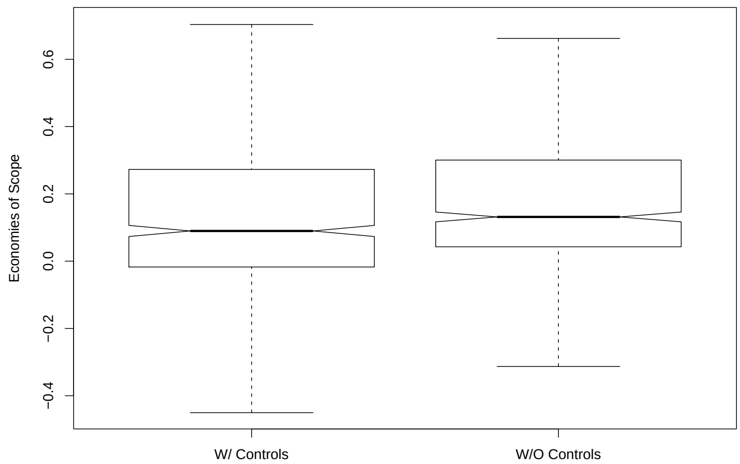
We further find that not all MFIs have scope economies. Roughly 25% of the MFIs in our dataset have diseconomies of scope. If we look at summary statistics for those MFIs with estimated scope diseconomies we see that they experience an increase of costs by 26% (12%) by offering both savings and loans when we control for environmental factors (excluding). Focusing instead on those firms that experience positive estimated scope economies, which makes up nearly three-quarters of our total sample, we see that the inclusion of our controls leads to a 33% increase in mean estimated scope economies, and a doubling of the spread of our estimates. The third quartile estimate of 40% suggests that there are more than a few MFIs which are(could be) experiencing substantial reductions in cost by offering both savings and loans.

In Table 5 scope economies in MFIs which actually provide loans and savings facilities are compared with those in MFIs which only lend. The results suggest that scope economies may not be the driving force behind decisions to collect savings. The reason is that when we ignore the cost environment it is clear that MFIs offering a product mix have higher scope economies, but when we control for the operating environment and its impact on the cost structure, we have exactly the opposite result, firms currently providing only loans would benefit more in terms of cost reduction by offering savings above the cost savings that firms currently offering savings and loans could achieve.

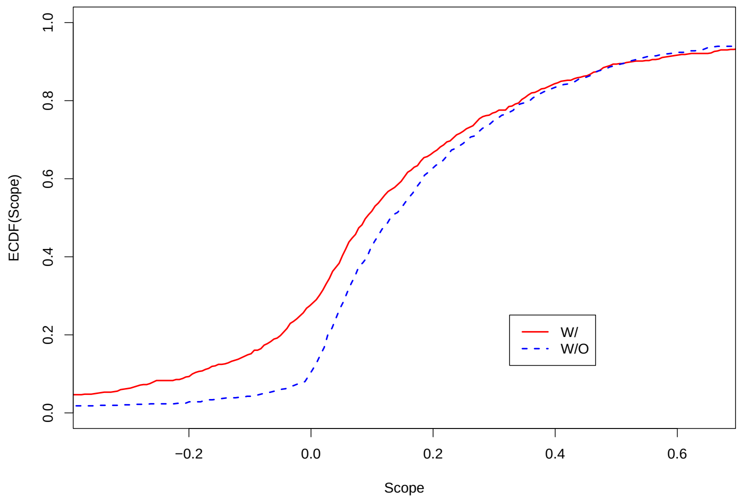
Given the magnitudes of our bandwidths when we include control variables in our smooth coefficients of our cost function, as well as the visual evidence provided by the distributional plots and the results of our stochastic dominance test, we choose to more carefully analyze scope economies estimated via inclusion of these variables directly into the smooth coefficients. All three of our key discrete variables, Region, Area and Method, are not smoothed out of our model and therefore suggest an impact on the smooth coefficients of our cost function and consequently our estimates of scope economies in a hierarchical fashion.

[6]A box-and-whisker plot (sometimes called simply a 'box plot') is a histogram-like method of displaying data, invented by J. Tukey. To create a box-and-whisker plot, draw a box with ends at the quartiles Q_1 and Q_3. Draw the statistical median M as a horizontal line in the box. Now extend the 'whiskers' to the farthest points that are not outliers (i.e., that are within 3/2 times the interquartile range of Q_1 and Q_3). Then, for every point more than 3/2 times the interquartile range from the end of a box, draw a dot.

FIGURE 1. Distributional Characteristics of Scope Economies.



(a) Boxplots of Estimated Scope Economies



(b) Empirical Distribution Functions of Estimated Scope Economies

4.1. **Scope Economies Across Control Variables.** Given the fact that nonparametric models deliver estimates which are observation specific, a quick and intuitive way to present results is to condition on specific levels of the variables on interest. To that end we present our estimates of scope economies here by focusing on how they vary across three of our most important controls, the region the MFI is located, the area in which the MFI provides services and the type of loan structure the MFI uses.

Table 5. Summary of Scope Economies by Output Structure

	Including Controls		Excluding Controls	
	Loans	Savings & Loans	Loans	Savings & Loans
Mean	0.16	0.02	0.19	0.21
Median	0.11	-0.01	0.11	0.20
Std. Dev.	0.30	0.46	0.24	0.36
Q1	0.02	-0.19	0.04	0.05
Q3	0.28	0.18	0.28	0.35

Scope economies are larger in countries with higher population density. For example, in a model without environmental factors, countries with diseconomies of scale have 73 persons per sq.km and those with scope economies have 92 persons per sq.km. In a model including the environmental factors the difference is even larger with population density in countries with scope diseconomies of 46 persons per sq.km, while it is 96 persons per sq.km in countries with scope economies. However, the level of financial development of a country measured by the financial depth variables does not significantly vary across model, consistent with our description of its bandwidth mentioned earlier.

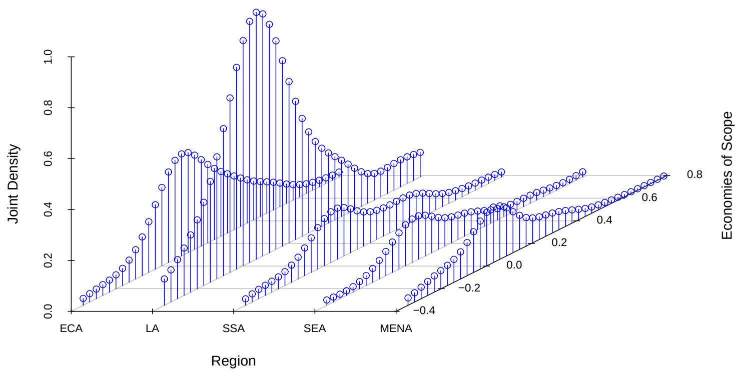
Table 6. Scope Economy Measures Across Regions

	ECA	LA	SSA	SEA	MENA
Mean	0.18	0.14	0.10	0.15	-0.04
Median	0.14	0.08	0.09	0.09	0.01
Std. Dev.	0.33	0.33	0.42	0.39	0.32
Q1	0.04	-0.03	-0.03	-0.002	-0.14
Q3	0.33	0.26	0.36	0.29	0.09

Focusing on the regions where our MFIs operate, Table 6 presents summary statistics of estimated scop economies based on our regional definitions from Table 1. We observe that the highest scope economies are realized ECA region and where also there are no MFIs with diseconomies of scope. MFIs in Latin America and in South East Asia scope economies of similar magnitude on average as well as some observations with diseconomies of scope. The only region with average diseconomies of scope is MENA although their magnitudes of 4% on average is relatively small. Scope economies in MFIs targeting mainly urban markets are highest at 22%, in those serving predominantly rural market are 17% and those without clear specialization are at % percent. MFIs using mainly village banking lending exhibit the largest economies of 21%, followed by MFIs using mainly individual loans

Among the most interesting results are those about magnitudes in scope economies by market served (Table 7). While a model without environmental factors would suggest that scope economies in MFIs targeting mainly rural markets are the smallest (0.16) and have the highest standard deviation, the model with environmental factors suggests that MFIs targeting mainly rural markets have the largest scope economies (0.38 on average) and their standard deviation is as large as that of MFIs targeting operating in rural and urban markets.

FIGURE 2. Density plots for Scope Economies based on Region.



(a) Scope Economies by Region (w/ Controls)

TABLE 7. Summary of Scope Economies by Main Market Served and Lending Method

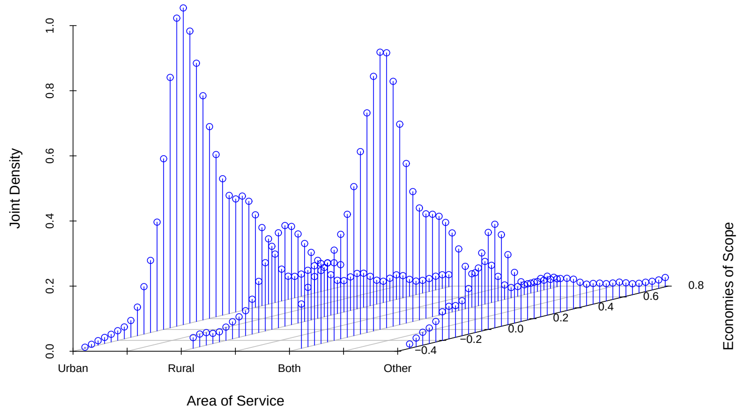
	Main Market Served			Lending Method		
	Urban	Rural	Both	Village	Solidarity	Individual
Mean	0.22	0.17	0.05	0.21	0.11	0.17
Median	0.15	0.11	0.03	0.14	0.05	0.11
Std. Dev.	0.27	0.33	0.37	0.24	0.22	0.34
Q1	0.06	0.01	-0.09	0.04	-0.01	0.03
Q3	0.34	0.36	0.19	0.34	0.22	0.25

The ranking of the size of scope economies is preserved however as Table 7 shows, when scope economies ate cross-tabulated by the type of lending methodology used. MFIs using mainly village banking exhibit the largest economies of 21%, followed by MFIs using mainly individual loans with 17% on average and solidarity groups with the lowest scope economies of 11%. While village banks are more likely to be cost effective in also mobilizing deposits, it seems that the solidarity group model does not offer the best opportunity to cost-effectively offer both loans and savings facility.

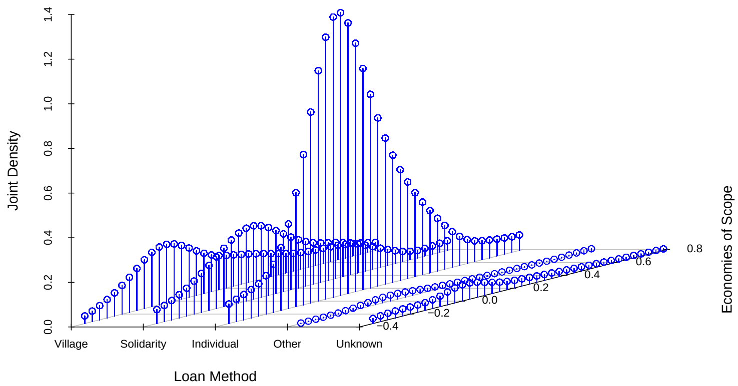
We present visual evidence of these estimates in Figure 3 where we plot the joint density of our estimated scope economies and area of service (panel (a)) and type of loan (panel (b)), respectively.[7] We note that our joint density of scope economies and area of service produces a rougher set of plots than those based on loan method.

[7]We used likelihood cross-validation to obtain the bandwidths for these joint densities.

FIGURE 3. Density plots for Scope Economies based on Loan Service and Method.



(a) Scope Economies by Service Area



(b) Scope Economies by Loan Method

5. Conclusions

While economies of scope of lending and mobilizing deposits in banking are justified theoretically (Diamond 1984) and found empirically (see Sounders 1999), in microfinance, the existence and magnitude of scope economies has not been investigated. We use a semiparametric smooth coefficient model to estimate these economies using a dataset put together from rated MFIs with over 750 annual observations from MFIs across the world. This model affords the researcher sufficient flexibility in incorporating zero valued input prices and outputs into the cost function without resorting to *ad hoc* data replacement techniques.

We estimate two models of scope economies, one where only variables typically used in a cost function approach are included (total cost, output values and relative input prices) and a model where in addition to these variables we include population density, a measure of financial sector development, type of market served (urban, rural or both) and the predominant loan methodology (village banking, solidarity groups and individual loans, as well as controls for time and region). We find that scope economies are substantial across both settings and, for either model that over 70% of the MFIs in our dataset have(or would) experience reductions in cost by offering both savins and loan services. We find that larger and more variable scope economies are present when we allow our cost function coefficients to vary depending on key environmental variables. We also find that scope economies would not necessarily come from direct decrease in the cost of capital if deposits are collected but it is through these costs interaction with the rest of the production technology.

Overall, our finding of scope economies is intuitive and relevant. Our results suggest that agencies providing funding to MFIs in the future require, or at a minimum suggest, that both savings and loan services be offered. Future work should attempt to obtain a more comprehensive data set for MFIs outside of Latin America and determine if other environmental variables are relevant in MFIs' cost structures.

References

Armendariz de Aghion, B. & Morduch, J. (2005), *Microeconomics of Microfinance*, MIT Press.

Asaftei, G., Parmeter, C. F. & Yuan, Y. (2009), Economies of scope in financial services: A semi-parametric approach. Virginia Tech Working Paper.

Baumol, W. J., Panzer, J. & Willing, R. (1982), *Contestable Markets and the Theory of Market Structure*, Harcourt, New York.

Diamond, D. (1984), 'Financial intermediation and delegated monitoring', *Review of Economic Studies* **51**, 393–414.

Hall, P., Li, Q. & Racine, J. S. (2007), 'Nonparametric estimation of regression functions in the presence of irrelevant regressors', *Review of Economics and Statistics* **89**, 784–789.

Hartarska, V. & Holtmann, M. (2006), 'An overview of recent developments in the microfinance literature', *Agricultural Finance Review* **66**(2), 147–165.

Hayfield, T. & Racine, J. S. (2008), 'Nonparametric econometrics: The np package', *Journal of Statistical Software* **27**(5).
*http://www.jstatsoft.org/v27/i05/

Li, Q., Huang, C. J., Li, D. & Fu, T.-T. (2002), 'Semiparametric smooth coefficient models', *Journal of Business & Economic Statistics* **20**, 412–422.

Li, Q. & Racine, J. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

Li, Q. & Racine, J. S. (2004), 'Cross-validated local linear nonparametric regression', *Statistica Sinica* **14**, 485–512.

Pulley, L. B. & Braunstein, Y. M. (1992), 'A composite cost function for multiproduct firms with an application to economies of scope in banking', *The Review of Economics and Statistics* **74**, 213–230.

Pulley, L. B. & Humphrey, D. (1993), 'The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals', *Journal of Business* **66**, 437–462.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org

Racine, J. S. & Li, Q. (2004), 'Nonparametric estimation of regression functions with both categorical and continuous data', *Journal of Econometrics* **119**(1), 99–130.

Racine, J. S., Parmeter, C. F. & Du, P. (2009), Constrained nonparametric kernel regression: Estimation and inference. Virginia Tech Working Paper.

Sounders, A. (1999), *Financial Institutions Management*, 3 edn, Irwin/McGraw-Hill.