Economies of Diversification in Microfinance: Evidence from Quantile Estimation on Panel Data

Malikov, Emir and Hartarska, Valentina and Mersland, Roy

University of Nevada, Las Vegas, Auburn University, University of Agder

2019

Economies of Diversification in Microfinance: Evidence from Quantile Estimation on Panel Data*

EMIR MALIKOV[1] VALENTINA HARTARSKA[2] ROY MERSLAND[3]

[1]University of Nevada, Las Vegas, United States
[2]Auburn University, United States
[3]University of Agder, Norway

June 25, 2019

Abstract

Prior research finds the presence of diversification-driven cost savings from the joint provision of credit and deposits in microfinance institutions (MFIs) and thus provides supportive evidence for policies aimed at promoting the transformation of credit-only MFIs into diversified loans-and-savings entities. However, these studies usually ignore the multi-way heterogeneity across MFIs which vary substantially in size, business model, target clientele and operate in diverse environments. We use a quantile panel data model with correlated effects that accommodates (i) technological heterogeneity across specialized credit-only and diversified loans-and-savings MFIs, (ii) distributional heterogeneity in the cost structure of MFIs along the scale of their costs and (iii) unobserved heterogeneity correlated with covariates in the cost function. Our results show that the typical measurement of economies of diversification at the conditional mean provides an incomplete and distorted picture of the magnitude and prevalence of diversification economies in the industry. While we find statistically significant diversification economies, they are modest for the majority of small-size MFIs but are quite substantial for large-scale institutions suggesting that the degree of diversification economies increases with the scale of operations. Thus, policy makers and donors should be cautious when encouraging smaller credit-only MFIs into the provision of savings deposits with the expectations of diversification-driven cost savings, since the latter seem to be more relevant and economically substantive mainly for the very large MFIs.

Keywords: cost, diversification, microfinance institutions, quantile regression

JEL Classification: G15, G21, O16, L33

* *Email*: emir.malikov@unlv.edu (Malikov), hartavm@auburn.edu (Hartarska), roy.mersland@uia.no (Mersland).

We would like to thank an anonymous referee for many insightful comments and suggestions that helped improve this paper. Any remaining errors are our own.

1 Introduction

The microfinance industry is comprised of heterogeneous institutions providing credit and savings to a variety of clients around the globe who are without access to other formal financial services. Varying in size, target clientele as well as business and regulatory environment in which they operate, these microfinance institutions (MFIs) include small-scale local operations as well as industry giants with millions of clients. Yet, most prior studies attempting to measure the degree of scope-driven cost savings ensuing from the joint provision of lending and deposits by diversified MFIs (e.g., Hartarska et al., 2010, 2011; Delgado et al., 2015) ignore potential technological heterogeneity and use a *common* multi-output cost function for both the integrated and specialized institutions. The focus has also been exclusively on the estimation of MFIs' cost structure at the conditional *mean* leaving open questions about distributional heterogeneity across MFIs with varying scales of operations.[1]

In this paper, we contribute to the literature by pursuing a more robust measurement of economies of diversification/scope and their prevalence in the microfinance industry by explicitly allowing for a multi-way heterogeneity amongst MFIs. Specifically, we estimate the microfinance cost structure using a quantile panel data model with correlated effects that accommodates the following kinds of heterogeneity: (i) technological heterogeneity across specialized credit-only and diversified loans-and-savings MFIs, (ii) distributional heterogeneity in the cost structure of MFIs along the scale of their costs, and (iii) unobserved heterogeneity correlated with covariates in the cost function (e.g., latent management quality or efficiency).

First, by allowing specialized loans-only and diversified loans-and-savings MFIs to have *technologically* heterogeneous cost structures, we improve upon the common approach pursued in most studies whereby all MFIs are presumed to have the same microfinance cost structure regardless of different mixes of financial services they offer to customers. Second, it is reasonable to expect that large- and small-scale MFIs exhibit different degrees of scope-driven potential for cost saving (if any), with the institution's cost of operation being a good direct measure of the scale of its operations. Estimates of the cost functions at the conditional mean can be uninformative from a policy perspective even after controlling for MFI-specific characteristics because an "average" may not be representative of actual MFIs, especially in the presence of thick tails of the cost distribution.[2] To accommodate this heterogeneity, we assess the degree of economies of diversification at different conditional quantiles of the microfinance cost distribution. Our quantile model, besides being more robust to the error distributions including the presence of outliers, also allows for *distributional* heterogeneity in the cost structure of MFIs. Third, the cost performance of MFIs naturally depends on many unobserved factors that are unique to an institution. For instance, an MFIs run by a skilled experienced manager is likely to operate at a lower cost than an MFIs managed by a novice, *ceteris paribus*. Ignoring this *unobserved* heterogeneity may produce inconsistent estimates

[1] The prior papers have instead focused more on minimizing the functional form misspecification by adopting a semiparametric approach to modeling the cost function thereby being able to better accommodate nonlinearities.

[2] Appendix provides some empirical evidence of "fatness" in the tails of conditional (log) cost distribution.

since unobserved heterogeneity is likely to be correlated with covariates present in the estimated equation. We address this issue by modeling unobservable MFI-specific heterogeneity in the form of correlated effects,[3] in addition to conditioning the cost relationships on *observable* controls.

Among prior studies of the revenue-diversification-driven cost savings in microfinance, Malikov & Hartarska (2018) is the most closely related to our paper. It is the only other study that explores technological heterogeneity between credit-only and credit-and-savings MFIs but, like all other papers, they focus on the conditional mean of cost, leaving important distributional heterogeneity unexplored.[4] More generally, our paper also contributes to a broader literature on MFIs and their performance; see Caudill et al. (2009), Mersland & Strøm (2009), Cull et al. (2011), Hermes et al. (2011), Hartarska et al. (2013), Delgado et al. (2015), D'Espallier et al. (2017), among many others.

Using the data on 344 MFIs operating in 65 countries during 1998–2013, we show that the typical measurement of economies of diversification at the conditional mean provides an incomplete and distorted picture of the magnitude and prevalence of diversification economies in the industry. While we find statistically significant diversification economies, they are modest for the majority of small-size MFIs but are quite substantial for large-scale institutions suggesting that the degree of diversification economies increases with the scale of operations. These differences are statistically significant distribution-wise. Thus, policy makers and donors should be cautious when encouraging smaller credit-only MFIs into the provision of savings deposits with the expectations of diversification-driven cost savings, since the latter seem to be more relevant and economically substantive primarily for the very large MFIs.

2 Measuring Economies of Diversification

Economies of diversification are said to exist if the firm's average cost declines with the increase in the scope of outputs it produces reflective of revenue diversification (see Panzar & Willig, 1981; Chavas & Kim, 2010). In the context of microfinance industry whereby MFIs are conceptualized as producing two outputs—(micro)loans and (micro)deposits—a financial institution exhibits economies of diversification if the cost of producing credit/loans (y_1) and savings/deposits (y_2) separately exceeds the cost of their joint production, i.e., $\mathcal{C}(y_1, 0) + \mathcal{C}(0, y_2) > \mathcal{C}(y_1, y_2)$, where we have suppressed all other arguments of the cost function besides outputs. This definition requires computation of counterfactual costs of producing each of the two outputs (loans and deposits) separately with the counterfactuals defined as fully specialized single-output MFIs. In practice, while there are credit-only MFIs, there are rarely savings-only institutions.[5] Although the projection of multi-

[3]We opt for the correlated-effects treatment of unobservable heterogeneity over admittedly less restrictive fixed-effects approach because the consistency of the latter estimator in the quantile setting requires both large N and T. Our data however are such that $N >> T$.

[4]However, in contrast to Malikov & Hartarska (2018), we abstract away from the issues concerning endogenous product-mix selection which, admittedly, is a shortcoming of our analysis. Extending our model to also controlling for selectivity may provide an avenue for fruitful future research.

[5]While there are government-owned institutions that only collect savings in Kenya and South Africa, these are not a part of our dataset.

product cost functions to the boundary single-output cases is a popular solution in such instances, it is notoriously known to suffer from "excessive extrapolation" (Evans & Heckman, 1984; Hughes & Mester, 1993). Therefore, we instead adopt Malikov et al.'s (2016) approach, whereby we (i) use *observable* specialized loans-only MFIs to identify their single-output cost relationship and thus decrease the number of counterfactuals necessary for the evaluation of diversification economies in the industry from two to one and (ii) replace the remaining non-existent counterfactual case of a *full* savings-only specialization with the closely related case of an MFI *highly* specialized in savings within the observable output-mix configuration in the data. A diversified loans-and-savings MFI is then said to exhibit economies of diversification if

$$\mathcal{C}_1\big(\varpi y_1\big) + \mathcal{C}_2\Big((1-\varpi)y_1, y_2\Big) > \mathcal{C}_2\big(y_1, y_2\big), \tag{2.1}$$

where $\mathcal{C}_1(\cdot)$ and $\mathcal{C}_2(\cdot)$ are respectively the single-output loans-only and multi-output loans-and-savings cost functions, and $0 \leq \varpi \leq 1$ is a distribution weight controlling the degree of specialization of the second counterfactual firm. We further ensure that each counterfactual MFI does not produce less of each output than MFIs actually do in the data. That is, we require that $\varpi y_1 \geq \underline{y}_1$ and $(1-\varpi)y_1 \geq \underline{y}_1$, where $\underline{y}_1$ denotes a sample minimum of y_1. With this, we define our *within-sample* measure of economies of diversification as

$$ED(\varpi) = \frac{\mathcal{C}_1\big(\varpi y_1^* + \underline{y}_1\big) + \mathcal{C}_2\Big((1-\varpi)y_1^* + \underline{y}_1, y_2\Big) - \mathcal{C}_2\big(y_1, y_2\big)}{\mathcal{C}_2\big(y_1, y_2\big)}, \tag{2.2}$$

where $y_1^* = y_1 - 2\underline{y}_1$. Positive/negative values of $ED(\varpi)$ correspond to economies/diseconomies of diversification, whereas a zero value indicates the cost invariance to output diversification. We choose ϖ (via a grid search) that yields the smallest value for $ED(\varpi)$ to arrive at the measure of "global" diversification economies (for each MFI):

$$GED = \min_{\varpi}\ ED(\varpi). \tag{2.3}$$

If economies of diversification of the smallest value are still significantly positive, we can then conclude that such economies are "globally" significant over the MFI's output space in a given year.

3 Estimation Details

For a given quantile index $\tau \in (0, 1)$, we specify the τth output-mix-specific conditional quantile function of the MFI's cost distribution as a flexible second-order expansion in logs, where we also allow for unobserved effects:

$$\mathbb{Q}_\tau[\log(C_{r,it})|\log(\mathbf{v}_{r,it}), \mu_{r,i}] = \alpha_{r,\tau} + \log(\mathbf{v}_{r,it})'\boldsymbol{\beta}_{r,\tau} + \frac{1}{2}\log(\mathbf{v}_{r,it})'\mathbf{B}_{r,\tau}\log(\mathbf{v}_{r,it}) + \mu_{r,i}, \tag{3.1}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where the index $r = \{1, 2\}$ differentiates between the loans-only ($r = 1$) and loans-and-savings ($r = 2$) cost relationships. Here, $\mathbf{v}_{r,it}$ $\forall r = 1, 2$ are the vectors of relevant (strictly exogenous) arguments to the loans-only and loans-and-savings dual cost functions, with the corresponding conformable first-order parameter vectors $\boldsymbol{\beta}_{r,\tau}$ and second-order symmetric parameter matrices $\mathbf{B}_{r,\tau}$. By our definition, $\mathbf{v}_{2,it}$ includes $y_{2,it}$ while $\mathbf{v}_{1,it}$ does not. Also, note that all unknown parameters in model (3.1) are indexed by r indicating that they may vary with the MFI type defined by their output mix. These parameters are also quantile-specific which allows us to accommodate heterogeneity across MFIs belonging to different (conditional) quantiles of the cost distribution.

Following Abrevaya & Dahl (2008), we let the unobserved heterogeneity correlate with regressors by treating $\mu_{r,i}$ as correlated effects, which is a popular device for modeling unit-specific effects in nonlinear models. More specifically, we can approximate unobserved effects $\mu_{r,i}$ conditional on covariates in the spirit of Chamberlain (1980) by specifying the latter in the form of a linear projection on lags and leads of observables: $\mu_{r,i} = \log(\mathbf{v}_{r,i1})'\boldsymbol{\delta}_{r,1} + \cdots + \log(\mathbf{v}_{r,iT})'\boldsymbol{\delta}_{r,T}$. However, given that in our case $N >> T$, we opt for a more parsimonious specification of correlated effects à la Mundlak (1978) by restricting $\boldsymbol{\delta}_{r,1} = \boldsymbol{\delta}_{r,t} = \boldsymbol{\delta}_{r,T}$ which implies the following formulation:

$$\mu_{r,i} = \overline{\log(\mathbf{v}_{r,i})}'\boldsymbol{\gamma}_r, \tag{3.2}$$

where $\overline{\log(\mathbf{v}_{r,i})} \equiv T^{-1}\sum_t \log(\mathbf{v}_{r,it})$ is a vector of time averages, and $\boldsymbol{\gamma}_r$ is a conformable vector of auxiliary parameters.[6]

We estimate the quantile model in (3.1) subject to (3.2) for loans-only and loans-and-savings MFIs separately. For each of these two MFI types, we estimate the model for six different (conditional) cost quantiles, namely for $\tau = \{0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$, with the last 0.95th quantile corresponding to exceptionally large-scale giants in the industry.[7] To investigate the consequences of overlooking distributional heterogeneity across MFIs, we also re-estimate cost functions via a more traditional conditional *mean* model.

4 Data and Results

Our data come from several rating agencies reports compiled by the Center for Research on Social Enterprise and Microfinance (CERSEM) at the University of Agder in Norway. Compared to the Mixmarket data, which are used in most studies on the performance of MFIs, the CERSEM dataset

[6]Since the Chamberlain (1980) formulation nests (3.2), we can formally discriminate between the two specifications by means of Koenker & Machado's (1999) likelihood-ratio test for quantile regressions based on the difference between the restricted (under the null) and unrestricted (under the alternative) sums of estimated check functions. We consistently fail to reject the null of the Mundlak-type specification across the two types of MFIs.

[7]Unlike Malikov & Hartarska (2018) who employ a least-squares estimator with a closed-form solution which facilitates an easy-to-implement restricted estimation, we do not impose regularity conditions onto the estimated cost functions because our quantile estimation requires numerical search and the restrictions are observation-specific. We check for violations of theoretical monotonicities after the estimation: they are only modestly prevalent across the estimated quantiles, averaging at 2.9% and 9.1% of observations for loans-only and the integrated MFIs, respectively.

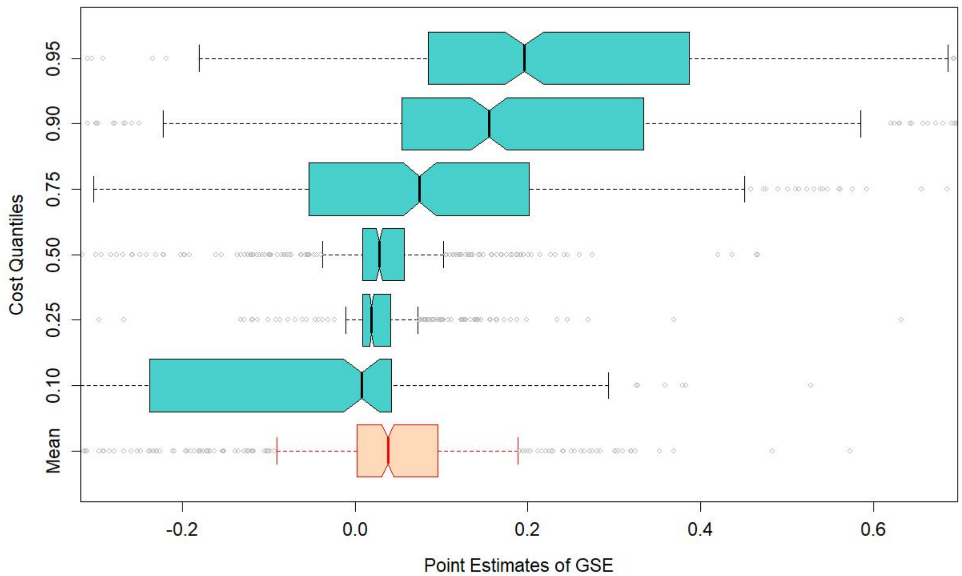

Figure 1. Estimates of Diversification Economies across Cost Quantiles

has some advantages (see Randøy et al., 2015). First, the data are verified by professional external rating agencies and is not self-reported as in the case of Mixmarket. Second, and of the particular importance to our study, the dataset has less of a large-firm bias typical for the Mixmarket dataset. This allows us to explore the economies of diversification with a sample that is more reflective of the microfinance industry with the large number of small institutions, which are most likely to benefit from learning about the cost consequences of offering savings together with loans. Lastly, accurate information on more variables is available in the CERSEM dataset which lets us control for important MFI characteristics such as the MFI's main lending methodology.

Our dataset is an unbalanced panel of 1,467 annual observations for 344 MFIs in 65 countries for the 1998–2013 period. Following the literature, the main cost function variables are as follows. The outputs are total loans (y_1) and, if offered, total savings (y_2). The input prices are those corresponding to physical capital, labor and total borrowed funds. The total cost is the sum of these operating and financial expenditures. In addition to these standard variables, we also include a risk proxy (a share of portfolio at risk for 30 days or more) as well as a fully interacted time trend meant to accommodate temporal shifts in the cost frontiers. Besides allowing for the unobserved heterogeneity being captured via correlated effects, we follow Caudill et al. (2009) in also conditioning the estimated cost relationships on observable controls for the MFI's lending methodology (village banking, individual lending, solidarity group) and legal status (non-governmental organization, cooperative, non-banking financial institution, bank).

Figure 1 summarizes point estimates of economies of diversification (GED) across different quantiles of the MFI cost distribution in the form of a box-plot. Comparing distributions of *observation-specific* estimates of diversification economies for different cost quantiles τ as opposed

Table 1. Statistically Significant Estimates of Diversification Economies

Model	*Diseconomies of Diversification (GED<0)*				*Economies of Diversification (GED>0)*			
	% Obs.	25th Perc.	Median	75th Perc.	% Obs.	25th Perc.	Median	75th Perc.
$\mathbb{Q}_{0.10}$	4.91	–0.607	–0.548	–0.464	8.04	0.041	0.090	0.155
$\mathbb{Q}_{0.25}$	0.67	–0.516	–0.465	–0.447	51.79	0.019	0.032	0.061
$\mathbb{Q}_{0.50}$	1.34	–0.590	–0.511	–0.474	37.95	0.029	0.046	0.099
$\mathbb{Q}_{0.75}$	3.35	–0.641	–0.520	–0.459	39.51	0.092	0.186	0.372
$\mathbb{Q}_{0.90}$	1.34	–0.715	–0.607	–0.556	47.10	0.135	0.295	0.644
$\mathbb{Q}_{0.95}$	0.67	–0.865	–0.749	–0.653	69.64	0.157	0.283	0.526
$\mathbb{M}$ean	0.89	–0.462	–0.418	–0.374	31.92	0.050	0.106	0.165

The "economies of diversification" (GED>0) and "diseconomies of diversification" (GED<0) categories are based on whether GED point estimates are statistically greater and less than zero, respectively. The remaining observations exhibit diversification invariance of cost, i.e., their GED point estimates are not statistically different from zero. For inference, we use wild residual block-bootstrap percentile confidence intervals estimated using 500 bootstrap iterations.

to merely focusing on their marginal moments, we are able to take a more holistic look at differences in the scope-driven cost saving potential across MFIs with different scales of operation. (We caution readers to not confuse quantiles of the cost distribution τ, for which our model is estimated, with quantiles of the fitted GED distribution across individual observations for each τ.) Here, we also include the results from the conditional mean model. In addition, Table 1 reports the results for statistically significant diversification economies estimates from all estimated models broken-down in two categories: positive economies and negative diseconomies of diversification. This classification is based on the GED point estimate being statistically less or greater than zero at the conventional 5% significance level (using bootstrap percentile confidence intervals), with the remaining statistically *in*significant point estimates corresponding to the invariance to output diversification. For inference, we employ Feng et al.'s (2011) wild bootstrap algorithm modified to suit the asymmetric loss function used in the quantile estimation. During implementation, we draw cross-sectional blocks to preserve the panel structure of data. We use 500 bootstrap iterations.

Our empirical results show that the measurement of economies of diversification at the conditional mean of MFI costs provides an incomplete and rather distorted picture of the magnitude of diversification economies and the prevalence thereof in the microfinance industry. The inability of the conditional mean model, which focuses on the microfinance cost structure "on average," to capture heterogeneity across small- and large-cost MFIs is vividly illustrated in Figure 1. More concretely, the results from the quantile estimation indicate that the primary beneficiaries of economies of diversification in the industry appear to be large-scale MFIs lying in the upper quartile of the cost distribution ($\tau \geq 0.75$), whereas the MFIs in the bottom three quartiles mainly have near-zero GED estimates. In contrast, the results from the least-squares conditional mean model are largely dominated by near-zero estimates.

When accounting for the sampling error in the diversification economies estimates, based the conditional mean model we find that only 31.9% of MFIs exhibit positive (and statistically significant) economies of diversification. The quantile results however provide a much stronger evidence

Table 2. p-Values for the Stochastic Dominance Tests of
Diversification Economies across Cost Quantiles

	Single Prospect					Multiple Prospects			
	$\mathbb{Q}_{0.90}$	$\mathbb{Q}_{0.75}$	$\mathbb{Q}_{0.50}$	$\mathbb{Q}_{0.25}$	$\mathbb{Q}_{0.10}$	$\{\mathbb{Q}_{0.90}, \mathbb{Q}_{0.75}, \dots\}$	$\{\mathbb{Q}_{0.75}, \mathbb{Q}_{0.50}, \dots\}$	$\{\mathbb{Q}_{0.50}, \mathbb{Q}_{0.25}, \dots\}$	$\{\mathbb{Q}_{0.25}, \mathbb{Q}_{0.10}\}$
$\mathbb{Q}_{0.95}$	0.939	0.472	0.552	0.105	0.698	0.130	0.105	0.105	0.105
$\mathbb{Q}_{0.90}$		0.331	0.507	0.050	0.763		0.050	0.050	0.050
$\mathbb{Q}_{0.75}$			0.000	0.000	0.482			0.000	0.000
$\mathbb{Q}_{0.50}$				0.000	0.427				0.000
$\mathbb{Q}_{0.25}$					0.341				

in support of the large prevalence of significantly positive economies of diversification in the industry, with up to 69.6% of MFIs in the upper three quartiles of cost distribution ($\tau \geq 0.25$) enjoying diversification economies. These results also differ magnitude-wise. The median value of significant diversification economies across individual MFIs from the conditional mean model is estimated at about 10.6%. However, based on the conditional quantile estimation, we find the median magnitudes to lie in a noticeably higher 18.6–29.5% range for the large-scale MFIs from the top cost quartile ($\tau \geq 0.75$) and to be only in single digits for MFIs of a more modest scale from the rest of the cost distribution. Thus, the quantile regression results indicate that, despite being statistically significant, diversification economies are very timid for the majority of MFIs but are quite material for larger MFIs suggesting that the degree of scope-driven economies of diversification increases with the scale of operations (also see Figure 1). Having said that, the mean and quantile models consistently point to little evidence of *dis*economies of diversification in microfinance.

We formally test for the existence of systemically higher diversification economies among larger-scale MFIs by testing for the (first-order) stochastic dominance of diversification economies exhibited by MFIs in the upper quantiles of the cost distribution over those exhibited in lower quantiles. We use Linton et al.'s (2005) generalization of the Kolmogorov-Smirnov test which permits testing dominance over multiple variables/prospects and allows these variables to be the estimated latent quantities as opposed to observables from the data (as in our case). We formulate the null hypotheses as corresponding to the stochastic dominance[8] and, to perform these tests, we employ the sub-sampling procedure suggested by Linton et al. (2005).[9] The left panel of Table 2 reports p-values for the tests of pair-wise dominance of GED from the "row" quantile model over GED from the "column" quantile model, whereas the right panel reports p-values for the tests of dominance of GED from the "row" quantile over a joint multi-dimensional set of GED from "column" quantiles. Consistent with the visual evidence in Figure 1, the data suggest that economies of diversification among the largest MFIs in the top cost distribution decile ($\tau \geq 0.9$) are *overall* statistically larger than those among smaller-scale MFIs in the rest of the cost distribution.

We conclude our analysis by also testing for the underlying technological and distributional heterogeneity across MFIs, which has served as the principal motivation for our analysis. More

[8]Thus, the failure to reject (when $p \geq 0.05$) implies that one variable stochastically dominates the other(s).

[9]We use $d_N = 199$ equidistant sub-sample sizes $B_n = \{b_1, \dots, b_{d_N}\}$, where $b_1 = [\log\log N]$ and $b_{d_N} = [N/\log\log N]$ with N being the sample size, each of which produces a p-value. The reported are the means of these d_N p-values.

concretely, we use the bootstrap Wald statistic to test if (i) for each cost quantile τ, the common cost function parameters are the same for loans-only ($r = 1$) and loans-and-savings ($r = 2$) MFIs, and (ii) for each MFI type r, the cost function parameters are the same across all and any pair of the estimated cost quantiles τ. The former nulls are consistently rejected with p-values at least as small as 10^{-5}, and the latter nulls are rejected with $p < 10^{-6}$, thereby strongly indicating the presence of non-negligible heterogeneity among MFIs.

5 Concluding Remarks

Our findings lend strong support in favor of a more flexible approach to modeling cost relationships in microfinance that is capable of accommodating distributional and other types of heterogeneity. Our quantile panel data estimates of economies of diversification reveal important differences across individual MFIs that remain unmasked in the more traditional conditional-mean results. In particular, we find that, despite being statistically significant, diversification economies are very modest for the majority of MFIs, which tend to be small in size, but are quite substantial for larger-scale institutions suggesting that the degree of diversification economies increases with the scale of operations. These differences are statistically significant distribution-wise. Therefore, it is imperative that that policy makers and donors be cautious when pushing smaller credit-only MFIs into the provision of savings deposits with the expectations of diversification-driven cost savings, since the latter seem to be more relevant and economically substantive primarily for the very large MFIs.

Appendix

This appendix reports some empirical evidence that the cost distribution of MFIs exhibits fat tails as opposed to what one would have expected of the log-normally distributed variable.

Given the non-negativity of cost and the convention to model cost relationships in logs (like we do in this paper as well), in what follows, we work with the logged cost of MFIs. To partial out influence of the cost determinants, we examine the (conditional) distribution of the log-cost net of its linear projection on the cost-function regressors. Essentially, we work with the residuals from a mean-regression counterpart of our quantile model in (3.1). We contrast this log-cost residual distribution with the normal distribution to see if it is fat-tailed.

We first consider the so-called Q-Q plot of the MFI log-cost (Figure A.1) where we plot empirical quantiles of the standardized log-cost residuals against theoretical quantiles of a standard normal variable. From the figure, we see that, while the cost data fall along a straight line in the middle of the plot, they curve off in the tails. This suggests that, after partialling out, the log-cost variable takes more extreme values than expected, had it been normally distributed. Next, we test for a positive excess kurtosis of the conditional log-cost. The unbiased measure of excess kurtosis (Joanes & Gill, 1998) is estimated at 0.39, with the one-sided p-value of 0.044. This kurtosis estimate is

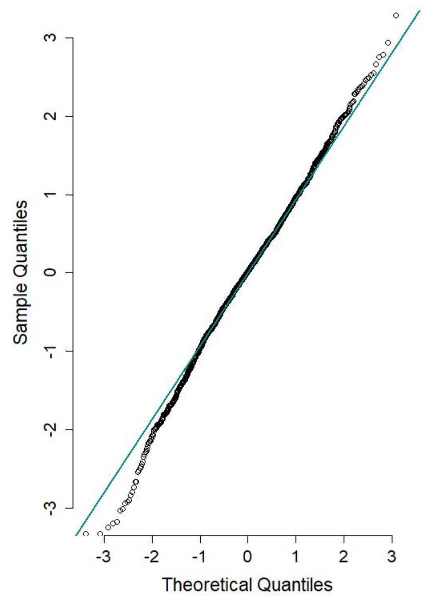

Figure A.1. The Q-Q Plot of the Standardized Log-Cost against the Normal Distribution

significantly *positive* at the conventional level, rejecting a zero-value null under normality, and suggests "heavy tails" of the conditional log-cost distribution.

References

Abrevaya, J. & Dahl, C. M. (2008). The effects of birth inputs on birthweight: Evidence from quantile estimation on panel data. *Journal of Business and Economic Statistics*, *26*, 379–397.

Caudill, S., Gropper, G., & Hartarska, V. (2009). Which microfinance institutions are becoming more cost-effective with time? Evidence from a mixture model. *Journal of Money, Credit and Banking*, *41*(4), 651–672.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, *47*, 225–238.

Chavas, J.-P. & Kim, K. (2010). Economies of diversification: A generalization and decomposition of economies of scope. *International Journal of Production Economics*, *126*, 229–235.

Cull, R., Demirguc-Kunt, A., & Morduck, J. (2011). Does regulatory supervision curtail microfinance profitability and outreach? *World Development*, *39*, 949–965.

Delgado, M. S., Parmeter, C. F., Hartarska, V., & Mersland, R. (2015). Should all microfinance institutions mobilize microsavings? Evidence from economies of scope. *Empirical Economics*, *48*(1), 193–225.

D'Espallier, B., Goedecke, J., Hudon, M., & Mersland, R. (2017). From NGOs to banks: Does institutional transformation alter the business model of microfinance institutions? *World Development*, *89*, 19–33.

Evans, D. S. & Heckman, J. J. (1984). A test for subadditivity of the cost function with an application to the Bell System. *American Economic Review*, *74*(4), 615–623.

Feng, X., He, X., & Hu, J. (2011). Wild bootstrap for quantile regression. *Biometrika*, *98*, 995–999.

Hartarska, V., Parmeter, C., & Nadolnyak, D. (2011). Economies of scope of lending and mobilizing deposits in rural MFIs: A semiparametric approach. *American Journal of Agricultural Economics*, *93*(2), 389–398.

Hartarska, V., Parmeter, C., Nadolnyak, D., & Zhou, B. (2010). Economies of scope for microfinance: Differences across output measures. *Pacific Economic Review*, *15*(4), 464–481.

Hartarska, V., Shen, X., & Mersland, R. (2013). Scale economies and elasticities of substitution in microfinance institutions. *Journal of Banking and Finance*, *37*(1), 118–131.

Hermes, N., Lensink, R., & Meesters, A. (2011). Outreach and efficiency of microfinance institutions. *World Development*, *39*(6), 938–948.

Hughes, J. P. & Mester, L. J. (1993). A quality and risk-adjusted cost function for banks: Evidence on the "too-big-to-fail" doctrine. *Journal of Productivity Analysis*, *4*, 293–315.

Joanes, D. N. & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society, Series D (The Statistician)*, *47*, 183–189.

Koenker, R. & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*, 1296–1310.

Linton, O., Maasoumi, E., & Whang, Y.-J. (2005). Consistent testing for stochastic dominance under general sampling schemes. *Review of Economic Studies*, *72*, 735–765.

Malikov, E. & Hartarska, V. (2018). Endogenous scope economies in microfinance institutions. *Journal of Banking and Finance*, *93*, 82–109.

Malikov, E., Zhao, S., & Kumbhakar, S. C. (2017). Economies of diversification in the US credit union sector. *Journal of Applied Econometrics*, *32*(7), 1329–1347.

Mersland, R. & Strøm, R. O. (2009). Performance and governance in microfinance institutions. *Journal of Banking and Finance*, *33*(4), 662–669.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46*, 69–85.

Panzar, J. C. & Willig, R. D. (1981). Economies of scope. *American Economic Review*, *71*(2), 268–272.

Randøy, T., Strøm, R. Ø., & Mersland, R. (2015). The impact of entrepreneur-CEOs in microfinance institutions: A global survey. *Entrepreneurship Theory and Practice*, *39*(4), 927–953.