Endogenous Scope Economies in Microfinance Institutions

Malikov, Emir and Hartarska, Valentina

Auburn University, Auburn University

25 March 2018

Endogenous Scope Economies in Microfinance Institutions*

Emir Malikov Valentina Hartarska

Auburn University, Auburn, AL

May 25, 2018

Abstract

Scope economies resulting from the joint offering of loans and savings accounts (as opposed to loans only) are customarily invoked to promote the transformation of credit-only microfinance institutions (MFIs) into integrated loans-and-savings entities. To ensure robust inference, we estimate scope economies for the microfinance industry using a novel approach which, among its other advantages, accommodates inherent heterogeneity across loans-only and loans-and-savings MFIs as well as controls for endogenous self-selection of institutions into the either type. For analysis, we use a large 2004–2014 Mixmarket dataset. Unlike earlier studies, we do *not* find prevalent scope economies in the microfinance industry. We find that the median degree of scope economies is statistically indistinguishable from zero and that scope economies are significantly positive for less than a half of loans-and-savings MFIs. For a non-trivial 14% of institutions, the empirical evidence suggests the existence of significantly negative *dis*economies of scope indicating that the separate production of loans and savings accounts actually has the potential to reduce an MFI's costs. We also find that the failure to account for endogenous selectivity dramatically overestimates the degree of scope economies resulting in the failure to detect scope diseconomies among MFIs. Thus, our findings call for caution when invoking scope economies as a blanket justification for universal expansion of the scope of financial operations by MFIs. Instead, promoting integrated loans-and-savings MFIs may be justifiable as a means to meeting the needs of the poor rather than as a way for the industry to save costs.

Keywords: microfinance institutions, scope economies, endogenous selection, financial intermediation, savings and lending

JEL Classification: G15, G21, O16, L33

* *Email*: emalikov@auburn.edu (Malikov), hartavm@auburn.edu (Hartarska).

We would like to thank the editor, associate editor and two anonymous referees for many insightful comments and suggestions that helped improve this paper. Any remaining errors are our own.

1 Introduction

The microfinance industry consists of diverse entities offering small-scale financial products to the poor who lack access to standard banking services. In the past two decades, a major trend in the industry has been for typical loans-only microfinance institutions (MFIs) engaged solely in lending activities to transform into loans-and-saving MFIs that also offer saving products to their customers. While a decade ago no more than a third of MFIs (including credit unions) offered voluntary savings accounts, by 2014 (last year in our dataset) the share of such loans-and-savings MFIs had increased dramatically to about 54%. Policymakers, donors and socially oriented investors have provided incentives for MFIs to get licensed to collect savings deposits or have preferred to fund mainly loans-and-savings MFIs in part to respond to the evidence that the poor demand more than just loans (Collins et al., 2009). The main objective however has been to promote MFIs' self-sustainability, to improve their access to commercial funds and thereby to decrease their subsidy-dependence. The promotion of integrated loans-and-savings MFIs has been justified on the grounds of their potential performance improvements of which there are two commonly invoked sources. First, a license to collect savings deposits from customers is usually associated with the ability of MFIs to overcome size barriers to entry and thus to capitalize on scale economies associated with larger size. Second and more importantly, improvements in MFIs' performance have been expected to emerge from scope economies stemming from the *joint* delivery of (micro)loans and (micro)savings.

Since significant resources are used to promote organizational transformation of MFIs as well as because preferential funding to loans-and-savings MFIs leaves loans-only MFIs with less resources available, it is imperative that the claims about existing scope economies be substantiated with robust empirical findings. Anecdotal evidence or stakeholders' believes are hardly sufficient to inform the choice of an industry serving over 170 million poor clients worldwide (Microfinance Information Exchange, 2015). Furthermore, providing financial services to the poor remains challenging and, in the absence of substantial scope and/or scale economies, loans-and-savings MFIs may end up with a "mission drift" away from serving their target customers. For instance, while the financial sustainability of Grameen Bank, a flagship MFI, had improved materially after it changed its business model to start offering microsavings, these improvements also coincided with a simultaneous abandonment of its poorest clients (Hulme, 2008).[1] Once licensed to collect savings deposits, loans-and-savings MFIs become subject to banking regulations, and there is substantial evidence that the new, more stringent supervisory environment entices profit-oriented MFIs to curtail outreach to women and, more generally, to costly-to-reach-customers (Cull et al., 2011). While scope economies have been studied for commercial banks engaged in the traditional financial intermediation (e.g., Berger et al., 1987, 1999; Saunders, 2000) as well as for those with more diversified activities (Elsas et al., 2010), the results from these bank-centered studies cannot be easily extended to MFIs owing to their specific outreach mission. This warrants a rigorous stand-alone measurement of scope economies in the microfinance industry based on recent data reflective of changes that the industry has gone through. In the absence of such scope economies, the simultaneous satisfaction of outreach and sustainability may remain a challenge.

In this paper, we seek to provide robust empirical evidence on the existence, magnitude and the distribution of scope economies (if any) in the microfinance industry worldwide during a more recent time period (2004–2014) relative to the existing work based on the data prior to 2006. Our contribution to the literature is as follows. First, when assessing the extent of scope economies in MFIs, we allow loans-only and loans-and-savings MFIs to have *heterogeneous* production tech-

[1] Similar results have also been found for other transformed MFIs worldwide; e.g., see Woller et al. (1999) and Wagenaar (2012).

nologies, which is starkly different from the approach pursued in virtually all prior studies of scope economies in the industry that estimate a single microfinance cost structure *a priori* presumed to be common to all MFIs with no regard to their heterogeneous mixes of financial services offered to customers. Second, our methodology explicitly recognizes that the above technological heterogeneity is an outcome of endogenous *self-selection* by institutions whereby offering deposit accounts as an additional financial service (and consequently adopting the appropriate production technology to keep costs at minimum) is the endogenous decision of MFIs. Hence, it would be econometrically inappropriate to treat the observed type of an MFI—loans-only versus loans-and-savings—as being exogenously/randomly assigned, likely resulting in inconsistent and potentially misleading estimates of scope economies. To our knowledge, no prior study has entertained this likely possibility. Third, we employ Kyriazidou's (1997) *semiparametric* kernel estimator to estimate microfinance technologies with selectivity, which allows for unobserved heterogeneity across MFIs and does *not* require distributional assumption or parametrization of the dependence between the outcome and selection equations. By using this model, we are able to strike a balance between mitigating potential misspecification (e.g., by avoiding a restrictive and potentially incorrect bivariate normality assumption in the popular Heckman's selection model) and alleviating the curse-of-dimensionality problem immanent in nonparametric methods.[2] Fourth, building on Malikov et al.'s (2016) work, in contrast to popular alternatives employed in the microfinance literature, our measurement of scope economies does *not* rely on a rather unrealistic assumption whereby specialized loans-only MFIs share the same technology with and incur the same fixed costs as do the integrated loans-and-savings MFIs. The latter substantially decreases the reliance of our scope economies estimates on counterfactuals thereby minimizing the "excessive extrapolation" problem (Evans & Heckman, 1984) that most studies of scope economies inherently suffer from. Lastly but not least importantly, we use more recent data on MFIs from around the world during the 2004–2014 period which describes an industry much different from that analyzed in earlier work. Altogether, we are therefore able to offer policy-makers and stakeholders a fresher and more robust perspective on benefits and costs of promoting integrated loans-and-savings MFIs on grounds of the cost saving potential due to scope economies.

To briefly preview the results, we find that the microfinance industry largely exhibits invariance to scope of outputs produced, with the median degree of scope economies estimated at statistically insignificant –0.06. After controlling for endogenous self-selection, scope economies are significantly positive for 46% of loans-and-savings MFIs only. Perhaps more importantly, for a non-trivial 14% of institutions, the empirical evidence suggests the existence of significantly negative *dis*economies of scope indicating that the *separate* production of loans and savings accounts actually has the potential to reduce an MFI's costs. The mean degree of scope diseconomies for these multi-output MFIs is estimated at a sizable –0.21 implying, on average, the potential for a 21% cost saving if the joint production of loans and savings is replaced with two single-output MFIs. However, note that the presence of such scope diseconomies in no way implies cost inefficiency or sub-optimality on the part of these loans-and-savings MFIs. Neither does it say that these MFIs may not reduce their costs by scaling up their operations to capitalize on (universally) significant scale economies. It does however suggest that it may be ill-advised to invoke scope economies as a blanket justification for *universal* expansion of the scope of financial operations by MFIs. After all, scope economies are significantly positive only for 46% of MFIs, with the average value estimated at a non-negligible 0.23. (The analysis of temporal dynamics in these scope economies shows that their magnitude as

[2]The problem refers to the phenomenon whereby the convergence rate of a nonparametric estimator worsens with an increase in the number of continuous covariates in the model. We circumvent this problem by imposing some parametric structure on the outcome equation; hence, we have a "semiparametric" model.

well as the prevalence in the industry have been steadily declining, especially in the aftermath of the global financial crisis.) For 50% of institutions, the costs exhibit scope invariance as indicated by statistically insignificant estimates of the degree of scope economies.

Among prior studies, Hartarska et al. (2011) and Delgado et al. (2015), who also estimate the degree of scope economies in microfinance, are perhaps the most closely related to our paper.[3] However, both of these studies estimate a single cost function for all MFIs regardless of financial services they offer (or do not offer) and consequently do not correct for MFIs' endogenous self-selection into the loans-and-savings type.[4] Also, these papers study the period before the financial crisis [Delgado et al. (2015) use data up to 2006, and Hartarska et al.'s (2011) data stop in 2008] when the loans-and-savings institutions comprised only about a fifth of the microfinance industry, whereas such MFIs are the majority in our 2004–2014 data. In contrast to our results, whereby the median degree of scope economies estimate is statistically indistinguishable from zero, Delgado et al. (2015) report significantly positive scope economies of about 10% at the median. As a matter of fact, they find that as many as 65% of MFIs in their sample exhibit statistically significant scope economies leading them to conclude that "in general MFIs realize positive and significant reductions in costs from offering both loans and savings" (Delgado et al., 2015, p.212) which starkly differs from the conclusion we arrive at using our robust methodology and more recent data. Our findings also differ from those by Hartarska et al. (2011) who too generally find empirical support for positive scope economies in the industry with the statistically significant median estimate of 19%. They also hardly find any evidence of diseconomies of scope among MFIs with most of their reported negative estimates being statistically insignificant instead suggesting scope invariance, whereas we find that the MFIs with significantly negative scope diseconomies actually constitute a non-negligible 14% of the microfinance industry. Other related studies of MFIs' performance have focused on scale economies and efficiency. The results generally point to prevalent increasing returns to scale (which we confirm as well) and non-ignorable heterogeneity in cost effectiveness of MFIs depending on their subsidy-reliance and target clientele (Caudill et al., 2009; Hartarska et al., 2013).[5] More recently, D'Espallier et al. (2017) have examined how the financial performance of MFIs changes with their transformation from non-governmental organizations into banks.

The rest of the paper unfolds as follows. Our framework for modeling the microfinance production technology and measuring endogenous scope economies in the industry is introduced in Section 2. Section 3 then presents our two-step semiparametric estimation methodology. We discuss the data in Section 4. The empirical results are reported in Section 5. Section 6 concludes.

2 Endogenous Scope Economies

In order to examine whether MFIs have access to the untapped cost saving potential due to scope and/or scale economies, we first need to formalize the model of their production costs.

Following the vast majority of banking and microfinance literature, we model the production technology of MFIs via dual cost function. Not only is the dual cost approach to modeling microfi-

[3]Two other studies of scope economies use the same methods but focus on differences attributable to alternative measures of output quantities (Hartarska et al., 2010) and on linking scope economies to the MFI governance attributes (Hartarska et al., 2013).

[4]In contrast, we find that the data consistently reject the null of common technology shared by loans-only and loans-and-savings MFIs and that the selection between the two in not exogenous/ignorable.

[5] Hartarska et al. (2013) find substantial increasing returns to scale for all but profitability-focused deposit-mobilizing MFIs. Caudill et al. (2009) show that MFIs, which get more cost effective over time, are less reliant on subsidies and are more likely to target women.

nance technology convenient because it facilitates the direct measurement of MFIs' costs necessary for the evaluation of their scope economies, but it also does not require the use of input quantities during the estimation (unlike in the primal production framework) which can lead to simultaneity problems since the input allocation is endogenous to MFIs. The structural identification of cost functions instead relies on the exogeneity of competitively set input prices along with output quantities being fixed (at least in the short run) as theoretically justified on the grounds of cost minimization premise widely accepted in the financial services literature [e.g., see Hughes & Mester (2015) for an excellent review]. The use of cost function is also advantageous over an alternative dual specification of production technology in the form of a profit function which requires information on output prices. First, the microfinance industry does not collect consistent output price data (i.e., interest rates charged on various loans). Furthermore, while being price takers in the input markets, MFIs have some market power in the output market because they serve poor clients whom other traditional lenders tend to avoid. In the market for inputs, the environment is however largely competitive: MFIs pay competitive salaries for relatively skilled labor such as loan officers, compete with peers worldwide for the access to financial capital (equity, loans and donations) and participate in a competitive market for physical infrastructure. Therefore, while input prices are widely accepted to be exogenous, the same cannot be easily afforded for output prices thereby making the estimation of dual profit functions problematic from an econometric point of view. Also, MFIs are organizationally diverse and, while some operate as for-profit entities, a larger fraction of them remain non-profit thus rendering the profit-maximizing objective somewhat questionable. Lastly but not least importantly, the stated objectives of microfinance are (i) outreach, or serving marginalized clients, and (ii) sustainability, which is widely understood as covering cost as opposed to maximizing profits.

Given the bank's core functions as a financial intermediary, most empirical studies of the performance of commercial banks apply the so-called "intermediation approach" when modeling their production technology. In this approach, deposits are conceptualized as inputs to the production of various types of loans and investments which are all treated as outputs. Unlike commercial banks, MFIs are however focused more on serving the marginalized clientele and not on intermediation *per se* (Cull et al., 2011). Specifically, MFIs offer costly, characteristically small-scale savings accounts to their clientele because the poor need such services (Collins et al., 2009); offering these microsavings meets the industry's larger overarching objective to expand the financial frontier and to offer a wide range of financial services to the marginalized clients excluded from a formal financial sector. This contrasts with the banks' goals of mobilizing deposits in order to decrease the cost of capital and extend loans. Indeed, the capital structure of MFIs primarily includes equity (if organized as a shareholder company) or quasi-equity (accumulated donations and start-up capital), loans from commercial banks, socially motivated lenders, governmental and intra-governmental organizations, and deposits, if an MFI collects them. To extend loans to their clients, MFIs rely on own capital and loans more than on the savings deposits, mainly because of the limited availability of savings among the poor living in urban slums or remote rural areas. This distinction is apparent in the capital structure of MFIs relative to that of banks. While banks worldwide have the deposits-to-assets ratio of over 80% (Barth & Wihlborg, 2017), in our sample of loans-and-savings MFIs, this ratio is only half of that at 43%.[6] The share of MFIs' "own" capital in the form of equity and quasi-equity comprises 26%, while the non-deposit borrowed funds take up 29%; together, these two sources are larger than savings. Thus, the production of microdeposits in itself is one of the MFI's goals aimed at meeting the demands of the target clientele. Therefore, following the bulk of

[6]Our data reveal that even the MFIs organized as banks have the deposits-to-assets ratio of 48% with other types of MFIs such as non-governmental organizations having the ratio of only 25%.

microfinance literature, we adopt a modified "production approach" to microfinance, whereby deposits are treated as outputs, which is better suited to study MFIs (Caudill et al., 2009; Hartarska et al., 2011, 2013; Hermes et al., 2011; Delgado et al., 2015). In fact, we formally test if our data corroborate such an approach by empirically verifying the output-like role of microdeposits before proceeding to the estimation of our main model (see Section 5).

Within this approach, the two outputs are the total dollar value of (micro)loans (y_1) and (micro)deposits (y_2), where the latter savings deposits are also sometimes interchangeably referred to as (micro)savings. The four inputs to the microfinance production are labor (x_1), physical capital (x_2), financial capital (x_3) and equity (e). The labor input is measured by the number of employees with the corresponding input price (w_1) being the average annual wage. Physical capital is defined as the number of physical offices (including mobile facilities) that the MFI operates to extend loans and collect savings deposits; the corresponding input price (w_2) is the non-labor operating cost per office. Financial capital is the non-deposit borrowed funds with the corresponding price w_3. The treatment of equity as an input to the microfinance production technology is consistent with the argument that MFIs may use it as a source of loanable funds and thus as a cushion against losses. However, due to the unavailability of the price of equity, we adopt the common practice in the banking literature by modeling it as a quasi-fixed input (the other three inputs are said to be freely varying). Lastly, the MFI's variable cost (C) is defined as the sum of expenditures on the above three variable inputs and thus measures the operating and non-deposit financial expenses.

2.1 Heterogeneous Microfinance Technologies

Virtually all prior attempts to empirically assess the extent of scope economies in the microfinance industry have done so by estimating a *single* multi-output cost function for all MFIs with no regard to their heterogeneous output mixes. Such an approach is problematic at least on two fronts.

First, it embodies a rather strong and unrealistic assumption that loans-only and loans-and-savings MFIs share the *same* underlying production technology despite apparent differences in the scope of their outputs. Under this assumption, the cost structure of the loan issuance operations by MFIs is *a priori* presumed to be the same for both loans-only and loans-and-savings institutions thereby effectively implying that the microfinance production technology is invariant to the scope of produced outputs. Essentially, such an approach takes the observed technological heterogeneity across MFIs in the form of their differing output mixes completely for granted, which is likely to result in the loss of information and the misspecification of an econometric model. Therefore, in our paper, we model the cost structure of loans-only and loans-and-savings MFIs separately.

Second and perhaps most importantly, the conventional single-equation approach fails to recognize that the above technological heterogeneity is likely to be an outcome of *endogenous* self-selection by institutions. That is, offering voluntary savings accounts as an additional financial service to their customers — and consequently adopting the appropriate production technology to produce all outputs at the minimum cost — is the endogenous decision that MFIs make based on the set of relevant internal and external factors. If so, it is econometrically inappropriate to treat the observed type of an MFI (loans-only versus loans-and-savings) as being exogenously/randomly assigned. We therefore explicitly model endogenous selection of the output scope and the corresponding output-mix-specific production technology by MFIs.

Specifically, we consider the following (output-mix-specific) heterogeneous dual variable cost functions for loans-only and loans-and-savings MFIs:

$$\mathbb{C}_1(\boldsymbol{w}, e, y_1) = \inf_{\boldsymbol{x}} \left\{ \boldsymbol{x}'\boldsymbol{w} \mid \mathbb{T}_1(\boldsymbol{x}, e, y_1) \leq 1 \right\} \qquad \text{if} \quad D = 1 \tag{2.1}$$

$$\mathbb{C}_2\left(\boldsymbol{w}, e, y_1, y_2\right)=\inf_{\boldsymbol{x}}\left\{\boldsymbol{x}'\boldsymbol{w} \mid \mathbb{T}_2\left(\boldsymbol{x}, e, y_1, y_2\right) \leq 1\right\}, \qquad \text{if} \quad D=2 \tag{2.2}$$

where $\boldsymbol{x}=(x_1,x_2,x_3)'$ is a vector of variable inputs with the corresponding vector of input prices $\boldsymbol{w}=(w_1,w_2,w_3)'$. The production transformation functions $\mathbb{T}_r(\cdot)$ and the associated variable cost functions $\mathbb{C}_r(\cdot)$ for $r=1,2$ are allowed to vary across the two types of MFIs not only in their dimensionality (five versus six covariates) but also in their underlying specification. We therefore index them differently. Only one of these two cost functions applies to a given MFI as captured by the indicator variable D which equals 1 for loans-only institutions (i.e., if $y_2=0$) and 2 for the more diversified loans-and-savings MFIs (i.e., if $y_2>0$). In turn, we can conceptualize the endogenous determination of the MFI type (i.e., whether $y_2>0$ or not) in the form of the following output-mix-selection equation:

$$D=1+\mathbb{1}\{D^*\geq 0\}, \tag{2.3}$$

where D^* is a latent variable representing the MFI's propensity to expand the scope of its outputs by adding voluntary savings accounts y_2 to the menu of its financial services; and $\mathbb{1}\{\cdot\}$ is the standard indicator function.

Note that, by modeling cost structures of loans-only and loans-and-savings MFIs separately, we are able to explicitly accommodate potential discontinuities (say, due to parameter instability) in the MFI's variable cost function across different financial service menus, which are likely to arise due to the change in dimensionality of the cost function's support across the two types of MFIs (since the number of outputs is not the same). Further, unlike the widely used single-equation model where a common cost relationship is assumed to be shared by all MFIs regardless of their output mixes, the output-mix-specific cost functions in (2.1)–(2.2) do not suffer from the problem of having to deal with zeros reported for y_2 that loans-only MFIs do not produce. In our case, the cost function contains only those outputs that MFIs actually produce and thus are non-zero. That is, having loans-only MFIs not offer voluntary savings accounts is being treated not as a simple matter of the variable y_2 taking a zero value but rather as an indication that this variable does not enter the loans-only MFI's cost optimization problem whatsoever.

2.2 Measuring Scope Economies

The output-mix-specific variable cost functions in (2.1)–(2.2) can readily be used to compute the degree of scope economies (if any) for the microfinance industry. Such scope economies are said to exist if the firm's average cost declines with the increase in the number of outputs it produces (Panzar & Willig, 1981). If the latter were the case for MFIs, one would expect loans-and-savings institutions to generally incur lower average costs than loans-only MFIs do *ceteris paribus*. This may occur for a number of reasons including the potential to spread fixed costs over the larger output mix as well as cost complementarities across different products. For instance, spreading fixed costs over the expanded product mix can contribute to scope economies of MFIs if the provision of loans and savings jointly (rather than separately) reduces institutions' excess capital capacity. On the other hand, cost complementarities may also lead to scope economies if consumer information developed in the production of either loans or savings is being reused by MFIs to reduce the monitoring requirements of the other product.

In what follows, we describe our methodology for measuring the degree of scope economies in the microfinance industry. To help us highlight its merits over the existing less robust alternatives, we first consider the more conventional approach to quantifying the degree of economies of scope that most prior studies have resorted to. Specifically, if we were to follow the literature and employ the traditional measurement of scope economies à la Baumol et al. (1982), then a (multi-output) loans-and-savings MFI would be said to enjoy economies of scope if the cost of producing loans and savings individually is greater than the cost of their joint production, i.e.,

$$\mathbb{C}_2(y_1, 0) + \mathbb{C}_2(0, y_2) > \mathbb{C}_2(y_1, y_2), \tag{2.4}$$

where we have suppressed all other arguments of the variable cost function besides outputs.

Naturally, the evaluation of (2.4) requires computation of the counterfactual cost of producing each of the two outputs separately with the counterfactual being defined as the case of fully specialized single-output production processes. As emphasized by Evans & Heckman (1984) and Hughes & Mester (1993), the approach expectedly suffers from the "excessive extrapolation" problem since the two needed counterfactuals require a projection of the estimated *multi*-output cost function $\mathbb{C}_2(\cdot,\cdot)$ to the case of complete *single*-output specializations $(y_1, 0)$ and $(0, y_2)$. Further, not only does the measure in (2.4) extrapolate the estimated $\mathbb{C}_2(\cdot,\cdot)$ to its boundary, but it also assumes that the specialized single-output MFIs share the same technology and incur the same fixed costs as the integrated loans-and-savings MFIs, which is rather unrealistic (see Fuss & Waverman, 2002).

Therefore, in order to alleviate the "excessive extrapolation" problem to the largest degree feasible, in this paper we seek to minimize the extent of extrapolation required for the measurement of scope economies. Building on Malikov et al. (2016), we do so in several ways.

First, we make explicit use of the observed technological heterogeneity across loans-only and loans-and-savings MFIs, which the literature has largely ignored. The existence of fully specialized loans-only institutions uniquely enables us to identify their single-output cost relationship, which we can use in order to decrease the number of counterfactuals necessary for the evaluation of scope economies in the industry from two to one. More specifically, we adapt the measure of scope economies for loans-and-savings MFIs to reflect the following more theoretically accurate cost comparison:

$$\mathbb{C}_1(y_1) + \mathbb{C}_2(0, y_2) > \mathbb{C}_2(y_1, y_2), \tag{2.5}$$

where the more appropriate single-output loans-only cost function $\mathbb{C}_1(\cdot)$ is now being used to evaluate the cost of producing y_1 separately. The definition in (2.5) is therefore "less counter" to *factum*. Nonetheless, while constituting an improvement over the more conventional (2.4), the modified scope economies measure in (2.5) still requires extrapolation of the multi-output cost function $\mathbb{C}_2(\cdot,\cdot)$. Indeed, it would be most desirable to also use the single-output savings-only cost function when computing the cost of producing y_2 separately. For instance, such an approach would be in line with Berger et al.'s (2000) study of scope economies in the U.S. insurance industry in which the authors also differentiate between heterogeneous specialized and joint production technologies (although, unlike us, without accounting for the endogenous selection thereof). Unfortunately, the latter is *in*feasible in our case owing to the non-existence of such savings-only MFIs in practice, which preclude us from identifying their cost function.[7] However, we can avoid the extreme extrapolation of $\mathbb{C}_2(\cdot,\cdot)$ to the *non-existent* counterfactual case of the full specialization in y_2 by instead focusing on its most closely related output-mix configuration observable in the data, namely the production of (y_1, y_2) exhibiting a *higher* degree of specialization in y_2.

Specifically, in line with Evans & Heckman (1984), a loans-and-savings MFI is then said to

[7] Berger et al. (2000) happen to observe "specialists" of both kinds which uniquely enables them to successfully identify cost functions for all specialized insurers. In the case of microfinance industry however, we only have data on one type of specialized firms, which ultimately necessitates extrapolation of the joint technology to the case of specialization in one of the two inputs.

enjoy economies of scope if

$$\mathbb{C}_1\big(\varpi y_1\big) + \mathbb{C}_2\Big((1-\varpi)y_1, y_2\Big) > \mathbb{C}_2\big(y_1, y_2\big), \tag{2.6}$$

where $0 \leq \varpi \leq 1$ is some distribution weight. Note that (2.6) does *not* require the computation of the cost for a non-existent *fully* specialized MFI producing exclusively y_2 (unless $\varpi = 1$). To operationalize our measure of scope economies in (2.6), we need to specify the distribution weight ϖ. To further avoid excessive extrapolation, we follow Evans & Heckman (1984) and restrict the choice of ϖ to the "admissible region". More concretely, we ensure that each counterfactual MFI does not produce less of each output than MFIs actually do in the data.[8] Therefore, we require that $\varpi y_1 \geq \underline{y}_1$ and $(1-\varpi)y_1 \geq \underline{y}_1$, where $\underline{y}_1$ denotes sample minimum of y_1. We make this constraint more operational by defining our measure of the degree of scope economies (DSE) as follows

$$DSE = \frac{\Big[\mathbb{C}_1\big(\varpi y_1^* + \underline{y}_1\big) + \mathbb{C}_2\Big((1-\varpi)y_1^* + \underline{y}_1, y_2\Big)\Big] - \mathbb{C}_2\big(y_1, y_2\big)}{\mathbb{C}_2\big(y_1, y_2\big)}, \tag{2.7}$$

where $y_1^* = y_1 - 2\underline{y}_1$. Thus, DSE measures the *within-sample* scope economies, positive (negative) values of which correspond to economies (*dis*economies) of scope. The instance of DSE=0 indicates invariance of costs to the scope of outputs.

Lastly, our scope economies measure is obviously dependent on the choice of the ϖ weight. To circumvent this problem, we adopt a conservative approach to measuring the degree of scope economies, whereby we choose ϖ (via a grid search) that yields the *smallest* value for DSE in the admissible region. Specifically, we measure the the degree of "global" economies of scope (for each MFI) as follows:

$$DGSE = \inf_{\varpi}\ DSE(\varpi). \tag{2.8}$$

The intuition is as follows. If scope economies of the smallest degree are still statistically significantly above zero, we can conclude that scope economies are "globally" significant over the loans-and-savings MFI's entire (admissible) output space in a given year. Combined with the fact that we employ the more theoretically appropriate cost function when evaluating the counterfactual cost of producing y_1, the latter renders our methodology significantly more robust to excessive extrapolation than that routinely employed in the literature.

3 Semiparametric Estimation

In this section, we describe an econometric model that we employ in order to estimate the cost structure of microfinance institutions while allowing for both the observed and unobserved technological heterogeneity across MFIs producing different output mixes, the choice of which is itself an outcome of endogenous self-selection. Readers not interested in the estimation details may jump directly to Section 4.

We let our model of output-mix-specific heterogeneous dual variable cost functions for loans-only and loans-and-savings MFIs subject to sample selection in (2.1)–(2.3) take the following form:

$$\log(C_{1,it}) = \log(\mathbf{v}_{1,it})'\boldsymbol{\beta}_1 + \tfrac{1}{2}\log(\mathbf{v}_{1,it})'\mathbf{B}_1\log(\mathbf{v}_{1,it}) + \mu_{1,i} + u_{1,it} \qquad \text{if} \quad D_{it} = 1 \tag{3.1}$$

[8] The admissible region in Evans & Heckman (1984) or Malikov et al. (2016) also regulates the range of values for the ratio of distributed outputs. In our case, such a constraint is redundant (and therefore omitted) because the computation of scope economies involves the distribution of a *single* output only, namely y_1.

$$\log(C_{2,it}) = \log(\mathbf{v}_{2,it})'\boldsymbol{\beta}_2 + \tfrac{1}{2}\log(\mathbf{v}_{2,it})'\mathbf{B}_2\log(\mathbf{v}_{2,it}) + \mu_{2,i} + u_{2,it} \qquad \text{if} \quad D_{it} = 2 \tag{3.2}$$

$$D^*_{it} = \mathbf{z}'_{it}\boldsymbol{\gamma} - v_{it} \quad \text{with} \quad v_{it} \equiv \zeta_i + \epsilon_{it}, \tag{3.3}$$

where we have specified unknown cost functions $\mathbb{C}_1(\cdot)$ and $\mathbb{C}_2(\cdot)$ as flexible second-order expansions in logs. Here, $\mathbf{v}_{1,it} \equiv (\boldsymbol{w}'_{it}, e, y_{1,it})'$ and $\mathbf{v}_{2,it} \equiv (\boldsymbol{w}'_{it}, e, y_{1,it}, y_{2,it})'$ are the vectors of (strictly exogenous) arguments to the cost function for loans-only and loans-and-savings loans, respectively, with their corresponding conformable first-order parameter vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ and second-order symmetric parameter matrices $\mathbf{B}_1$ and $\mathbf{B}_2$. The MFI type indicator variable $D_{it} = 1 + \mathbb{1}\{D^* \geq 0\}$ is as defined earlier, with the latent D^*_{it} being modeled as a function of some relevant contextual variables $\mathbf{z}_{it}$ capturing the institution's internal and external environment. Lastly, $u_{1,it}$, $u_{2,it}$ and ϵ_{it} are the *i.i.d.* random errors.

Note that the unknown parameters differ across (3.1) and (3.2) thereby allowing for (observed) technological heterogeneity across loans-only and loans-and-savings MFIs. We also allow for unobserved heterogeneity amongst individual institutions as captured by MFI-specific unobserved effects $\mu_{1,i}$, $\mu_{2,i}$ and ζ_i. We treat the unobserved effects in the outcome equations $\mu_{1,i}$ and $\mu_{2,i}$ as capturing the fixed-effect-type heterogeneity by allowing these effects to correlate with any of the right-hand-side covariates in an arbitrary way.[9] However, given that such fixed effects are difficult to accommodate in a semiparametric distribution-free estimation of nonlinear selection equations (which we seek to employ in our paper for robustness purposes), we assume that ζ_i is a random effect, implying that the composite error term in the selection equation v_{it} is also *i.i.d.*

Essentially, system in (3.1)–(3.3) is a panel data endogenous switching regressions model, where (3.1)–(3.2) are the outcome equations of interest and (3.3) is a selection equation governing the self-selection of MFIs into either the loans-only or loans-and-savings type. We assume that random disturbances $(u_{1,it}, u_{2,it}, v_{it})$ are *i.i.d.* over i with zero means and finite variances and are strictly mean-independent of right-hand-side covariates, i.e., $\mathbb{E}[u_{r,it}|\boldsymbol{\chi}_{r,i}] = 0$ and $\mathbb{E}[v_{it}|\boldsymbol{\chi}_{r,i}] = 0$ with $\boldsymbol{\chi}_{r,i} \equiv [\log(\mathbf{v}_{r,i1})', \ldots, \log(\mathbf{v}_{r,iT})', \mathbf{z}'_{i1}, \ldots, \mathbf{z}'_{iT}, \mu_{r,i}]\ \forall\ r = 1, 2$. However, the distributions of these random errors are allowed to correlate, i.e., $\mathbb{E}[u_{r,it}v_{it}|\boldsymbol{\chi}_{r,i}] \neq 0\ \forall\ r = 1, 2$. That is, we allow for the correlation between the disturbances in outcome equations (3.1)–(3.2) and that in the selection equation (3.3). Ignoring this correlation would naturally lead to selection bias in the estimates of outcome equations of interest (the cost functions) and hence the estimates of scope economies.

To see why the estimation of cost functions in (3.1)–(3.2) suffers from selectivity bias, consider the following. First, for convenience, define $\boldsymbol{\omega}_{r,i} \equiv (\log(\mathbf{v}_{r,it})', \log(\mathbf{v}_{r,is})', \mathbf{z}'_{it}, \mathbf{z}'_{is}, \mu_i)'$ for $r = 1, 2$ and $t \neq s$. One may naturally be inclined to estimate unknown parameters in (3.1)–(3.2) by first-differencing respective "selected" samples of loans-only and loans-and-savings MFIs in order to remove unknown fixed effects, i.e., using the observations for which $D_{it} = D_{is}$ $(t \neq s)$ are jointly equal to either 1 or 2. The procedure yields the cost function in first differences:

$$\begin{aligned}\log(C_{r,it}) - \log(C_{r,is}) = {}& \big[\log(\mathbf{v}_{r,it}) - \log(\mathbf{v}_{r,is})\big]'\boldsymbol{\beta}_r\ + \\ & \frac{1}{2}\Big[\text{vec}\{\log(\mathbf{v}_{r,it})\log(\mathbf{v}_{r,it})'\} - \text{vec}\{\log(\mathbf{v}_{r,is})\log(\mathbf{v}_{r,is})'\}\Big]'\text{vec}\{\mathbf{B}_r\}\ + \\ & u_{r,it} - u_{r,is} \qquad \text{if} \quad D_{it} = D_{is} = r\ (t \neq s),\end{aligned} \tag{3.4}$$

where "vec$\{\cdot\}$" denotes the column-wise vectorization operator. However, one generally should *not* expect $\mathbb{E}[u_{r,it}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}] = 0$ or $\mathbb{E}[u_{r,it}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}] = \mathbb{E}[u_{r,is}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}]$. Therefore, if we do not control for selection explicitly, the first-differenced error term $(u_{r,it} - u_{r,is})$ in (3.4) is likely to correlate with the right-hand-side covariates thereby

[9] Note that nothing precludes $\mu_{1,i}$ and $\mu_{2,i}$ from being the same, i.e., the special case of $\mu_{1,i} = \mu_{2,i}$.

resulting in inconsistent estimates of unknown coefficients $\boldsymbol{\beta}_r$ and $\mathbf{B}_r$ for $r = 1, 2$ (essentially, due to the "omitted variable bias"). We therefore explicitly recognize the presence of the selectivity effect and, for $r = 1, 2$, rewrite outcome equations in (3.1)–(3.2) as partially linear equations of the following form:

$$\log(C_{r,it}) = \log(\mathbf{v}_{r,it})'\boldsymbol{\beta}_r + \tfrac{1}{2}\log(\mathbf{v}_{r,it})'\mathbf{B}_r\log(\mathbf{v}_{r,it}) + \mu_{r,i} + \theta_{r,it} + e_{r,it} \qquad \text{if} \quad D_{it} = r, \tag{3.5}$$

where $\theta_{r,it} \equiv \mathbb{E}[u_{r,it}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}]$ is an *unknown* sample selection bias term, and $e_{r,it} \equiv u_{r,it} - \theta_{r,it}$ is a new random disturbance such that $\mathbb{E}[e_{r,it}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}] = 0$ by construction.

Intuitively, we could consistently estimate the outcome equation in (3.5) via some procedure that removes not only the unknown fixed effect $\mu_{r,i}$ but also the unknown selectivity effect $\theta_{r,it}$. We accomplish the latter by following a semiparametric approach proposed by Kyriazidou (1997). Specifically, under the "conditional exchangeability" assumption, according to which $(u_{r,it}, u_{r,is}, v_{it}, v_{is})$ and $(u_{r,is}, u_{r,it}, v_{is}, v_{it})$ for $t \neq s$ are identically distributed conditional on $\boldsymbol{\omega}_{r,i}$,[10] for some ith MFI of type $r = 1, 2$ such that $\mathbf{z}_{it}'\boldsymbol{\gamma} = \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s)$, we have that[11]

$$\begin{aligned}
\theta_{r,it} &= \mathbb{E}[u_{r,it}|D_{it} = D_{is} = r\ (t \neq s), \boldsymbol{\omega}_{r,i}] \\
&= \begin{cases} \mathbb{E}[u_{r,it}|v_{it} > \mathbf{z}_{it}'\boldsymbol{\gamma}, v_{is} > \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s), \boldsymbol{\omega}_{r,i}] & \text{for } r = 1 \\ \mathbb{E}[u_{r,it}|v_{it} \leq \mathbf{z}_{it}'\boldsymbol{\gamma}, v_{is} \leq \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s), \boldsymbol{\omega}_{r,i}] & \text{for } r = 2 \end{cases} \\
&= \begin{cases} \mathbb{E}[u_{r,is}|v_{is} > \mathbf{z}_{it}'\boldsymbol{\gamma}, v_{it} > \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s), \boldsymbol{\omega}_{r,i}] & \text{for } r = 1 \\ \mathbb{E}[u_{r,is}|v_{is} \leq \mathbf{z}_{it}'\boldsymbol{\gamma}, v_{it} \leq \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s), \boldsymbol{\omega}_{r,i}] & \text{for } r = 2 \end{cases} \\
&= \theta_{r,is}.
\end{aligned} \tag{3.6}$$

The result in (3.6) says that the sample selection bias term for the ith r-type unit such that $\mathbf{z}_{it}'\boldsymbol{\gamma} = \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s)$, i.e., for an MFI with the same likelihood of self-selecting into either the loans-only or loans-and-savings type in period t and s, is the same for the two time periods and hence can be removed for these observations along with μ_i via first differencing across periods t and $s (\neq t)$. That is, in theory, the outcome equation of interest could be consistently estimated for each output-mix type of MFIs via conventional least squares performed on the first-differenced "selected" sample consisting of MFIs of a given type for which $\Delta\mathbf{z}_{it}'\boldsymbol{\gamma} \equiv (\mathbf{z}_{it} - \mathbf{z}_{is})'\boldsymbol{\gamma} = 0$, where "$\Delta$" designates the first-difference operator. In practice, such an estimator would however be infeasible for two reasons: (i) $\boldsymbol{\gamma}$ is normally *un*known, and (ii) $\Pr[\Delta\mathbf{z}_{it}'\boldsymbol{\gamma} = 0] = 0$ so long as $\mathbf{z}_{it}$ contains at least one continuous variable rendering $\Delta\mathbf{z}_{it}'\boldsymbol{\gamma}$ a continuous variable too.

To overcome these two problems and thereby to facilitate a practical implementation of (3.5), we do the following. First, we can replace unknown $\boldsymbol{\gamma}$ with its consistent estimate $\widehat{\boldsymbol{\gamma}}$ obtained from selection equation (3.3) estimated in the first stage. Second, under the assumption of $\theta_r(\cdot)$ being sufficiently smooth, the unit for which $\Delta\mathbf{z}_{it}'\boldsymbol{\gamma} \cong 0$ should also have $\theta_{r,it} \cong \theta_{r,is}$. Thus, following Kyriazidou's (1997) idea, we can use data on the MFIs for which $\mathbf{z}_{it}'\widehat{\boldsymbol{\gamma}}$ is "close" to $\mathbf{z}_{is}'\widehat{\boldsymbol{\gamma}}\ (t \neq s)$ and weigh these data on the basis of the "closeness" of $\Delta\mathbf{z}_{it}'\widehat{\boldsymbol{\gamma}}$ to 0. Intuitively, the cross-sectional units, for which the self-selection likelihoods are close to being the same in both time periods t and

[10]That is, if $\mathbb{F}(u_{r,it}, u_{r,is}, v_{it}, v_{is}|\boldsymbol{\omega}_{r,i}) = \mathbb{F}(u_{r,is}, u_{r,it}, v_{is}, v_{it}|\boldsymbol{\omega}_{r,i})$ with $\mathbb{F}(\cdot)$ being a multivariate joint distribution.

[11]Note that we could have obtained the same results under a much stronger assumption of $(u_{r,it}, v_{it})$ being independently and identically distributed over *both* i and t. In that case, $\theta_{r,it} = \mathbb{E}[u_{r,it}|D_{it} = D_{is} = r\ (t \neq s)] = \mathbb{E}[u_{r,it}|D_{it} = r\ (t \neq s)] = \mathbb{E}[u_{r,it}|v_{it} > \mathbf{z}_{it}'\boldsymbol{\gamma}$ if $r = 1$ or $v_{it} \leq \mathbf{z}_{it}'\boldsymbol{\gamma}$ if $r = 2] \equiv \theta_r(\mathbf{z}_{it}'\boldsymbol{\gamma})$, and it is easy to see that $\theta_{r,it} = \theta_{r,is}$ if $\mathbf{z}_{it}'\boldsymbol{\gamma} = \mathbf{z}_{is}'\boldsymbol{\gamma}\ (t \neq s)$. However, the advantage of the "conditional exchangeability" assumption is that it allows marginal distributions of errors $(u_{r,it}, v_{it})$ and hence the function $\theta_r(\cdot)$ to vary over i.

s ($t \neq s$), ought to be given heavier weights. Note that our approach somewhat differs from Ahn & Powell's (1993), who difference out sample selection bias by subtracting one cross-sectional unit from another cross-sectional unit "matched" on the basis of similarity in their likelihoods of being selected into the sample. In our paper, we eliminate the sample selection effect by "matching" observations for the *same* cross-sectional unit across the two time periods, which allows us to also purge the unknown unit-specific fixed effects.

Thus, for each MFI type $r = 1, 2$, we can consistently estimate the outcome equation of interest in (3.4) along the lines of Kyriazidou (1997) from the following kernel-weighted first-differenced least-squares problem based on any two time periods t and $s(\neq t)$:

$$\min_{\boldsymbol{\beta}_r, \mathbf{B}_r} \quad \sum_i \left[\Delta \log(C_{r,it}) - \Delta \log(\mathbf{v}_{r,it})'\boldsymbol{\beta}_r - \tfrac{1}{2}\Delta \text{vec}\{\log(\mathbf{v}_{r,it}) \log(\mathbf{v}_{r,it})'\}'\text{vec}\{\mathbf{B}_r\}\right]^2 \mathcal{D}_{r,it}\widehat{\mathcal{K}}_{r,it}, \quad (3.7)$$

where $\mathcal{D}_{r,it} \equiv \mathbb{1}\{D_{r,it} = D_{r,is} = r \ (t \neq s)\}$ is an indicator variable selecting the same ith r-type unit that has observations in both periods t and $s(\neq t)$; and $\widehat{\mathcal{K}}_{r,it} = k\left(\Delta \mathbf{z}_{it}'\widehat{\boldsymbol{\gamma}}/h\right)$ is the kernel weight with $k(\cdot)$ being a symmetric kernel function with the corresponding bandwidth parameter $h \to 0$ slowly as $N \to \infty$. Letting $\boldsymbol{\Delta}\mathbf{m}_{r,it} = \left(\Delta \log(\mathbf{v}_{r,it})', \tfrac{1}{2}\Delta \text{vec}\{\log(\mathbf{v}_{r,it}) \log(\mathbf{v}_{r,it})'\}'\right)'$, we then obtain a semiparametric least-squares estimator of unknown $\boldsymbol{\beta}_r$ and $\mathbf{B}_r$, i.e.,

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}}_r \\ \text{vec}\{\widehat{\mathbf{B}}_r\} \end{bmatrix} = \left[\sum_i \boldsymbol{\Delta}\mathbf{m}_{r,it}\boldsymbol{\Delta}\mathbf{m}_{r,it}'\mathcal{D}_{r,it}\widehat{\mathcal{K}}_{r,it}\right]^{-1} \left[\sum_i \boldsymbol{\Delta}\mathbf{m}_{r,it}\Delta \log(C_{r,it})\mathcal{D}_{r,it}\widehat{\mathcal{K}}_{r,it}\right]. \quad (3.8)$$

So long as the first-stage selection-equation estimator $\widehat{\boldsymbol{\gamma}}$ converges at the rate faster than $N^{-2/5}$,[12] the estimator in (3.8) is consistent and asymptotically normal with the asymptotic bias, which results from the $\Delta \mathbf{z}_{it}'\boldsymbol{\gamma} \cong 0$ approximation and goes to zero as $h \to 0$ (Kyriazidou, 1997). The bias can however be tackled via the bias-correction procedure. The convergence rate of the (second-stage) estimator of $\boldsymbol{\beta}_r$ and $\mathbf{B}_r$ is expectedly slower than that of $\widehat{\boldsymbol{\gamma}}$, namely $N^{-1/2+\nu/2}$ with $1-2p < \nu < p/2$, where N^{-p} is the convergence rate of $\widehat{\boldsymbol{\gamma}}$, and the bandwidth is such that $h \propto N^{-\nu}$. In fact, we can maximize the convergence rate of $\left(\widehat{\boldsymbol{\beta}}_r', \text{vec}\{\widehat{\mathbf{B}}_r\}'\right)'$ by setting $\nu = 1/(\tau+1)$, where τ is the order of kernel function $k(\cdot)$, provided that $\boldsymbol{\gamma}$ is estimated fast enough, i.e., when $p > \tau/(2\tau+1)$. For more details, see Kyriazidou (1997).

Note that the above estimator makes use of *two* time periods only. However, in practice, panels can extend well beyond $T = 2$. To make use of the entire data, one can estimate (3.8) for $\mathcal{C}(T, 2) = \frac{T!}{2!(T-2)!}$ unique pairs of the time periods separately and then combine the obtained estimates of unknown coefficients using some minimum distance measure. Naturally, for efficiency purposes, it is preferable to combine such estimates using optimal weights which, in turn, necessitates the estimation the covariance matrix of the estimators for different pairs of the time periods. However, Charlier et al. (2001) show that the covariances between the estimators for different pairs of the time periods converge to zero, and one therefore may combine the estimates simply using the inverses of their corresponding variances as weights. Alternatively, the parameter estimates may be merged via simple unweighted averaging, as for instance, suggested by Kyriazidou (1997). Another practical way to go about having $T > 2$ is to stack all unique pairs of time periods together and estimate (3.8) for each MFI type only once a là Ahn & Powell (1993), which we do in this paper. A few more remarks about the implementation of our estimator in (3.8).

First-Stage Estimation of $\boldsymbol{\gamma}$. To ensure the fastest attainable convergence rate for the second-stage estimator of cost functions, which are of the primary interest to our analysis, we employ Klein

[12]Naturally, with the parametric $\sqrt{N}$-rate being the upper bound.

& Spady's (1993) semiparametric single-index estimator for binary response models to estimate selection equation (3.3) in the first stage. More concretely,

$$\begin{aligned}\Pr[D_{it}=1|\boldsymbol{\chi}_{r,i}] &= \Pr[v_{it} > \mathbf{z}_{it}'\gamma|\boldsymbol{\chi}_{r,i}] = 1-\mathcal{F}(\mathbf{z}_{it}'\gamma)\\ \Pr[D_{it}=2|\boldsymbol{\chi}_{r,i}] &= \Pr[v_{it} \le \mathbf{z}_{it}'\gamma|\boldsymbol{\chi}_{r,i}] = \mathcal{F}(\mathbf{z}_{it}'\gamma),\end{aligned} \tag{3.9}$$

where $\mathcal{F}(\cdot)$ is an *unspecified* cdf of the random error v_{it} which is estimated, along with the unknown parameter vector $\boldsymbol{\gamma}$, via (kernel-based) semiparametric maximum likelihood.

Our choice of this estimator for $\boldsymbol{\gamma}$ (over other available alternatives) has been primarily motivated by two considerations. First, owing to its single-index treatment of $\mathbf{z}_{it}$, the estimator delivers a parametric $\sqrt{N}$ rate of convergence for $\widehat{\boldsymbol{\gamma}}$, i.e., $p=1/2$. When using the popular second-order kernel function ($\tau=2$) in the second-stage estimation like we do, the latter facilitates the attainment of the maximal rate of convergence for our main estimator in (3.8). Specifically, as discussed earlier, condition $p=1/2>\tau/(2\tau+1)$ is satisfied when $\tau=2$, and we attain the familiar nonparametric rate $N^{-2/5}$ for $\left(\widehat{\boldsymbol{\beta}}_r', \mathrm{vec}\{\widehat{\mathbf{B}}_r\}'\right)'$. Second, Klein & Spady's (1993) estimator does not require the distributional assumption about v_{it} thereby significantly reducing the risk of model misspecification which, however, the alternative $\sqrt{N}$ fully parametric estimators for binary response models (such as the logit or probit) are inherently subject to. Alternative semiparametric binary-response estimators such as Manski's (1987) maximum score, Horowitz's (1992) smoothed maximum score, Khan's (2013) sieve-based probit and Blevins & Khan's (2013) local kernel-based probit estimators, while also distribution-free, however converge at slower rates which fail to satisfy the desired condition $p>2/5$.[13]

Bias Reduction. We tackle the presence of asymptotic bias in estimator (3.8) by employing a bias-reduction procedure which preserves the estimator's maximal rate of convergence, as suggested by Kyriazidou (1997). Specifically, consistent with our discussion above, let $\left(\widehat{\boldsymbol{\beta}}_r', \mathrm{vec}\{\widehat{\mathbf{B}}_r\}'\right)'$ be the estimator with bandwidth $h \propto N^{-\nu}$ such that $\nu = 1/(\tau+1)$, and $\left(\breve{\boldsymbol{\beta}}_r', \mathrm{vec}\{\breve{\mathbf{B}}_r\}'\right)'$ be the estimator with bandwidth such that $\nu=\delta/(\tau+1)$ where $0<\delta<1$. Then, the bias-corrected estimator of $\left(\boldsymbol{\beta}_r', \mathrm{vec}\{\mathbf{B}_r\}'\right)'$ is given by

$$\begin{bmatrix}\widetilde{\boldsymbol{\beta}}_r\\ \mathrm{vec}\{\widetilde{\mathbf{B}}_r\}\end{bmatrix} = \left(\begin{bmatrix}\widehat{\boldsymbol{\beta}}_r\\ \mathrm{vec}\{\widehat{\mathbf{B}}_r\}\end{bmatrix} - N^{-(1-\delta)\tau/(2\tau+1)}\begin{bmatrix}\breve{\boldsymbol{\beta}}_r\\ \mathrm{vec}\{\breve{\mathbf{B}}_r\}\end{bmatrix}\right) \Big/ \left(1-N^{-(1-\delta)\tau/(2\tau+1)}\right) \tag{3.10}$$

and converges at the same rate as does the "un-corrected" estimator $\left(\widehat{\boldsymbol{\beta}}_r', \mathrm{vec}\{\widehat{\mathbf{B}}_r\}'\right)'$.

Bandwidth Selection. We select the optimal bandwidth for estimator (3.8) via a data-driven leave-one-*cross-section*-out cross-validation a là Malikov et al. (2016) meant to accommodate the panel structure of the data. Specifically, for each MFI type $r=1,2$, we select the optimal bandwidth by minimizing the following cross-validation objective function:

$$\min_{h_r} \sum_i \left[\Delta\log(C_{r,it}) - \Delta\log(\mathbf{v}_{r,it})'\widehat{\boldsymbol{\beta}}_{r,-i}(h_r) - \tfrac{1}{2}\Delta\mathrm{vec}\{\log(\mathbf{v}_{r,it})\log(\mathbf{v}_{r,it})'\}'\mathrm{vec}\{\widehat{\mathbf{B}}_{r,-i}(h_r)\}\right]^2\mathcal{D}_{r,it}, \tag{3.11}$$

[13] To be concrete, their convergence rates are $N^{-1/3}$ $N^{-2/5}$, $N^{-2/5}$ and $N^{-1/3}$, respectively, although the convergence rate of Horowitz's (1992) estimator can be made arbitrarily close to $N^{-1/2}$ by making a stronger assumption about a higher degree of smoothness for the underlying distribution (and thereby using a higher-order kernel function).

where $\left(\widehat{\boldsymbol{\beta}}_{r,-i}(h_r)', \text{vec}\{\widehat{\mathbf{B}}_{r,-i}(h_r)\}'\right)'$ is the following leave-one-cross-section-out estimator of unknown $\left(\boldsymbol{\beta}_r', \text{vec}\{\mathbf{B}_r\}'\right)'$:

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}}_{r,-i}(h_r) \\ \text{vec}\{\widehat{\mathbf{B}}_{r,-i}(h_r)\} \end{bmatrix} = \left[\sum_{j\neq i} \Delta\mathbf{m}_{r,jt}\Delta\mathbf{m}_{r,jt}'\mathcal{D}_{r,jt}\widehat{\mathcal{K}}_{r,jt}(h_r)\right]^{-1} \left[\sum_{j\neq i} \Delta\mathbf{m}_{r,jt}\Delta\log(C_{r,jt})\mathcal{D}_{r,jt}\widehat{\mathcal{K}}_{r,jt}(h_r)\right]. \tag{3.12}$$

Note that, for notational consistency, both (3.11) and (3.12) are formulated for a single pair of time periods t and $s(\neq t)$. In the spirit of our earlier remarks, when $T > 2$ (like in our case), multiple unique pairs of time periods can be accommodated by stacking them together and then cross-validating the model for each MFI type only once.

4 Data

The unbalanced micro-level panel data on MFIs come from the Mixmarket database. Our data are different from those previously analyzed in the literature in that we start with a larger proportion of deposit-collecting MFIs. After data cleaning[14] and given the availability of the matched country-level variables used to control for external factors influencing MFIs' output mix selection, the usable sample contains the total of 4,692 observations on 1,003 MFIs from 76 countries during the 2004–2014 period. The composition of the sample by MFI type is as follows: 45% loans-only and 55% loans-and-savings MFIs. Also note that, unlike some previous studies, we are unable to distinguish between compulsory and voluntary savings deposits and therefore use total savings in our analysis.

As discussed in Section 2, the outputs are the total loans (y_1) and, if offered by an MFI, total savings deposits (y_2), both measured in thousands of real USD. The input prices are constructed as follows. The price of labor w_1 is the average annual wage measured in thousands of USD. The price of physical capital (w_2) is computed as the ratio of non-labor operating expenses to the number of offices and thus is measured in thousands of USD per office/branch. The price of financial capital (w_3) is constructed by dividing the corresponding financial expenses by the borrowed funds and is a unit-free interest rate. The total equity e (a quasi-fixed input) is the MFI's equity in thousands of real USD. Total variable cost (C), also measured in thousands of USD, is the sum of expenditures on the three variable inputs. More specifically, we compute the variable cost as follows. First, from the MFI's total financial expense, which in its original form contains the expense on both borrowed funds and deposits, we subtract the cost of deposits obtained by multiplying the volume of deposits by the deposit (interest) rate reported by the Mixmarket. Next, we add the obtained (netted-out) financial expense on non-deposit borrowed funds to the operating expense (on physical capital and labor) to arrive at the *restricted* variable cost. Thus, the deposits are excluded from the variable cost measure, consistent with labor, physical capital and borrowed (non-deposit) funds being the three variable inputs to the production.[15]

[14] We remove observations with missing or negative values for cost, output quantities, input prices and the risk proxy. We also drop few observations for which the ratio of loans delinquent by more than 30 days to the MFI's total portfolio (our risk proxy) exceeds one, since these are likely to be the result of erroneous data reporting. To minimize the distortionary effects of outliers on the cost function estimates (and hence the estimates of the degree of scope economies), we also exclude observations from the 1st and 99th percentiles of distributions of the variable cost function covariates.

[15] We thank an anonymous reviewer for suggesting that we model microfinance technology using such a "restricted" formulation of the variable cost function.

Table 1. Data Summary Statistics

Variable	1st Qu.	Median	Mean	3rd Qu.	1st Qu.	Median	Mean	3rd Qu.
		Loans-Only				**Loans-and-Savings**		
C	712.90	1,971.86	5,019.51	5,347.48	836.12	2,730.18	10,399.72	10,187.69
w_1	6.14	10.03	10.87	14.33	3.74	7.54	8.98	13.08
w_2	26.82	54.21	77.17	99.64	22.92	67.15	125.14	174.66
w_3	7.37	10.21	11.50	13.40	5.03	6.97	7.83	9.62
e	906.54	2,489.58	6,293.26	7,016.12	831.33	2,841.91	11,265.29	10,619.91
y_1	2,313.74	6,522.80	19,918.06	20,787.60	3,342.61	10,959.64	56,986.70	46,919.96
y_2	—	—	—	—	1,170.04	4,530.52	41,710.83	28,311.48
$risk$	0.008	0.028	0.047	0.058	0.018	0.042	0.060	0.080
Size	2,678.02	7,575.13	22,198.97	24,086.61	4,513.94	13,724.41	71,487.18	58,831.26
Non-Profit			0.613				0.527	
Age: New			0.084				0.068	
Age: Young			0.193				0.156	
Age: Mature			0.723				0.776	
Target Market: Broad			0.469				0.554	
Target Market: High-End			0.029				0.066	
Target Market: Low-End			0.487				0.330	
Target Market: Small Business			0.015				0.050	
GDP per capita	4.64	7.88	8.77	10.90	2.57	5.28	6.79	9.31
HDI	0.61	0.70	0.67	0.73	0.54	0.65	0.62	0.71
National Saving Rate	16.51	21.34	23.19	28.13	18.24	24.92	25.79	33.90
National Remittance Rate	2.10	4.21	7.52	9.99	1.86	3.74	6.38	9.83
Financial Sector Depth	32.22	44.68	51.59	69.72	32.35	40.85	45.28	58.97
Branches per 100,000 adults	7.16	11.00	16.79	19.63	5.06	7.67	12.94	14.00
Rural Population Share	26.25	43.03	42.92	60.62	35.88	53.27	51.62	69.41
Internet Subscription per 100 people	0.65	1.56	3.30	4.62	0.14	0.68	1.97	2.53
WGI: Control of Corruption	–0.76	–0.51	–0.53	–0.29	–0.83	–0.68	–0.64	–0.40
WGI: Government Effectiveness	–0.71	–0.39	–0.37	–0.05	–0.74	–0.49	–0.43	–0.11
WGI: Political Stability	–0.98	–0.67	–0.67	–0.31	–1.33	–0.81	–0.85	–0.44
WGI: Rule of Law	–0.86	–0.59	–0.55	–0.21	–0.90	–0.65	–0.64	–0.43
WGI: Regulatory Quality	–0.46	–0.28	–0.19	0.23	–0.66	–0.34	–0.34	–0.11
WGI: Voice and Accountability	–0.59	–0.11	–0.21	0.15	–0.45	–0.14	–0.23	0.09
% Obs.			44.88				55.12	

C – total variable costs; w_1 – price of labor; w_2 – price of physical capital; w_3 – price of financial capital; e – total equity; y_1 – total loans, y_2 – total savings deposits; $risk$ – portfolio at risk (30 days). Variables C, y_1, y_2, y_3, w_1, w_2, e, size and GDP per capita are in thousands of real USD. Variable w_3, National Saving and Remittance Rates, Financial Sector Depth and Rural Population Share are in % points. The $risk$ variable is a unit-free proportion. The HDI and WGI indices are unit-free (for more details, see their original sources).

We also condition the estimated variable cost functions on a measure of credit risk ($risk$). Like for any other financial institution, credit risk is an important factor for MFIs which, following the tradition in the literature, we proxy using the information on non-performing loans. It is imperative to account for such a risk when modeling the cost structure of financial institutions because lower-quality assets (reflected in a higher non-performing loans ratio) generally require more resources to manage a higher-level risk exposure thereby raising the costs for MFIs. Failure to account for riskiness/quality of loans may therefore produce misleading results (e.g., see Hughes & Mester, 2013). In this paper, we proxy the level of risk that an MFI is exposed to via the ratio of loans delinquent by more than 30 days to the MFI's total portfolio, which is a unit-free proportion.

We next proceed to discussion of the variables entering the output mix selection equation. Here, we seek to control for factors that may potentially affect the MFI's endogenous decision to operate as a loans-only or a loans-and-savings institution. The vector $\mathbf{z}$ includes covariates capturing both internal and external factors that may influence the MFI's propensity to offer savings accounts as an additional financial service to their customers. The internal (MFI-specific) factors include (i) age of the MFI captured via three dummy variables (New for age less than 4 year, $Young$ for age between 4 and 8 years and $Mature$ for age above eight years),[16] (ii) the non-profit status indicator, (iii) size of the MFI defined as the total value of its portfolio and measured in thousands of USD, and (iv) four dummies for the target market: $Low\ End$, $High\ End$, $Broad$ and $Small\ Businesses$. The latter categories are defined based on the poverty level of target customers using the "depth of clientele" ratio of average loan balance per borrower to the country's GNI per capita. The "low-end", "broad", "high-end" and "small business" target market categories respectively correspond to the depth ratio of less than 20% or the average loan size being less than \$150, 20–149%, 150–250% and above 250%.

We also include the following external (country-specific) factors, the data on which we obtain from the World Bank's World Development and Worldwide Governance Indicators database as well as from the United Nations Development Programme: (i) GDP per capita (ii) national saving rate, (iii) national remittance rate, (iv) a measure of the depth of financial sector, (v) the number of branches per 100,000 adults, (vi) the share of rural population, (vii) the broadband internet subscription per 100 people, (viii) six governance indicators measuring institutional quality in the country including control of corruption, government effectiveness, political stability, rule of law, regulatory quality and voice and accountability, and (ix) the human development index (HDI). In what follows, we describe each of these variables.

We include GDP per capita, in thousands of PPP-adjusted USD, to proxy for level of economic activity in the country. The national gross savings rate, the ratio of total savings to GNI, captures population's willingness to save and thus availability of savings that an MFI can collect. The ratio of total remittances to GDP is included because significant remittances from abroad necessitate the use of savings accounts thereby making adoption of the loans-and-savings output mix more attractive for MFIs. The depth of financial sector is defined as the ratio of the M2 monetary aggregate, including currency and deposits, to the country's GDP and measures the level of banking sector development. It captures how difficult/easy it is to collect deposits from marginalized clients. In a similar vein, the number of branches per 100,000 people reflects the "density" of banking services in the country. The share of rural population is the proportion of the country's total population living in rural areas. The self-selection of MFIs into deposit-taking is likely to be influenced by transaction costs associated with distance and the lack of infrastructure in less urbanized regions, which directly affect the cost of serving the rural poor. Hence, we include the rural population share variable in the selection equation to proxy for the latter. Since many microfinance services

[16]Unfortunately, no continuous measure of the MFI's age is available in the Mixmarket database.

Table 2. Cost Elasticity of Deposits in the Loans-and-Savings MFIs

	Point Estimate	One-Sided Lower Bound
10th Percentile	0.024	–0.047
1st Quartile	0.052	–0.006
Median	0.084	0.041
3rd Quartile	0.119	0.077
90th Percentile	0.148	0.109
Mean	0.085	0.041

Reported are the estimates of $\partial \log \mathrm{C}/\partial \log y_2$ from the translog $\mathrm{C}(\boldsymbol{w}, e, y_1, y_2)$ function with C being a sum of expenses on $\boldsymbol{x}$ less the cost of deposits. The lower bounds are for the *one*-sided 95% percentile block-bootstrap confidence intervals, with the corresponding upper bounds being $+\infty$.

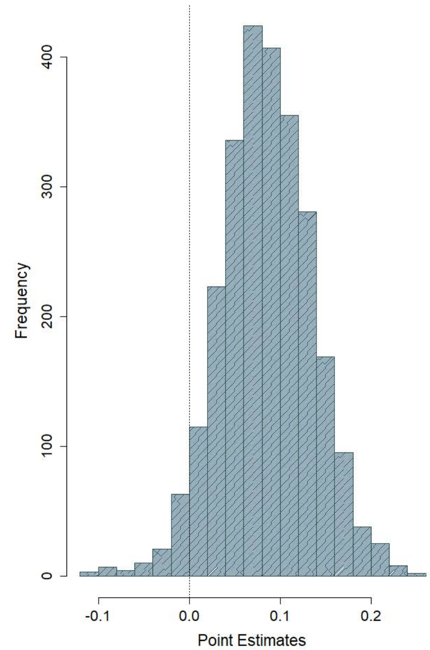

Figure 1. Distribution of Cost Elasticity of Deposits in the Loans-and-Savings MFIs

rely on wireless technology to mobilize deposits, especially in remote area (e.g., cell phone mobile banking), the level of internet subscription in the country is also included in the selection equation. Lastly, in light of growing evidence of the importance of institutions for both the overall economic development and the performance of individual organizations, we condition the MFI's output mix choice on several measures of institutional quality and human development (e.g., Eicher & Leukert, 2009; Mersland & Strøm, 2009; Yasar et al., 2011). Table 1 reports the summary statistics of the variables used in the analysis.

5 Empirical Results

Deposits as Outputs. We first assess the empirical validity of treating deposits as an output to the microfinance production. Specifically, motivated by Hughes & Mester's (1993) arguments, we verify the output-like role of microdeposits empirically by estimating a variable cost function for loans-and-savings MFIs, which we condition on the level/quantity of savings deposits with their costs subtracted from the financial expenditures entering the left-hand-side cost variable, and then formally testing the sign of cost (log)derivative with respect to the level of deposits. A positive sign would indicate that deposits are an output, whereas a negative sign would indicate that they are an input in the production process.[17]

Table 2 summarizes point estimates of the cost elasticity of deposits obtained from the first-difference estimator of $\mathrm{C}(\boldsymbol{w}, e, y_1, y_2)$ for loans-and-savings MFIs using the translog specification subject to the Slutsky symmetry and linear homogeneity in $\boldsymbol{w}$.[18] The point estimates are positive

[17]We thank an anonymous reviewer for this insight; also, see Braeutigam & Daughety (1983).

[18]Here and in what follows, when estimating cost functions, we also include a credit risk proxy, the time trend (along

for 96% of observations, as can also be seen in Figure 1 which plots the histogram of these elasticity estimates. The average estimate is 0.085, and it is statistically positive at the conventional level. Namely, its one-sided 95% bootstrap lower bound clustered at the MFI level is 0.041 (see Table 2) which is well above zero thereby letting us reject the null in favor of the ">0" alternative. We formally test for the variable cost elasticity of deposits being positive at each observation and find it to be (one-sidedly) statistically greater than zero for two thirds of the sample (68%). We also perform additional tests of similar fashion applied to the *non-interest* variable cost function estimated treating all financial assets—deposits and other borrowed capital—as quasi-fixed with the left-hand-side variable cost measuring operational expenses on labor and physical capital only. The results are reassuring, with the non-interest cost elasticity of deposits being statistically positive for 90% of observations (for more details, see Appendix A). Thus, we find strong empirical evidence in support of the output-like treatment of microdeposits.

Complementarity of Deposits and Loans. Before proceeding to our main analysis of scope (dis)economies in the industry, as a preliminary investigation we examine the extent of output complementarity in the integrated loans-and-savings microfinance production. Following Baumol et al. (1982), such a complementarity is said to exist when, owing to "positive synergies," an increase in the level of one output leads to a decline in the marginal cost of another output (also see Chavas & Kim, 2010). In our context, the latter is equivalent to the condition that the variable cost function for loans-and-savings MFIs be such that $\partial^2 \mathrm{C}(\cdot)/\partial y_1 \partial y_2 < 0$. It is informative to investigate this complementary between deposits and loans because, if present, it would be one of the sources of positive scope economies in integrated MFIs.

We recover cross-output derivatives of the loans-and-savings variable cost function in levels from the already estimated translog specification of $\mathrm{C}(\boldsymbol{w}, e, y_1, y_2)$ in logs from above as follows:

$$\frac{\partial^2 \mathrm{C}(\cdot)}{\partial y_1 \partial y_2} = \left(\frac{\partial^2 \ln \mathrm{C}(\cdot)}{\partial \ln y_1 \partial \ln y_2} + \frac{\partial \ln \mathrm{C}(\cdot)}{\partial \ln y_1} \frac{\partial \ln \mathrm{C}(\cdot)}{\partial \ln y_2} \right) \times \frac{\mathrm{C}}{y_1 y_2}, \tag{5.1}$$

where the obtained second-order cross-partial is observation-specific, and its sign is governed by the sign of estimated term in parentheses. Table 3 reports the summary of point estimates of this cross-output cost derivative. The sample mean and median estimates are both statistically negative at the conventional level, but the two values visibly diverge (–2.75 and –0.058, respectively) owing to an expectedly strong positive skew in the $\mathrm{C}/(y_1 y_2)$ term. Figure 2 scatter-plots point estimates of the same cross-output cost derivative scaled back by the inverse of $\mathrm{C}/(y_1 y_2)$ to remove skewness.[19] The figure also plots the *one*-sided 95% upper percentile bootstrap bound (solid line) for each observation-level estimate, which we have sorted by the bound as well as color-coded depending on if it is statistically negative or not. Formally testing for the cross-output cost derivate being less than zero at each observation, we find that it is (one-sidedly) statistically negative for 40.3% of the sample only. This empirical evidence suggests that complementarity between loans and microdeposits is not exhibited uniformly across all integrated MFIs. The latter however does not necessarily imply that technological conditions for substantive economies of scope in the industry are lacking, since output complementarity is only one of potential sources of scope economies when defined as in (2.6) allowing for partial specialization à la Evans & Heckman (1984). Other sources of scope-driven cost savings include scale economies, effects of the cost convexity as well as discontinuities between heterogeneous loans-only and loans-and-savings cost functions. For more on

with its expansion terms) meant to accommodate temporal shifts in the cost frontiers as well as allow for shifts during and after the 2008 financial crisis.

[19] Effectively, plotted is the term appearing inside parentheses in (5.1).

Table 3. Cross-Output Cost Derivative in the Loans-and-Savings MFIs ($\times 10^5$)

	Point Estimate	One-Sided Upper Bound
10th Percentile	–2.927	–0.969
1st Quartile	–0.463	–0.090
Median	–0.058	–0.001
3rd Quartile	–0.004	0.030
90th Percentile	0.001	0.220
Mean	–2.750	–0.992

Reported are the estimates of $\partial^2 \mathrm{C}/\partial y_1 \partial y_2 \times 10^5$ recovered from the translog $\mathrm{C}(\boldsymbol{w}, e, y_1, y_2)$ function. The upper bounds are for the *one*-sided 95% percentile block-bootstrap confidence intervals, with the corresponding lower bounds being $-\infty$.

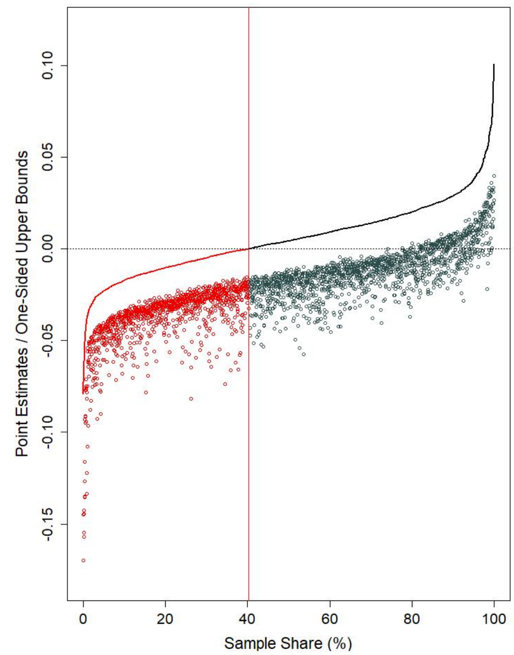

Figure 2. The One-Sided 95% Percentile Block-Bootstrap Upper Bounds (Solid Line) of the Cross-Output Cost Derivative Estimates (Scatter Points) in the Loans-and-Savings MFIs, scalled by $\frac{y_1 y_2}{\mathrm{C}}$

necessary and sufficient conditions for and the decomposition of economies of scope (also sometimes referred to as economies of diversification) defined in a fashion analogous to that of our paper, see Chavas & Kim (2010). Therefore, to comprehensively assess the prevalence of scope economies in the industry, we proceed to the estimation of our main model and the direct measurement of the degree of scope economies in the microfinance production.

Endogenously Heterogeneous Technologies. We estimate endogenously switching heterogeneous cost functions in (3.1)–(3.3) subject to not only the linear homogeneity in input prices and the Slutsky symmetry but also the monotonicity regularity conditions in order to ensure that our results are economically meaningful because, to paraphrase Barnett (2002, p.199), without theoretical regularity conditions imposed, the conditions for optimal behavior may fail prompting the duality to also fail making the policy and value functions (such as the cost function) become invalid.[20] The linear homogeneity property is imposed by dividing the cost and input prices by one of the three input prices, the choice of which the model is invariant to. We use the price of labor w_1 as a normalizing input price. The imposition of the Slutsky symmetry is trivial. We impose the positive monotonicity of $\mathbb{C}_r(\cdot)$ [and hence of $\log(\mathbb{C}_r(\cdot))$ which we actually estimate in practice] in output quantities, input prices and risk and the negative monotonicity in quasi-fixed input by "tilting" our estimator in (3.8) using the procedure proposed by Hall & Huang (2001) and further extended by Du et al. (2013). The procedure essentially mutes or magnifies the impact of any given

[20]On the importance of incorporating economic theory in the econometric measurement, also see Barnett et al. (1991), Heckman & Serletis (2014, 2015) and Serletis & Feng (2015).

Table 4. First-Stage Estimates of the Selection Equation

Covariate	Coefficient Estimate	Two-Sided Lower Bound	Two-Sided Upper Bound	Median Marginal Effect
Non-Profit	–1.0000	—	—	–0.5060
Age: Young	0.0544	0.0376	0.0667	0.0275
Age: Mature	0.0546	0.0359	0.0671	0.0276
Broad Target Market	–0.1710	–0.1931	–0.1562	–0.0865
High-End Target Market	–0.1653	–0.2593	–0.1507	–0.0836
Low-End Target Market	–0.2154	–0.2419	–0.2062	–0.1090
log(Size)	0.0039	0.0016	0.0057	0.0020
log(GDP per capita)	–0.3543	–0.3790	–0.3424	–0.1793
HDI	–0.8077	–0.9072	–0.7167	–0.4087
Gross Saving Rate	0.0059	0.0054	0.0069	0.0030
Depth of Financial Sector	0.0004	0.0001	0.0007	0.0002
Share of Rural Population	0.0277	0.0268	0.0284	0.0140
Remittances-to-GDP Ratio	–0.0049	–0.0059	–0.0039	–0.0025
Branches per 100K people	–0.0058	–0.0064	–0.0054	–0.0029
Internet Subscription per 100 people	0.0119	0.0094	0.0138	0.0060
WGI: Control of Corruption	–0.1674	–0.1973	–0.1471	–0.0847
WGI: Government Effectiveness	0.2801	0.2518	0.3177	0.1417
WGI: Political Stability	–0.1526	–0.1615	–0.1458	–0.0772
WGI: Rule of Law	–0.3224	–0.3418	–0.2930	–0.1631
WGI: Regulatory Quality	0.0269	0.0123	0.0473	0.0136
WGI: Voice and Accountability	0.1005	0.0844	0.1115	0.0509
Financial Crisis Year	–0.0054	–0.0184	–0.0005	–0.0027
Post-Financial Crisis Period	0.0111	0.0051	0.0237	0.0056

For identification, the Non-Profit coefficient is normalized to negative unity. The two middle columns report the two-sided 95% percentile block-bootstrap confidence bounds. *New*, *for-profit* MFIs targeting *small businesses* in the *pre-financial crisis period* are the reference group.

data point used in the estimation by reweighing observations so that the regularity conditions are satisfied at each data point. This allows us to impose the monotonicity restrictions *post*-estimation via a quadratic programming technique (for more details, see Appendix B).

To empirically assess sensitivity of the results to controlling for endogenous self-selection of MFIs into either of the two types of institutions (loans-only and loans-and-savings), we also estimate an alternative model whereby heterogeneous cost functions in (3.1)–(3.2) are fitted while completely ignoring the technology selection process in (3.3). In this case, to guarantee as much comparability across the models as possible, the cost functions are estimated via simple first-difference least-squares estimator. Essentially, this alternative model estimates heterogeneous microfinance technologies under the rather strong assumption of exogenous/ignorable selection. Comparing the results across the two models allows us to examine the degree to which the technological estimates such as scope and scale economies may get distorted as a result of the potential model misspecification due to the ignored endogenous self-selection of MFIs.

We begin by briefly discussing the first-stage estimates of the MFI-type selection equation described in (3.3) and (3.9). Table 4 reports estimates of the $\boldsymbol{\gamma}$ coefficient parameters along with their two-sided 95% block-bootstrap percentile confidence intervals from Klein & Spady's (1993) semiparametric maximum likelihood estimator. For identification purposes, the coefficient in front of one of the covariates (the Non-Profit indicator, in our case) is normalized to negative unity.[21] Also, to account for potential short- and long-lasting structural changes in the microfinance industry

[21] The identification also requires the presence of at least one continuous covariate among $\mathbf{z}_{it}$: we have multiple.

in the wake of the 2008 global financial crisis, in addition to the $\mathbf{z}$ variables described in Section 4, we also include dummies for the year 2008 and the 2008-onward period.[22] The model has a good fit with the correct classification ratio (using the 0.5 threshold) being equal to 0.78 overall and 0.81 and 0.73 by loans-only and loans-and-savings choice outcome. Further note that, unlike traditional parametric binary-choice models (probit or logit), semiparametric single-index models like the one we employ in this paper do not impose monotonicity on unknown $\mathcal{F}(\cdot)$ since they assume no particular distribution. The latter implies that the sign of estimated coefficients does not necessarily coincide with the sign of marginal effects of covariates at each observation. This however poses no problem for us since the sole purpose of the first-stage estimation is to facilitate a selection-bias correction during the (second-stage) estimation of cost functions which are of primary interest to us. To facilitate interpretation, we therefore report the median marginal effects of covariates on the conditional propensity to select the loans-and-savings MFI type (see the last column in Table 4).

The results show that all selection variables are statistically significant. Unsurprisingly, we find that the deposit-collecting MFIs tend to be for-profit and larger in size, possibly because obtaining a license to collect savings deposits is associated with the size entry barriers that larger for-profit entities are more likely to satisfy. We find that, relative to the reference group of new MFIs (less than 3 years in operation), older MFIs classified as "young" and "mature" are more likely to collect savings but the median partial effect on the probability is only 3% which is quite timid. We also find that loans-and-savings MFIs are less likely to target poorer clients. More concretely, consistent with the view that it is difficult to meet the savings needs of a poorer clientele, the median MFIs targeting borrowers other than the largest client group of small businesses (our reference group) are found to be 8–11% less likely to self-select into offering savings. Deposit-collecting MFIs are also less likely to operate in richer countries and countries with higher levels of human development plausibly because, in less developed countries, the marginalized clients have a higher demand for the savings services from institutions like MFIs due to being under-served by more traditional financial intermediaries such as commercial banks. This interpretation is supported by our results whereby the deposit-collecting MFIs are more likely to operate in countries with higher proportions of rural population and fewer bank branches per 100K people. While remittances have been a big target and motivation for MFIs to start offering savings deposits and payment services, the empirical results show that the deposit-collecting MFIs are more likely to operate in countries with *lower* remittance rates. It is therefore possible that regular commercial banks have already captured the remittances flows. At the same time, the positive link to a higher internet subscription density suggests that MFIs have maintained an edge in the innovative technology to reach rural borrowers. We also find that the institutional environment matters significantly for the MFIs' ability to collect deposits. Finally, we find a positive but small partial effect on the selection propensity in the years following the financial crisis which suggests that, after the financial crisis (which is estimated to have had a negative effect in its peak year of 2008), MFIs were more likely to be loans-and-savings rather than loans-only. This is however likely due to the already existing trend toward commercialization and transformation in the industry rather than due to the financial crisis itself. These results attest to the timeliness of our work for policy considerations because all previous studies of scope economies derive their findings from the older pre-crisis data. In what follows, we now focus on the main results of our paper concerning technological metrics of microfinance.

We also examine the second-stage translog parameter estimates (omitted to conserve space) of the variable cost functions for loans-only and loans-and-savings MFIs, with and without selection. Overall, the results suggest two findings. First, the cost parameters are non-negligibly heteroge-

[22] As footnoted earlier, these dummies are also included in the outcome equations (in addition to the time trend along with its expansion terms which are meant to allow for continuous temporal shifts in the cost frontiers).

Table 5. Returns to Scale Estimates

Model	*Point Estimates* Mean	1st Qu.	Median	3rd Qu.	*Categories, %* **DRS**	NDRS	**CRS**	NCRS	**IRS**	NIRS
			Loans-Only							
With Selection	1.552	1.425	1.536	1.650	**0.00**	100.00	**0.00**	100.00	**100.00**	0.00
	(1.506, 1.608)	(1.390, 1.482)	(1.491, 1.590)	(1.591, 1.733)						
Without Selection	1.738	1.446	1.678	1.941	**0.00**	100.00	**0.43**	99.57	**99.67**	0.33
	(1.648, 1.852)	(1.384 1.514)	(1.597, 1.757)	(1.824, 2.089)						
			Loans-and-Savings							
With Selection	1.601	1.397	1.561	1.771	**0.00**	100.00	**0.46**	99.54	**99.61**	0.39
	(1.529, 1.681)	(1.335, 1.466)	(1.495, 1.630)	(1.658, 1.887)						
Without Selection	1.906	1.549	1.805	2.154	**0.00**	100.00	**0.50**	99.50	**99.57**	0.43
	(1.727, 2.105)	(1.408, 1.693)	(1.620, 1.970)	(1.902, 2.383)						

The left panel summarizes RTS point estimates with the corresponding two-sided 95% percentile block-bootstrap confidence intervals in parentheses. Each MFI is classified as exhibiting decreasing/constant/increasing returns to scale (DRS/CRS/IRS) if the point estimate of its returns to scale is statistically less than/equal to/greater than one at the 5% significance level using appropriate one- or two-sided percentile confidence bounds. The right panel reports sample shares for each category and for its corresponding negating alternative: non-decreasing/non-constant/non-increasing returns to scale (NDRS/NCRS/NIRS), respectively. Percentage points sum up to a hundred within binary groups only.

neous across loans-only and loans-and-savings MFIs thereby buttressing our concerns about the validity of assumption that MFIs share the same production technology regardless of their output mixes embedded in a popular single-equation approach to modeling cost relationships for MFIs. We formally test the null of non-heterogeneous cost functions across the two types of MFIs via the cross-equation multiple-restriction Wald test on the parameters of common regressors in (3.1)–(3.2) and strongly reject the null in favor of heterogeneous cost functions with the block-bootstrap p-value of 4×10^{-6}, when the endogenous technology selection is being controlled for, and the p-value of 0.002, when selectivity is ignored. Thus, the data strongly indicate that it is more appropriate to model *different* cost functions for loans-only and loans-and-savings MFIs not only to avoid the "log-of-zero" problem but, more importantly, to allow for potential parameter instability and discontinuities across MFI-type-specific cost functions. The second finding concerns the importance of explicitly controlling for MFIs' endogenous selection of the output mix and, by extension, of the appropriate microfinance technology. Rather expectedly, we find that the failure to account for selectivity is capable of significantly distorting the results: the non-negligibly *over*estimated returns to scale, as is evident from Table 5, are perhaps the most exemplary of such distortions.

Here, we use the measure of returns to scale that takes into account quasi-fixity of the equity input (Caves et al., 1981): $RTS = (1 - \partial \log \mathbb{C}_r / \partial \log e) / \sum_m \partial \log \mathbb{C}_r / \partial \log y_m$. While the results reported in Table 5 are qualitatively similar, overwhelmingly pointing to statistically significant increasing returns to scale for all MFIs, the quantitative magnitude of untapped potential for *scale*-driven cost reduction is however significantly *over*estimated when one does *not* control for self-selection among institutions. Generally, such cost savings are attributed to fixed costs being spread over more units of output as MFIs radially expand the scale of their operation, which leads to a decline in the cost per unit of output. However, we would like to emphasize that such scale-driven cost savings are being measured for a *given* output mix and exclude any other potential cost reduction associated with the provision of a more diversified output mix. As discussed in Section 2.2, to measure the latter cost saving potential, we look at the degree of scope economies in MFIs.

We formally test the null of exogenous output mix selection among MFIs. The Hausman-type test based on the distance between the first-difference least-squares estimator (ignoring selection)

Table 6. Degree of Scope Economies Estimates

Model	*Point Estimates* Mean	1st Qu.	Median	3rd Qu.	*Categories, %* **SD**	SND	**SI**	SNI	**SE**	SNE
With Selection	0.072	–0.041	0.056	0.169	13.46	86.54	50.27	49.73	45.59	54.41
	(0.001, 0.119)	(–0.148, 0.024)	(–0.022, 0.085)	(0.101, 0.196)						
Without Selection	0.152	0.043	0.108	0.219	0.00	100.00	41.76	58.24	69.53	30.47
	(0.086, 0.189)	(–0.015, 0.063)	(0.034, 0.141)	(0.150, 0.279)						

The left panel summarizes $DGSE$ point estimates with the corresponding two-sided 95% percentile block-bootstrap confidence intervals in parentheses. Each MFI is classified as exhibiting scope diseconomies/invariance/economies (SD/SI/SE) if its point estimate of $DGSE$ is statistically less than/equal to/greater than zero at the 5% significance level using appropriate one- or two-sided percentile confidence bounds. The right panel reports sample shares for each category and for its corresponding negating alternative: scope non-diseconomies/non-invariance/non-economies (SND/SNI/SNE), respectively. Percentage points sum up to a hundred within binary groups only.

and Kyriazidou's (1997) estimator (controlling for selection) decisively rejects the null of ignorable selection with block-bootstrap p-values being no larger than 10^{-4} for both loans-only and loans-and-savings MFIs. Failure to accommodate this endogenous selectivity is therefore likely to lead to biased and inconsistent estimates of technological metrics for MFIs and thus misleading conclusions about the industry.

Degree of Scope (Dis)Economies. To measure the degree of scope economies for individual MFIs, which is a main focus of our paper, we use the fitted MFI-type-specific variable cost functions (from the two models: with and without selection). As discussed in Section 2, since our measure of the degree of within-sample scope economies (DSE) depends on the distribution [controlled by the choice of the ϖ weight] of a more diversified loans-and-savings MFI's output quantities across more specialized production units, we employ a conservative approach to measuring the degree of scope economies by focusing on the *smallest* DSE estimates in the data-defined admissible region. Specifically, for each loans-and-savings MFI in a given year, we perform a grid search (over ϖ) of the lowest value of DSE. The grid search is performed over the permissible range of ϖ between 0 and 1 at the 0.01 increments. We refer to the located smallest DSE estimate as the measure of the degree of MFI's "global" scope economies, i.e., $DGSE$. The reported $DGSE$ estimates are obtained using actual data for the multi-output loans-and-savings MFIs which gives us the distribution of estimates for the entire industry.

The point estimates of the degree of "global" scope economies $DGSE$ from two models are summarized in Table 6, the right panel of which also breakdowns MFIs into three main categories—scope diseconomies, scope invariance and scope economies (SD/SI/SE)—each of which has a negating alternative: scope non-diseconomies, scope non-invariance and scope non-economies (SND/SNI/SNE), respectively. The classification into these three binary groups is performed on the basis of a $DGSE$ point estimate being statistically less than/equal to/greater than zero at the 5% significance level using appropriate one- or two-sided percentile bootstrap confidence bounds. More concretely, given the "less than" ("greater than") definition of scope diseconomies (economies), an MFI is classified as exhibiting significantly negative scope diseconomies (positive scope economies) if its *one*-sided 95% upper (lower) bound is less (greater) than zero; if not, an MFI is said to exhibit scope non-diseconomies (non-economies). In the case of scope invariance, the notion of which is based on the "equal to" definition, an MFI is classified as exhibiting invariance to scope if its *two*-sided 95% confidence interval contains zero; if not, it is said to show significant scope non-invariance. Clearly, the SD/SI/SE categorization is *not* mutually exclusive by construction.[23] For a graphical

[23]The assignment into one of the three SD/SI/SE categories *and* their corresponding negating alternative—

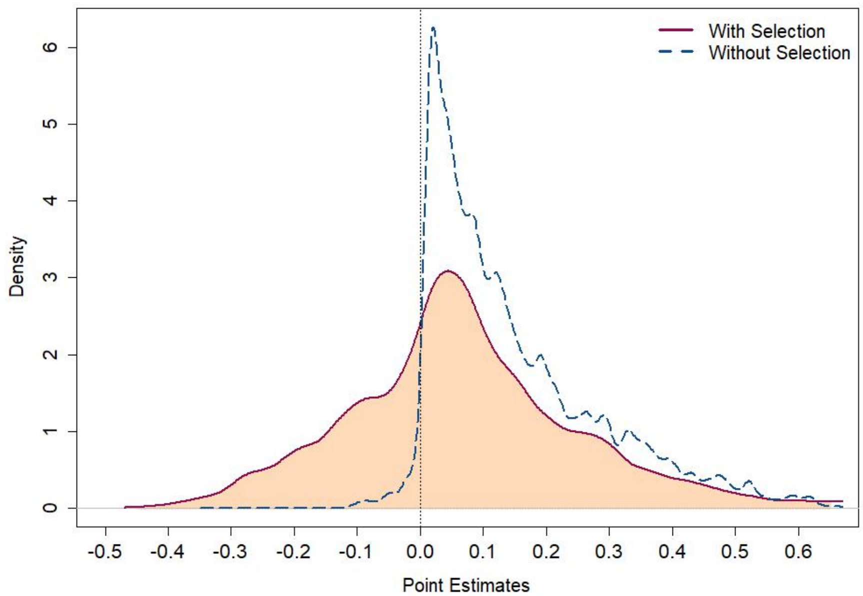

Figure 3. Distribution of the Degree of Scope Economies Estimates

summary of the results, also see Figure 3 which plots cross-validated kernel densities of the $DGSE$ point estimates across individual loans-and-savings MFIs.

Based on our preferred model, which controls for the endogenous selection of output mix/scope, we find that, *at the median*, the microfinance production technology largely exhibits invariance to scope of outputs produced, with the median estimate of the degree of scope economies at statistically insignificant –0.06. In contrast, the median estimate from the alternative model that fails to correct for endogenous self-selection is positive and statistically significant, pointing to non-negligible economies of scope at around 0.11. However, the median point estimates may be rather uninformative because they mask heterogeneity across individual institutions the presence of which can be vividly seen from the empirical distribution of the $DGSE$ estimates in Figure 3. The figure indicates that a third of microfinance industry (32%) is comprised of the institutions with *negative* point estimates of the degree of scope economies. Indeed, we find that, after controlling for endogenous selection, only 46% of loans-and-savings MFIs in our sample enjoy significantly positive scope economies. More importantly, for a non-trivial 14% of institutions, the empirical evidence suggests the existence of significant *dis*economies of scope indicating that the *separate* production of loans and savings accounts has the potential to *reduce* an MFI's costs. For concreteness, the mean $DGSE$ value for these multi-output MFIs is a sizable –0.21 implying, on average, the potential for a 21% cost saving if the joint production of loans and savings is replaced with two single-output MFIs. Note that the presence of such scope diseconomies however in no way implies cost inefficiency or sub-optimality on the part of these multi-output MFIs. Neither does it suggest that these MFIs may not reduce their costs by scaling up their operations to capitalize on (universally) significant *scale* economies.[24] It does however suggest that it may be inappropriate to invoke scope economies

SND/SNI/SNE, respectively—is however mutually exclusive.

[24] As mentioned earlier, the increasing returns to scale are in fact a contributor to positive economies of scope (see Chavas & Kim, 2010). This is readily seen from our definition of scope economies in (2.6) where the magnitudes of y_1 used in the construction of a counterfactual sum of $\mathbb{C}_1(\cdot)$ and $\mathbb{C}_2(\cdot)$ on the left-hand side are the scaled-down

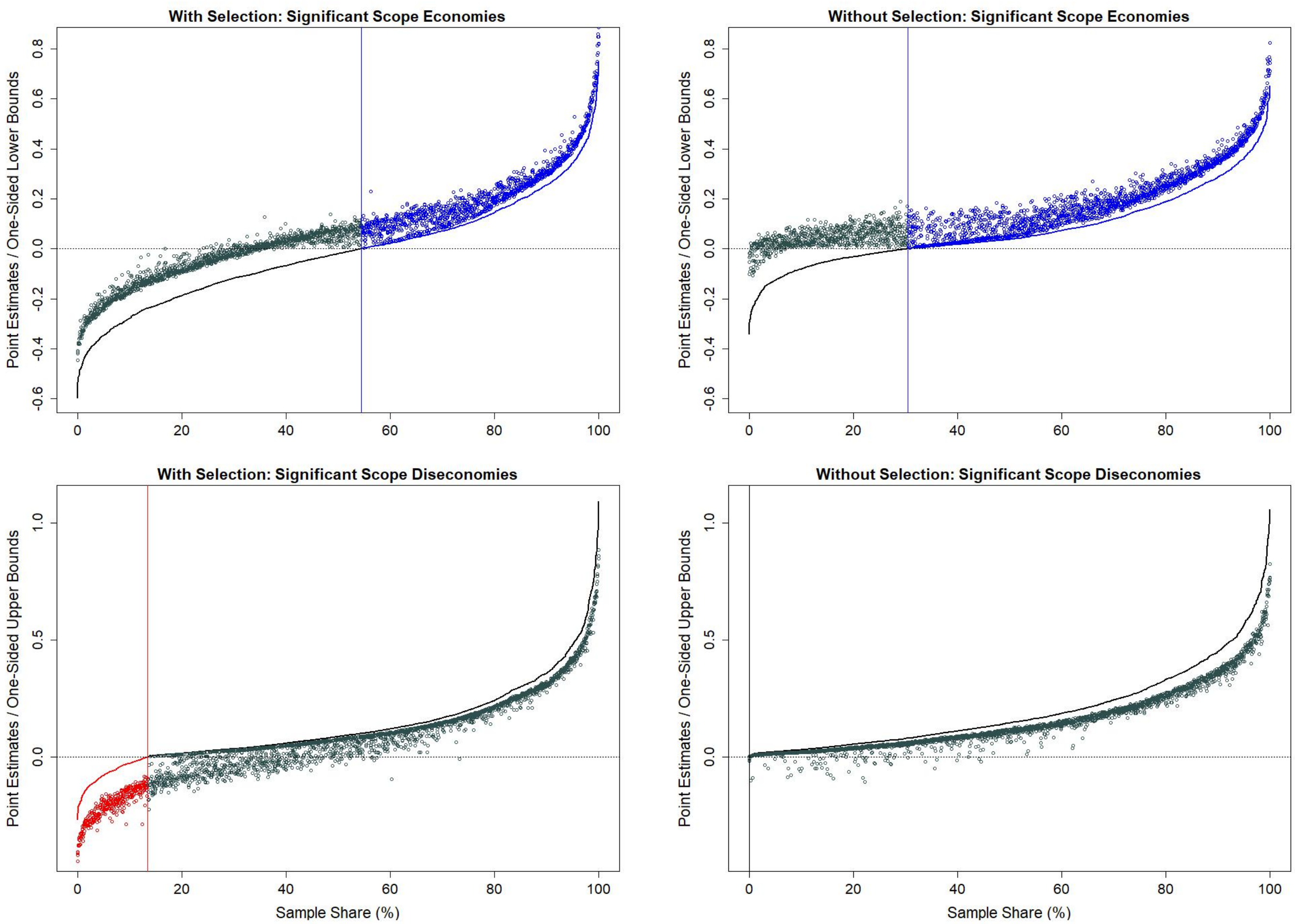

Figure 4. The One-Sided 95% Percentile Block-Bootstrap Lower and Upper Bounds (Solid Lines) of the Degree of Scope Economies Estimates (Scatter Points) [*Note:* $DGSE < 0$ and $DGSE > 0$ are in red and blue, respectively]

as a justification for a *universal* expansion of the scope of financial operations by MFIs. That is, it may be ill-advised to urge loans-only MFIs to start offering savings accounts to their customer base unconditionally on the grounds of potential cost savings due to the expanded scope of services. After all, scope economies are significantly positive for *less* than a half of MFIs only, although with the average value estimated at a non-negligible 0.23. For 50% of loans-and-saving MFIs, the operating costs exhibit scope invariance as indicated by statistically insignificant degree of scope economies estimates (see Table 6).

The above findings from our preferred model differ from those based on the alternative model, which takes the selectivity issue for granted. More specifically, from Figure 3, it is easy to see that the failure to account for selection leads to the *over*estimation of degree of scope economies for MFIs, with the entire $DGSE$ distribution visibly shifted to the right. As a result, the number of MFIs for which one finds significantly negative scope economies is misleadingly decreased to none at all with the overwhelming majority of institutions (70%) said to be exhibiting significant scope

shares of y_1 used on the right-hand side. Hence, the presence of scale economies would be one of the factors contributing to a positive gap between the left- and right-hand sides of inequality.

economies. Figure 4 vividly illustrates tendency of the alternative model to overestimate (as a result of ignoring the output scope selection) the magnitude of scope economies in the industry with a consequent mischaracterization of microfinancing technology. The figure scatter-plots the degree of scope economies ($DGSE$) estimates for each MFI-year along with their *one*-sided 95% lower and upper percentile bootstrap bounds (solid lines) from the two models, based on which the right panel of Table 6 is partly populated. The observation-level estimates are sorted by their confidence bounds and color-coded depending on whether they are above/below zero or not. The estimated one-sided intervals are visibly shifted upwards if the endogenous selection is left unaccounted for and, as a result, many MFIs are misleadingly classified as enjoying significantly positive scope economies: 70% of institutions against 46% as suggested by our more robust preferred model.

Our selection-corrected results differ from those previously reported in Delgado et al. (2015) who document the 10% scope economies in the industry overall. In fact, they find that as many as 65% of MFIs in their sample exhibit statistically significant scope economies leading them to conclude that the industry generally enjoys scope-driven cost savings, which significantly differs from our conclusions. Our findings also differ from those by Hartarska et al. (2011) who too generally find empirical support for positive scope economies in the industry with the statistically significant median estimate of 19%. They also hardly find any evidence of diseconomies of scope among MFIs with most of their reported negative estimates being statistically insignificant instead suggesting scope invariance,[25] whereas we find that the MFIs with significantly negative scope diseconomies actually constitute a non-negligible 14% of the microfinance industry. Furthermore, based on Hartarska et al.'s (2011) estimates, the institutions that do have statistically significant scope economies enjoy much larger (almost trice as large) scope-driven cost savings than what we find in this paper. It is noteworthy however that these previous studies do *not* control for endogenous self-selection of MFIs and nor do they estimate separate cost functions to allow for (discontinuous) technological heterogeneity across the specialized and more integrated MFIs.[26] Instead, prior studies usually accommodate the heterogeneity across MFIs by including control variables directly in the cost function common to all MFIs. Such an approach leads to qualitatively different implications of controlling for the MFI heterogeneity whereby adding controls leads to an increase in the magnitude of scope economies resulting in a rightward shift of the distribution of their estimates. That is, in previous studies, the estimated degree of scope diseconomies becomes smaller (in absolute value) and the degree of scope economies becomes larger when controls are included in the model. In contrast, controlling for (endogenous) technological heterogeneity across MFIs using our approach shifts the estimates of degree of scope economies to the left: they becomes smaller in magnitude with a much larger share of them being significantly negative (see Figures 3–4).

Next, we provide a more in-depth analysis of our findings about scope (dis)economies in the industry based on the results from our preferred model of endogenously heterogeneous technologies.

We begin by examining the evolution of scope (dis)economies over time, which is of particular interest given that our sample period includes both the pre- and post-financial-crisis years. In Table 7, we document yearly changes in the breakdown of MFIs exhibiting (statistically significant) scope economies and diseconomies along with their corresponding mean $DGSE$ estimates. Up until 2008, the share of MFIs with significantly positive scope economies had been on the rise with a half (52%) of the deposit-collecting MFIs experiencing substantial economies of scope, on average, of about 23% scope-driven cost reduction. The arrival of the global financial crisis appears to have reversed

[25] While Hartarska et al. (2011) do not report shares of MFIs with scope economies/invariance/diseconomies, the statistical significance or insignificance of the quartile estimates of scope economies that they report suggests that only few, if not none at all, MFIs exhibits scope diseconomies.

[26] Prior studies also do not impose theoretical regularity conditions on the estimated technology like we do.

Table 7. Scope (Dis)Economies over Time

Year	*Scope Economies* (SE) % Obs	Mean Est.	*Scope Diseconomies* (SD) % Obs	Mean Est.
2004	34.67	0.287	26.67	–0.246
2005	40.37	0.280	22.36	–0.218
2006	48.99	0.311	17.81	–0.225
2007	50.76	0.269	16.41	–0.220
2008	52.37	0.231	15.88	–0.208
2009	46.40	0.211	14.40	–0.195
2010	45.71	0.204	11.96	–0.196
2011	45.75	0.162	8.82	–0.191
2012	47.18	0.164	4.62	–0.186
2013	33.12	0.151	3.82	–0.189
2014	8.93	0.117	3.57	–0.164
2004–2014	45.59	0.225	13.46	–0.209

The estimates are from our preferred model of heterogeneous technologies with selectivity. The "scope economies" and "scope diseconomies" categories are respectively based on whether $DGSE$ point estimates are statistically greater or less than zero.

this trend. Since then, the average magnitude of (significantly positive) scope economies has been steadily decreasing and was only 12% in 2014. The share of MFIs exhibiting these significant scope economies has declined as well, to 33% in 2013 with a consequent collapse to mere 9% in 2014. Meanwhile, the proportion of MFIs exhibiting significantly negative scope *dis*economies has also been shrinking steadily throughout the years (from 27 to 4% of the industry) reinforced by a simultaneous decline in the absolute magnitude of such scope diseconomies from 25% to 16%. These findings suggest that many deposit-collecting MFIs, which might used to enjoy positive scope economies, may have since seen them erode following the turbulent times of turmoil in global financial markets. Also, the MFIs that have entered the deposit collection only recently, post the financial crisis, may not have the advantages of earlier entrants. Thus, with the rapidly declining shares of institutions exhibiting both the economies and diseconomies of scope, the cost structure of more and more MFIs appears to have generally become invariant to scope (i.e., they have statistically insignificant $DGSE$ estimates). Figure 5 provides a convenient illustration of these industry trends across *all* MFIs, those with statistically significant and insignificant (positive and negative) $DGSE$ point estimates. The pictured box-plot of the distribution of $DGSE$ point estimates across the years tells a familiar story, consistent with the already documented findings, whereby the microfinance industry has experienced a notable reversal in the pre-crisis upward trend in the magnitude of scope economies accompanied by the growing dominance of near-zero estimates in the last years of the sample period.

To systemically analyze how the magnitude of scope (dis)economies varies across different institutions, we estimate several regressions of the degree of scope economies on attributes of MFIs (also see Berger et al., 2000). The results are reported in Table 8. The first column reports the results from a simple OLS regression where the left-hand-side variable is the $DGSE$ point estimates (both statistically significant and insignificant) for all loans-and-savings MFIs, whereas the second column contains results from the same regression estimated using a sample of statistically significant scope (dis)economies estimates only (i.e., statistically positive and negative $DGSE$ point estimates). The third column presents results from the probit model with an indicator outcome variable correspond-

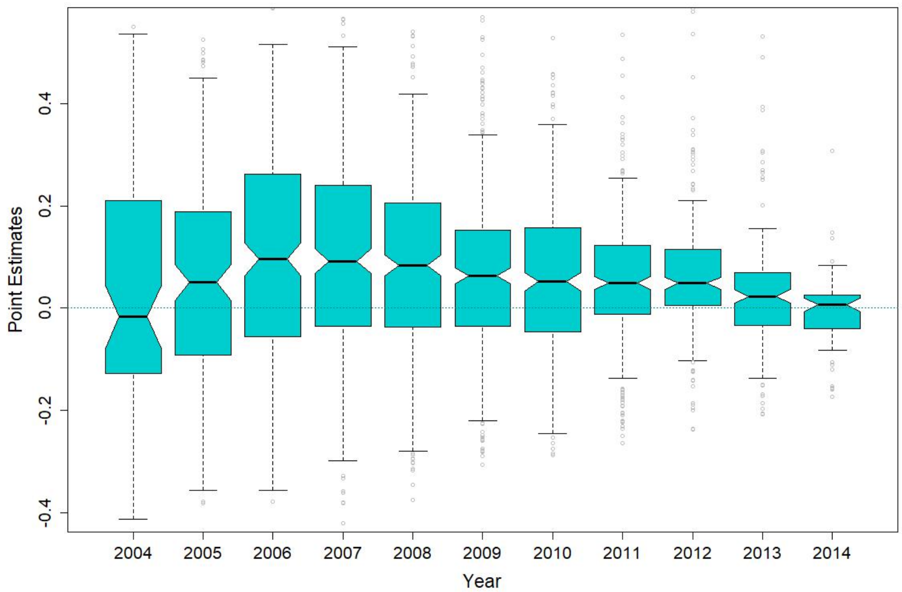

Figure 5. Yearly Distribution of the Degree of Scope Economies Estimates

ing to significantly positive "scope economies" and its negating alternative "scope non-economies," otherwise. The two categories are based on whether $DGSE$ point estimates are statistically greater than zero or not. Since probit coefficients cannot be readily interpreted, the last column reports estimates of the corresponding average marginal effects on the propensity that an MFI exhibits significant scope economies (over non-economies).

We begin by examining the relationship between scope economies and the size of MFIs. While one generally expects a negative relationship between the degree of scale economies (returns to scale) and the firm size, economic theory is however far less clear on such relationship in the case of scope economies. The results in Table 8 indicate that larger MFIs are less likely to exhibit scope economies with, on average, 1% increase in the size being associated with about 0.10% point lower $DGSE$ point estimates (overall and statistically significant only). Looking at the MFIs' propensity of exhibiting significantly positive scope economies, we also find that larger MFIs are less likely to enjoy scope economies (the average marginal effect is –0.21). Figure 6 provides additional insights into the said relationship. It shows the distribution of $DGSE$ point estimates across sample quintiles of the MFI size in the form of box-plots. Based on our preferred model, the data point to a clear inverse relationship suggesting that smaller loans-and-savings MFIs are the ones capitalizing on positive scope economies.[27] Therefore, while deposit-collecting MFIs often must meet the size entry barriers to obtain a license to start collecting savings, such larger MFIs are less likely to benefit from scope economies associated with jointly lending and mobilizing deposits.

Relative to the reference group of newly established MFIs (with fewer than 3 years since inception), "young" and "mature" MFIs generally have 3% point smaller $DGSE$ estimates although there is no statistically significant relationship with the associated probability of significantly posi-

[27] In our data, we are unable to distinguish between mandatory and voluntary savings and the proportion thereof within the "savings" variable because we only have a total measure of savings. Therefore, we evaluate scope economies for different MFIs categorized by the total savings to assets ratio: 0.1 or more, 0.2 or more, and 0.3 or more. We do not find significant differences in the magnitude and statistical significance of the estimated scope economies across these categories.

Table 8. Scope Economies Regressions

	OLS (1)	*OLS* (2)	*Probit* Coeff.	*Probit* AME
log(Size)	–0.0988***	–0.1144***	–1.6665***	–0.2110***
	(0.0023)	(0.0027)	(0.1014)	(0.0040)
Age: Young	–0.0358***	–0.0336***	0.3298	0.0417
	(0.0110)	(0.0113)	(0.2185)	(0.0277)
Age: Mature	–0.0315***	–0.0306***	0.2127	0.0269
	(0.0109)	(0.0114)	(0.2234)	(0.0284)
Equity-to-Assets Ratio	0.0055	–0.0066	–0.8672**	–0.1098**
	(0.0181)	(0.0200)	(0.4193)	(0.0531)
log(Price of Fin. Capital)	–0.0273***	–0.0407***	0.0402	0.0051
	(0.0049)	(0.0048)	(0.0988)	(0.0125)
Risk	0.1260***	0.0388	1.6620**	0.2104**
	(0.0384)	(0.0449)	(0.8325)	(0.1060)
Non-Profit	–0.0051	0.0094	–0.2259	–0.0286
	(0.0086)	(0.0103)	(0.2503)	(0.0318)
Small Biz. Target Market	–0.0267**	–0.0307*	–0.1981	–0.0251
	(0.0117)	(0.0168)	(0.3306)	(0.0420)
Broad Target Market	–0.0010	–0.0024	0.1629	0.0206
	(0.0057)	(0.0064)	(0.1559)	(0.0197)
High-End Target Market	–0.0051	–0.0051	0.2280	0.0289
	(0.0109)	(0.0121)	(0.3643)	(0.0464)
NGO	–0.0010	–0.0286*	0.4686	0.0593
	(0.0133)	(0.0173)	(0.3335)	(0.0417)
Rural Bank	–0.0166	–0.0212	–0.6050*	–0.0766*
	(0.0113)	(0.0150)	(0.3150)	(0.0408)
NBFI	0.0146*	–0.0111	0.3227	0.0409
	(0.0087)	(0.0109)	(0.2898)	(0.0362)
Credit Unions/Coops	–0.0064	–0.0445***	–0.0402	–0.0051
	(0.0124)	(0.0164)	(0.3408)	(0.0432)
Other Legal Status	–0.0687***	–0.1237***	–1.3534***	–0.1713***
	(0.0209)	(0.0193)	(0.4180)	(0.0540)
Africa	0.0070	0.0012	–0.2104	–0.0266
	(0.0077)	(0.0090)	(0.1964)	(0.0248)
E. Asia and the Pacific	0.0200**	–0.0023	0.2215	0.0280
	(0.0097)	(0.0107)	(0.2960)	(0.0377)
E. Europe & C. Asia	0.0240***	0.0206**	–0.4114	–0.0521
	(0.0083)	(0.0092)	(0.3483)	(0.0437)
Middle East & N. Africa	0.0025	–0.0368***		
	(0.0138)	(0.0106)		
S. Asia	0.0039	–0.0211**	–0.1208	–0.0153
	(0.0084)	(0.0094)	(0.2233)	(0.0282)
Year Effects	✓	✓	✓	✓
Obs	2,586	1,527	2,572	
R^2	0.853	0.912		
Pseudo-R^2			0.676	

The regressions are based on the *DGSE* point estimates from our preferred model. The first two OLS models use all [in (1)] or only statistically positive and negative *DGSE* estimates [in (2)] as a left-hand-side variable. The probit model uses a binary indicator variable categorizing each *DGSE* point estimates as corresponding to significant "scope economies" or "scope non-economies". The two categories are based on whether *DGSE* point estimates are statistically greater than zero or not. The last column report the average marginal effects (AMEs) on the probability of an MFI exhibiting significantly positive scope economies. MFIs with the legal status of a *bank* targeting *"low-end"* clients in *Latin America* are the reference group. The dummy for MENA (along with the 14 corresponding observations) is omitted from the probit model due to being a perfect predictor. Clustered (at the individual level) standard errors in parentheses: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

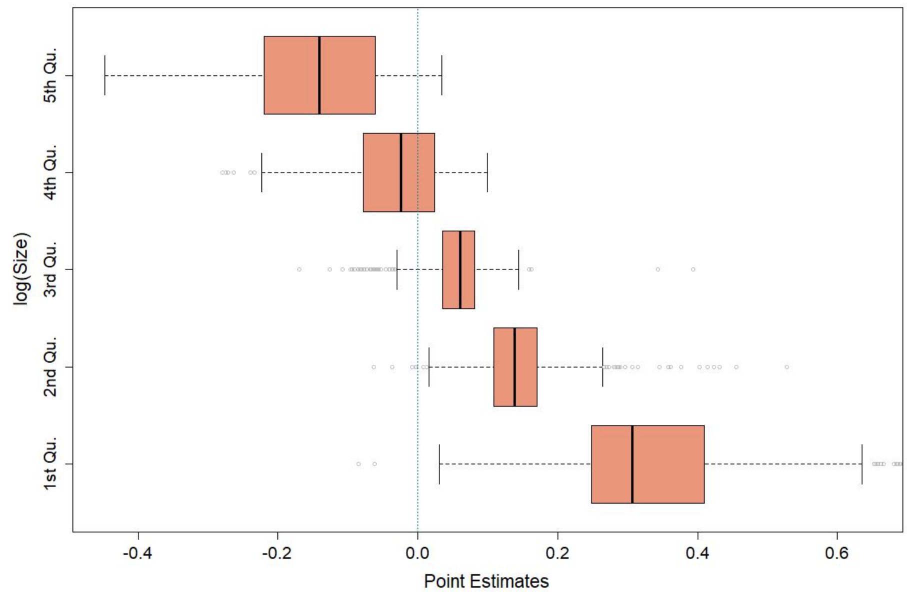

Figure 6. Degree of Scope Economies Estimates by (log) Size Quintiles

tive scope economies. These results suggest no improvement in the potential for scope-driven cost savings with age. On the contrary, the OLS results indicate that MFIs that start off as loans-and-savings generally enjoy larger cost savings from offering deposits as well as loans, but these cost savings may not be preserved later in their life cycle.

The results in Table 8 further show that the level of capitalization (the equity-to-assets ratio) does not systematically associate with the magnitude of the $DGSE$ estimates. However, the likelihood of these point estimates being statistically positive corresponding to scope economies is (significantly) inversely, albeit weakly, related to the level of capitalization, with a 10% point increase in the equity-to-assets ratio associated with a 0.011 lower likelihood of positive scope economies. Along these lines of inquiry, we find that the cost of borrowed capital is negatively associated with the degree of scope economies, whereby a 1% increase in the cost of financial capital is associated with 2.7% point smaller overall $DGSE$ estimates and 4.1% point smaller statistically significant $DGSE$ estimates, but there is no link to the probability of exhibiting significantly positive scope economies. Since the cost of funds is inversely related to the degree of scope economies and equity serves as a cushion to protect against risk, the magnitude and prevalence of scope economies among MFIs is likely sensitive to the loan default risk. Indeed, we find that the riskiness of lending activity exhibits a significantly positive association with both the $DGSE$ point estimates (coefficient of 0.13) and the conditional probability of significant scope economies with a non-negligible average marginal impact on the latter of about 0.02 per 10% point increase in the portfolio at risk.

One of the most crucial discussions in microfinance today is the possibility of a "mission drift" whereby, as MFIs go through commercialization (including the transformation into deposit-collecting institutions) and diversify their products, the depth of outreach or the MFIs' willingness to serve the very poor may be sacrificed. The results from a selection equation have already provided evidence that the MFIs targeting the least poor microfinance clients are more likely to be offering deposit accounts. The results from our scope economies regressions reported in the second column of Table 8 show that the deposit-collecting MFIs targeting the least poor borrowers (small businesses) have about 2.7–3.1% point smaller degree of scope economies estimates relative to the MFIs targeting the poorest (our reference group). There is no statistically significant difference between the degree of scope economies exhibited by the MFIs targeting the poorest and those targeting other clientele categories as classified by the loan size. There is also no significant difference in the probability of having positive scope economies (when accounting for the sampling error in their estimates) across the target market groups, thereby suggesting a certain degree of neutrality and no statistically significant evidence of a mission drift motivated by cost savings due to scope economies. This interpretation is also supported by the finding that the degree of scope economies is largely unrelated to the MFI's profit status.

Next, we look at the differences across organizational (legal) types of MFIs. We use banks as the reference group because the deposit-collecting MFIs are most likely to seek to emulate banks' behavior, and thus the MFIs organized as banks would have scope economies similar to those found in regular commercial banks. The two groups of MFIs most similar to banks are rural banks and credit unions (CUs, also known as financial cooperatives). We find that, relative to banks, credit unions have smaller statistically significant $DGSE$ estimates by 4.5% points, whereas rural banks are less likely to exhibit scope economies, on average, by 7.7%. Since we control for the institution's size, capitalization, age and the target market, it appears that the MFI's organizational form itself is associated with the variation in scope economies. While the findings pertaining to rural banks are somewhat expected because serving remote rural populations is harder, it is interesting that the same is observed for credit unions which typically serve clients in smaller geographic areas. The results for the MFIs operating as non-governmental organizations (NGOs), which is a legal form most different from banks, only show that, relative to banks, NGOs on average have a 2.9% point smaller magnitude of statistically significant $DGSE$ estimates but generally are not more likely to have significantly positive scope economies. The results for MFIs operating as non-bank financial institutions (NBFIs), an organizational form used when a country's regulatory environment allows non-bank, non-CU and non-NGO organizations to operate as deposit-collecting MFIs, show that they enjoy somewhat higher degree of scope economies relative to banks overall (by 1.5% point) but the differences in magnitudes of significant $DGSE$ estimates are insignificant. Finally, for the 15 organizations in our sample registered as "other" as the best alternative under their country's regulations permitting them to collect deposits, we observe that they have about 6.8% point smaller $DGSE$ estimates, and 12.4% point smaller statistically significant $DGSE$ estimates than banks, and are about 17% more likely on average to have scope diseconomies. This may be plausibly attributed to the inhospitable regulatory environments in which these MFIs might be operating.

We also examine geographical differences in scope economies, and our results are largely consistent with previous microfinance studies showing a regional variation in the MFI performance, prevalence of scope economies and the interaction thereof with the overall economy (Lopatta & Tchikov, 2016; Hartarska et al., 2013). As a reference group, we choose the MFIs in the Latin America because they include some of the oldest as well as recently established institutions; they also have capital and organizational structures very similar to the industry average. We find no significant differences in the estimates of the degree of scope economies across the Latin American MFIs and those in Africa. However, MFIs in the Eastern Europe and East Asia, on average, have about 2% point higher $DGSE$ point estimates, but only those in the Eastern Europe continue to have higher estimates when we account for the sample error in the $DGSE$ estimates. The regression results also show that MFIs in the MENA region have on average 3.7% point, and those in the Southeast Asia have 2.1% point, smaller estimates of statistically significant $DGSE$ estimates. We find no systematic differences across regions in terms of the probability of exhibiting scope economies.

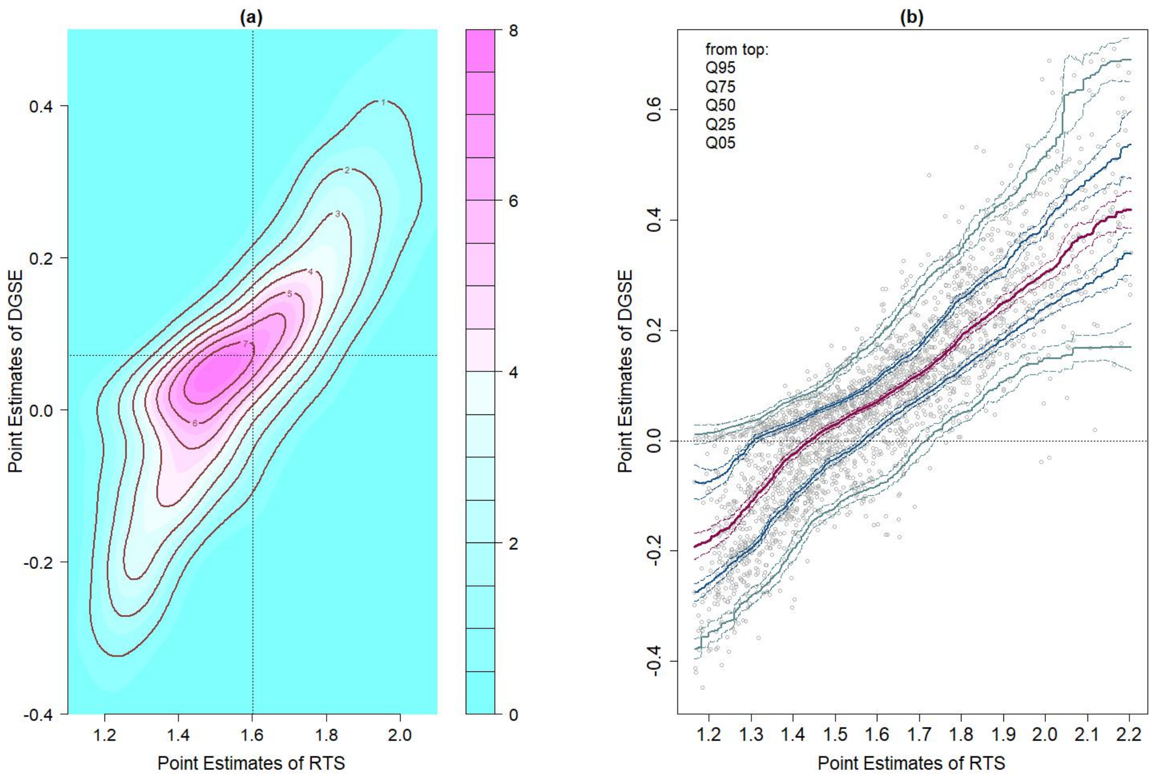

Figure 7. Bivariate Relationship between the Degree of Scope Economies and Returns to Scale Estimates: (a) Contours of the Estimated Joint Density; (b) Fitted Quantiles (solid) with the Corresponding Two-Sided 95% Bootstrap Confidence Bounds (dashed)

Overall, the regression results point to limited differences in scope economies across organizational types and geographic locations of MFIs. We do document consistent differences in terms of capitalization and the cost of borrowed funds as well as in terms of the size with smaller MFIs found to be more likely to have scope economies. This suggests that smaller loans-and-savings MFIs are the primary beneficiaries of scope economies but, since they are small, they may also have not taken a full advantage of their scale economies. (Recall that all MFIs are found to almost universally exhibit increasing returns to scale.) The untapped potential to drive the costs of these institutions further down by capitalizing on significant increasing returns to scale via expanding their operations can also be inferred from Figure 7. It depicts a contour plot of the bivariate density of $DGSE$ and RTS point estimates across MFIs, which enables us to assess the relationship between the two not just at some fixed moment but distribution-wise; it is complemented by the plot of the same relationship at different quantiles. Both sub-figures point to a positive relation between the degrees of scope and scale economies with the MFIs exhibiting diseconomies of scope also having increasing returns to scale but of smaller magnitudes than the MFIs enjoying scope economies. These findings compare well to Delgado et al.'s (2015) who too find that, prior to 2006, the majority of MFIs enjoying scope economies also exhibited increasing returns to scale. In contrast to our finding of almost universal scale economies in the industry, they however find that diseconomies of scope tend to go hand-in-hand with decreasing returns to scale.

6 Conclusion

Scope economies ensuing from the joint offering of loans and savings accounts (as opposed to loans only) are customarily invoked to promote the transformation of credit-only microfinance institutions (MFIs) into integrated loans-and-savings entities. To ensure robust inference, this paper estimates scope economies for the microfinance industry using a novel approach which accommodates inherent heterogeneity across loans-only and loans-and-savings MFIs as well as controls for endogenous self-selection of institutions into the either type. In contrast to popular alternatives employed in the microfinance literature, our measurement of the degree of scope economies does not rely on a rather unrealistic assumption whereby specialized loans-only MFIs share the same technology with and incur the same fixed costs as do the integrated loans-and-savings MFIs, which substantially decreases the reliance of our estimates on counterfactuals and minimizes the "excessive extrapolation" problem. We therefore are able to offer policy-makers and stakeholders a fresher and more robust perspective on benefits and costs of promoting integrated loans-and-savings MFIs on grounds of the cost savings potential due to scope economies.

Using recent data on about a thousand MFIs from 76 countries operating in 2004–2014, we find that the microfinance industry largely exhibits scope invariance with the median degree of scope economies being statistically indistinguishable from zero, and scope economies are significantly positive for less than a half of loans-and-savings MFIs. For a non-trivial 14% of institutions, the empirical evidence suggests the existence of significant *dis*economies of scope indicating that the separate production of loans and savings accounts actually has the potential to reduce an MFI's costs. This suggest that it may be ill-advised to invoke scope economies as a blanket justification for *universal* expansion of the scope of financial operations by MFIs. We also find that failure to account for self-selection dramatically overestimates the degree of scope economies resulting in the failure to detect scope diseconomies among MFIs. A more in-depth analysis shows very small differences in the estimates of the degree of scope economies among MFIs based on their geographic region, organizational type or age. The results most prominently suggest a negative relation between the extent of scope economies and the MFI's size or cost of capital and a positive relation with the level of capitalization. The analysis of temporal dynamics in scope economies shows that their magnitude as well as the prevalence in the industry have been steadily declining, especially in the aftermath of the global financial crisis.

Appendix

A Deposits as Outputs: Additional Evidence

Table A.1 reports point estimates of the cost elasticity of deposits obtained from the first-difference estimator of $\mathrm{C}(w_1, w_2, x_3, e, y_1, y_2)$ for loans-and-savings MFIs using the translog specification. The left-hand-side C variable is a sum of expenses on x_1 and x_2 only. The point estimates are overwhelmingly positive, as can be seen in Figure A.1 which plots the histogram of these elasticity estimates. The mean elasticity is estimated at the (one-sidedly) statistically positive 0.095. Overall, the non-interest cost elasticity of deposits is statistically greater than zero for 90% of the sample.

Table A.1. Non-Interest Cost Elasticity of Deposits in the Loans-and-Savings MFIs

	Point Estimate	One-Sided Lower Bound
10th Percentile	0.046	0.007
1st Quartile	0.071	0.033
Median	0.098	0.060
3rd Quartile	0.125	0.084
90th Percentile	0.144	0.106
Mean	0.095	0.059

Reported are the estimates of $\partial \log \mathrm{C}/\partial \log y_2$ from the translog $\mathrm{C}(w_1, w_2, x_3, e, y_1, y_2)$ function with C being a sum of expenses on x_1 and x_2. The lower bounds are for the *one*-sided 95% percentile block-bootstrap confidence intervals, with the corresponding upper bounds being $+\infty$.

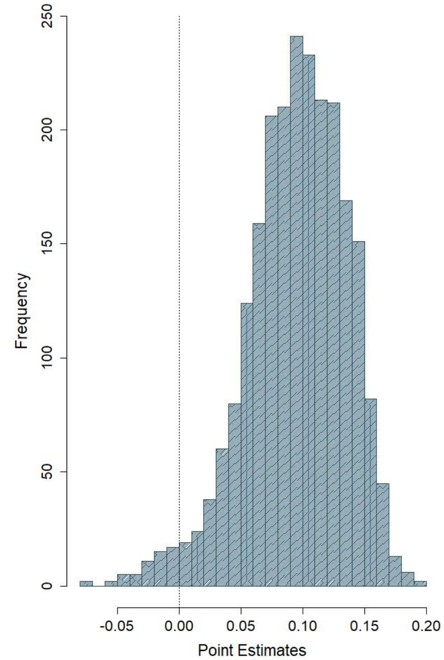

Figure A.1. Distribution of Non-Interest Cost Elasticity of Deposits in the Loans-and-Savings MFIs

B Monotonicity Regularity Conditions

To impose the monotonicity conditions on the cost functions, we rewrite the estimator in (3.8) as a weighted average of the outcome variable, i.e.,

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}}_r \\ \mathrm{vec}\{\widehat{\mathbf{B}}_r\} \end{bmatrix} = \sum_i \pi_i \Delta \log(C_{r,it}), \tag{B.1}$$

where $\pi_i = \left[\sum_i \mathbf{\Delta m}_{r,it} \mathbf{\Delta m}'_{r,it} \mathcal{D}_{r,it} \widehat{\mathcal{K}}_{r,it}\right]^{-1} \times \mathbf{\Delta m}_{r,it} \mathcal{D}_{r,it} \widehat{\mathcal{K}}_{r,it}$ is the data-dependent weight.

Following Hall & Huang (2001), we can rewrite (B.1) in a more general form as follows

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}}_r(\boldsymbol{p}) \\ \mathrm{vec}\{\widehat{\mathbf{B}}_r(\boldsymbol{p})\} \end{bmatrix} = N \sum_i p_i \pi_i \Delta \log(C_{r,it}), \tag{B.2}$$

where $\boldsymbol{p} = (p_1, \ldots, p_N)'$ is the sequence of additional weights such that $\sum_{i=1} p_i = 1$. Note that p_i equals $1/N$ (i.e., uniform weights) in the case of an *un*constrained estimator in (B.1).

When necessary, we can impose the monotonicity conditions by selecting $\boldsymbol{p}$ that minimizes the L_2-type metric $D(\boldsymbol{p}) = (1/N\mathbf{i}_N - \boldsymbol{p})'(1/N\mathbf{i}_N - \boldsymbol{p})$ subject to $\mathbf{i}'_N\boldsymbol{p} = 1$ and the monotonicity conditions that all can be written in the form of $\mathbf{R}\big[\widehat{\boldsymbol{\beta}}_r(\boldsymbol{p})', \mathrm{vec}\{\widehat{\mathbf{B}}_r(\boldsymbol{p})\}'\big]' > \mathbf{0}_N$, where $\mathbf{R}$ is the data-dependent design matrix, $\mathbf{i}_N$ is an $N \times 1$ vector of ones, and $\mathbf{0}_N$ is an $N \times 1$ vector of zeros. The objective function $D(\boldsymbol{p})$ is the sum of squared deviations of p_i from the unrestricted value of $1/N$. In our choice of the distance metric, we follow Du et al. (2013) over Hall & Huang (2001) because it allows $\boldsymbol{p}$ to be both positive and negative. The minimization problem is solved via a standard quadratic programming technique. Let $\widehat{\boldsymbol{p}}$ be the solution to this optimization problem. Then, the constrained estimator of unknown parameters is given by $\left(\widehat{\boldsymbol{\beta}}_r(\widehat{\boldsymbol{p}})', \text{vec}\{\widehat{\mathbf{B}}_r(\widehat{\boldsymbol{p}})\}'\right)'$ for $r = 1, 2$.

References

Ahn, H. & Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, *58*(1), 3–29.

Barnett, W. A. (2002). Tastes and technology: Curvature is not sufficient for regularity. *Journal of Econometrics*, *108*, 199–202.

Barnett, W. A., Geweke, J., & Wolfe, M. (1991). Seminonparametric Bayesian estimation of the Asymptotically Ideal Production Model. *Journal of Econometrics*, *49*, 5–50.

Barth, J. R. & Wihlborg, C. (2017). Too big to fail: Measures, remedies, and consequences for efficiency and stability. *Financial Markets, Institutions and Instruments*, *26*, 175–245.

Baumol, W., Panzar, J., & Willig, R. (1982). *Contestable Markets and the Theory of Industry Structure*. San Diego: Harcourt, Brace & Jovanovich.

Berger, A. N., Cummins, J. D., Weiss, M. A., & Zi, H. (2000). Conglomeration versus strategic focus: Evidence from the insurance industry. *Journal of Financial Intermediation*, *9*, 323–362.

Berger, A. N., Demsetz, R. S., & Strahan, P. E. (1999). The consolidation of the financial services industry: Causes, consequences and implications for the future. *Journal of Banking and Finance*, *23*, 135–194.

Berger, A. N., Hanweck, G. A., & Humphrey, D. B. (1987). Competitive viability in banking: Scale, scope, and product mix economies. *Journal of Monetary Economics*, *20*, 501–520.

Blevins, J. R. & Khan, S. (2013). Local NLLS estimation of semi-parametric binary choice models. *Econometrics Journal*, *16*, 135–160.

Braeutigam, R. R. & Daughety, A. F. (1983). On the estimation of returns to scale using variable cost functions. *Economics Letters*, *11*, 25–31.

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.

Caves, D. W., Christensen, L. R., & Swanson, J. A. (1981). Productivity growth, scale economies, and capacity utilization in US railroads, 1955-74. *American Economic Review*, *71*(5), 994–1002.

Charlier, E., Melenberg, B., & van Soest, A. (2001). An analysis of housing expenditure using semiparametric models and panel data. *Journal of Econometrics*, *101*(1), 71–107.

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.

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). *Portfolios of the Poor: How the World's Poor Live on $2 a Day*. Princeton: Princeton University Press.

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.

Du, P., Parmeter, C. F., & Racine, J. S. (2013). Nonparametric kernel regression with multiple predictors and multiple shape constraints. *Statistica Sinica*, *23*(3), 1347–1371.

Eicher, T. S. & Leukert, A. (2009). Institutions and economic performance: Endogeneity and parameter heterogeneity. *Journal of Money, Credit and Banking*, *41*, 197–219.

Elsas, R., Hackethal, A., & Holzhäuser, M. (2010). The anatomy of bank diversification. *Journal of Banking and Finance*, *34*, 1274–1287.

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.

Fuss, M. A. & Waverman, L. (2002). Econometric cost functions. In M. E. Cave, S. K. Majumdar, & I. Vogelsang (Eds.), *Handbook of Telecommunications Economics*, volume 1. Elsevier.

Hall, P. & Huang, L.-S. (2001). Nonparametric kernel regression subject to monotonicity constraints. *Annals of Statistics*, *29*(3), 624–647.

Hartarska, V., Mersland, R nad Nadolnyak, D., & Parmeter, C. (2013). Scale economies and elasticities of substitution in microfinance institutions. *International Journal of Corporate Governance*, *4*(1), 74–86.

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.

Heckman, J. J. & Serletis, A. (2014). Introduction to *Internally consistent modeling, aggregation, inference, and policy*. *Journal of Econometrics*, *183*, 1–4.

Heckman, J. J. & Serletis, A. (2015). Introduction to Econometrics with theory: A special issue honoring William A. Barnett. *Econometric Reviews*, *34*, 1–5.

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

Horowitz, J. L. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica*, *60*(3), 505–531.

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.

Hughes, J. P. & Mester, L. J. (2013). Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation*, *22*(4), 559–585.

Hughes, J. P. & Mester, L. J. (2015). Measuring the performance of banks: Theory, practice, evidence, and some policy implications. In A. Berger, P. Molyneux, & J. Wilson (Eds.), *Oxford Handbook of Banking* (2 ed.). (pp. 247–270). Oxford: Oxford University Press.

Hulme, D. (2008). The story of the Grameen Bank: From subsidised microcredit to market-based microfinance. Brooks World Poverty Institute Working Paper No. 60.

Khan, S. (2013). Distribution free estimation of heteroskedastic binary response models using Probit/Logit criterion function. *Journal of Econometrics*, *172*, 168–182.

Klein, R. W. & Spady, R. H. (1993). An efficienct semiparametric estimator for binary response models. *Econometrica*, *61*, 387–421.

Kyriazidou, E. (1997). Estimation of a panel data sample selection model. *Econometrica*, *65*(6), 1335–1364.

Lopatta, K. & Tchikov, M. (2016). Do microfinance institutions fulfill their promise? Evidence from cross-country data. *Applied Economics*, *48*(18), 1655–1677.

Malikov, E., Kumbhakar, S. C., & Sun, Y. (2016). Varying coefficient panel data model in the presence of endogenous selectivity and fixed effects. *Journal of Econometrics*, *190*(2), 233–251.

Malikov, E., Zhao, S., & Kumbhakar, S. C. (2016). Economies of diversification in the US credit union sector. *Journal of Applied Econometrics*. forthcoming.

Manski, C. F. (1987). Semiparametric analysis of random effects linear models from binary panel data. *Econometrica*, *55*(2), 357–362.

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

Microfinance Information Exchange (2015). Microfinance Bulletin No. 18.

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

Saunders, A. (2000). *Financial Institutions Management* (3 ed.). New York: McGraw-Hill.

Serletis, A. & Feng, G. (2015). Imposing theoretical regularity on flexible functional forms. *Econometric Reviews*, *34*, 198–227.

Wagenaar, K. (2012). Institutional transformation and mission drift in microfinance. University of Cambridge, Centre of Development Studies.

Woller, G. M., Dunford, C., & Woodworth, W. (1999). Where to microfinance? *International Journal of Economic Development*, *1*(1), 29–64.

Yasar, M., Morrison Paul, C. J., & Ward, M. R. (2011). Property rights institutions and firm performance: A cross-country analysis. *World Development*, *39*, 648–661.