A PANEL ANALYSIS OF MICROFINANCE EFFICIENCY MEASURES: EVIDENCE ON THE EFFECTS OF UNOBSERVED MANAGERIAL ABILITY

François Seck Fall, Hubert Tchakoute Tchuigoua, Anne Vanhems, Léopold Simar

LIDAM Discussion Paper ISBA
2024 / 20



UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

A panel analysis of microfinance efficiency measures: Evidence on the effects of unobserved managerial ability

F.S. Fall*
Francois-seck.Fall@ut-capitole.fr

H. Tchakoute Tchuigoua†
hubert.tchakoute@kedgebs.com

A. Vanhems‡
a.vanhems@tbs-education.fr

L. Simar§
leopold.simar@uclouvain.be

August 29, 2024

Abstract

This study analyzes the impact of unobserved heterogeneity on microfinance efficiency. Using panel data for 168 microfinance institutions (MFIs) over the period 2010_2015, we examine the persistence of the effect of unobserved heterogeneity on microfinance efficiency. Using recent nonparametric and robust techniques, we identify a latent heterogeneity factor related to the ability of MFI managers to promote efficiency, independent of MFI size, and analyze its impact on MFI inefficiency measures over time. We then assess the robustness of our results to several factors: the MFI status of the MFI (for-profit or nonprofit), the definition of the efficiency measure (social and financial) and an observed degree of heterogeneity captured by the percentage of women on the board. Finally, we analyze the relationship between our unobserved heterogeneity factor and institutional and socio-economic indicators.

Keywords: OR in developing countries, Microfinance, Efficiency, Unobserved heterogeneity, Panel data, Nonparametric robust frontier models

*Lereps, UT2J, Sciences Po Toulouse, UT1 capitole

†Kedge Business School

‡Corresponding author, TBS Business School, 1 place Alphonse Jourdain, 31068 Toulouse

§Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), LIDAM, UCLouvain, Voie du Roman Pays 20, B 1348 Louvain-la-Neuve, Belgium

1 Introduction

Around the world, microfinance aims to fight poverty and inequality while being profitable. This dual purpose has led to increased interest in analyzing the efficiency of the sector. The profound changes that microfinance has undergone in recent years and its important role in financial inclusion have brought efficiency analysis to the forefront of policymakers' and regulators' attention (Fall et al. (2018)). In particular, efficiency analysis can quantify gaps in production and resource allocation, identify best practices and sources of waste, and enable benchmarking microfinance institutions (MFIs). However, to effectively guide public and management policy on MFI efficiency, and to avoid bias and noise in the definition of the efficiency frontier, studies need to use robust methodologies and include factors that are not limited to those under the control of managers. In particular, when estimates are biased, it is difficult to explain differences in the efficiency of the firms under comparison (Berger and Mester (1997)), which can lead managers and policymakers to make incorrect decisions (Bos et al. (2009); Mester (1997)). Therefore, in this study, we examine the persistence of the effect of unobserved heterogeneity on efficiency frontiers using robust estimation techniques.

The literature on microfinance efficiency has focused on factors under the control of managers, benchmarking MFIs, or comparing how different types of MFIs allocate resources to achieve a given level of output, using either stochastic frontier analysis (SFA) or data envelopment analysis (DEA). While the impact of observed and unobserved heterogeneity on the microfinance frontier has been examined recently (e.g.Fall et al. (2021); Fall et al. (2023)), these studies are based on cross-sectional data, analyze MFI decision-making units (DMUs), and focus only on social outcomes, which is only one dimension of microfinance performance. Indeed, focusing only on the social component neglects the dual objective of MFIs, which is to alleviate poverty while being profitable. Estimating the efficiency frontier of MFIs while accounting for unobserved heterogeneity and time is a significant technical challenge that needs to be addressed. We attempt to address this challenge by analyzing the impact of unobserved heterogeneity on the overall efficiency of microfinance over time, including the social and financial dimensions, in terms of the level of the achievable efficiency frontier and inefficiency scores.

We build our model on the results of Fall et al. (2021) and Fall et al. (2023). In particular, Fall et al. (2021) consider the effect of observed heterogeneity, such as the percentage of women on the board, on both the frontier and the inefficiency level. Fall et al. (2023) investigate an additional effect of unobserved heterogeneity, which may be related to the ability of managers to reduce inefficiency and achieve higher levels of efficiency. In this paper, we analyze both the impact of unobserved heterogeneity

on MFI performance and the persistence of this impact over time, accounting for the potential observed heterogeneity effect.

Our study is in line with the microfinance literature that examines the efficiency of MFIs over time using either a stochastic frontier analysis (SFA) or data envelopment analysis (DEA) on panel data to investigate economies of scope or cost efficiency in microfinance (e.g., Cozarenco et al. (2022); Hartarska et al. (2014); Malikov and Hartarska (2018)), or efficiency gaps between private MFIs and microfinance NGOs over time using a pseudo-panel Malmquist index approach (Piot-Lepetit and Tchakoute Tchuigoua (2022)). These longitudinal studies allow for the possibility of a time effect, but consider unobserved heterogeneity as exogenous to the definition of the efficiency frontier. Moreover, our study accounts for the double bottom line objective of MFIs by aggregating social and financial outcomes and constructing an index that captures an MFI's effectiveness in reducing poverty while remaining profitable. In doing so, our study goes beyond single output-based studies (e.g., Fall et al. (2021), Fall et al. (2023); Piot-Lepetit and Tchakoute Tchuigoua (2022)).

Methodologically, our study lies at the intersection of two approaches to characterizing unobserved heterogeneity. Some studies on this topic are based on the approach of Demerjian et al. (2012), which quantifies unobserved heterogeneity to assess its impact on accounting and financial outcomes for financial and non-financial organizations (e.g., De Franco et al. (2017); Bui et al. (2018); Demerjian et al. (2013)). This approach attributes unobserved heterogeneity to managerial ability and is derived from the residuals of regression models that estimate efficiency scores based on firm characteristics using techniques such as DEA or SFA. However, these studies assume that unobserved heterogeneity is exogenous to the construction of the efficiency frontier. As in some studies, we build on Simar et al. (2016)and adopt an alternative hypothesis assuming that unobserved heterogeneity or unobservable characteristics of managers may affect how MFIs efficiently mobilize inputs to achieve a given level of output (Fall et al. (2023)). While this approach typically uses cross-sectional data, we use a longitudinal sample to examine whether unobserved heterogeneity allows MFIs to maintain high production efficiency over time. In particular, a longitudinal sample enables examining whether unobserved heterogeneity allows MFIs to maintain high production efficiency over time. Unobserved heterogeneity is quantified by the residual of the regression of our index combining social and financial aspects on size, following the approach of Racine (1997) and Daraio and Simar (2014).

We isolate the effects of unobserved heterogeneity from observed environmental factors and temporal trend effects captured by board gender diversity, using the recent methodology developed by Florens et al. (2014) and implemented in Mastromarco and

Simar (2021). By cleaning the inputs and outputs from the effects of common temporal trends and observed heterogeneity, our study goes beyond Fall et al. (2021) and Fall et al. (2023), who highlight a potential positive effect of observed heterogeneity on MFI efficiency. The application of this location-scale model allows us to construct both an unconditional and a conditional frontier using pure inputs and outputs, stripped of common temporal trends. The identification of the unobserved heterogeneity effect is based on Simar et al. (2016) and we use a robust conditional nonparametric method to assess the impact of this unobserved effect on both the frontier and the inefficiency level (for order-α conditional FDH methods, see Aragon et al. (2005), Daouia and Simar (2007), Daraio and Simar (2007a)). In a second step, we compare the conditional and unconditional bounds over time. Our results show that an increase in the value of unobserved heterogeneity is persistently associated with higher efficiency and lower inefficiency. This may reflect the ability of managers to use inputs effectively to achieve a given level of output.

The remainder of the paper is organized as follows. The next section reports the data and variables. We then present the methodology to assess the persistence of the unobserved heterogeneity effect on the inefficiency estimation in a nonparametric robust way, and we comment on the results obtained. The last section concludes by discussing the benefits of analyzing the effect of unobserved heterogeneity over time, with implications for the microfinance sector.

2 Data

2.1 Sampling

The data in this study come from a well-known database used in many microfinance studies (Fall et al. (2023); Goodell et al. (2020); Tadele et al. (2018)), namely the Microfinance Information Exchange (MIX). While MIX is not fully representative of the global microfinance market, it is currently one of the few available and accessible databases that allows for longitudinal studies. MIX contains more than 19,700 MFI-year observations over the period 1999–2019, with each MFI's annual data scored from 0 to 5 to classify MFIs according to their level of transparency (Goodell et al. (2020); Tadele et al. (2018)). Therefore, we removed MFIs with a transparency score of 0. In addition, we account for observed heterogeneity, which we capture using board gender diversity. Since data on the percentage of women on MFI boards are only available from 2007 onwards, we restricted our sample to the period 2007–2019. Finally, given that the Mastromarco and Simar (2018) model we use was developed for a dynamic framework,

specifically a balanced panel data structure, we balanced our data panels and retained the period for which we had the maximum number of observations, namely from 2010 to 2015. Therefore, our final sample consists of 1,008 MFI-year observations and includes 168 MFIs over the period 2010 -2015[1].

2.2 Variables

Inputs. As in previous works on MFIs efficiency (Fall (2018); Gutiérrez-Nieto et al. (2009); Piot-Lepetit and Nzongang (2014); Servin et al. (2012); Wijesiri et al. (2015); Wijesiri and Meoli (2015); Fall et al. (2021) and Fall et al. (2023)), we use physical inputs such as labor and costs: number of employees, which measures the number of personnel actively employed by the institution; operating expenses, including all personnel costs, depreciation, amortization, and administrative expenses expressed in US dollars.

Outputs. In this study, we account for the hybrid nature of MFIs, as they are decision-making units (DMUs) that combine a financial (profitability) and a social (poverty alleviation) logic, with two outputs. The first is financial output, measured by MFI income, which includes all financial and other operating income generated from non-financial services. While we could have assumed that the loan portfolio is the main source of MFI income by measuring financial output through loan income (Bassem (2014); Piot-Lepetit and Nzongang (2014); Piot-Lepetit and Tchakoute Tchuigoua (2022); Wijesiri (2016)), we assume that MFI income can also come from their nonfinancial activities.

The second is social output that we measure by the number of poor borrowers (Gutiérrez-Nieto et al. (2009); Piot-Lepetit and Tchakoute Tchuigoua (2022)), an indicator that combines both depth of outreach (average loan balance per borrower scaled by GNI per capita) and breadth of outreach (number of poor borrowers). Following Gutiérrez-Nieto et al. (2009) and Piot-Lepetit and Tchakoute Tchuigoua (2022), for a given year t, we first estimate the share of poor borrowers reached by a given MFI i as $p_i = (Depth_i - Min(Depth))/(Range(Depth))$ where $Min(Depth)$ is the minimum of depth of outreach for all DMUs i during the year t, and $Range(depth)$ is the difference between the minimum and maximum value of the depth of all DMUs i in the year t. A value close to 0 indicates that a given DMU is reaching the poorest borrowers. However, since we are using an output-maximizing approach, we subtract the previously calculated number from 1 to obtain a new index $p_i^0 = 1 - p_i$. A value close to 1 indicates that an MFI is meeting its goal of reaching the poor. Thus, we calculate the number of poor borrowers by multiplying p_i^0 by the number of

[1]We use classical multiple imputation to deal with 34 missing values for 21 MFIs, or about 0.4% of the total dataset.

active borrowers. To determine the number of poor borrowers, this p_i value is then multiplied by the breadth of outreach (number of active borrowers) of each DMU: Number of poor borrowers $= p_i^0$(Number of active borrowers $)_i$.

We use logarithm transformation for both inputs and outputs to reduce the weight of extreme values.

Additional external variables. We also control for environmental factors (observed heterogeneity) that are not under the control of managers and may affect an MFI's ability to efficiently achieve its double bottom line performance (Fall et al. (2021)). The environmental factor we control for is board gender diversity, measured by the percentage of female board members.

At last, in order to capture MFIs unobserved heterogeneity, we include an auxiliary variable that captures a global common feature of MFIs: the number of offices (in logarithm).

2.3 Inputs and outputs aggregation

Following Daraio and Simar (2007b) and Wilson (2018), we aggregate inputs and outputs to mitigate dimensionality issues. First, to account for the double bottom line objective of MFIs, we aggregate the social and financial outputs and construct an index that captures an MFI's effectiveness in reducing poverty while being profitable. Since the two outputs are correlated (66.34%), we aggregate both outputs and check that the correlation between the initial outputs and the aggregated output is strong enough (84.28% and 96.18%) as well as the percentage of inertia (99.61%).

Similarly, we aggregate the two inputs, namely the number of employees and operating expenses, which are also strongly correlated (83.82%). The aggregated input is again strongly correlated with the initial inputs (90.88% and 98.93%) with a percentage of inertia of 99.57%.

Aggregating inputs and outputs then allows reducing the curse of dimensionality and also visualizing frontiers and efficiencies on 2-D plots.

3 Methodology

We combine several recent nonparametric methods to account for the panel structure of our dataset and to deal with the time dimension and cross-sectional dependence. This is critical for analyzing productivity over time and across different entities, while accounting for spatial dependence and heteroscedasticity. Following Florens et al. (2014), we use nonparametric location-scale models to address the issues of model specification

uncertainty, potential endogeneity, and cross-sectional dependence in modeling technical efficiency. This approach allows for a more flexible modeling structure that is less sensitive to outliers, the inclusion of time factors, and other observed environmental variables in the modeling process.

In the context of nonparametric efficiency analysis, the location-scale model approach can be described in two steps (Mastromarco and Simar (2018)). In a first step, inputs and outputs are cleaned of temporal dynamics and observed environmental factors, allowing for a more nuanced understanding of how inputs are transformed into outputs over time and under different conditions. In a second step, the cleaned inputs and outputs are used to estimate nonparametric, robust efficiency measures that are less sensitive to outliers. There are several methodological advantages to this combination of recent nonparametric techniques. First, robustness to model specification: unlike parametric models, which require the specification of a particular functional form (which may be incorrect), nonparametric models are more flexible and can adapt to the actual shape of the data, making them robust to model misspecification. Second, dealing with endogeneity and cross-sectional dependence: the flexibility of the model allows endogeneity to be taken into account, ensuring that the estimates of the relationship between inputs and outputs are not biased by omitted variables or measurement error. It can also handle cross-sectional dependence, which is important when dealing with data from multiple entities (such as firms or countries) where unobserved factors may affect all entities simultaneously. The inclusion of time and environmental factors allows for a more nuanced understanding of how inputs are transformed into outputs over time and under different conditions. Third, robustness to outliers: the nonparametric approach is typically less sensitive to outliers than parametric methods, making the analysis more robust and reliable, especially when the data contain extreme values or anomalies. Finally, unobserved heterogeneity: the model can be extended to account for unobserved heterogeneity, as in Mastromarco and Simar (2021).

3.1 Trend patterns and observed heterogeneity

Consider an input $X \in \mathbb{R}_+$, an output $Y \in \mathbb{R}_+$ and an environmental variable $Z \in \mathbb{R}$. Our sample is denoted by (X_{it}, Y_{it}, Z_{it}), where $i = 1, ..., n$ is the MFI index and $t = 1, ..., s$ is the time index. As proposed in Pesaran (2006) and Bai (2009), we consider $F_t = (t, X_{.t}, Y_{.t})$ as a proxy for the unobserved trend patterns and the common time factor, where a dot in a subscript indicates that we have averaged over that index. We

assume that the data are generated by a nonparametric location-scale regression model:

$$\begin{cases} X_{it} & = \mu_x(Z_{it}, F_t) + \sigma_x(Z_{it}, F_t)\epsilon_{x,it} \\ Y_{it} & = \mu_y(Z_{it}, F_t) + \sigma_y(Z_{it}, F_t)\epsilon_{y,it}, \end{cases}$$

where ϵ_x and ϵ_y have a mean of 0 and standard deviation of 1 and (ϵ_x, ϵ_y) is independent of (Z, F_t). Therefore, ϵ_x and ϵ_y are cleaned versions of the input and output where the dependence on common time effects (captured by F_t) and environmental factor Z have been removed. Note that this cleaning step allows reducing the curse of dimensionality of (Z, F_t) in the frontier estimation step. Following the two step estimation procedure of Florens et al. (2014), we derive estimated versions of μ_x, μ_y, σ_x and σ_y using nonparametric techniques (local constant and local linear). The estimated counterparts $\hat{\mu}_x$, $\hat{\mu}_y$, $\hat{\sigma}_x$ and $\hat{\sigma}_y$ allow us to define the cleaned estimated input and output as follows:

$$\begin{cases} \hat{\epsilon}_{x,it} & = \frac{X_{it} - \hat{\mu}_x(Z_{it}, F_t)}{\hat{\sigma}_x(Z_{it}, F_t)} \\ \hat{\epsilon}_{y,it} & = \frac{Y_{it} - \hat{\mu}_y(Z_{it}, F_t)}{\hat{\sigma}_y(Z_{it}, F_t)}, \end{cases}$$

which are then be used to derive inefficiency measures cleaned of both the time and observed heterogeneity effects.

3.2 Unobserved heterogeneity

In addition to the observed heterogeneity factor Z, we also capture the unobserved heterogeneity effect V, which may affect the inefficiency measure but is not directly observable by the econometrician. Fall et al. (2023) highlight a latent heterogeneity factor that assesses the ability of MFI managers to implement credit policies in favor of women's inclusion and empowerment, and is related to the output variable Y. Following their analysis, we define unobserved heterogeneity V as the residual from the nonparametric non-separable regression $Y = \varphi(W, V)$ of output Y on an auxiliary variable W that captures the size effect (such as the logarithm of offices). Assuming that W is independent of V, that φ is monotonic with respect to V, and that V is uniformly distributed on $[0, 1]$, this unobserved heterogeneity term can be identified as $V = F_{Y|W}$. Given a sample of observations (W_{it}, Y_{it}), a nonparametric counterpart $\hat{V}_{it}$ can be derived as $\hat{V}_{it} = \hat{F}_{Y|W}(Y_i|W_i)$, where $\hat{F}_{Y|W}$ is the nonparametric counterpart of the conditional distribution of Y given W. Then, similar to the input and output variables, the unobserved heterogeneity term can be cleaned of its dependence on common time effects F_t and the environmental factor Z using a location scale model to obtain:

$$\hat{\epsilon}_{v,it} = \frac{\hat{V}_{it} - \hat{\mu}_v(Z_{it}, F_t)}{\hat{\sigma}_v(Z_{it}, F_t)},$$

where $\hat{\mu}_v(Z_{it}, F_t)$ and $\hat{\sigma}_v(Z_{it}, F_t)$ are the estimated location and scale functions from the location scale regression: $V_{it} = \mu_v(Z_{it}, F_t) + \sigma_v(Z_{it}, F_t)\epsilon_{v,it}$.

3.3 Robust conditional efficiency measure

After this cleaning step, the initial dataset $(X_{it}, Y_{it}, Z_{it})_{i,t}$ is transformed into $(\hat{\epsilon}_{x,it}, \hat{\epsilon}_{y,it}, \hat{\epsilon}_{v,it})_{i,t}$ to estimate the robust conditional inefficiency measures. The main advantage is that these measures are cleansed of the effect of (F, Z) and thus allow a fairer comparison of the performance of each MFI. for calculating both the unconditional and conditional robust efficiency measures using $\hat{\epsilon}_{v,it}$ as the conditioning environmental factor and the order-α method to account for potential extreme values. Following Cazals et al. (2002), the unconditional frontier function can be defined as:

$$\tau(e_x) = \sup\{e_y | S_{\epsilon_y|\epsilon_x}(e_y|\epsilon_x \leq e_x) > 0\},$$

where function S is the survival function defined by $S_{\epsilon_y|\epsilon_x}(e_y|\epsilon_x \leq e_x) = P(\epsilon_y \geq e_y|\epsilon_x \leq e_x)$. The associated unconditional pure inefficiency measure for a unit with values (e_x, e_y) is given by

$$\rho(e_x, e_y) = \tau(e_x) - e_y.$$

Similarly, the conditional frontier function is defined as (see Mastromarco and Simar (2021)):

$$\tau(e_x|e_v) = \sup\{e_y | S_{\epsilon_y|\epsilon_x,\epsilon_v}(e_y|\epsilon_x \leq e_x, \epsilon_v = e_v) > 0\},$$

where the associated conditional pure inefficiency measure for a unit with values (e_x, e_y) and conditional on the pure latent factor with value e_v is given by:

$$\rho(e_x, e_y|e_v) = \tau(e_x|e_v) - e_y.$$

Since some extreme values in the data may confound the evaluation of the effect of ϵ_v on the frontier and the inefficiency measure, we also compute robust versions of order-α for both the frontier and the inefficiency measure (see Aragon et al. (2005)) as follows:

$$\begin{aligned} \tau_\alpha(e_x) &= \sup\{e_y | S_{\epsilon_y|\epsilon_x}(e_y|\epsilon_x \leq e_x) > 1-\alpha\}, \\ \tau_\alpha(e_x|e_v) &= \sup\{e_y | S_{\epsilon_y|\epsilon_x,\epsilon_v}(e_y|\epsilon_x \leq e_x, \epsilon_v = e_v) > 1-\alpha\}, \end{aligned}$$

where the robust conditional and unconditional inefficiency measures are given by $\rho_\alpha(e_x, e_y|e_v) = \tau_\alpha(e_x|e_v) - e_y$ and $\rho_\alpha(e_x, e_y) = \tau_\alpha(e_x) - e_y$. As in Fall et al. (2021) and Fall et al. (2023), we calibrate the tuning parameter α to leave a fixed percentage of points (about 8%) above the frontier, and the robust conditional and unconditional inefficiency measures are given by $\rho_\alpha(e_x, e_y|e_v)$ and $\rho_\alpha(e_x, e_y)$. The nonparametric counterparts

of $\rho_\alpha(e_x, e_y|e_v)$ and $\rho_\alpha(e_x, e_y)$ are obtained by plugging nonparametric estimators of $S_{\epsilon_y|\epsilon_x,\epsilon_v}$ and $S_{\epsilon_y|\epsilon_x}$ into the previous formulas. As recalled in Mastromarco and Simar (2021), the estimator of the survival function $S_{\epsilon_y|\epsilon_x}$ is simply an empirical counterpart, whereas the estimator of $S_{\epsilon_y|\epsilon_x,\epsilon_v}$ involves smoothing in the variable ϵ_v and requires bandwidth selection (for more details, see Simar et al. (2016) and Daraio and Simar (2014)).

3.4 Impact of the unobserved heterogeneity and time effect

The analysis of the effect of the unobserved heterogeneity factor ϵ_v is captured by analyzing the difference between the unconditional and conditional inefficiency measures $\rho_\alpha(e_x, e_y|e_v) - \rho_\alpha(e_x, e_y)$ and the conditional inefficiency level $\rho_\alpha(e_x, e_y|e_v)$. We perform a nonparametric local linear regression of both inefficiency differences and inefficiency levels on ϵ_v. A positive effect of ϵ_v on the inefficiency difference can be interpreted as a shift above the frontier, and a negative effect of ϵ_v on the inefficiency level as a decrease in inefficiency.

Note that we can also recover the conditional inefficiency measures in the original inputs-outputs as follows:

$$\begin{aligned} \delta_\alpha(x, y|f, z) &= \sigma_y(f, z)\rho_\alpha(e_x, e_y), \\ \delta_\alpha(x, y|f, z, v) &= \sigma_y(f, z)\rho_\alpha(e_x, e_y|e_v). \end{aligned}$$

This last calculation also allows analyzing the impact of (F, Z) and unobserved heterogeneity V in the original framework.

4 Results and discussion

In this section, we present and discuss the main results on the effects of unobserved heterogeneity and time on the achievable frontier and on inefficiency. In particular, the effects are observed at the global frontier (of the full sample) and at the level of MFI subsamples by status. Therefore, we attempt to identify the factors influencing unobserved heterogeneity using macroeconomic variables related to good governance. We first analyze the effect of the unobserved heterogeneity variable after cleaning to remove the global time effect and the observed heterogeneity effect. In a second step, we examine the persistence of the effect of unobserved heterogeneity over time on the raw dataset. In both settings (cleaned vs raw) and for both the full sample and the subsamples, we first examine the effect of the unobserved heterogeneity variable on the shape of the production set. This effect is captured by analyzing the relationship between inefficiencies (conditional and unconditional) and the unobserved factor. If the relationship

is positive and significant, it means that accounting for the unobserved heterogeneity factor allows for higher levels of efficiency and shifts the achievable frontier upwards. We then examine the effect of this unobserved heterogeneity factor on conditional managerial efficiency. This second effect is captured by analyzing the relationship between the conditional measure of inefficiency and the unobserved heterogeneity factor. If the relationship is negative, it means that accounting for the environmental factor allows reducing the level of inefficiency.

In both cases, we compute robust estimators of order-α to be able to apply the double bootstrap methodology developed by Racine (1997) and to test the significance of the association. In all figures, the value of the parameter α is set so that about 8% of the units are above the unconditional frontier, and the pvalue is computed from the double bootstrap procedure following Daraio and Simar (2014). Since all pvalues are less than 5%, all derivatives of the local linear approximation with respect to unobserved heterogeneity are significantly different from 0, and we can interpret the shape of the local linear regression.

4.1 The effect of unobserved heterogeneity on the frontier and level of inefficiency in the full sample

First, we consider the setting where the inputs, outputs, and unobserved heterogeneity are cleaned from the global time effect and observed heterogeneity (percentage of women on the board).

In Figure 1, we observe a positive effect of the cleaned unobserved heterogeneity on the efficiency frontier (top left) and a negative effect on the cleaned microfinance inefficiency scores (top right). These two effects are estimated from pure inputs and outputs. Thus, by considering pure inputs and outputs, we find that the level of the achievable frontier increases with the level of unobserved heterogeneity. Accounting for unobserved heterogeneity in the frontier estimation allows MFIs to achieve higher levels of efficiency. Similarly, accounting for unobserved heterogeneity reduces the level of microfinance inefficiency of microfinance, although the effect is less pronounced.

Unobserved heterogeneity is a latent variable that reflects managerial ability or discretionary power (Fall et al. (2023)). Managerial ability allows microfinance to achieve higher levels of efficiency and thus to push the efficiency frontier. It also allows microfinance to reduce inefficiencies. Thus, this result confirms the conclusion of the previous analysis and shows the robustness of the effect of unobserved heterogeneity once time patterns and observed heterogeneity are taken into account.

Second, we return to the original inputs, outputs and unobserved heterogeneity (i.e.,

not cleaned of time and observed heterogeneity). Figure 2 shows that the overall effect of unobserved heterogeneity on the frontier is positive (both figures at the top). The effect is similar for all years, from 2010 to 2015. Therefore, as in the cleaned dataset, accounting for unobserved heterogeneity allows MFIs to achieve higher levels of efficiency, and this effect is persistent over the observation period.

The two figures at the bottom of Figure 2 show the impact of unobserved heterogeneity on microfinance inefficiency over time . We find a negative impact of unobserved heterogeneity on MFI inefficiency, which is consistent over the 6 years. Accounting for unobserved heterogeneity allows MFIs to reduce their level of inefficiency, and again this effect is persistent over the period.

4.2 Examining unobserved heterogeneous effects bt MFI profit status

Figures 3 and 4 zoom in on the MFI subsamples by profit status.

For both subsamples, we find a positive effect of unobserved heterogeneity on the efficiency frontier and a negative effect on microfinance inefficiency scores. These two effects are observed for both the cleaned and the raw data. The frontier estimate increases with the level of unobserved heterogeneity for commercial and non-commercial MFIs. Again, accounting for unobserved heterogeneity allows MFIs with either for-profit status or nonprofit status to achieve higher levels of efficiency. Similarly, the inefficiency scores decrease with the level of unobserved heterogeneity for both nonprofit and for-profit MFIs. Thus, accounting for unobserved heterogeneity reduces the overall inefficiency, regardless of the profit status of MFIs. Accounting for managerial discretionary allows for higher efficiency frontiers and also reduces the microfinance inefficiency scores.

4.3 Interpreting unobserved heterogeneity

Our goal is to analyze the relationship between our unobserved heterogeneity variable and 6 observed macroeconomic factors in order to enrich the analysis of this variable. Table 2 presents some descriptive statistics and Table 3 provides the correlation matrix for these 6 factors. To simplify, we perform principal components analysis (PCA) to summarize the information provided by these factors. The results of this PCA are presented in Tables 4 and 5. We mainly consider the first two principal components that explain more than 70% of the total variance. The first principal component (PC1) is highly positively correlated with 5 out of 6 of the macro variables and reflects the quality of country governance. The second component (PC2) is highly positively correlated with the remaining factor and thus reflects country political stability.

Figure 5 shows that unobserved heterogeneity is positively correlated with both PC1

and PC2. Thus, the higher the political stability, the more important the discretionary power of managers. Similarly, the better the country's governance quality, the more important the discretionary power of managers.

5 Conclusion

Our study addresses the observed gap on the effect of unobserved heterogeneity on MFI efficiency over time. To fill this gap, we examine a longitudinal sample of MFIs consisting of 1008 observations on 168 MFIs over the period 2010 to 2015. Using a location-scale model (Florens et al. (2014); Mastromarco and Simar (2018)), we first cleaned the inputs and outputs from both common time trends and observed heterogeneity effects. In our study, observed heterogeneity is captured by board gender diversity. Our results show that unobserved heterogeneity has a positive effect on the shape of the conditional production set and on the conditional inefficiency level. This suggests that as the level of unobserved heterogeneity increases, the conditional frontier level may increase, and the conditional inefficiency level may decrease. Moreover, by accounting for time effects, we consistently find a positive effect of unobserved heterogeneity on both conditional efficiency and conditional inefficiency over time. This unobserved heterogeneity may reflect managerial ability and other unobserved attributes. Our results have important implications for the microfinance sector. In particular, they suggest that policies and practices that focus on improving managerial skills and other unobserved factors can lead to substantial improvements in MFI efficiency, political stability, and the quality of country governance.

References

Aragon, Y., Daouia, A., and Thomas-Agnan, C. (2005). Nonparametric frontier estimation: a conditional quantile-based approach. *Econometric Theory*, pages 358–389.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Bassem, B. S. (2014). Total factor productivity change of mena microfinance institutions: A malmquist productivity index approach. *Economic Modelling*, 39:182–189.

Berger, A. N. and Mester, L. J. (1997). Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of banking & finance*, 21(7):895–947.

Bos, J. W., Koetter, M., Kolari, J. W., and Kool, C. J. (2009). Effects of heterogeneity on bank efficiency scores. *European Journal of Operational Research*, 195(1):251–261.

Bui, D. G., Chen, Y.-S., Hasan, I., and Lin, C.-Y. (2018). Can lenders discern managerial ability from luck? evidence from bank loan contracts. *Journal of Banking & Finance*, 87:187–201.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1):1–25.

Cozarenco, A., Hartarska, V., and Szafarz, A. (2022). Subsidies to microfinance institutions: how do they affect cost efficiency and mission drift? *Applied Economics*, 54(44):5099–5132.

Daouia, A. and Simar, L. (2007). Nonparametric efficiency analysis: a multivariate conditional quantile approach. *Journal of Econometrics*, 140(2):375–400.

Daraio, C. and Simar, L. (2007a). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Springer Science & Business Media.

Daraio, C. and Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28(1-2):13–32.

Daraio, C. and Simar, L. (2014). Directional distances and their robust versions: Computational and testing issues. *European Journal of Operational Research*, 237(1):358–369.

De Franco, G., Hope, O.-K., and Lu, H. (2017). Managerial ability and bank-loan pricing. *Journal of Business Finance & Accounting*, 44(9-10):1315–1337.

Demerjian, P., Lev, B., and McVay, S. (2012). Quantifying managerial ability: A new measure and validity tests. *Management science*, 58(7):1229–1248.

Demerjian, P. R., Lev, B., Lewis, M. F., and McVay, S. E. (2013). Managerial ability and earnings quality. *The accounting review*, 88(2):463–498.

Fall, F. (2018). L'efficience technique des institutions de microfinance en zone UEMOA. *Revue d'économie politique*, 128(4):667–689.

Fall, F., Akim, A.-m., and Wassongma, H. (2018). DEA and SFA research on the efficiency of microfinance institutions: A meta-analysis. *World Development*, 107:176–188.

Fall, F. S., Tchakoute Tchuigoua, H., Vanhems, A., and Simar, L. (2023). Investigating the unobserved heterogeneity effect on outreach to women: lessons from microfinance institutions. *Annals of Operations Research*, pages 1–22.

Fall, F. S., Tchuigoua, H. T., Vanhems, A., and Simar, L. (2021). Gender effect on microfinance social efficiency: A robust nonparametric approach. *European Journal of Operational Research*, 295(2):744–757.

Florens, J.-P., Simar, L., and Van Keilegom, I. (2014). Frontier estimation in nonparametric location-scale models. *Journal of Econometrics*, 178:456–470.

Goodell, J. W., Goyal, A., and Hasan, I. (2020). Comparing financial transparency between for-profit and nonprofit suppliers of public goods: Evidence from microfinance. *Journal of International Financial Markets, Institutions and Money*, 64:101146.

Gutiérrez-Nieto, B., Serrano-Cinca, C., and Mar Molinero, C. (2009). Social efficiency in microfinance institutions. *Journal of the Operational Research Society*, 60(1):104–119.

Hartarska, V., Nadolnyak, D., and Mersland, R. (2014). Are women better bankers to the poor? evidence from rural microfinance institutions. *American Journal of Agricultural Economics*, 96(5):1291–1306.

Malikov, E. and Hartarska, V. (2018). Endogenous scope economies in microfinance institutions. *Journal of Banking & Finance*, 93:162–182.

Mastromarco, C. and Simar, L. (2018). Globalization and productivity: A robust nonparametric world frontier analysis. *Economic Modelling*, 69:134–149.

Mastromarco, C. and Simar, L. (2021). Latent heterogeneity to evaluate the effect of human capital on world technology frontier. *Journal of Productivity Analysis*, 55:71–89.

Mester, L. J. (1997). Measuring efficiency at US banks: Accounting for heterogeneity is important. *European Journal of Operational Research*, 98(2):230–242.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Piot-Lepetit, I. and Nzongang, J. (2014). Financial sustainability and poverty outreach within a network of village banks in cameroon: A multi-DEA approach. *European Journal of Operational Research*, 234(1):319–330.

Piot-Lepetit, I. and Tchakoute Tchuigoua, H. (2022). Ownership and performance of microfinance institutions in latin america: A pseudo-panel malmquist index approach. *Journal of the Operational Research Society*, 73(5):1167–1180.

Racine, J. (1997). Consistent significance testing for nonparametric regression. *Journal of Business & Economic Statistics*, 15(3):369–378.

Servin, R., Lensink, R., and Van den Berg, M. (2012). Ownership and technical efficiency of microfinance institutions: Empirical evidence from latin america. *Journal of Banking & Finance*, 36(7):2136–2144.

Simar, L., Vanhems, A., and Van Keilegom, I. (2016). Unobserved heterogeneity and endogeneity in nonparametric frontier estimation. *Journal of Econometrics*, 190(2):360–373.

Tadele, H., Roberts, H., and Whiting, R. H. (2018). Microfinance institutions' transparency in sub-saharan africa. *Applied Economics*, 50(14):1601–1616.

Wijesiri, M. (2016). Weathering the storm: ownership structure and performance of microfinance institutions in the wake of the global financial crisis. *Economic Modelling*, 57:238–247.

Wijesiri, M. and Meoli, M. (2015). Productivity change of microfinance institutions in kenya: A bootstrap malmquist approach. *Journal of Retailing and Consumer Services*, 25:115–121.

Wijesiri, M., Viganò, L., and Meoli, M. (2015). Efficiency of microfinance institutions in sri lanka: a two-stage double bootstrap DEA approach. *Economic Modelling*, 47:74–83.

Wilson, P. W. (2018). Dimension reduction in nonparametric models of production. *European Journal of Operational Research*, 267(1):349–367.

6 Appendix: Tables

	MFI institutional forms			
	Shareholder or for-profit MFIs	**Nonprofit MFIs** Credit Unions/ Cooperatives	NGO	**Total**
Total	486	120	402	1008

Table 1: MFI Distribution by Institutional Form.

	Mean	Standard Deviation	Minimum	Maximum
Input 1: Operating expenses	10,424,369	35,860,815	43,705	570,000,000
Input 2: Number of employees	755	1,996	1	21,719
Output 1: Number of poor borrowers	133,651	464,680	1	5,373,375
Output 2: Financial revenues	21,566,969	72,560,170	103,015	1,000,000,000
% of female board members	67%	24%	0%	100%
Offices	73	200	1	2,647
Control of corruption	-0.63	0.37	-1.64	0.79
Government effectiveness	-0.46	0.41	-2.04	0.61
Political stability	-0.65	0.67	-2.81	0.79
Regulatory quality	-0.40	0.51	-1.54	0.93
Rule of law	-0.65	0.44	-1.90	0.56
Voice and accountability	-0.24	0.51	-1.68	1.15

Table 2: Descriptive Evidence.

7 Frontier and inefficiency effects

7.1 Full sample

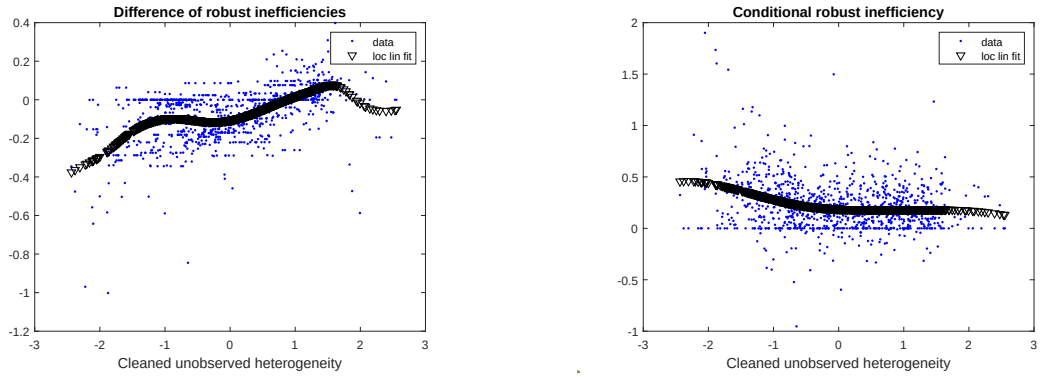


Figure 1: Cleaned unobserved heterogeneity effect for the full sample for cleaned inputs, output and unobserved heterogeneity. The two figures correspond to the effect of the cleaned unobserved heterogeneity on the frontier and the efficiency measures computed from the cleaned input and output.

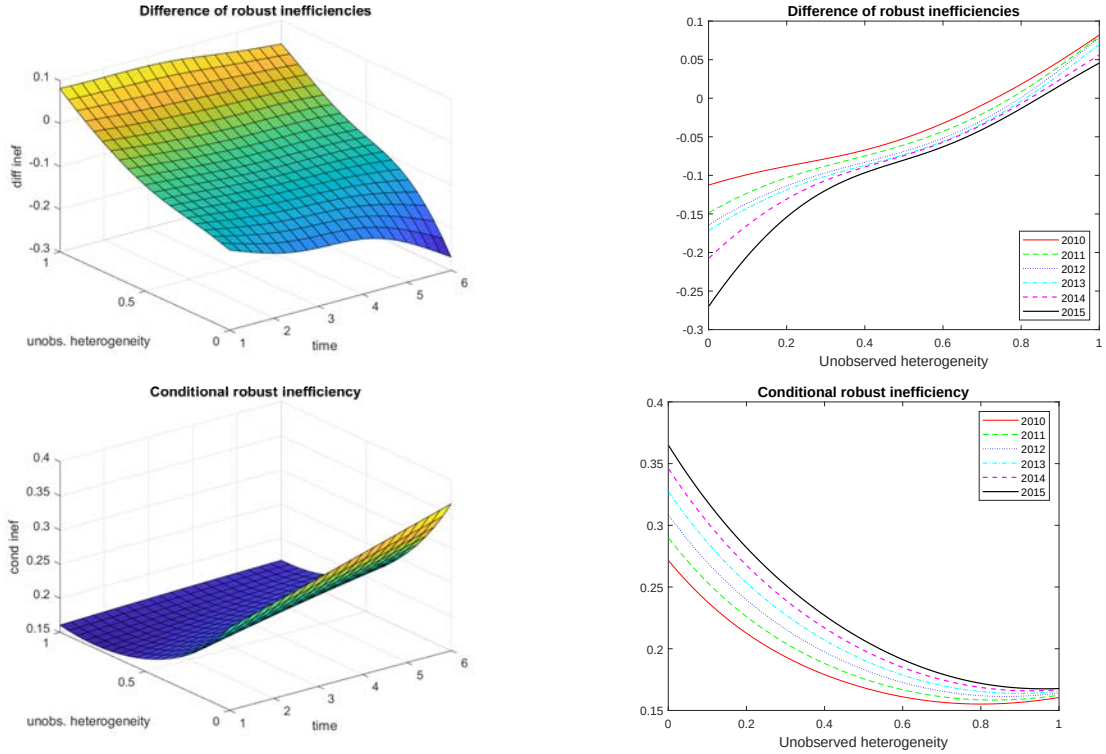


Figure 2: Unobserved heterogeneity effect for the full sample. The four figures correspond to the effect of unobserved heterogeneity on the frontier and inefficiency level computed from the raw input and output. The two top figures show the effect of unobserved heterogeneity on the frontier for the 6 years, and the two bottom figures the effect of unobserved heterogeneity on the inefficiency level for the 6 years.

7.2 Zoom on subsamples 'for-profit' and 'non-profit'

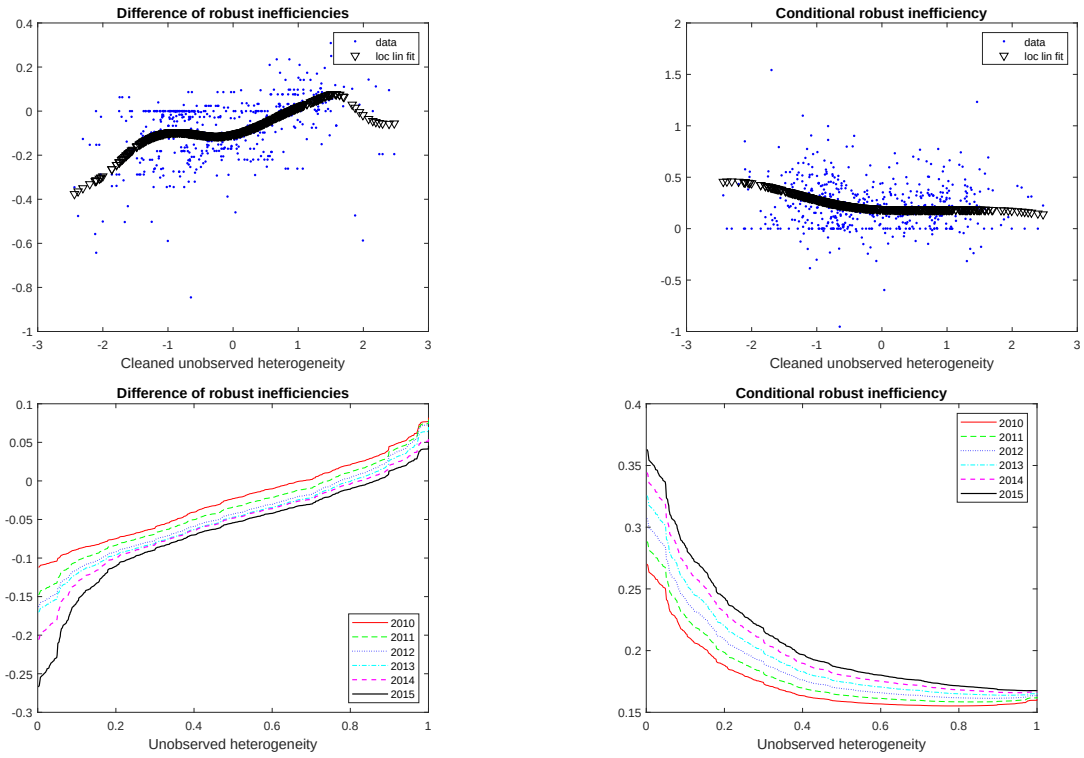


Figure 3: Zooming in on the nonprofit subsample to analyze the unobserved heterogeneity effect. The two top figures correspond to the effect of cleaned unobserved heterogeneity on the cleaned frontier and cleaned inefficiency. The two bottom figures correspond to the effect of raw unobserved heterogeneity on the raw frontier and inefficiency for the 6 years.

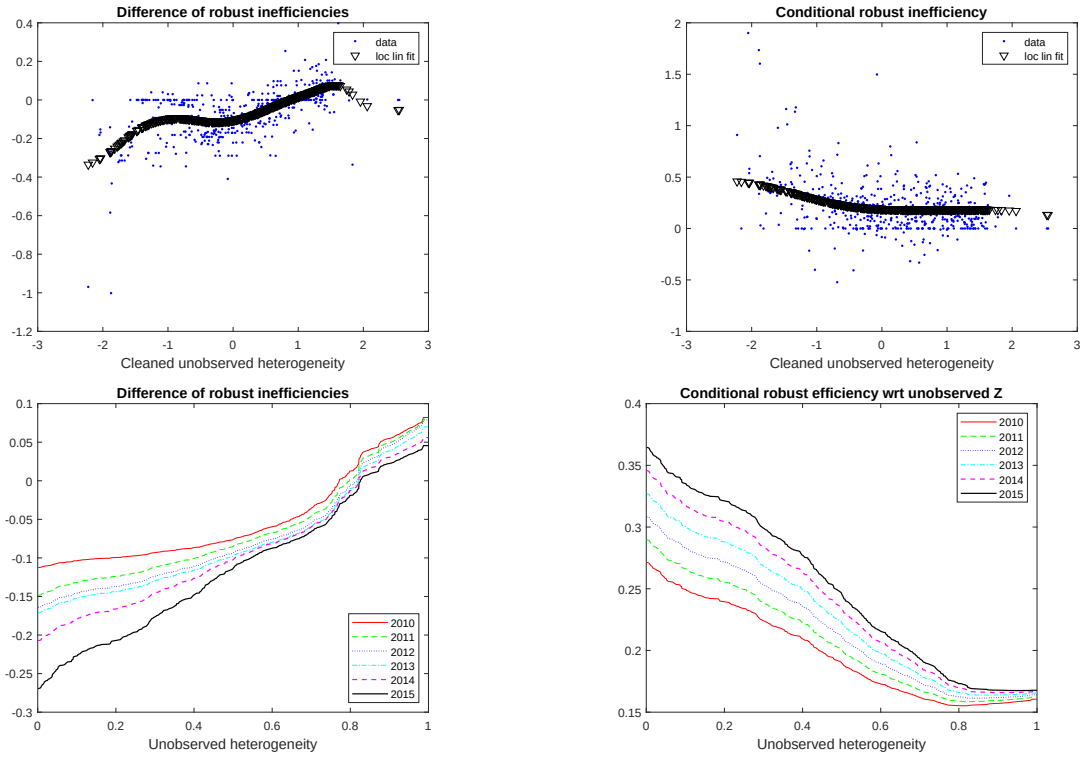


Figure 4: Zooming in on the for-profit subsample to analyze the unobserved heterogeneity effect. The two top figures correspond to the effect of cleaned unobserved heterogeneity on the cleaned frontier and inefficiency. The two bottom figures correspond to the effect of the raw unobserved heterogeneity on raw frontier and inefficiency for the 6 years.

8 Analysis of unobserved heterogeneity

	Control of corruption	Government effectiveness	Political stability	Regulatory quality	Rule of law	Voice and accountability
Control of corruption	1					
Government effectiveness	0.78	1				
Political stability	0.39	0.26	1			
Regulatory quality	0.64	0.66	0.29	1		
Rule of law	0.76	0.77	0.18	0.71	1	
Voice and accountability	0.69	0.60	0.30	0.40	0.62	1

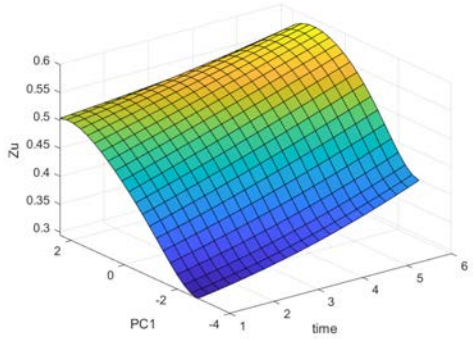
Table 3: Correlation Matrix between Macro Variables.

	PC1	PC2	PC3	PC4	PC5	PC6
% of Variance	55.82	17.59	12.32	6.00	4.44	3.8

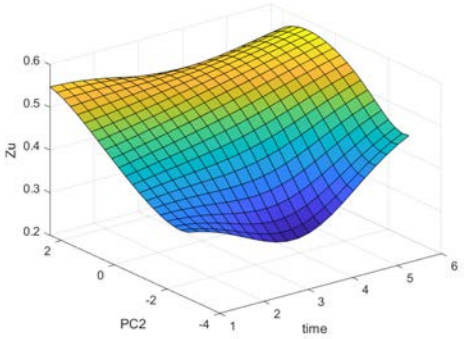
Table 4: Percentage of Variance of Principal Components.

	PC1	PC2
Control of corruption	0.86	0.10
Government effectiveness	0.87	-0.06
Political stability	0.12	0.98
Regulatory quality	0.73	0.09
Rule of law	0.85	-0.25
Voice and accountability	0.73	-0.01

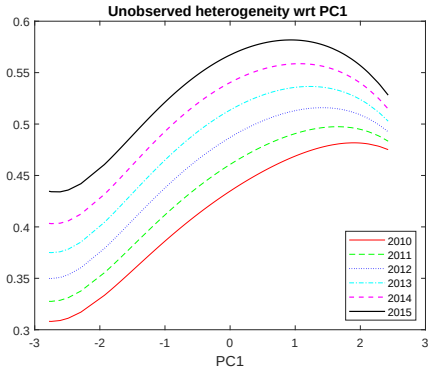
Table 5: Correlations between the First Two Principal Components and the Macro Variables.



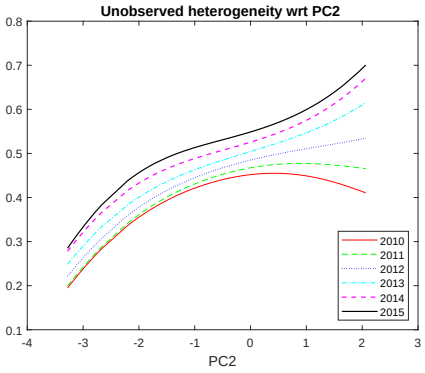
Unobs heterogeneity wrt PC1 and time



Unobs heterogeneity wrt PC2 and time



Unobs heterogeneity wrt PC1



Unobs heterogeneity wrt PC2

Figure 5: .