



Replication, Reproduction and the Credibility of Micro-econometric Studies of the Impact of Microfinance and Informal Sector Borrowing in Bangladesh

Maren Duvendack^{1, 2}

Richard Palmer-Jones¹

13 January 2012

1. UEA, 2. ODI

Introduction : Replication

- Replication is important to assess the robustness of findings, they may be weaker than they appear
 - Not guaranteed by peer review process
 - Nor is robustness guaranteed by academic status of authors
- Many good reasons to conduct replications
 - To fully understand what the authors have done
 - To check for errors
 - Of data cleaning and variable construction
 - Of estimation processes (specification or data mining....)

Replication continued

- ❑ To use other operationalisation of concepts into variables to see if this makes a difference to findings
- ❑ To apply different estimation methods including different software
- ❑ To check against different comparable data sets!!!
 - Including different but relevant contexts for external validity

Link replication to sub-group analysis

- Importance of exploring impact heterogeneity:
 - Crucial for drawing appropriate policy conclusions, confounds causal effects because results only hold for the groups identified, and not others
 - Neglecting impact heterogeneity misleading by inappropriately merging groups which respond quite differently to the treatment and indeed experience different treatments
 - Sub-groups known feature of MF context
 - Replication should seek to know the conditions under which the results hold, often by repeating the experiments
 - Thus, heterogeneity and replication are closely linked. We want to know what works, for whom, under what circumstances.

Introduction: Microfinance evaluations

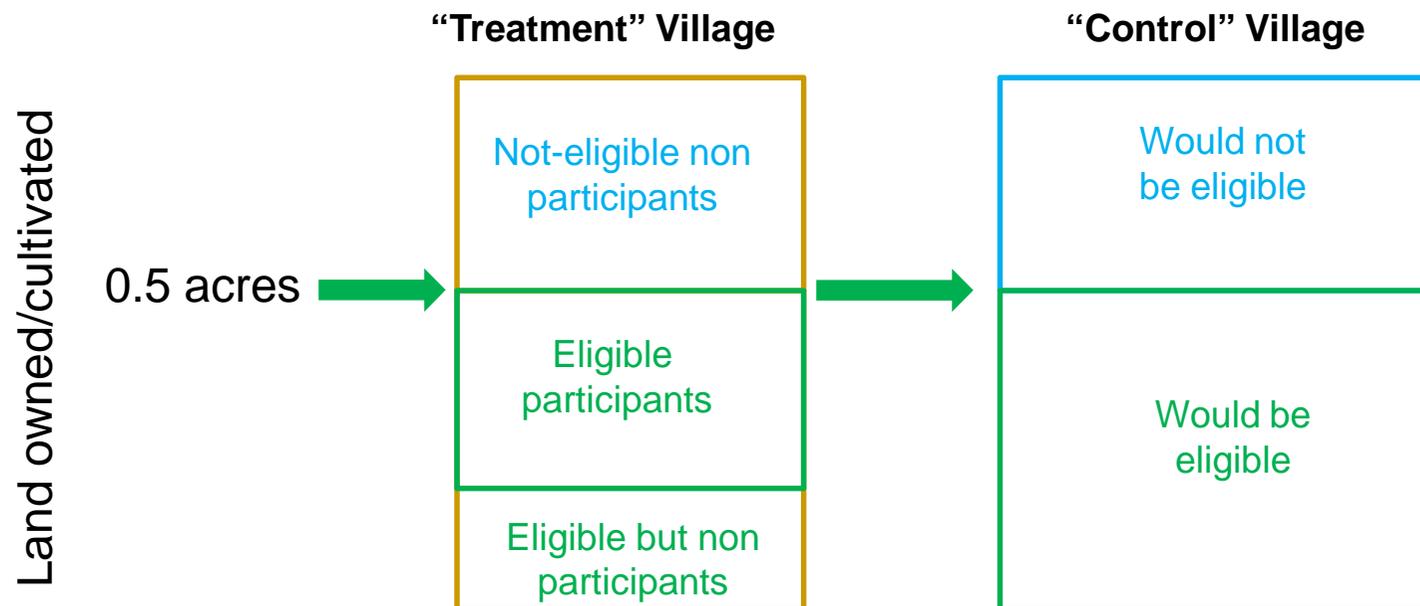
- Microfinance hype: MF has long been seen as silver bullet for alleviating poverty and empowering women through providing financial services to the poor
- Studies suggesting social and economic benefits:
 - Hulme and Mosley (1996), Coleman (1999), Pitt and Khandker (1998), Khandker (1998 and 2005), Rutherford (2001) and Morduch and Haley (2002)
- Critical voices:
 - Roodman and Morduch (2009), Bateman (2010) and Dichter and Harper (2007), Roy, 2010)
 - First two RCTs in the sector (Banerjee et al, 2009; Karlan and Zinman, 2009) raising doubts about the causal link between MF and poverty alleviation.
- Most influential MF IE to date: Pitt and Khandker (1998)

Why is Pitt and Khandker so important?

- Methodologically innovative
 - Large original World Bank survey in 1991-2
 - With follow up panel in 1998-9
 - Complex and sophisticated analysis (WESML-LIML)
 - Most rigorous impact evaluation of microfinance
- Key work of main academic author(s)
 - Key paper in top economics journal & 2 books
 - More than 21 papers mainly in peer reviewed journals sole or co-authored with collaborators and or PhD students
 - Second most highly cited paper in Pitt's career
- Widely cited by high level MF advocates such as M. Yunus

Introduction to Pitt and Khandker (1998)

- Iconic study finding positive impacts of MF especially when lending to women (male and female groups)
- Quasi-experimental design & eligibility condition used to identify impact
 - Primary eligibility criterion: landownership (0.5 acres = 50 decimals)
 - Overall sample: 1,798; 1,538 households from treatment villages, 260 from controls



Pitt and Khandker contested

Study	Method	Findings
PnK 1998	Weighted Exogenous Sampling Maximum Likelihood-Limited Information Maximum Likelihood- Fixed Effects (WESML-LIML-FE)	Significantly positive impacts on per capita expenditure, women's non-landed assets, labour supply and school enrolment
Morduch 1998	Differences-in-Differences (DiD)	Largely no impacts
Pitt 1999	Expansion of PnK model and comparison to Morduch using a simulation-based approach	Defence of original claims, i.e. significantly positive impacts
Chemin 2008	Propensity Score Matching (PSM)	Positive impacts for some outcomes but smaller than those claimed by PnK
RnM 2009	2-stage (Weighted, Truncated) Maximum Likelihood Estimation using cmp ¹	Reproduces original PnK claims but doubt causality between MF and poverty reduction
Duvendack 2010, D & P-J, 2011	PSM, DID, cmp, & Panel Random Effects	Mixed/Small negligent impacts, cannot reproduce Chemin

Sub-groups & Microfinance

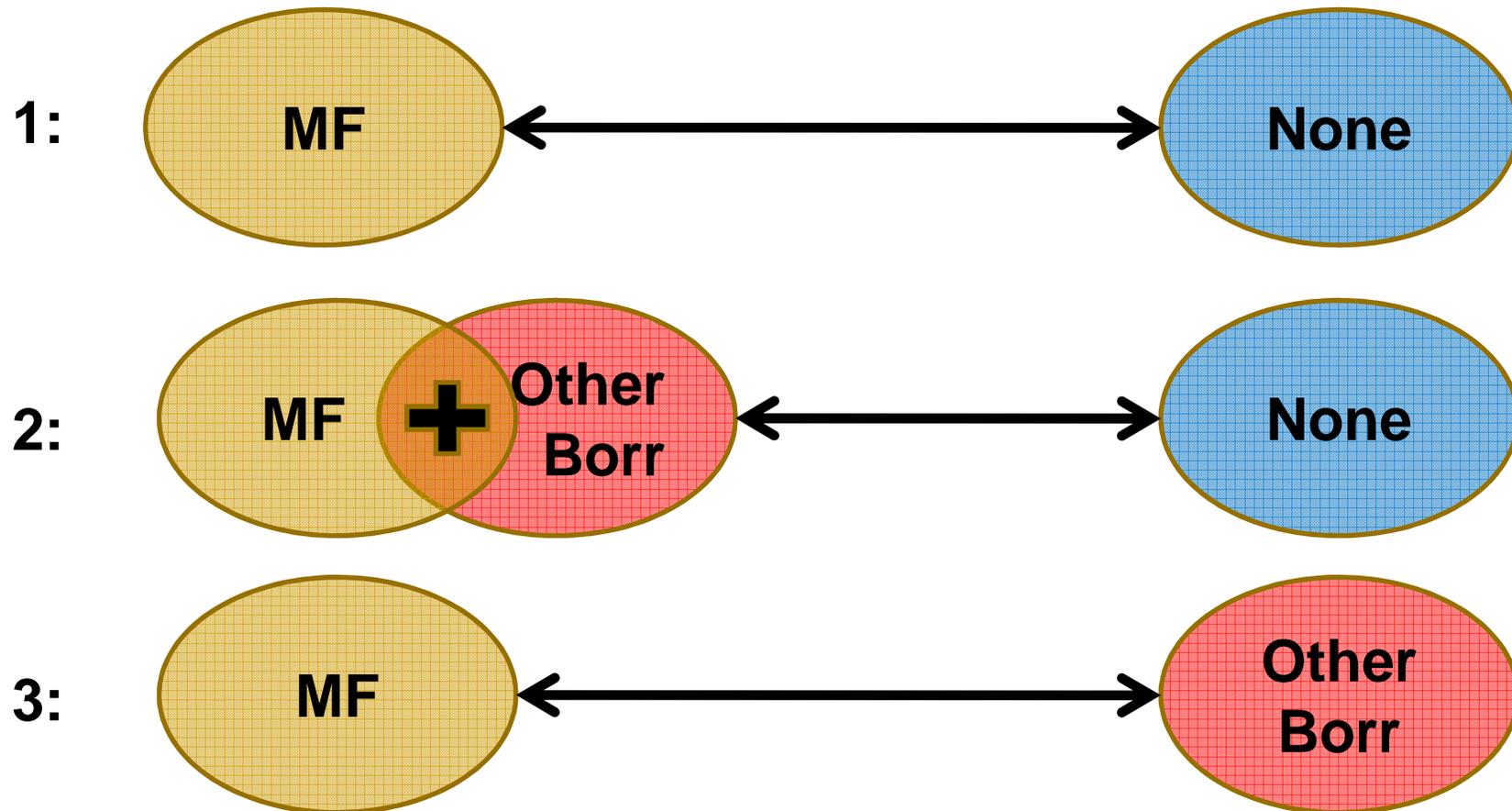
- PnK simply ignore alternative sources of finance but they appear in their data:

Status	Treatment villages, no. of individuals with multiple sources (in %)	Control villages, no. of individuals with multiple sources (in %)
Eligible	4	7
Not eligible	2.5	8

- Khandker (2000) argues MF substituting other forms of borrowing
- Sub-groups should have been in the design to explore impact heterogeneity
- One should test against next best alternative, research design neglected alternatives
- Sub-groups little analysed and not vs MF

Sub-group comparisons

- Lack of sub-group heterogeneity undermines claim that MFIs make unique contribution to poverty reduction.



Sub-group PSM results

Outcome variables	Y^{MF} vs Y^{None}	Y^{MF+} $Y^{Multiple+}$ Y^{Borr} vs Y^{None}	Y^{MF} vs Y^{Borr}
Comparison	1	2	3
	Kernel matching, 0.05		
Log per capita expenditure (Taka)	-0.011	0.019	-0.089**
Log women non-landed assets (Taka)	0.498***	0.349**	-0.022
Girl school enrolment, aged 5-17 years	0.060*	0.061**	0.077
Boy school enrolment, aged 5-17 years	0.035	0.060**	-0.011

- Comparison 3: Not obvious MF better than alternative sources
- Results by gender: No obvious advantage of female borrowing
- Sensitivity analysis: Estimates are sensitive to unobservables
- Summary: Sub-group analysis undermines PnK's claims, no advantage of MF vs other sources

External validity: Zeller et al (2001)

- Introduction to Zeller et al (2001):
 - IFPRI study, conducted in rural B/d in 1994/95
 - Quasi-experimental design, 120 villages (not all reached by credit programme) for community and group-level surveys, household survey (n=350) in 7 villages, 3-rounds accounting for B/d's agricultural seasons, re-survey in 2006/07
 - Using econometric approach: Instrumental Variables
- Findings: Positive and significant impacts on income generation and food & calorie consumption

Comparability PnK – Zeller

- Differences in actual locations, time differences and village selection
 - Actual locations:
 - PnK thana not known, all 5 regions in B/d represented
 - Zeller 5 random districts out of 19
 - Both geographically disbursed and broadly representative
 - Time difference (3 year gap 1991/- cyclone in Chittagong - vs 1994/5 – agricultural depression?)
 - Difference in selection, PnK villages (87) with programmes and Zeller select thanas (12) in which there are programmes and randomly select villages (120) within

Comparability PnK – Zeller continued....

- ❑ Quasi-experimental design, eligibility criterion landownership used
- ❑ Outcomes differ:
 - PnK: Expenditure, women's non-landed assets, labour supply, school enrolment
 - Zeller: Participation on household resource allocation, income generation, food and non-food consumption, social attitudes and capacities
- ❑ Similar interventions: rural B/d, group lending, credit plus
 - PnK: BRAC, BRDB, GB
 - Zeller: BRAC, ASA, RDRS

Sub-groups in Zeller

- Immense importance of informal sector borrowing
 - 350 households had 338 formal and 2,567 informal loan → 9 times more informal loans!
- Puzzling that in PnK only about 8% of all survey participants had informal sector borrowing
- Incidence of multiple borrowing in Zeller a lot higher than in PnK
- Multiple borrowing appears to be a normality

Sub-group results - Zeller

- Greater proportion of informal sector loans are used to finance food, health, social events, education and consumption expenditure:
 - 34.5% (informal) vs 15.8% (formal)
- Wealthier households (esp non-participants owning more than 0.5 acres) have higher demand for credit
- MF participants vs eligible non-participants: Amounts borrowed from informal sources are not that different between the 2 groups → contrary to Khandker (2000)
- MF participants borrow from informal sources to manage the strict repayment schedules of MFIs
- Supported by Fernando (1997)

Conclusions

- Sub-group analysis crucial → undermines PnK's original claims, supported by other data sources
- Award prestige only if public deposit of original data and code allowing replication and reproduction using alternative data construction, software, etc.
- Too easy to believe MF is beneficial without considering evidence in balanced way
- Dangers of readily available statistical tools, multiple journal outlets, etc. leads to loss of original thought as people extend what others have done making small extensions, adjustments (Marc Pagel)
- Repeat studies in different locations, or same location different sample, etc. to check for external validity
- Continuing need for high quality observational studies with ethical reporting and publication practices (enabling replication)

Q & A Session

For further questions or comments please email:

m.duvendack@odi.org.uk; or

r.palmer-jones@uea.ac.uk

Appendix: PSM Results – By Gender

Outcome variables			Y ^{MF} vs Y ^{None}	Y ^{MF} + Multiple+Y ^{Borr} vs Y ^{None}	Y ^{MF} vs Y ^{Borr}
Comparison			1	2	3
Kernel matching, 0.05					
Log per capita expenditure (Taka)	1	Women	-0.013	0.012	-0.126*
	2	Men	-0.046**	0.015	-0.079**
Log women non-landed assets	3	Women	0.754***	0.561***	-0.848
	4	Men	-0.000	0.244	0.249
Girl enrolment, aged 5-17 years	5	Women	0.067*	0.061**	-0.216
	6	Men	0.032	0.060**	0.133
Boy enrolment, aged 5-17 years	7	Women	0.045	0.050	-0.128
	8	Men	-0.001	0.054**	-0.017

Source: Authors calculations.

Notes: *statistically significant at 10%, **statistically significant at 5%, ***statistically significant at 1%. Results refer to the differences in the mean values between matched samples. t-tests before and after matching employed to investigate the differences in the mean values for each covariate X across matched samples; the test provided conclusive results.

Appendix: Sensitivity Analysis

- PSM estimate for log of **women non-landed assets** for Y^{MF} : 0.498*** (comparison 1) - sensitive to selection on unobservables?

Gamma (Γ)	Significance levels		Hodges-Lehmann point estimates		95% Confidence intervals	
	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
1	< 0.0001	< 0.0001	0.886	0.886	0.315	1.317
1.2	< 0.0001	< 0.0867	0.465	1.218	-0.245	1.570
1.3	< 0.0001	< 0.2329	0.274	1.341	-0.532	1.694
1.4	< 0.0001	< 0.4422	0.065	1.439	-0.710	1.796
1.5	< 0.0001	< 0.6547	-0.159	1.533	-0.886	1.891

Source: Authors calculations.

Sub-group comparisons

- Lack of sub-group heterogeneity undermines claim that MFIs make unique contribution to poverty reduction.
 1. Y^{MF} versus Y^{None} - MF participants versus all other individuals across treatment and control villages that do not have other sources of borrowing.
 2. $Y^{MF} + Y^{Multiple} + Y^{Borr}$ versus Y^{None} - individuals that participate in either MF, MF and other non-MF borrowing and other non-MF borrowing only versus all other individuals across treatment and control villages that do not have other sources of borrowing.
 3. Y^{MF} versus Y^{Borr} - MF participants versus individuals that have other non-MF borrowing across treatment and control villages.