UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Essays on Finance and Schooling Reforms in Asia

Inaugural-Dissertation

zur Erlangung des Doktorgrades

des Fachbereichs Wirtschaftswissenschaften

der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Min Xie

Heidelberg

Juni 2020

****

to my parents

****

Acknowledgements

My dissertation would never have come about without the guidance and support from my supervisor, committee members, colleagues, friends and family.

I am most indebted to my supervisor, Prof. Dr. Stefan Klonner, for his continuous mentoring over the years. I would like to express my deepest gratitude for his patient, insightful and encouraging guidance not only on my dissertation but also on my career. My sincere thanks extend to Prof. Christina Gathmann, Ph.D. and Prof. Dr. Rahul Mukherji, for their immediate readiness to serve as the referees for this dissertation.

I am thankful to the members of the Alfred-Weber-Institute for Economics for their constructive comments and advice on the research upon which this dissertation is based. I am also grateful to my colleagues in South Asia Institute who have created a stimulating and supporting atmosphere for research. In particular, I would like to thank Micha for his guidance at the early stage of my PhD. I thank Christina for her administrative support, to Paula for her extensive feedback on the draft, and to Kafeel, my coauthor, for his invaluable input in our research.

Finally, my parents and my brother's family have played important roles in this process, and I am forever grateful to them for their unconditional love.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction to the Essays

This dissertation is a collection of three self-contained essays that evaluate the impacts of three finance and schooling reforms in Pakistan and Thailand.

Improving access to finance and education are two building blocks in pursuit of economic development, and more generally, in achieving a higher standard of living. The association of financial expansion and economic growth has been well established in the empirical literature (King and Levine 1993, Jayaratne and Strahan 1996, Rajan and Zingales 1998, Levine and Zervos 1998, Bekaert et al. 2005). These studies have focused on the importance of credit in facilitating business growth. Recently, a growing body of rigorous empirical research has revealed other channels, through which access to finance improves the welfare of the poor. First, consumer credit that provides consumption smoothing mechanisms has shown welfare-improving effects (Karlan and Zinman 2010, Fink et al. 2014). In contrast, traditional microcredit that seeks to promote business growth does not yield transformative effects for an average household (Banerjee et al., 2015c). Second, access to savings products, especially with features that help people overcome behavioral biases, demonstrate large improvements in various outcomes, including business investment, health and education expenditures, and consumption. Access to savings products is more promising to deliver transformative impacts than microcredit (Karlan et al., 2014b). Third, although the research on micro-insurance is still in its infancy, current evidence shows that providing index-based crop insurance encourages higher productivity investments and behavior (Karlan et al. 2014a, Cole et al. 2017). Finally, digital payments significantly improve the risk-sharing within networks of friends and relatives, and lead to higher savings and consumption (Jack and Suri, 2014). The use of digital payments also reduces leakage and corruption in social protection programs (Muralidharan et al. 2016, Banerjee et al. 2016).

Regarding education, the theoretical literature has long attributed economic growth to education. Augmenting the Solow growth model by including human capital, Mankiw et al. (1992) argue that education can increase human capital and lead to a higher equilibrium level of output. Other authors focus on the nexus between education and technological innovation, arguing that education can increase the innovative capacity of the economy, which facilitates technological innovation. In turn, the innovative technologies promote economic growth (Lucas 1988, Romer 1990). Furthermore, Nelson and Phelps (1966) argue that education facilitates the diffusion of knowledge that is needed to implement innovative technologies. A recent empirical literature has provided evidence supporting these theories. Using a natural experiment and instrumental variable technique, Duflo (2001) shows that an additional year of schooling increases earnings in the labor market by 1.5 to 2.7 percent in Indonesia. In the context of India, Foster and Rosenzweig (1995) show that the year of schooling is positively correlated with the adoption of new agricultural technologies among rural households. Skinner and Staiger (2005) examine the adoption of new technologies across U.S. states over the course of the twentieth century. They find that education (measured by high-school enrollments) and social networks were the only variables positively associated with the adoption of innovations. Education is also argued to improve socioeconomic outcomes other than incomes, e.g., health (Strauss and Thomas 1995, Schultz 1997, Schultz 2002). Sen (2001) has regarded education itself as an intrinsic good.

Policymakers also enthusiastically support financial inclusion and education. United Nation has positioned financial inclusion as a prominent target in achieving eight out of the seventeen "Sustainable Development Goals (SDGs)".[1] Moreover, the World Bank has put forward an ambitious goal to reach universal financial access by 2020, i.e., globally all adults are able to have access to a transaction account or electronic instrument to store, send and receive money. Today, more than 55 countries have made commitments to financial inclusion, and more than 60 have either launched or are developing a national

[1]The eight relevant SDGs are SDG-1, on eradicating poverty; SDG-2 on ending hunger, achieving food security and promoting sustainable agriculture; SDG-3 on profiting health and well-being; SDG-5 on achieving gender equality and economic empowerment of women; SDG-8 on promoting economic growth and jobs; SDG-9 on supporting industry, innovation, and infrastructure; SDG-10 on reducing inequality; and SDG-17 on strengthening the means of implementation.

financial inclusion strategy (World Bank, 2018). As for education, low- and middle-income countries have massively expanded their education systems and steadily increased the government expenditure on education to around 4% of their GDP in the last two decades.[2] More recently, one of SDGs is set out to ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

Accompanying the expansion of these development policies, the fraction of the adult population with an account in a financial institution has increased from 51 to 69 percent since 2011 (Demirguc-Kunt et al., 2018). The global number of out-of-school children has decreased from 377 to 263 million since 2000 (UNESCO, 2018). On the other hand, there are persistently a third of financial account owners, who do not make any use of their accounts (Demirguc-Kunt et al., 2018). Students' learning quality has also stagnated. Today, still more than 617 million children and adolescents are not achieving minimum proficiency levels in reading and mathematics (UNESCO, 2017).

The crucial question, whether or not the development policies have contributed to these changes or stagnation, often remains unanswered. Answering this question addresses the immediate need of the policymakers on deciding whether an intervention achieves its intended outcomes, and accumulates knowledge about what works and what does not. Furthermore, it assists in comparing the effectiveness of alternative policies, and thus provides insight on what works better, a policy question that has received escalating interest as the policymakers become increasingly focused on understanding how to gain value for money (Gertler et al., 2016). Overall, it promotes accountability in resource allocation across public policies (Khandker et al., 2010). However, program managers and policymakers have long focused on measuring and reporting the inputs and immediate outputs of a program, e.g., how much money is spent, how many textbooks are distributed and how many people participate in a microcredit program (Gertler et al., 2016). As the focus of project evaluation has shifted from measuring project inputs and outputs to measuring outcomes in the last decade, the demand for evaluating policies' impacts is on the rise (World Bank, 2012).

An impact evaluation addresses these crucial needs. It assesses the causal effects of a policy on the individuals' well-beings. More formally, I formulate the causal effect with

[2] Data is extracted from World Bank's World Development Indicators on June 5th 2020. https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS?end=2017&locations=XO&start=2000

the help of the Rubin Causal Model (Rubin, 1974).[3] The Rubin Causal Model assumes that each individual can be in one of the two treatment states, i.e., treated or untreated by a policy. Each state is associated with a potential outcome. The causal effect of a policy for an individual is the difference between the two potential outcomes particular to that individual. However, we cannot observe an individual being treated and untreated at the same time. It is thus impossible to obtain the causal effect directly. Instead, we can only infer the average causal effect by comparing a group of individuals that are treated to those that are not. When the treated and untreated groups are different in characteristics in addition to the treatment status, a naïve comparison between them would yield a biased estimate of the causal effect. An example is the study of the impacts of microcredit on microentrepreneurial activities. The difference in business profits between borrowers and non-borrowers is a biased estimate of the impact of microcredit on business profits. The reason is that the differences in the characteristics between borrowers and non-borrowers may also account for the difference in business profits. It is difficult to control for these confounding factors, especially those that are unobservable to the researchers. For example, a microcredit borrower might also have a higher entrepreneurial ability, which is often unobservable and difficult to measure. A higher entrepreneurial ability could contribute to higher profitability, regardless of microcredit (de Mel et al., 2008). To eliminate the biases, the ideal approach is to use "Randomized Controlled Trials (RCTs)", which randomly assign the treatments to the individuals (Angrist and Pischke, 2008).

Using RCTs is the golden standard for drawing causal inference. However, most development policies are implemented in non-experimental settings due to financial, ethical or political concerns (Athey and Imbens, 2017). For example, it would be unethical to prevent households from receiving remittances in order to study the causal effect of remittances on consumption smoothing. Despite the lack of purposeful randomization, the implementations of policies often create variations in treatment assignments. For example, large-scale policies are often rolled out in phases, creating geographical and temporal variations. Some policies target individuals with a measure of pre-determined characteristics above a certain cutoff, providing opportunities to compare the individuals whose measures are just above and below the cutoff. Some policies are implemented in a one-

[3]The Rubin Causal Model is a formal framework for causal inference, around which the modern research on impact evaluation has been developed. Important references are Rubin (1974), Rubin (1977) and Holland (1986). For a systematic review of the model, please refer to Imbens and Rubin (2010)

size-fits-all manner, creating variation in treatment intensity among individuals. When these variations are exogenous, i.e., not threatened by confounding factors, they constitute natural experiments (for the discussion on natural experiments, see Meyer 1995 and Rosenzweig and Wolpin, 2000). Depending on the rules that govern the variations of interest in natural experiments, researchers employ various econometric techniques to explore these exogenous variations and draw causal inference. These techniques are often referred to as identification strategies or empirical strategies (Angrist and Krueger, 1999). For example, difference-in-differences methods exploit the geographical and temporal variations at the same time.[4] Regression discontinuity designs explore the randomness of treatment around the cutoff of the "assignment variable" (also referred to in the literature as the "forcing variable" or the "running variable").[5] An instrumental variable induces exogenous change in the explanatory variable of interest, to which a corresponding change of outcome is attributed (for a recent review, see Imbens 2014). Matching methods pair each treated individual with untreated individual(s) with similar covariates (see Imbens and Wooldridge 2009 for an overview). More recently, the research on the application of machine learning in impact evaluation has been very active (for recent developments in this literature, see Athey and Imbens 2017).

My dissertation consists of three essays that evaluate the impacts of three development policies, which have been implemented in non-experimental settings. Using various identification strategies tailored to the settings, I address the critical policy question of whether or not a policy has achieved its intended outcomes on access to finance, education and poverty reduction. The first essay evaluates the effects of a large state-led microcredit program on households' non-agricultural entrepreneur activities and welfare in rural Thailand. The second essay investigates a reform on the bank branch expansion policy in Pakistan and estimates its impacts on the rural population's access to formal finance and poverty reduction. Finally, the third essay evaluates the impacts of a school funding reform on school infrastructures and education outcomes in Punjab Pakistan. In

[4]Difference-in-differences methods have been an important tool for empirical researchers since the early 1990s. The most prominent study using this technique is Card and Krueger (1994) on the effects of minimum wages on employment. Imbens and Wooldridge (2009) review this technique in detail. The recent development in difference-in-differences has focused on the construction of synthetic control (Abadie and Gardeazabal 2003, Abadie et al. 2010, Abadie et al. 2015)

[5]This method was first introduced by Thistlethwaite and Campbell (1960) in their work in psychology, and has become popular in economics literature since the early 2000s. Recent reviews on its applications to empirical economics research include Imbens and Lemieux (2008) and Lee and Lemieux (2010).

what follows, I provide a brief introduction to each of the three essays of this dissertation.

The Regressive Effects of a Rural Credit Program: Evidence from Thailand's Village Fund Program

In chapter 2, "The Regressive Effects of a Rural Credit Program: Evidence from Thailand's Village Fund Program", a joint work with Stefan Klonner, we study the effects of Thailand's Village Fund Program with particular focuses on non-agricultural entrepreneurship, household welfare in the longer run, and heterogeneous program effects. Thailand's Village Fund is one of the most ambitious state-led microcredit intervention in the world (Boonperm et al., 2012). Aiming at stimulating economic growth in rural Thailand, in 2001, the Government of Thailand started to distribute one million baht (approximately USD 24,000 in 2001) to each of the 79,000 villages in Thailand, regardless of village size. By the end of 2002, the government had spent around 2.4 percent of Thailand's GDP of that year to fund this program (Pungprawat, 2012). Meanwhile, 99 percent of the targeted villages had set up the fund by 2002 (Pungprawat, 2012). By the end of 2004, the active borrowers of the Village Fund added up to 11 percent of Thailand's population (Boonperm et al., 2012).

For the empirical identification of program effects, we exploit the variation in the Fund's intensity per village household, which is reciprocal to village size, within a difference-in-differences approach using a household-year panel data set. This variation is argued to be exogenous for two reasons. First, villages are administrative units whose sizes have stayed relatively constant before and after the program. Second, the program forbids lending to nonlocals. This identification strategy delivers marginal program effects evaluated at the program's large average intensity of 10,000 baht (650 PPP US dollars) per household, an amount roughly equals to the pre-intervention average of formal household borrowings.

We find an immediate and substantial increase in access to credit along the entire income distribution without crowding-out of other forms of credit. For the real economy, we document, first, that the program increased business activity as measured by the share of households operating a non-agricultural business by 20 percent. Second, there is significant heterogeneity in program effects along the wealth distribution. While we find no effect on business activity for initially poor households, initially wealthy households expand the number of businesses and employees by 21 percent. Additionally, the value of their business assets increases by 73 percent.

Consistent with these heterogeneous effects on entrepreneurship, we find that welfare, as measured by expenditures on high-frequency consumption items, increases by 10 percent among initially wealthy households, while we find no such effect among the initially poor. This heterogeneity persists over the ten years after the program's inception. As a result, the Village Fund sharpens village-level consumption inequality, as measured by initially wealthy households' share in aggregate consumption, by about half a percentage point per year.

We argue that business expansion is likely to be the mechanism behind these heterogeneous welfare gains. In support of our argument, we find that wealthy households accumulate business assets by adapting preexisting house and vehicle to business use, rather than acquiring new assets. Precisely, we find that the Village Fund increases the expenditure on house and vehicle repairs among the wealthy households. On the contrary, the initially poor households do not achieve similar gains regarding business expansion, even though they benefit from the Village Fund in terms of financial inclusion.

In contrast to the conventional argument, this essay shows that initial wealth complements rather than substitutes the access to credit when it comes to business expansion. This might explain why most microcredit interventions fail to bring welfare gains to ultra-poor households, who are deprived of household assets. Therefore, to improve welfare for the ultra-poor through fostering entrepreneurship, interventions, such as a graduation program (Banerjee et al., 2015b), might be more effective than microcredit.

A Financial Inclusion Policy Gone Wrong: Pakistan's Bank Branch Expansion Reform

In the last two decades, the number of microfinance institutions whose primary clientele are relatively poor households has skyrocketed. This observation is in line with the belief that traditional banks are not suitable for lending to these households. Accompanying the expansion of microfinance institutions, the literature has witnessed a sharp increase in the number of studies that evaluate the ability of microfinance institutions to cater to the poor. However, these studies need to be supplemented with others that address the extent to which other banking institutions can also serve the poor. The essay in chapter 3, "A

Financial Inclusion Policy Gone Wrong: Pakistan's Bank Branch Expansion Reform", serves this objective, evaluating whether expanding commercial bank branches to rural areas improves rural households' access to formal finance, and whether it helps reduce poverty.

In 2004, the State Bank of Pakistan (SBP), Pakistan's central bank, implemented a reform in its "Branch Expansion Policy" with an objective to enhance the outreach of the banking services to the rural and underserved areas of the country. This reform mandated the banks with a network size of no less than 100 branches to open 20% of their new branches in "Rural and Underserved Areas (RUA)", which were vaguely defined as "villages, small towns or unbanked tehsil headquarters". This policy was expanded to include all banks regardless of their network sizes effectively in 2007.

I draw on several sources for the data in this analysis, including the State Bank of Pakistan, Pakistan Bureau of Statistics, European Commission and the United States National Geospatial-Intelligence Agency. Using bank-branch-level geographical data, I track the expansion of bank networks along the policy timeline. I establish the policy's effect on the networks' rural population coverage using a simple trend break model. On an aggregate level, I identify the policy's effects on branch accessibility, agricultural credit disbursement and poverty by estimating a difference-in-differences trend break model. This model exploits the change in banks' selection on less remote rural locations for their RUA branches after the policy.

I find that the banks have complied with the policy by reserving 20% new branches for RUA locations. However, the bulk of new RUA branches are in less remote locations, i.e., locations that are closer to district capitals and more likely to have other branches within 5 kilometers radius, than the RUA branches opened before the reform. As a result, the policy failed to reach to unbanked rural areas. Precisely, the fraction of the country's rural population living within 5 kilometers radius of a bank branch (banked rural population) stagnated at 57 percent. A counterfactual analysis reveals that the banked rural population would have reached to 70 to 72 percent had the banks chosen locations with a higher "Rural Market Potential (RMP)", i.e., locations with a larger rural population per branch or unbanked locations with a larger rural population.

Consistent with the actual branch-opening pattern, the aggregate level analysis shows that the size of the RUA branch network has grown slower in districts with a larger rural population share after the policy. In other words, the growth rate of RUA branch network in a more rural district (e.g., "Lower Dir" whose rural population share is 97 percent) was surpassed by the less rural one (e.g., "Karachi" whose rural population share is 7 percent) by 0.86 branches per million persons per year after the reform. This accounts for an accumulative difference of 5.16 RUA branches per million persons between a more and less rural district by 2012, a value that is 66% of the average RUA network size in 2000.

However, the number of borrowers and the amount of credit disbursed to agricultural sector did not grow differently between a more and less rural district in the wake of the reform. These results suggest that the policy has failed to expand credit disbursement to the rural sector, consistent with the branch network's stagnant rural population coverage. Finally, poverty has also evolved in parallel between more and less rural districts along the policy timeline, suggesting a lack of impact on poverty.

Taken together, I conclude that the rural expansion policy has failed to enhance the banking sector's rural outreach and reduce poverty. The reason is that the new branches reserved by the 20 percent quota were opened in less remote rural locations where a branch network has already existed. More unfortunately, the reform has sharpened the geographical inequality in the number of RUA branches. My findings demonstrate the need for more carefully designed government interventions for improving financial access than Pakistan's arguably naïve bank branch expansion policy.

School Grants and Education Outcomes: The Impacts of the Non-Salary Budget Reform in Punjab Pakistan

Regarding education, policymakers have become more and more interested in providing decentralized grants to schools in the hope of improving education outcomes. The belief is that decision-makers at the local level have a better understanding of the schools' needs than those at the national level. Therefore, they are in a better position to identify schools' deficiencies and to use the resources more efficiently. However, evidence from rigorous evaluations are still limited. In this connection, in chapter 4, "School Grants and Education Outcomes: The Impacts of the Non-Salary Budget Reform in Punjab Pakistan", co-authored with Kafeel Sarwar, we investigate a decentralized school grant program in Punjab Pakistan. We address two questions. First, does providing decentralized school grants have any impacts on human capital (measured by teacher's attendance rate) and school's physical infrastructure? Second, does this decentralized grant improve the education quantity and quality?

Public schools in Punjab Pakistan have suffered from underfunding of non-salary expenditure. Prior to the reform, they received a fixed grant of Rs. 20,000 (approximately 200 USD in 2013) or 40,000 per annum, depending on the school level, for the non-salary expenditures. However, this was highly insufficient even to cover the maintenance fee of the basic infrastructures, which was estimated to be about Rs. 70,000. The underfunding for the non-salary expenditure has persistently hindered public schools' ability to attract and retain students.

Against this backdrop, the Punjab government implemented the "Non-salary Budget (NSB)" reforms in 2013, which altered the funding allocation rule from a fixed to need-based rule. The need-based rule accounts for school level, student enrollment, furniture deficiency and building condition. Under this need-based rule, the average annual non-salary budget per school amounts to Rs. 220,000, which is about ten times the amount received under the erstwhile fixed allocation rule. The reform has also devolved the financial decision-making to the school councils, which consist of parents and headteachers. The NSB reform was rolled out in three phases to cover the whole province. In 2013, nine districts were selected for phase one. In 2014, additional nine districts were added in phase two. Since 2015, the NSB reform was extended to the whole Punjab province.

To identify the reform's impacts, we use a difference-in-differences method, exploring the reform's staggered rollout schedule, which has created temporal and geographical variations in policy treatment. This identification strategy delivers the average effects of entering the NSB reform. We combine administrative data and independent household surveys from various sources. We collect administrative data on schools' financial accounts, infrastructure conditions, teacher and student attendance rate, and standardized test scores from the School and Education Department of Punjab and the Punjab Examination Commission. We measure children's school enrollment rate and basic reading and mathematical ability measures using two independent household surveys, namely "Multiple Indicator Cluster Survey Punjab" and "Annual Status of Education Report Pakistan".

We report four sets of results. First, regarding school's financial account, the NSB reform has significantly increased an average school's annual income and expenditure by

Rs. 180,000 and Rs. 135,000, respectively. Second, the reform has significantly improved the conditions of the school infrastructures by 0.045 standard deviations. This effect is driven by the improvements in the existing infrastructures, e.g., the fraction of functional toilets, complete boundary wall and safe school building condition. In terms of human capital input, we find that teacher's attendance has increased slightly by 0.9 percentage points, a value that is only 1% of the average at the baseline. Third, we do not find any effect on education quantity as measured by student's attendance rate and enrollment rate among children of age 5 to 16. Finally, we do not find any impacts on students' numeracy and literacy test scores, or basic reading and mathematical ability scores.

We interpret the lack of effects on education outcomes as a result of insufficient capacity among the school council members. In fact, a report on the procurement in the education sector of Punjab, I-SAPS (2014), has revealed that the majority of school council members do not understand the procedure of school procurement. More importantly, the school councils have limited power to hold the teachers accountable for delivering quality education, as the teachers in public schools of Pakistan have political patronage. Simply increasing the funding to schools does not change the lack of accountability relationship between the teachers and the school councils. Overall, our findings are consistent with Mbiti et al. (2019) and Beasley and Huillery (2017), who show that the effectiveness of decentralized school grants depends on the capacity of the local planners who are responsible for grant management.

CHAPTER 2

The Regressive Effects of a Rural Credit Program: Evidence from Thailand's Village Fund Program

**Joint work with Stefan Klonner*

Abstract We study the impacts of an expansion of credit supply in the context of Thailand's Village Fund Program on entrepreneurship and household welfare. We use exogenous variation in the program's intensity across villages of different sizes to identify the effects of a credit supply expansion beyond the program's mean intensity. While credit reached out to poor and wealthy households alike, we find that credit deepening causes an expansion in micro-entrepreneurship and business assets only among initially wealthy households. Similarly, for household consumption, we find sustained increases in the long run only among initially wealthy households, implying an increase in consumption inequality, as measured by the initially wealthy households' share in aggregate consumption, of half percentage point per year. Our results challenge the view that universal access to credit enhances equity, at least for supply expansions of the Village Fund's scale.

2.1 Introduction

Rural households in low-income countries have traditionally faced large impediments to access to formal finance (Banerjee and Duflo, 2007). High interest rates and credit constraints are commonly believed to impinge on both entrepreneurial activity and consumption smoothing. In response, expanding access to institutional credit has become an important development goal with the potential to bring gains in efficiency as well as equity. The expansion of access to institutional credit among rural populations has largely been promoted by government banking and lending programs and, more recently, microfinance institutions (Karlan and Morduch, 2010).

Lately, plenty of empirical evidence on the effects of microfinance on entrepreneurship and welfare has accumulated. Banerjee et al. (2015c) summarize the following stylized facts. First, microfinance has small but measurable positive effects on micro-entrepreneurship. Second, there is great heterogeneity in these effects with respect to certain baseline characteristics of households, in particular the presence of a business. Third, given common practical limitations in the design of field experiments, there are rarely measurable welfare gains. For large-scale government banking and lending programs, the evidence is more limited. Some prominent studies find a sizable positive effect on welfare (Boonperm et al. 2013, Kaboski and Townsend 2012, Burgess and Pande 2005, Cuong 2008), while the evidence for entrepreneurship is more mixed (Burgess and Pande 2003, Kaboski and Townsend 2012).

In this paper, we revisit Kaboski and Townsend's (2012) study of the effects of Thailand's Village Fund with particular focuses on non-agricultural entrepreneurship, household welfare in the longer run, and heterogeneous program effects. Sanctioned by the Thai parliament in 2001, this government program allocated one million baht (approximately USD 65,000 in purchasing-power parity terms) to each of the country's 79,000 villages, regardless of village size. We use data from the Townsend Thai Project (Townsend, 2013) for the years 1998 to 2011. For the empirical identification of program effects, we exploit the variation in the Fund's intensity per village household, which is reciprocal to village size, within a difference-in-differences approach. This identification strategy delivers marginal program effects evaluated at the program's large average intensity of 10,000 baht (650 PPP US dollars) per household, an amount roughly equal to the pre-intervention average of formal household borrowings.

Consistent with Kaboski and Townsend (2012) as well as other authors' (Menkhoff and Rungruxsirivorn 2011, Boonperm et al. 2013) results, we find an immediate and substantial increase in access to credit along the entire income distribution without crowding-out of other forms of credit. Departing from previous findings, for the real economy, we document, first, that the program increased business activity as measured by the share of households operating a non-agricultural business by 20 percent. Second, there is significant heterogeneity in program effects along the wealth distribution. While we find no effects on business activity for initially poor households, initially wealthy households expand the number of businesses and employees by 21 percent while their business assets increase by 73 percent.

Consistent with these heterogeneous effects on entrepreneurship, we find that welfare, as measured by expenditures on high-frequency consumption items, increases by 10 percent among initially wealthy households, while we find no such effect among the initially poor. This heterogeneity persists over the ten years after the program's inception for which we analyze household data. As a result, the Village Fund sharpens village-level consumption inequality, as measured by initially wealthy households' share in aggregate high-frequency consumption, by about half a percentage point per year.

Our paper contributes to the empirical literature on the effects of credit programs in developing countries. Our first contribution is that we demonstrate important heterogeneities in program effects. As Banerjee et al. (2015c) point out, although average impacts of access to institutional credit seems to be moderate, some subgroups tend to benefit disproportionately. Similar to our results, for a microfinance expansion in Bosnia and Herzegovina, Augsburg et al. (2015) show that it is primarily households with a pre-existing business that expand business activities when obtaining access to credit. Similarly, for an urban area in India, Banerjee et al. (2015a) find that business expansion is mostly driven by households who owned a business before a microfinance intervention. Our main finding in this connection is that wealth is an important predictor of benefits from the program, regarding both entrepreneurship and welfare.[1]

Second, we add to the evidence on the longer-term effects of access to finance. Recent

[1]Our finding contrasts, to some extent, the results of Boonperm et al. (2013), who use matching methods and find that average program effects on consumption and income are concentrated at lower quantiles of the income distribution.

randomized impact evaluations of microcredit have found only small and insignificant welfare effects in the short to medium run, with a negative sign of the effect in some cases (Angelucci et al. 2015, Attanasio et al. 2015, Augsburg et al. 2015, Banerjee et al. 2015a, Tarozzi et al. 2015, Crépon et al. 2015, Karlan and Zinman 2011). The two explanations given by several of these authors are the low to moderate loan take-up rates in these studies, which reduces the precision of intent-to-treat estimates, and the possibility of long lags for welfare gains to materialize. Other non-experimental evaluations of large-scale government programs tend to find positive effects on welfare, but rarely investigate impact heterogeneity and inequality, even when long-run data is available (Burgess and Pande 2005, Cuong 2008, Boonperm et al. 2013, Kaboski and Townsend 2012). Our innovations here are that we document persistent welfare effects together with an increase in consumption inequality over a time horizon of ten years. Perhaps most interestingly in this connection, our results demonstrate that the poor appear not to reap long-term welfare gains from credit deepening despite sizable gains in access to finance.

Given that our study is closely related to Kaboski and Townsend (2012) (hereinafter referred to as "KT"), in particular regarding the identification strategy and data, we shall close this introduction by highlighting the differences between their and our analysis. First, regarding heterogeneous effects, KT only consider heterogeneity by the household initial income, gender or education of household head. In contrast, guided by recent evidence from randomized controlled trials and an interest in the distributional properties of the program's effects, we explore initial wealth as source of heterogeneity. Second, regarding welfare, KT as well as our study find only a short-term effect on consumption in the sample of all households. In addition, however, we find large long-run effects for initially wealthy households as well as long-run effects on village-level consumption inequality. Finally and perhaps most importantly, regarding the mechanism of the Village Fund's impact on welfare, KT conclude that the program's main effect was a transitory consumption spike triggered by households' running-down of buffer stocks. We challenge this interpretation by showing that village fund credit leveraged expenditures on durable items like housing and vehicle improvements or repairs accompanied by a transfer of the respective assets from the consumption to the entrepreneurial sphere of the household. Put differently, in our view the transitory consumption spike in KT veils large hidden investments in non-agricultural household enterprises not explicitly captured by the Townsend

Thai Project's entrepreneurship questionnaire.

The remainder of this paper is structured as follows. Section 2.2 describes the Village Fund in detail. Section 2.3 presents our empirical methodology and data. Section 2.4 contains the results. Section 2.5 provides some robustness checks. We conclude in section 2.6.

2.2 Thailand's Village Fund Program

2.2.1 Institutional Background

The Village and Urban Community Fund Program, commonly known as the Village Fund Program, was launched in 2001 soon after Thaksin Shinawatra was elected as the prime minister of Thailand. The government borrowed the idea of the Village Fund Program from the Grameen Bank. Aiming at distributing one million baht (approximately USD 24,000) to each of the 79,000 villages in rural Thailand, this program is considered as one of the most ambitious financial access expansion program in the world (Boonperm et al., 2012). By the end of year 2002, the central government had spent around 80 billion baht (USD 2.4 billion) to fund this program, amounting to 2.4 percent of Thailand's GDP in 2002, and 99 percent of the targeted villages had set up the fund (Pungprawat, 2012). By the end of 2004, the active borrowers of the Village Fund added up to 11 percent of Thailand's population (Boonperm et al., 2012). The objectives of the Village Fund, according to the "Act of National Village and Urban Community Fund B.E.2547", are as follows: (1) to be used as a revolving fund for investment, job creation, income generation, welfare improvement and reducing expenses; (2) to be used as emergency fund to cope with urgent problems; (3) to empower and stimulate the grassroots rural economies. The government claimed that this program enabled the underserved and poor people to have better access to capital (Menkhoff and Rungruxsirivorn, 2011).

2.2.2 Organization and Lending Policy

Figure 2.1 demonstrates the organization of the Village Fund. It consists of three levels of administration. First and foremost is the national committee who administers the fund at the national level and sets regulations and guidelines. Second, the sub-committees report to the national committee and administrate the program at the district level. And third, the local committees at the village level directly handle the transactions with their local borrowers. While complying with the Village Fund Act, the local committee is entitled, to some degree, to regulate the interest rate and repayment procedure (Tangpianpant, 2010).

Any village wishing to obtain the fund needed to satisfy the following conditions as regulated by the national committee. First, a local committee at the village level had to be formed. The local committee had to consist of 9 to 15 members that had been fairly elected from the villagers.[2] Half of the local committee members had to be women (Boonperm et al., 2013). The local committee approved loan applications based on the borrower's ability to repay, the purpose of borrowing and the loan size (Tangpianpant, 2010). The close relationship between the local committee and borrowers mitigates the problem of asymmetric information. Second, in terms of managing the fund, the local committee needed to agree with the lending policy regulated by the national committee, which states the following: the loan size could not exceed 20,000 baht per borrower, however, in an emergency case, it could be extended to 50,000 baht. Moreover, the interest rate was not allowed to exceed 15 percent per year. And repayment had to be made within one year. Third, the local committee needed to submit to the national committee an application, in which the exact interest rate and repayment procedure have to be stated.

When these conditions have been fulfilled, the national committee proceeded to evaluate the village's application. After passing the evaluation, the village would receive one million baht transfer from the central government regardless of the village size. Notice that, the size of village is exogenous to the village's receipt of the fund. This generates fairly exogenous variations in terms of fund intensity across villages (Kaboski and Townsend, 2012). We explore this variation later in our estimation framework.

[2]These local committee members must have lived in the their villages for at least 2 years (Tangpianpant, 2010)

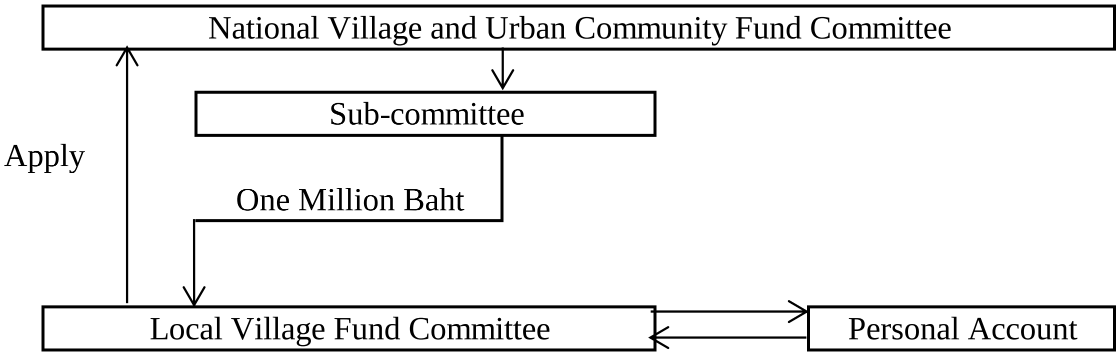


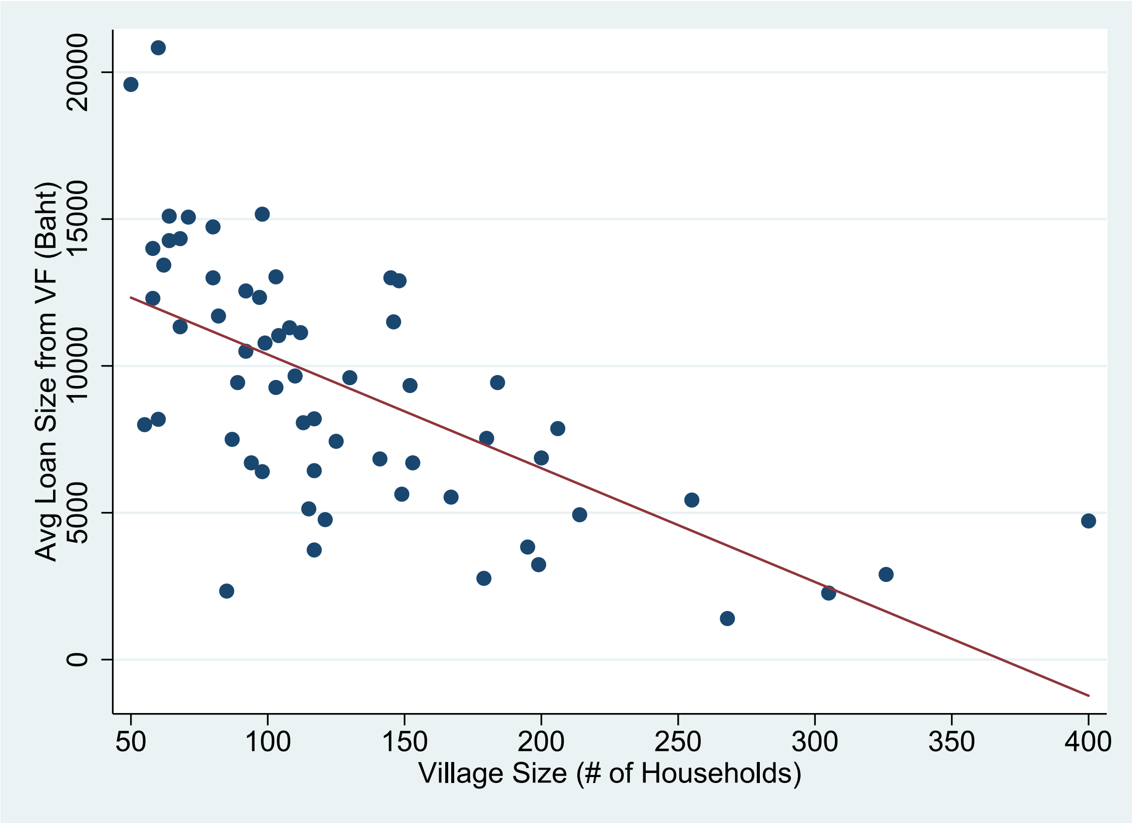
Figure 2.1: Organization of the Village Fund

2.3 Empirical Methodology

2.3.1 Quasi-experimental Setting

We use difference-in-differences method to identify treatment effects of the Village Fund. The first difference describes the fund intensity variation across villages, and the second difference depicts the discrepancy between pre- and post-inception of the Village Fund. To justify this identification strategy, we show the two features of the Village Fund program, which give the fund intensity and the timing of the intervention's inception, respectively, a fair degree of exogeneity.

First, each village received the same amount of transfer irrespective of the village size. This generates a variation in fund intensity across villages, i.e., the larger the village size, the less the Village Fund credits are available to each household in that village. This is supported by figure 2.2, which plots the average credits borrowed from the Village Funds against the village size. The fund intensity is determined by the village size, thus the exogeneity of village size will justify that of fund intensity. One potential confounder is given by the argument that the Village Fund drives the households to move to smaller villages where the fund is more intensive. However, as the Village Fund policy states, loans can only be lent to local villagers who have lived in the village for at least two years. Migration with the purpose to borrow from other villages is not likely, at least not in the short run. To further support this argument, we conduct a test and find that the temporal change in village size does not correlate with the village size at the baseline. We present the test in detail in Appendix 2.A.



Note: The data is from the Townsend Thai Project Household Annual Resurvey 2003. We aggregate the credit borrowed from the village fund from household- to village-level. Therefore, each point represents a village.

Figure 2.2: Average Loan Size and Village Size

Second, the Village Fund was unanticipated and implemented rapidly. This gives exogeneity to the timing of the intervention's inception, which assures that households did not have adjusted their behaviors before the Village Fund was implemented. In fact, the rapid implementation was assisted by the government's prerogative over the budgeting system. The prerogative allows the government to dispense money from the "Central Fund" budget without providing any details of the use. And most importantly, the prime minister has the full control to set the limit, to which the Bureau of the Budget usually agrees (Pungprawat, 2012). Table 2.1 shows the decomposition of the "Central Fund" budget expenditure before and during Thaksin's premiership. The share of the "Central Fund" budget in the annual budget was 10% between 1997 and 2001, it increased to 19% during Thaksin's Premiership. Regarding the "Central Fund" expenditure, as shown in the column under "Policy Implementation", after Thaksin was elected, his government started to use the "Central Fund" to finance his numerous policies, including the Village Fund program. On the contrary, the "Central Fund" was not used for policy implementation before Thaksin's premiership. Pungprawat (2012) documents that, the Village Fund, inaugurated at the end of 2001, had already reached to 99 percent of targeted villages by the end of 2002. This fast roll-out is also confirmed by our sample, where all the sample villages received the fund in the survey of 2002, while none of the villages received the fund in the survey of 2001.

Table 2.1: Decomposition of the "Central Fund" Budget Expenditure Before and During Thaksin's Premiership

	Fiscal Year	Regular Items (%)	Policy Implementation(%)	Extraordinary Budget Items(%)	Total (%)
Pre-Thaksin	1997	100	0	0	100
	1998	100	0	0	100
	1999	100	0	0	100
	2000	100	0	0	100
	2001	100	0	0	100
Thaksin's	2002	61	7.4	31.6	100
Premiership	2003	80.1	8.7	11.2	100
	2004	67.3	4.3	28.4	100
	2005	76.4	8.3	15.3	100
	2006	61.2	12.5	26.3	100

Note: The data source is from Pungprawat (2012). The share of "Central Fund" budget in the annual budget was 10% between 1997 and 2001, it increased to 19% during Thaksin's Premiership. Policy Implementation Items include: Village and Urban Community Fund (2002-present), Inheritance Pension (2004-2006) and Small-Medium-Large Village Development Fund (2005-2006)

2.3.2 Data

We use household-year panel data compiled from the Townsend's Thai Project. The panel consists of 960 households from 64 villages that are selected from 4 rural provinces. The sampling procedure assures the representativeness for rural households in Thailand. The survey is administrated annually in May. In case of the interviewed households emigrating to other regions, replacements are randomly selected in the same villages. We use 14 waves of the household survey from 1998 to 2011 (Townsend, 2013). The survey of 2011 was the most recent by the time we started this study. We focus on households, who live in villages with a size between 50 to 400 households.[3] For the short-run analysis, we use

[3]We trim down the 1 percent outliers. Originally, the range of the village size is 24 to 3245 households.

the sample from 2000 to 2003, and for the long-run analysis, we use the sample from 2000 to 2011. The sample from 1998 to 2000 is reserved for placebo tests. The detailed summary statistics are shown in table 2.2.

The village size is measured in terms of number of households. We define "fund intensity" as the reciprocal of the village size in 2001 (i.e., the year just prior to the inception of the intervention). For the sake of readability, we rescale the "fund intensity" to the unit of 10,000 baht per household. The baseline characteristics of an average household has 1.6 children of age younger than 18 years old, 1.5 adult male and 1.6 adult female. An average household head is 54 years old and attended 4 years of school. We also report incidence of "initially wealthy". We define a household as initially wealthy if it possesses a level of wealth above its village's median in 2001. Thus the sample is nearly equally partitioned, with 53 percent wealthy and 47 percent poor households. The household wealth is calculated by the sum of the value of four asset categories, namely household assets, agricultural assets, business assets, and land holdings. Except for the business assets module, each asset module provides a list of pre-identified items for the surveyed households to choose and fill in the corresponding quantity and self-estimated value. Regarding the business assets module, an open list is provided to the surveyed households to fill. In case of missing information on self-estimated value of an item, we impute the value by multiplying the item's median price in the province and year with its quantity reported by the households.

By accident, a "village fund" borrowing channel is listed in the questionnaires in all waves of survey. However, it does not specifically refer to the Village Fund Program, but rather generally to any village financial institutions. This explains the fact that both averages of the incidence of "borrow from Village Fund" and "VF credit" are visually close but not equal to zero in the pre-intervention sample in table 2.2. However, these averages are sharply higher in post-intervention sample, suggesting an average take-up rate of 57 percent and average loan size of 9,000 baht. Apart from the Village Fund, an average household also borrowed 11,000 baht from formal financial institutions, i.e., banks and agricultural cooperatives, and 3,000 baht from informal sources which include moneylenders, friends and relatives.

In our analysis, business refers to non-agricultural business, which includes the following categories: general shop, mechanic/repair service, hair salon/barber, rice miller, shrimp farmer and fish farmer, vehicle rental, trader and restaurant. Before the intervention, 41 percent of the sample households owned at least one such business. These businesses are typically small in scale. Prior to the intervention, the average number of employees (including the owner) is 1.3, where only 0.14 people are paid. In other words, an average household in our sample owns a business that has only household members working for it. The average business capital of a household amounts to 33,000 baht. We define the business investment expenditure as the value of the business assets purchased in the last 12 months. An average household spends 3,500 baht per year on the business investment. In the context of rural Thailand, it is not uncommon to see a business operating in the household's residence house and using household's vehicles. A typical example is the street food vendor, who would use part of the residence house as cooking and serving area, or use a motorcycle, that is otherwise used for school run, to transport the cooking utensils and ingredients to the sale spot. 11% and 4% of total households report house and vehicle, respectively, as business asset. This implies that 26% and 10% business owning households use their household's residence and vehicles for business.[4] Such business operating behavior has important implication in our analysis. We discuss the implication in further detail in the results section.

The expenditure module surveys a list of pre-identified items (self-produced and purchased) where the households occur expenditure. These items are: grain, dairy product, meat, alcohol, tobacco, gasoline, ceremonies, education, clothing, food eaten away from home, house and vehicle repairs. The total yearly expenditure of these items was 9,700 baht per capita before the intervention. Our measure of welfare is the regular consumption which is total consumption net of the house and vehicle repair expenditures because house and vehicle have both welfare and investment aspects. The regular consumption of an average household was 8,100 baht per capita. As we discussed previously, in the context of small business in rural Thailand, the house and vehicle are common capital for production. Hence, occurring larger expenditures on the repairs of these two items would imply the increase in investment. Due to the structure of the survey, it is not possible to disentangle the welfare and investment aspects of the repair expenditures. An average household spent 1,600 baht per capita in a year repairing house and vehicles. Compare to regular consumption, not only the value but also the incidence of house and vehicle re-

[4]Dividing 11% and 4% by 0.42, the fraction of business owning households, gives 26% and 10%.

pair expenditures is small. Only 21 and 55 percent of the households have ever occurred any repair expenditures on house and vehicle, respectively, before the intervention. In comparison, all households have spent on regular consumption.

Table 2.2: Summary Statistics

	Pre-intervention		Post-intervention	
Statistic	Mean	St. Dev.	Mean	St. Dev.
Fund Intensity (10,000 baht per household)	0.965	0.451	0.966	0.463
Household Characteristics				
Number of Children (< 18 year-old)	1.551	1.192	1.442	1.160
Number of Male Adults	1.467	0.939	1.433	0.859
Number of Female Adults	1.553	0.762	1.558	0.764
Household Head Age	54.029	13.401	55	13
Head Education (in year)	3.950	2.709	4.050	2.773
Wealth (in 10,000 baht)	141	397	138	424
Initially Wealthy (Incidence)	0.53	0.50	0.51	0.50
Borrowing				
Borrow from VF (Incidence)	0.003	0.057	0.573	0.495
VF Credit (in baht)	44	1,042	9,091	9,550
Formal Credit (in baht)	11,087	34,572	10,202	27,364
Informal Credit (in baht)	2,977	18,638	2,097	11,332
Total Credit (in baht)	14,109	40,762	21,392	33,581
Non-agricultural Business				
Business Owner (Incidence)	0.407	0.491	0.499	0.500
Number of Business	0.588	0.871	0.733	0.919
All Labor	1.266	2.298	1.653	2.564
Household Labor	1.175	1.962	1.496	2.092
Market Labor	0.135	0.938	0.153	1.269
Business Asset (Stock, in baht)	33,177	216,003	32,634	203,893
Investment Expenditure (in baht)	3,588	45,694	3,766	33,922
House as Business Asset (Incidence)	0.115	0.319	0.162	0.369
Vehicle as Business Asset (Incidence)	0.044	0.205	0.045	0.206
Consumption				
Total Consumption P.C. (in baht)	9,666	10,096	10,109	13,098
Regular Consumption P.C. (in baht)	8,077	6,854	8,066	6,681
House and Vehicle Rep. Expenditure P.C. (in baht)	1,588	6,559	2,042	10,437
House Repair Expenditure (Incidence)	0.208	0.406	0.221	0.415
Vehicle Repair Expenditure (Incidence)	0.551	0.498	0.606	0.489
House Rep. Exp. P.C. (in baht)	1,188	6,417	1,547	10,034
Vehicle Rep. Exp. P.C. (in baht)	400	1,198	495	1,906
Agricultural Assets				
House as Agri. Asset (Incidence)	0.026	0.159	0.031	0.174
Vehicle as Agri. Asset (Incidence)	0.014	0.119	0.010	0.099
Agri. Asset (Stock, in baht)	48,838	82,299	57,291	95,535

2.3.3 Econometric Approach

Short-run Analysis

Following Kaboski and Townsend (2012), we use the following econometric specification to identify the the Village Fund's effects in the short-run (i.e., two years after the inception of the program),

$$y_{ivpt} = \beta * Intensity_v * Post_t + X'_{ivpt}\gamma + \eta_{ivp} + \alpha_t + \theta_{pt} + u_{ivpt}, \tag{2.1}$$

where y_{ivpt} is an outcome of household i from village v and province p in year t. $Intensity_v$ is the inverse of village v's size in 2001.[5] We scale up the intensity to the unit of 10,000 baht per household, which is approximately the average intensity in post-program years. $Post_t$ is a dummy variable for post-intervention period (i.e., year 2002 and 2003. The reference period is from year 2000 to 2001).[6] X_{ivpt} is a vector of household demographic controls, including numbers of children, female and male adults, household head's age and age squared, head's years of education. Household fixed effects, η_{ivp}, capture the unobserved household time-invariant characteristics, and encapsulate the village fixed effects which include village size. Year fixed effects, α_t, captures other secular changes unobserved by the researcher. We also include the province-year fixed effects, θ_{pt}, which control for the time-variant factors at the province level, e.g. purchasing power. We cluster the standard error at the village level.

In this estimation framework, the coefficient β identifies the marginal intent-to-treat (MITT) effect of the program, i.e., the effect of increasing the fund intensity by 10,000 baht per household. The effect identified here differs from the average intent-to-treat effect (AITT), which, for its identification, requires a control group of untreated villages. However, untreated villages do not exist in our sample.

For heterogeneous program effects by initial wealth, we use the following specifica-

[5] We do not use contemporaneous village size to avoid endogeneity issues regarding village size in post-program years. Nonetheless, all our results are robust to using contemporaneous village size.

[6] Our results are robust to alternative choices of period. Precisely, our eight alternative choices are, 1998-2003, 1998-2004, 1998-2005, 1999-2003, 1999-2004, 1999-2005, 2000-2004, 2000-2005. Compared to these alternatives, our results yield the most conservative estimates for most of the outcome variables.

tion,

$$\begin{aligned} y_{ivpt} &= \delta * Intensity_v * Post_t * Highwealth_{ivp} + (X'_{ivpt}\sigma) * Highwealth_{ivp} \\ &+ \beta * Intensity_v * Post_t + X'_{ivpt}\gamma + \eta_{ivp} + \alpha_t + \theta_{pt} + u_{ivpt}, \end{aligned} \tag{2.2}$$

where $Highwealth_{ivp}$ is a dummy variable equal to one if a household i's initial wealth is above the median wealth in its village v. Thus the coefficient δ captures the program effect heterogeneity, i.e., the additional treatment effect for the initially wealthy households, relevant to the initially poor households.

Long-run Analysis

In our long-run analysis, we group every two years into a period. Thus we have one pre-intervention period (i.e., 2000-2001) and five post-intervention periods (i.e., 2002-2003, 2004-2005,2006-2007,2008-2009 and 2010-2011). We use the following specification,

$$y_{ivpt} = \sum_{\tau=1}^{5} \beta_\tau * Intensity_v * PERIOD_t^\tau + X'_{ivpt}\gamma + \eta_{ivp} + \alpha_t + \theta_{pt} + u_{ivpt}, \tag{2.3}$$

where $PERIOD_t^\tau$ are dummy variables that indicate year t belongs to period τ, with τ indicating one of the five post-intervention periods. The reference period is the pre-intervention period (i.e., from 2000 to 2001). Other variables' definitions remain unchanged. β_τ thus identifies the marginal program effects in period τ.

Village-level Analysis

We also conduct a village-level analysis of the program effect on the intra-village inequality. The specification takes the following form,

$$WRCShare_{vpt} = \sum_{\tau=1}^{5} \beta_\tau * Intensity_v * PERIOD_t^\tau + X'_{vpt}\gamma + \eta_{vp} + \alpha_t + \theta_{pt} + u_{vpt}, \tag{2.4}$$

where $WRCShare_{vpt}$ is the wealthy households' share of regular consumption in village v, province p and year t. Precisely,

$$WRCShare_{vpt} = \frac{\sum_{i=1}^{15} RC_{ivpt} * Highwealth_{ivp}}{\sum_{i=1}^{15} RC_{ivpt}}$$

where RC_{ivpt} refers to the regular consumption in baht of the household i in village v province p and year t. There are 15 households per village. Model (2.4) is practically the same as model (2.3), with X_{vpt} being a vector of village level controls, which are the village means of the X_{ivpt} in model (2.3).

2.4 Results

Using model (2.1), we report the program effects, β, on a list of outcomes for an average household in column (1) of table 2.3. Separating the full sample by initial wealth, column (2) and (3) report the effects for an initially wealthy and poor household, respectively. Column (4) reports the corresponding effect heterogeneity, δ, using model (2.2).

2.4.1 Financial Inclusion

We begin by showing the effects on households' borrowings. Consistent with Kaboski and Townsend (2012) as well as other authors' (Menkhoff and Rungruxsirivorn 2011, Boonperm et al. 2013) results, we find a substantial increase in access to credit along the entire income distribution without crowding-out of other forms of credit. As pointed out previously, the "village fund" channel was rarely used by the households for obtaining credit. We estimate an average take-up of the program of 57 percentage points by the simple pre- and post-program difference in the incidence of borrowing from the village fund channel as indicated in summary statistics (Table 2.2). In comparison, other studies of smaller-scale microcredit programs typically find average take-up rates of around 20 percent (Tarozzi et al. 2015, Banerjee et al. 2015a, Crépon et al. 2015, Angelucci et al. 2015). In terms of the marginal effect, column (1) shows that a marginal increase of 10,000 baht per household in the program fund increases the Village Fund take-up rate by 20 percentage points (row 1). Regarding the loan size, an average household borrows 6,640 baht from the program for every 10,000 baht per household injected in the village fund (row 2). Comparing column (2) to (3), the wealthy households have a slightly lower take-up but a larger loan size from the program. However, as shown in column (4), the difference in the point estimates between the two samples are insignificant, indicating that the program reaches out to the wealthy and poor alike. Furthermore, the program does not crowd out informal or formal sources of borrowing (row 3 and 4).[7] Overall, the program has significantly increased the total borrowings of an average household by 6,600 baht (row 5), a value that is 47% of the total borrowings of an average household before the program.

2.4.2 Entrepreneurship

One of the key promises of microfinance is that access to microcredit promotes growth among the small and medium-scale entrepreneurs in developing countries. This argument is based on the hypothesis that the small and medium-scale entrepreneurs in this setting have large returns to capital but are financially constrained (de Mel et al., 2008). When relaxing this constraint, these enterprises prosper by harnessing the large returns to capital. Microcredit thus provides a solution to relax the constraint. However, the empirical evidence regarding the effects of access to credit in this connection is not conclusive (Angelucci et al. 2015, Tarozzi et al. 2015, Banerjee et al. 2015a Kaboski and Townsend 2012, Burgess and Pande 2003).

In our analysis, we focus on the non-agricultural businesses, which include general shop, mechanic/repair service, hair salon/barber, rice miller, shrimp farmer and fish farmer, vehicle rental, trader and restaurant. Households can operate multiple businesses at the same time, and we define a business owner as a household owning at least one non-agricultural business. We analyze the effects on the entrepreneurial activity on both extensive and intensive margin.

On the extensive margin, we find that the program has improved the share of business owning households by 8 percentage points (row 6), a 20 percent increase relative to the baseline level. On the other hand, we do not find effects on any intensive margin measures, i.e., number of businesses, number of workers, value of business asset, investment expenditure (row 7 to 12).

However, the heterogeneity analysis reveals that the initially wealthy households experience a significant business expansion while the poor households remain silent (column 2 and 3). Precisely, we find that, in response to a marginal increase of 10,000 baht per household in the Village Fund program, the initially wealthy households raised the num-

[7] "Formal credits" include the credits borrowed from agricultural cooperatives, production credit group, commercial banks and rice bank. "Informal credits" include the credits from neighbor, relative, moneylender, store owner, supplier, purchaser, landlord and other unspecified.

ber of businesses by 0.157, a value that is 21% of the average number of businesses among wealthy households before the program (row 7). They also increased the total labor by 0.355 persons, a 20% improvement from the baseline value (row 8).[8] This expansion of labor employment is driven by household labor (row 9), rather than market labor (row 10). Moreover, the program increased the total value of business assets of the wealthy households by 73 percentage points (row 11).

While the wealthy households manage to accumulate their business assets stock, they do not achieve this by purchasing new business assets, as the effect on "Investment Expenditure" is not significant (row 12). There are two possible explanations. First, the increase in business asset holdings is leveraged by wealthy households. For example, if a household has been saving toward the purchase of a truck, the loan from the Village Fund facilitates the purchase together with the accumulated savings. This pattern, however is not consistent with the lack of impact on asset purchases ("Investment Expenditure").

The second possibility is that the increase in business assets results from re-dedicating preexisting household assets (e.g. vehicles or part of the household's real estate) that were not used for business purposes previously. And the Village Fund facilitates the adaptation of these assets to business purpose by financing the cost of adaptation (e.g. house renovation). As discussed previously, it is not uncommon to see a business operating in the household's residence house and using household's vehicles (e.g. motorcycle). A typical example is the street food vendor, who would use part of the residence house as cooking and serving area, or use a motorcycle to transport the cooking utensils and ingredients to the sale spot. To explore this possibility, we investigate whether wealthy households are more likely to report house or vehicle as business assets. Column (2) of row (13) and (14) strongly support this hypothesis. They show that the Village Fund increases wealthy households' propensity of declaring real estate or vehicles as business assets by 7 and 3 percentage points, respectively, which corresponds to a 44 percent and a 55 percent increase from pre-intervention levels.[9] Accordingly, only wealthy households witnessed significant program effects on vehicle and house repairs, for which expenditures almost double (an additional 65% increase, to be exact) as shown in row (15).

Similar effects are lacking on agricultural production activity, as we do not observe

[8] 0.756 and 1.735 are the means of business assets value and all labor employed, respectively, among the initially wealthy households at the baseline.

[9] The means of reporting house and vehicle as business assets are 0.158 and 0.06, respectively, for the wealthy households.

any agricultural assets accumulation or re-dedication of house and vehicle to agricultural activity (row 16 to 18).

Overall, our results on entrepreneurship suggest that non-agricultural business expansion activities are concentrated among wealthy households. They expand their household business by adapting their preexisting house and vehicle to business use and thus occur higher expenditure on house and vehicle repair, which are essentially hidden business investments.

2.4.3 Consumption

To analyze the impact on welfare, we focus on the effect on regular consumption, which is total consumption net of house and vehicle repair expenditures. Measured in logarithmic term, we do not find significant effect on regular consumption per capita for an average household. However, consistent with the pattern of heterogeneity regarding entrepreneurial activity, we find that the program significantly improves the welfare only for the wealthy households. Specifically, as shown in row (20) of Table 2.3, an marginal increase in the fund intensity by 10,000 baht per household improves the wealthy households' per capita regular consumption by 10% (column 2). In comparison, the effect is mute for the poor households (column 3). To make the comparison more robust, the heterogeneous effect on regular consumption between the wealthy and poor households are also statistically significant (column 4).

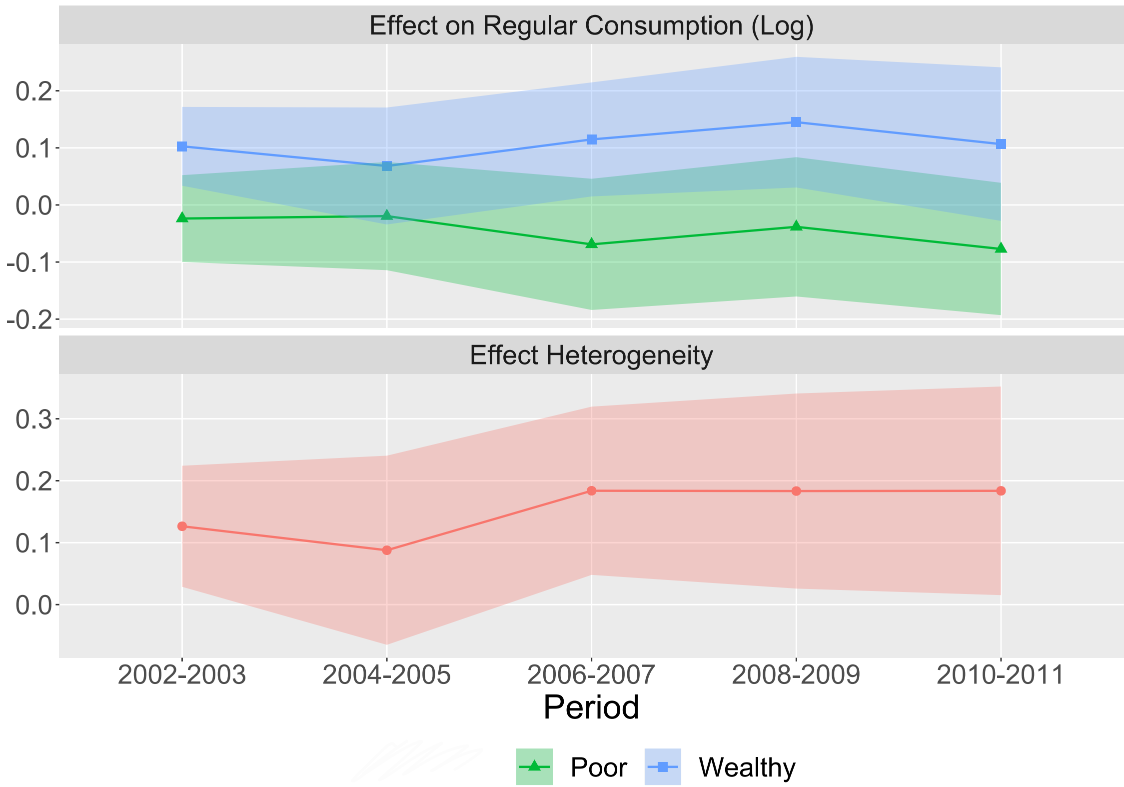
We extent the welfare analysis to long-term, i.e., to year 2011. We group every two years into a period, thus we have one pre-intervention period (i.e., 2000-2001) and five post-intervention periods. Using model (2.3), we estimate the effects on the regular consumption in each of the five post-intervention periods, for the wealthy and poor households separately. For ease of reading, we plot the point estimates of β_τ and the corresponding 90% confidence intervals in the upper panel of Figure 2.3. As we can read in the plot, the pattern of impact heterogeneity from the short-term analysis persists to 10 years after the inception of the program. While the poor households' welfare remains stagnant (line with triangular markers), the wealthy households experience a sustained increase of 10 percent in regular consumption (line with square markers). Moreover, the gap between these two lines slightly widens along the years. More formally, we test whether the differ-

Table 2.3: The Village Fund Program's Short-term Effects

	β from model (2.1)			δ from model (2.2)
Sample	Full	Wealthy	Poor	Full
	(1)	(2)	(3)	(4)
Financial Inclusion				
(1) Borrow from VF (Incidence)	0.201∗∗∗	0.192∗∗∗	0.227∗∗∗	-0.034
	(0.053)	(0.066)	(0.068)	(0.084)
(2) VF Credit (in Baht)	6641∗∗∗	7790∗∗∗	5508∗∗∗	2282
	(1063)	(1371)	(1193)	(1550)
(3) Formal Credit (in Baht)	2196	6921	-2657	9579∗∗
	(3111)	(4623)	(2240)	(4238)
(4) Informal Credit (in Baht)	-2269	-4625	715	-5341
	(2462)	(4322)	(1014)	(4359)
(5) Total Credit (in Baht)	6567∗	10086∗	3566	6520
	(3736)	(6039)	(2462)	(6224)
Entrepreneurship				
(6) Business Owner (Incidence)	0.083∗	0.081	0.088	-0.006
	(0.047)	(0.058)	(0.063)	(0.074)
(7) Number of Business	0.093	0.157∗	0.02	0.136
	(0.07)	(0.088)	(0.095)	(0.117)
(8) All Labor	0.136	0.355∗	-0.138	0.493∗
	(0.136)	(0.187)	(0.192)	(0.262)
(9) Household Labor	0.125	0.305∗	-0.101	0.407∗
	(0.132)	(0.166)	(0.193)	(0.238)
(10) Market Labor	-0.008	-0.017	0.003	-0.02
	(0.038)	(0.075)	(0.013)	(0.082)
(11) Business Asset (Stock, in Baht, in log)	0.36	0.728∗	-0.083	0.811
	(0.284)	(0.42)	(0.364)	(0.561)
(12) Investment Expenditure (in Baht, in Log)	-0.073	-0.181	0.026	-0.208
	(0.241)	(0.337)	(0.252)	(0.36)
(13) House as Business Asset (Incidence)	0.036	0.07∗∗	-0.009	0.079∗∗
	(0.023)	(0.03)	(0.03)	(0.039)
(14) Vehicle as Business Asset (Incidence)	0.02	0.033∗	0.003	0.029
	(0.013)	(0.017)	(0.018)	(0.024)
(15) House & Vehicle Rep. Expend P.C.	0.329	0.65∗∗	-0.107	0.756∗
(in Baht, in log)	(0.23)	(0.321)	(0.293)	(0.411)
Agricultural Assets				
(16) House as Agri. Asset (Incidence)	0.007	-0.005	0.021	-0.026∗
	(0.011)	(0.013)	(0.013)	(0.015)
(17) Vehicle as Agri. Asset (Incidence)	0.006	0.013	-0.001	0.014
	(0.005)	(0.01)	(0.004)	(0.011)
(18) Agri. Asset (Stock, in Baht, in log)	0.299	0.219	0.387	-0.168
	(0.222)	(0.213)	(0.43)	(0.478)
Consumption				
(19) Total Consumption P.C. (in Baht, in log)	0.061	0.173∗∗∗	-0.081	0.254∗∗∗
	(0.045)	(0.055)	(0.061)	(0.074)
(20) Regular Consumption P.C. (in Baht, in log)	0.033	0.099∗∗	-0.048	0.147∗∗
	(0.033)	(0.042)	(0.051)	(0.062)

Note: Standard errors are clustered at the village level and given in parentheses. * Significant at 10-percent level. ** Significant at 5-percent level. *** Significant at 1-percent level.

ences in these point estimates between the wealthy and poor households are significantly different from zero. The point estimates of the differences and their corresponding 90% confidence intervals are plotted in the lower panel. It shows that the wealthy households benefit more than the poor by about 10 percentage points in the period of 2002-2003. This difference increases to about 20 percentage points in the period of 2010-2011. Moreover, these point estimates display an ascending trend of approximately 1.2 percentage points per period.



Note: The upper panel plots the point estimates of β_τs of model (2.3) using wealthy and poor households sample separately. The standard errors are clustered at the village level. The ribbons display the 90% confidence intervals of the corresponding point estimates. The effect heterogeneity is the estimate of the difference in effect on regular consumption (log) between wealthy and poor households, with the poor being the reference group.

Figure 2.3: Long-term Effects on Regular Consumption

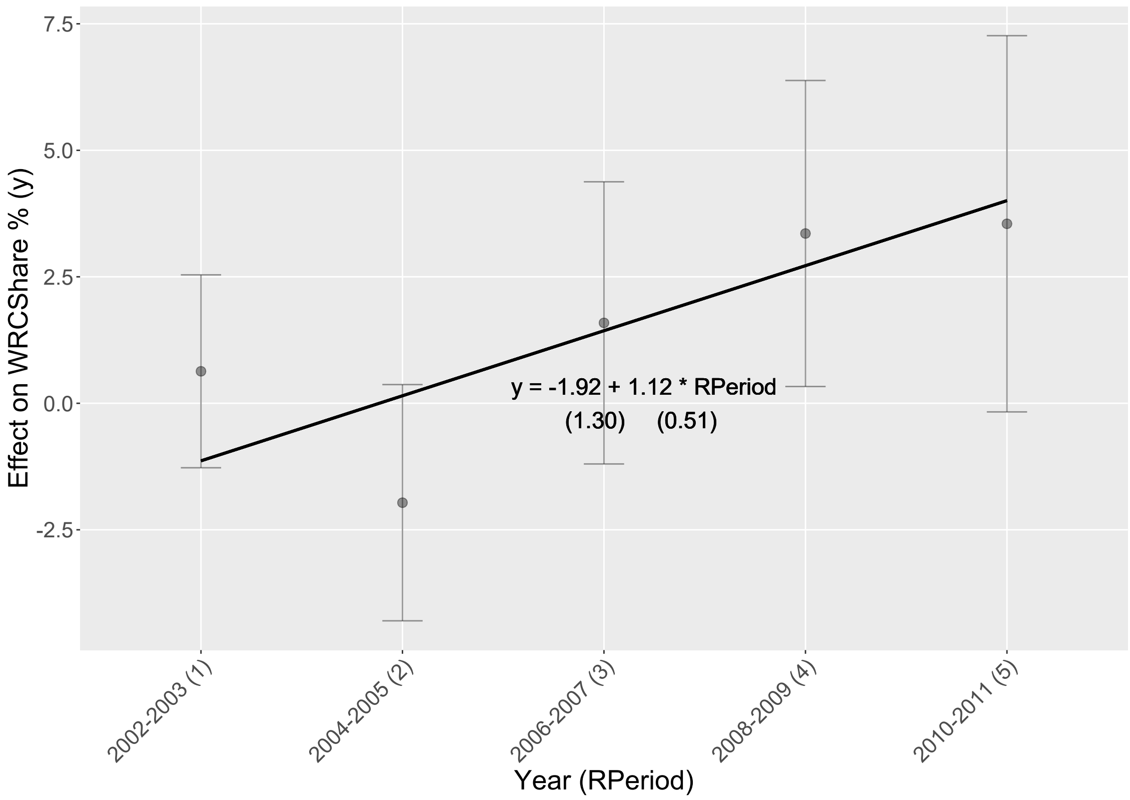
The persisting widening of heterogeneity of effect on welfare between the wealthy and poor households motivates us to investigate the effect on intra-village welfare inequality in the long-run. We measure the welfare inequality of a village by the wealthy households' share of the aggregate regular consumption in the village. Using model (2.4), we estimate the effects on this inequality indicator in each of the five post-intervention periods. The point estimates and their 90% confidence intervals are plotted in Figure 2.4.

Overall, the point estimates indicate that the wealthy households' share of regular consumption increase by around 1.4 percentage points for an average period. Although the point estimates are not precisely estimated, they display an obvious ascending trend. To test whether the trend is significantly ascending, we estimate a linear trend model as follow,

$$\begin{aligned} WRCShare_{vpt} = {} & \delta * Intensity_v * Post_t \\ & + \lambda * Intensity_v * RPeriod_t \\ & + X'_{vpt}\gamma + \eta_{vp} + \alpha_t + \theta_{pt} + u_{vpt} \end{aligned} \tag{2.5}$$

where $RPeriod_t$ is an integer variable whose range is between 0 and 5, with 0, 1, ..., 5 corresponding to the pre-intervention period, 2002-2003, ..., 2010-2011. Other variables' definitions remain unchanged. Intuitively, model (2.5) fits a line on the five β_τ estimates of model (2.4) with a slope of λ and intercept of δ. The estimated line is plotted in Figure 2.4. Confirming with the pattern of long-run regular consumption, the fitted line shows that the effects on the inequality exhibits an ascending trend of 1.12 percentage points per period (significant at 5% level), or approximately half percentage point per year, against a secular decrease during the ten post-intervention years (-1.92 percentage points). Overall, the long-run analysis on welfare suggests that the program sharpens the intra-village inequality.

To sum up, our results suggest that, by expanding business activities, wealthy households are able to enjoy a long term improvement in regular consumption. On the other hand, for poor households, we find neither an effect on business expansion nor an effect on consumption, despite of sizable gains regarding financial inclusion. More broadly speaking, our findings suggest that wealth and access to credit appear to be complements rather than substitutes for business expansion and long-term welfare gains.



Note: The standard errors are clustered at the village level. For model (4), the point estimates of the βs are plotted in grey dots, with their 90% confidence intervals displayed as error bars. For the linear trend model, model (2.5), the standard errors of the intercept and slope are shown in the parentheses.

Figure 2.4: Long-term Effects Consumption Inequality

2.5 Robustness Checks

2.5.1 Placebo Test

The difference-in-differences method of our analysis assumes a parallel trend among the compared groups, i.e., villages with different sizes (fund intensities) would witness same growth of outcome variables in the absence of the intervention. To test the validity of the parallel trend assumption, we perform a placebo test by running models (2.1) and (2.2) using only pre-intervention data, i.e., from 1998 to 2001. We treat the latter two years (i.e., 2000 and 2001) as pseudo-post-intervention years. The hypothesis that, if there is parallel trend, the estimates of β of model (2.1) and δ of model (2.2) should be zero. I report the placebo test results in Table 2.4. This placebo test finds no statistically result for all outcome variables. In fact, the point estimates also lack economic significance. For example, the estimate for "Borrow from VF" is only -0.004 and statistically insignificant, indicating that in the absence of the intervention a larger fund intensity (i.e., smaller village size) does not predict a larger propensity to borrow from the village fund channel. This confirms that the significant estimate of 0.2 in row (1) of Table 2.3 is driven by the intervention. Similarly, the placebo effects for the entrepreneurship and consumption variables among the wealthy households (column 2 of Table 2.4) are much smaller and closer to zero than their counterparts in Table 2.3. Overall, the placebo test results reinforce that the econometric model (2.1) is capable of identifying the program effects.

2.5.2 Impact Heterogeneity regarding Village Size

Village size could be a source of impact heterogeneity. For example, the smaller the village, the less competitive the supply side, and therefore, the more intensely the Village Fund affects business expansion. On the other hand, a smaller village has less demand, and thus the Village Fund affects business expansion less intensely. Therefore, whether there is impact heterogeneity by the village size is an empirical question. And one concern is that this impact heterogeneity biases the estimate of MITT for an average household. To address this concern, we now prove that even there exists impact heterogeneity regarding village size, our estimate of MITT still remains unbiased for households from small villages. For simplicity's sake, we specify the model as follows,

$$y_{ivt} = f(IP_{ivt}, S_{iv}) + X'_{ivt}\gamma + \alpha_{iv} + \eta_t + \varepsilon_{ivt} \tag{2.6}$$

where $IP_{ivt} = Intensity_{iv} * Post_t$ and $f(IP_{ivt}, S_{iv})$ is a general response function of a outcome variable with respect to fund intensity and village size for the household i from village v in year t. To approximate this response function, we apply a second degree taylor expansion around the average point (IP_{ivt}, S_{iv}), i.e., $(\overline{IP}, \overline{S})$. This is because we want to assess the treatment effect for an average household. For ease of readability, we drop the subscripts i and v. The approximated response function is thus:

Table 2.4: Placebo Test

Sample	β from model (2.1)			δ from model (2.2)
	Full	Wealthy	Poor	Full
	(1)	(2)	(3)	(4)
Financial Inclusion				
(1) Borrow from VF (Incidence)	-0.004	-0.003	-0.006	0.003
	(0.006)	(0.01)	(0.011)	(0.016)
(2) VF Credit (in Baht)	-114	-177	-44	-133
	(84)	(155)	(40)	(163)
(3) Formal Credit (in Baht)	1241	1588	818	770
	(2052)	(3081)	(1865)	(3143)
(4) Informal Credit (in Baht)	1790	3826	-569	4396
	(2759)	(5298)	(1323)	(5903)
(5) Total Credit (in Baht)	2917	5237	203	5033
	(3564)	(6541)	(2164)	(7167)
Entrepreneurship				
(6) Business Owner (Incidence)	-0.014	-0.006	-0.026	0.02
	(0.042)	(0.052)	(0.051)	(0.058)
(7) Number of Business	0.037	0.086	-0.023	0.109
	(0.058)	(0.082)	(0.063)	(0.09)
(8) All Labor	0.059	0.074	0.012	0.062
	(0.121)	(0.192)	(0.141)	(0.237)
(9) Household Labor	0.052	0.083	-0.003	0.087
	(0.113)	(0.168)	(0.143)	(0.218)
(10) Market Labor	0.008	-0.002	0.009	-0.011
	(0.036)	(0.066)	(0.022)	(0.074)
(11) Business Asset (Stock, in Baht, in log)	0.005	-0.042	0.055	-0.097
	(0.274)	(0.359)	(0.371)	(0.48)
(12) Investment Expenditure (in Baht, in Log)	0.09	0.034	0.113	-0.079
	(0.217)	(0.309)	(0.2)	(0.32)
(13) House as Business Asset (Incidence)	-0.01	-0.025	0.004	-0.03
	(0.034)	(0.043)	(0.033)	(0.04)
(14) Vehicle as Business Asset (Incidence)	-0.003	0.011	-0.021	0.031
	(0.013)	(0.02)	(0.013)	(0.021)
(15) House and Vehicle Rep. Expenditure P.C.	-0.156	-0.194	-0.106	-0.088
(in Baht, in log)	(0.266)	(0.403)	(0.433)	(0.633)
Agricultural Assets				
(16) House as Agri. Asset (Incidence)	0.003	0.015	-0.013	0.028
	(0.009)	(0.013)	(0.01)	(0.017)
(17) Vehicle as Agri. Asset (Incidence)	-0.003	-0.002	-0.004	0.002
	(0.007)	(0.011)	(0.007)	(0.013)
(18) Agri. Asset (Stock, in Baht, in log)	-0.079	-0.067	-0.133	0.067
	(0.196)	(0.304)	(0.229)	(0.384)
Consumption				
(19) Total Consumption P.C. (in Baht, in log)	-0.01	-0.04	0.017	-0.057
	(0.05)	(0.069)	(0.07)	(0.098)
(20) Regular Consumption P.C. (in Baht, in log)	0.01	0.001	0.014	-0.013
	(0.056)	(0.088)	(0.058)	(0.103)

Note: This placebo test is based on the data from 1998 to 2001 with 2000 and 2001 regarded as pseudo-post-intervention period. Standard errors are clustered at the village level and given in parentheses. * Significant at 10-percent level. ** Significant at 5-percent level. *** Significant at 1-percent level.

$$
\begin{aligned}
f(IP,S) &= f(\overline{IP},\overline{S})+f_1(\overline{IP},\overline{S})*(IP-\overline{IP})+f_2(\overline{IP},\overline{S})*(S-\overline{S}) \\
&+ \frac{1}{2}f_{11}(\overline{IP},\overline{S})*(IP-\overline{IP})^2+\frac{1}{2}f_{22}(\overline{IP},\overline{S})*(S-\overline{S})^2 \\
&+ f_{12}(\overline{IP},\overline{S})*(IP-\overline{IP})*(S-\overline{S}) \qquad (2.7)
\end{aligned}
$$

where $f_1(\overline{IP},\overline{S})$ and $f_2(\overline{IP},\overline{S})$ are the first partial derivatives with respect to IP and S, respectively. Similarly, $f_{11}(\overline{IP},\overline{S})$ and $f_{22}(\overline{IP},\overline{S})$ are the second partial derivatives with respect to IP and S, respectively. And $f_{12}(\overline{IP},\overline{S})$ is the heterogeneous effect by village size. All derivatives are evaluating at $(\overline{IP},\overline{S})$.

When we plug equation (2.7) in (2.6), $f(\overline{IP},\overline{S})$ is immediately captured by time fixed effects η_t. Moreover, $f_2(\overline{IP},\overline{S})$ and $f_{22}(\overline{IP},\overline{S})$ are encapsulated in household fixed effects.

Then the real MITT with respect to IP is:

$$
MITT=\frac{\partial f(IP,S)}{\partial IP}=f_1(\overline{IP},\overline{S})+f_{11}(\overline{IP},\overline{S})*(IP-\overline{IP})+f_{12}(\overline{IP},\overline{S})*(S-\overline{S}) \qquad (2.8)
$$

However, we are unable to disentangle IP and S as we defined S=$\frac{1}{IP}$. We would actually estimate the following response function, which is obtained by plugging in S=$\frac{1}{IP}$ in equation (2.7).

$$
\begin{aligned}
\widetilde{f}(IP,S) &= f(\overline{IP},\overline{S})+f_1(\overline{IP},\overline{S})*(IP-\overline{IP})+f_2(\overline{IP},\overline{S})*(S-\overline{S}) \\
&+ \frac{1}{2}f_{11}(\overline{IP},\overline{S})*(IP-\overline{IP})^2+\frac{1}{2}f_{22}(\overline{IP},\overline{S})*(S-\overline{S})^2 \\
&+ f_{12}(\overline{IP},\overline{S})*(1+\overline{IPS}-\overline{IP}S-IP\overline{S}) \qquad (2.9)
\end{aligned}
$$

Therefore, the estimated MITT becomes:

$$
\widetilde{MITT}=\frac{\partial \widetilde{f}(IP,S)}{\partial IP}=f_1(\overline{IP},\overline{S})+f_{11}(\overline{IP},\overline{S})*(IP-\overline{IP})-f_{12}(\overline{IP},\overline{S})*\overline{S} \qquad (2.10)
$$

We can immediatelt see that,

$$
\begin{cases}
\widetilde{MITT}=MITT, & \text{if } f_{12}=0 \text{ or } S=0 \\
\widetilde{MITT}\neq MITT, & \text{otherwise}
\end{cases}
$$

That is, the estimated effect, $\widetilde{MITT}$, is unbiased when there exist no impact heterogeneity by village size (i.e., $f_{12} = 0$) or when village size is small (i.e., S=0).

2.5.3 Nonlinear Treatment Effects

Another interest of estimate is the average intent-to-treat effect (AITT). However, our linear identification strategy cannot distinguish AITT from marginal intent-to-treat effect (MITT). Therefore, we augment the model to a quadratic form to analyze the nonlinear treatment effects for an average household. We will discuss the implied AITT and MITT as well as whether there is significant difference between them. The augmented model is specified as follows:

$$y_{ivpt} = \beta_1 Intensity_v * Post_t + \beta_2 Intensity_v^2 * Post_t + X_{ivpt}'\gamma + \eta_{ivp} + \alpha_t + \theta_{pt} + \varepsilon_{ivpt} \quad (2.11)$$

We scale up the unit of "Intensity" to "10,000 baht per household". Given the average fund intensity being approximately 10,000 baht per household (Table 2.2), this nonlinear model implies the following estimates:

$$\begin{cases} AITT = \beta_1 + \beta_2 & \text{,i.e., ITT at "Intensity"=1 (in 10,000 baht per household)} \\ MITT_0 = \beta_1 & \text{,i.e., MITT at "Intensity"=0} \\ MITT_1 = \beta_1 + 2 * \beta_2 & \text{,i.e., MITT at "Intensity"=1} \end{cases}$$

Moreover, the statistical significance of β_2 tells the significance of the difference between $AITT$ and $MITT_1$, i.e., whether there is difference between the average and marginal effect.[10] Given that our main results are driven by the wealthy households, we test the nonlinearity of effects only for the wealthy households. The results are presented in table 2.5. We also report the p-value of the joint-F test on β_1 and β_2.

For those variables whose estimates are significant in Table 2.3, we also find non-zero treatment effects using the nonlinear model, as suggested by the p-values of the joint-F tests being smaller than 0.1. Robustly, the estimates of $MITT_1$ of the nonlinear model are very close to the MITT (i.e., column (2) in table 2.3) estimated using our main specification (model 2.1)

[10]It also shows the difference between $MITT_0$ and $MITT_1$, but it is not the focus of our discussion in this section.

Regarding the existence of nonlinear effects, the estimates of β_2 are not significant and the magnitudes are rather small for the financial inclusion and consumption variables, suggesting that the correlation between fund intensity and village fund credit (row 1 and 2) and consumption (row 20) are fairly linear. Both marginal and average effects are significantly different from zero (row 1 and 2). In terms of the entrepreneurship variables, there is also a lack of nonlinear effects on business ownership, number of businesses owned and using vehicle for business purpose (the estimates of β_2 are also insignificant and small for row 6, 7 and 14). In comparison, the β_2 are less precisely estimated for business assets and labor, as indicated by the large standard errors (row 8, 11, 13 and 15). Moreover, the magnitudes are also large, which may imply zero average effects (the $AITT$ are very close to zero in row 11, 13 and 15).

Overall, we find that the intervention's impacts are fairly linear for wealthy households' borrowings, business ownership and regular consumption. However, the nonlinear effects might exist for business assets and labor.

2.6 Conclusion

This paper evaluates the impacts of Thailand's Village Fund program, a large scale decentralized government lending program, on entrepreneurship and welfare. We find that, while the impacts for an average household appears to be moderate, there is substantial impact heterogeneity along the wealth distribution.

More precisely, while initially wealthy households improve their consumption both in the short- and long-run, initially poor households do not reap off similar welfare gains. This explains the sharpening of intra-village inequality in consumption in the long-run. We argue that business expansion is likely to be the mechanism behind these heterogeneous welfare gains. In support of our argument, we find that wealthy households accumulate business assets by adapting preexisting house and vehicle to business use, rather than acquiring new assets. Precisely, we find that the Village Fund increases the expenditure on house and vehicle repairs among the wealthy households. On the contrary, the initially poor households do not achieve similar gains regarding business expansion, even though they benefit from the Village Fund in terms of financial inclusion. One obvious question is why there is a lack of impacts on business expansion for the poor. One expla-

Table 2.5: Nonlinear Treatment Effects for Wealthy Households

	β_1	β_2	Joint-F	$MITT_1$	$AITT$
Financial Inclusion					
(1) Borrow from VF (Incidence)	0.44	-0.11		0.22***	0.33**
	(0.28)	(0.12)	[0.00]	[0.00]	[0.05]
(2) VF Credit (in Baht)	11969**	-1892		8184***	10076***
	(5293)	(2373)	[0.00]	[0.00]	[0.00]
(3) Formal Credit (in Baht)	-7545	6552		5558	-993
	(20524)	(8991)	[0.16]	[0.22]	[0.93]
(4) Informal Credit (in Baht)	15347*	-9045**		-2743	6301
	(8008)	(4154)	[0.10]	[0.40]	[0.17]
(5) Total Credit (in Baht)	19770	-4385		10999*	15384
	(23262)	(10142)	[0.15]	[0.06]	[0.26]
Entrepreneurship					
(6) Business Owner (Incidence)	0.22	-0.06		0.09*	0.15
	(0.19)	(0.09)	[0.18]	[0.07]	[0.16]
(7) Number of Business	0.15	0.004		0.16*	0.15
	(0.36)	(0.16)	[0.13]	[0.07]	[0.47]
(8) All Labor	1.26	-0.41		0.44***	0.85*
	(0.78)	(0.35)	[0.04]	[0.01]	[0.06]
(9) Household Labor	0.20	0.05		0.30**	0.25
	(0.65)	(0.29)	[0.11]	[0.05]	[0.50]
(10) Market Labor	0.53	-0.25		0.03	0.28
	(0.61)	(0.26)	[0.49]	[0.73]	[0.42]
(11) Business Asset (Stock, in Baht, in log)	-0.83	0.70		0.58*	-0.12
	(1.28)	(0.60)	[0.11]	[0.10]	[0.87]
(12) Investment Expenditure (in Baht, in Log)	-0.40	0.10		-0.20	-0.30
	(1.19)	(0.58)	[0.73]	[0.44]	[0.64]
(13) House as Business Asset (Incidence)	-0.11	0.08*		0.05**	-0.03
	(0.10)	(0.04)	[0.00]	[0.05]	[0.61]
(14) Vehicle as Business Asset (Incidence)	0.01	0.01		0.03**	0.02
	(0.07)	(0.03)	[0.08]	[0.05]	[0.55]
(15) House and Vehicle Rep. Expenditure P.C.	-0.47	0.51		0.54*	0.04
	(1.16)	(0.50)	[0.03]	[0.07]	[0.95]
Agricultural Assets					
(16) House as Agri. Asset (Incidence)	0.17**	-0.08**		0.01	0.09*
	(0.08)	(0.03)	[0.02]	[0.45]	[0.06]
(17) Vehicle as Agri. Asset (Incidence)	0.04	-0.01		0.02	0.03
	(0.05)	(0.02)	[0.28]	[0.18]	[0.39]
(18) Agri. Asset (Stock, in Baht, in log)	-0.03	0.11		0.20	0.09
	(0.76)	(0.33)	[0.47]	[0.32]	[0.85]
Consumption					
(19) Total Consumption P.C. (in Baht, in log)	0.25	-0.04		0.18***	0.22
	(0.25)	(0.10)	[0.00]	[0.00]	[0.14]
(20) Regular Consumption P.C. (in Baht, in log)	0.16	-0.03		0.10**	0.13
	(0.19)	(0.08)	[0.02]	[0.02]	[0.24]

Note: To analogize to the analysis in column (2) table (2.3), we use data of the wealthy households from 2000 to 2003. The standard errors are clustered at the village level and given in parentheses. β_1 and β_2 refer to the coefficients of the linear and quadratic terms, respectively, of model (2.11). The "joint-F" test tests for the joint significance of β_1 and β_2 with its p-value given in brackets. * Significant at 10-percent level. ** Significant at 5-percent level. *** Significant at 1-percent level.

nation is that the poor use their loan for consumption smoothing rather than investment, or simply they are untalented entrepreneurs who exit their business within one year, thus the change in assets is not picked up by the annual survey.

This paper sheds light on the importance of initial wealth in channeling the effects of access to credit. Contrast to the conventional argument, initial wealth complements rather than substitutes the access to credit when it comes to business expansion. This might explain, why most of microcredit interventions fail to bring welfare gains to ultra-poor households, who are deprived of household assets. Therefore, to improve welfare for the ultra-poor through fostering entrepreneurship, interventions, such as a graduation program (Banerjee et al., 2015b), might be more effective than microcredit.

At the end, another point worth mentioning is that, the credit from Village Fund has a consumption dividend of 10 percent for a wealthy household. The question followed is, why households, especially wealthy households, did not reap off such a high consumption dividend by obtaining credit from other financial institutions? This might imply a high borrowing cost from other financial institutions. Further studies on this question would be desirable.

Appendix

2.A Test for the Change in Village Size

In this section, we test whether households migrate to smaller villages where fund is more intensive. If the statement is true, it implies that villages with a larger size at the baseline grows slower than those with a initially smaller size. This implication can be tested with the following model,

$$Size_{vt} = \beta * Size_v^{1998} * Post_t + \alpha_v + \eta_t + u_{vt} \tag{2.12}$$

where $Size_{vt}$ is the size of village v in year t. $Size_v^{1998}$ measures the initial size of village v as its size in 1998. $Post_t$ is a dummy variable that equals to 1 if year t is post the inception of the intervention, i.e., year 2002. α_v and η_t are village and year fixed effects. u_{vt} is the error term. β delivers the correlation between the initial village size and the temporal change in village size. We cluster the standard errors at the village level. We use the sample from 1998 to 2003.

The estimate of β is presented in Table 2.6. The point estimate is -0.031, meaning that an addition of one household in 1998 is associated with a smaller change in the village size by 0.031 households. However, the magnitude of this point estimate is rather nominal. Furthermore, the point estimate is not statistically significant. Taken together, this result suggests that there is no significant correlation between the initial village size and the temporal change in village size. This supports our argument that there is no migration to smaller villages.

Table 2.6: Test for the Change in Village Size

	Dependent variable:
	Village Size
Initial Village Size * Post	−0.031
	(0.072)
Observations	384
Adjusted R^2	0.877

Note: An observation is a village. Standard errors are clustered at the village level and reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

CHAPTER 3

A Financial Inclusion Policy Gone Wrong: Pakistan's Bank Branch Expansion Reform

Abstract Government interventions for expanding bank branch networks into rural areas have played a prominent role in the development strategies of numerous low-income countries. Several studies have documented beneficial effects of such policies on financial access and poverty reduction, especially for India. In this paper I examine the effects of Pakistan's rural bank branch expansion policy, by which commercial banks had to open at least 20 percent of their new branches in "Rural and Underserved Areas (RUA)". I show that banks complied with the policy but that the bulk of additional RUA branches is located on the fringes of urban centers, where dense branch networks already existed. This has sharpened spatial inequalities in bank branch coverage between more and less remote rural areas. As a result, the policy failed to reach out to unbanked rural areas and the fraction of the country's banked rural population stayed stagnant at 57%. Consistent with this branch-opening pattern, I also show that the policy has failed to expand credit disbursement to the agricultural sector and reduce poverty. My results are consistent with commercial banks' reluctance to expand their branch networks into remote rural areas. They demonstrate the need for more carefully designed government interventions for improving financial access than Pakistan's arguably naïve bank branch expansion policy.

3.1 Introduction

Access to finance has been seen as a key factor in economic growth as it provides facilities for mobilizing savings and accumulating capital stock, spurring investment and mitigating risk (Karlan and Morduch, 2010). A positive correlation between financial development and economic growth has been documented in both cross- and within-country studies (King and Levine 1993, Jayaratne and Strahan 1996, Rajan and Zingales 1998, Levine and Zervos 1998, Bekaert et al. 2005). However, it is well known that rural households in low-income countries have traditionally faced large impediments to access to formal finance (Banerjee and Duflo, 2007). In response, financially including the rural households into formal finance can thus bring efficiency and equity gains. The importance of financial inclusion has also been echoed in the UN's 2030 agenda which has prioritized financial inclusion as an enabler for achieving the "Sustainable Development Goals" (United Nations, 2015).

These arguments have often provided justification for widespread financial inclusion interventions. While the last two decades have seen the rise of microfinance promoted by NGOs, the majority of financial inclusion interventions remain state-led. These state-led interventions have mostly taken the forms of subsidized credit, no-frills savings accounts, mobile banking and rural bank network expansion (Demirgüç-Kunt et al. 2017). Rural bank network expansion has been a popular policy in developing countries since the 60s.[1] Such a policy mandates banks to open branches in underserved locations. The rationale is backed by a line of empirical studies that have established a positive correlation between closer geographical proximity to financial institutions and higher financial services take-up (Petersen and Rajan 1994, Degryse and Ongena 2005, Agarwal and Hauswald 2010, Brevoort et al. 2012, Nguyen 2019, Brown et al. 2016, Carletti et al. 2018). On the other hand, rural banking policies have also been widely believed to be ineffective as the political, rather than economic, considerations determine the flow of resources across sectors and individuals, rendering the banking sector more susceptible to elite capture and even undermine rural development (La Porta et al. 2002, Sapienza 2004). Despite its popularity and these scepticisms, rural expansion policy has mostly escaped rigorous evaluation regarding its effectiveness on financial access and poverty reduction. An exception is the

[1] Kenya in the 1970s, India in the 1970s, Bangladesh in the 1980s, Pakistan in the 2000s and Zambia in the 2010s

social banking intervention in India between 1977 and 1990. Burgess and Pande (2005) find that the social banking expansion in rural India has reduced rural poverty during the same period through a simultaneous expansion in credit disbursement to and saving mobilization from the rural sector. On the other hand, using microdata, Kochar (2011) criticizes that the rural banking expansion exacerbated consumption inequality as the benefit has been captured mainly by the rich.

Pakistan has a long history of rural financing policies which cost billions of rupees, however, she still remains one of the worst countries in terms of access to finance (State Bank of Pakistan, 2004). This dilemma is believed to be linked to the concentration of bank branches on urban areas. With bank branches being an important vehicle for delivering rural finance services and policies, e.g., subsidized credit for farmers and rural enterprises, having an extensive branch network is critical to include Pakistan's rural population, which make up 2/3 of the total population, into these services. Therefore, Pakistan has implemented the "Branch Expansion Policy (BEP)" reform since 2004 in the hope of enhancing the rural outreach of bank branch network. Being one of the largest interventions in the banking sector since its privatization in the 1990s, it is surprising that no evaluation has been done. This paper fills this gap and seeks to add to the thin evidence regarding the effectiveness of rural expansion policy on financial access and poverty reduction in the context of Pakistan. It also provides insights on Pakistan's disappointing status of financial inclusion.

In 2004, the State Bank of Pakistan (SBP), Pakistan's central bank, implemented the BEP which mandated the banks with a network size of no less than 100 branches to open 20% of their new branches in "Rural and Underserved Areas (RUA)" which were vaguely defined as "villages, small towns or unbanked tehsil headquarters".[2] This policy was expanded to include all banks regardless of their network sizes effectively in 2007.

In this paper, I draw on several sources for the data in this analysis, including the State Bank of Pakistan, Pakistan Bureau of Statistics, European Commission and the United States National Geospatial-Intelligence Agency. Using bank-branch-level geographical data, I track banks' network expansion along the policy timeline. I establish the policy-effect on the networks' rural population coverage using a simple trend break model. On an aggregate level, I identify the policy effects on branch accessibility, agricultural credit

[2]Pakistan has four levels of administrative division. They are, with descending order, province, district, tehsil and union council.

disbursement and poverty by estimating a difference-in-differences trend break model, which exploits the change in banks' selection on less remote rural locations for their RUA branches after the policy.

I find that the banks have complied with the policy by reserving 20% new branches for RUA locations, but the bulk of new RUA branches are in less remote locations, i.e., locations that are closer to district capitals and more likely to have other branches within 5 kilometers radius, than RUA branches opened prior to the reform. As a result, the policy failed to reach out to unbanked rural areas and the fraction of the country's rural population living within 5 kilometers radius of a bank branch (banked rural population) stagnated at 57 percent. A counterfactual analysis reveals that the banked rural population would have reached to 70 to 72 percent had the banks chosen locations with a higher "Rural Market Potential (RMP)", i.e., locations with a larger rural population per branch or unbanked locations with a larger rural population.

Consistent with the actual branch-opening pattern, the aggregate level analysis shows that the size of the RUA branch network has grown slower in districts with a larger rural population share after the policy. In other words, the growth rate of RUA branch network in a more rural district (e.g., "Lower Dir" whose rural population share is 97 percent) was surpassed by the less rural one (e.g., "Karachi" whose rural population share is 7 percent) by 0.86 branches per million persons per year after the reform. This accounts for an accumulative difference of 5.16 RUA branches per million persons between a more and less rural district by 2012, a value that is 66% of the average RUA branches intensity in 2000.

However, the number of borrowers and the amount of credit disbursed to agricultural sector, a sector that dominates rural Pakistan, did not grow differently between a more and less rural district in the wake of the reform. These results suggest that the policy has failed to expand credit disbursement to the rural sector, consistent with the branch network's stagnant rural population coverage. Finally, poverty has also evolved in parallel between more and less rural districts along the policy timeline. I conclude that the "Branch Expansion Policy" of Pakistan has failed to enhance the formal bank network's outreach to the rural population and reduce poverty because the new branches reserved by the 20 percent quota were opened in less remote rural locations where a branch network has already existed. More unfortunately, the reform has sharpened the geographical inequality in the number of RUA branches.

This study contributes to the literature on the importance of financial access in the process of economic development, more specifically, the effectiveness of rural branch expansion policies. While branch expansion policy is popular in developing countries, only the policy in India has received rigorous evaluations. Under India's policy, which was in effect between 1977 and 1990, banks, which had been nationalized in 1969, were required to open four branches in "unbanked" locations for every one branch opened in a "banked" location. Exploiting this rule, Burgess and Pande (2005) identify the poverty reduction effect of the policy by estimating the reduction in poverty in unbanked states between 1977 and 1990 relative to other periods and compare this to the corresponding change in poverty in banked states. They conclude that the expansion in banks resulted in a significant reduction in poverty through expanding credit provision to and saving mobilization from the rural sector. Using a regression discontinuity design, Young (2017) evaluates the policy that was revised in 2005 when the banks were required to open branches in district with a "branch insufficiency" measure higher than the national average (measured by population per branch) and finds that the policy has significantly increased the number of bank branches, bank accounts and crop yields in districts with baseline branch insufficiency just above the threshold. On the other hand, this policy also received criticism for its failure in reaching out to the ultra poor. Confining the analysis to the state of Uttar Pradesh, Kochar (2011) exploits the variation in "unbankedness" at the district level and find that the expansion of rural network under the India's policy has increased consumption inequality between households in the bottom and top quantile of the wealth distribution.

I evaluate the Pakistan's policy, which, to the best of my knowledge, has never been evaluated before. While the policy is similar to the one in India, my study differs from the studies in India in several important respects. First, in addition to using aggregate level data, I use branch-level geographical data which allows me to precisely track the branches' geographical locations and the resulting population coverage over the years and reveal the geographical inequality that is unobservable in aggregate data used in studies on India's policy. Second, Pakistan's policy was implemented in a privatized banking environment and thus reflects more closely South Asia today, where private sector banks play an increasingly important role in the economy. Consistent with Burgess and Pande

(2005) and Young (2017), my study reveals that the absence of a poverty reduction effect was associated with a lack of credit expansion to the rural sector. In line with Kochar (2011), I show that the inequality sharpening effect also existed on a more aggregate level.

The results of this study also contribute to the broader debate on how best to deliver credit to the poor (Pitt and Khandker 1998, Morduch 1999, Karlan and Zinman 2008, Karlan and Morduch 2010, Banerjee et al. 2015c). Recent years have witnessed an explosion of microfinance institutions whose primary clientele are relatively poor farmers, which is guided in part by the belief that traditional banks are ill suited for lending to such households. While the number of studies that evaluate the ability of microfinance institutions to cater to the poor is growing, these studies need to be augmented by others that address the extent to which other banking institutions can also serve this objective. This study serves this objective, evaluating the state-led policies that remain popular in many low-income countries.

Finally, this study contributes to the literature on the branching strategy adopted by commercial banks. Without regulation on branches' locations, commercial bank's entry and exit depend highly on the economic development of the location. Morgan et al. (2016) show that commercial banks were more likely to exit from low- than high-income locations in the USA in response to the 2008 financial crisis. Studying a microfinance bank's expansion in South-East Europe, Brown et al. (2016) document that the bank tended to enter rural locations with larger economic demands. Using a bank-country panel, Qi et al. (2019) demonstrate that information sharing spurs banks to open branches in localities that are new to them, but that are already well served by other banks. My study adds to this strand of literature by showing that commercial banks adopt the same branching strategy even under restrictions on branch openings. It demonstrates the need for more carefully designed interventions for improving financial access than Pakistan's arguably naïve branch expansion policy.

The remainder of this paper is structured as follows. In Section 3.2, I provide background information on access to formal finance in rural Pakistan and the "Branch Expansion Policy" reform. Section 3.3 describes the data sources and how I structure them for the analysis. Section 3.4 presents the evidence on banks' compliance with the policy. Section 3.5 discusses the empirical strategy and policy effect on the branch network's rural population coverage. Section 3.6 presents the empirical approach and policy effects on credit disbursement and poverty. Finally, section 3.7 concludes.

3.2 Institutional Background and the Policy Reform

3.2.1 Access to Formal Finance in Pakistan

According to the Global Findex Database by the World Bank, Pakistan is ranked the second to the bottom among the South Asian countries in terms of access to finance. In 2011, only 10% of Pakistani reported owning an account at a bank or another financial institution. In comparison, the account owning population in India and Sri Lanka in 2011 was 40% and 70% respectively. In fact, the disappointing financial access situation in Pakistan was not a recent incident but has persisted over the last six decades regardless of a rich experience in the provision of rural financial services (State Bank of Pakistan, 2004).

Within Pakistan, there is sharp inequality in financial access between the rural and urban sector. According to SBP's banking statistics, while the rural population comprised 2/3 of the country's population, they were served only by 1/5 of the bank branch network in 2004. The absence of access to the bank network in rural Pakistan has resulted in the following consequences. First, there is a dire lack of access to formal credit. According to SBP's Development Finance Review, the rural sector only accounted for 20% of the total credit advance in terms of the number of transactions and 7% in terms of credit value.[3] The farmers, the small and landless particularly, had to depend on informal sources for meeting most of their credit requirements (State Bank of Pakistan, 2004). Although the SBP has mandated banks to disburse targeted amount of credit to the agricultural sector, a sector that dominates rural Pakistan, only 30% of the disbusement target has been met in the decade of the 1990s (State Bank of Pakistan, 2000). This low fulfillment rate has also assisted in charging of extremely high (50% to 100% annually) interest rates by the informal agricultural credit providers from the needy farmers in every province of Pakistan (State Bank of Pakistan, 2001). The status is further saddened by the fact that the majority of the agricultural credit has been captured by large farm owners and the num-

[3]The calculation is obtained by aggregating statistics from multiple reviews, which can be accessed via http://www.sbp.org.pk/SME/DFG.htm

ber of borrowers has stagnated at 15% farmers in the country (State Bank of Pakistan, 2004). Second, the rural sector is capable of mobilizing savings far and beyond what is being captured by the formal sector at present. Past interventions like Rural Support Programmes have shown that whenever there were innovative or fresh saving products being offered, the rural sector responded immediately and with fervor. The lack of relevant saving products in the rural sector has led the rural sector to concentrate its savings primarily in the form of livestock (State Bank of Pakistan, 2001).

While a relatively lower demand may account for part of this low penetration of formal banks in rural areas, one particular supply side factor, i.e., the banking culture, has been frequently raised by the SBP. The culture of banking in Pakistan is highly urbanized and industry-oriented. The owners, senior managers and other bank officers are uncomfortable in rural settings (State Bank of Pakistan, 2001). The SBP believed that this banking culture, when tailored properly to the rural setting, can be harnessed to include the rural population into formal finance and attack the poverty in rural areas. Against this backdrop, SBP has rolled out the reform in her "Branch Expansion Policy (BEP)" in two phases in 2004 and 2007 in order to enhance the outreach of the banking services to rural and underserved areas of the country (BPD Circular No. 34 of 2004, BPRD Circular No. 15 of 2007).

3.2.2 The Branch Expansion Policy

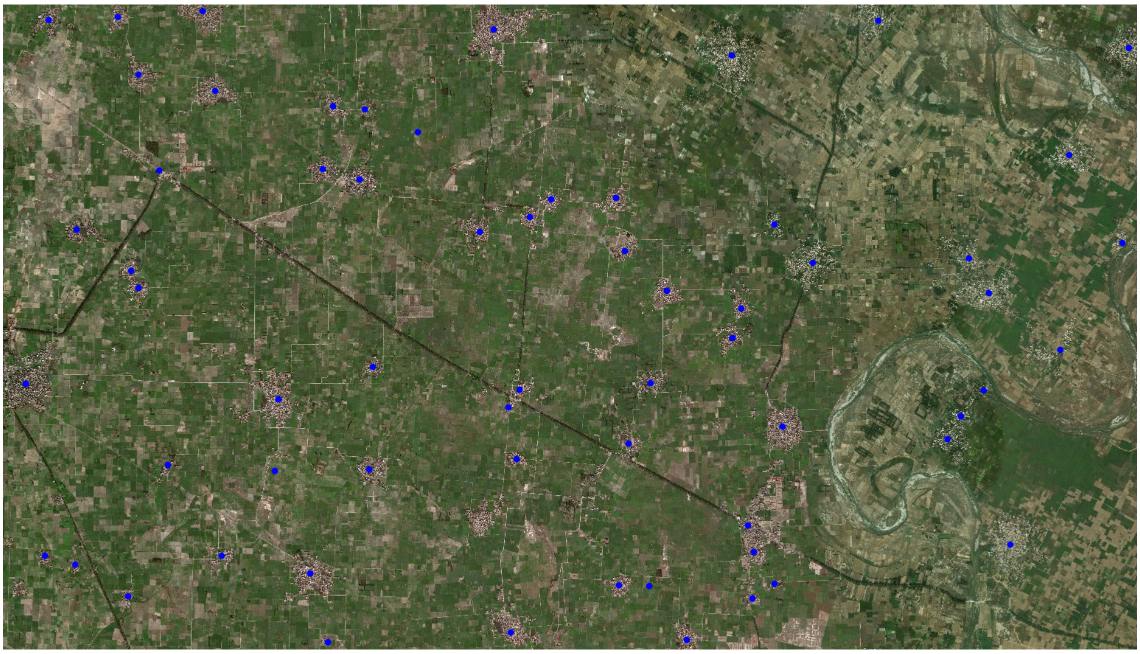
The banking sector in Pakistan does not permit free entry of banking firms (Banking Companies Ordinance 1962). New banking firms are granted entry infrequently by the SBP through special campaigns with recent waves in the early 1990s following the privatization of the banking sector. Banks must also acquire licenses prior to opening any branches and receiving permissions to close or shift branches on an annual basis through the "Annual Branch Expansion Plan (ABEP)", in which banks propose a set of branches to be opened, closed and shifted over the next year. The SBP reviews the proposals centrally and granted the set of licenses and permissions. There was no broad directive existed regarding the composition of markets served by the banks until the reform on the expansion policy (henceforth, the "reform") was introduced. The reform has changed this regulatory environment in a fundamental way in that it required the banks to reserve 20 percent of branch openings proposed in ABEP to locations termed as "Rural and Underserved Areas (RUA)", which were defined by the SBP as villages, small towns or unbanked tehsil headquarters. It is noteworthy that this definition of RUA is rather vague for two reasons. First, it is not consistent with the population census's definition of rural locations, which refers to locations that are neither within the boundary of a municipality nor cantonment. While tehsils (cities) are clear municipalities and thus urban according to the census, it is unclear whether a town or a village should be urban. Second, there is no quantifiable threshold to justify a town to be large or small in the definition of RUA. Regarding the policy rollout, in 2004, the first phase of the reform was rolled out and the 20 percent RUA quota was only applied to banks with a branch network of no less than 100 branches. And effectively since 2007, all banks regardless of their network size were subject to the RUA quota.

Although the quota rule is similar to that from the India's policy between 1977 and 1990 (Burgess and Pande, 2005), important differences exist and need to be discussed. First, the quota in Pakistan's reform is less demanding in that Pakistan's quota was set at 20 percent while India's was at 80 percent. Second, the RUA locations under Pakistan's policy are financially less backward than India's. India's policy coerced banks to open branches in locations that were "unbanked", while in Pakistan, these locations only needed to be justified as "rural or underserved". More importantly, unlike the India's policy during which a list of unbanked locations was provided centrally by the central bank, the SBP only provided a vague definition of RUA. This definition was not directly linked to any banking statistics of the locations (except for the tehsil headquarter which would be qualified as RUA if there was no bank branch existing within the tehsil headquarter boundary). While it is clear to define a location as "unbanked" because the term provides an implicit quantified threshold, it is less transparent to justify a location as RUA, especially when the SBP did not provide a clear quantifiable threshold. Therefore the banks enjoyed a certain degree of freedom to claim the proposed location as RUA.

3.3 Data

This study uses both branch- and district-level data for the analysis. The branch-level data is summarized in Table 3.1. First, from the SBP, I obtain a dataset which records each bank branch that has ever existed between 2000 and 2012, the dates it was opened and closed, whether it was situated in a RUA location and the physical address of the location. I exclude the branches opened in Azad Jammu and Kashmir as it is a disputed region and not subject to the reform. In total, I have 11,293 branches. I geocode the branches' physical addresses using the Google Maps Platform, which returns a pair of longitude and latitude (i.e., geolocation) for each physical address queried to the platform. I also obtain a layer of population raster data at 1 kilometer by 1 kilometer resolution covering the whole Pakistan (excluding the dispute region) as shown in Figure 3.10. This population raster data is estimated by the "Global Human Settlement" dataset from the European Commission using the 1998 Pakistan Population Census. I map out the rural and urban regions by calibrating the rural population share to the share reported by the population census. The calibration is done at the district level and the detailed procedure is described in Appendix 3.B. I only use the population data of 1998 as a basis for any population related measures used in my analysis because the most recent population census prior to the reform was in 1998 and the subsequent census was in 2017 which is beyond my analysis timeline. Based on the branches' geolocations and the population raster data, I then calculate, for each branch, the distance to the district's capital city and the rural and urban population coverage within 5 kilometers radius buffer. The detailed calculation for population coverage is demonstrated in Panel A and B of Figure 3.3.

Second, I obtain a list of populated locations of Pakistan from the "NGA Gazetteer" database by the United States National Geospatial-Intelligence Agency (NGA). For Pakistan, "NGA Gazetteer" offers 122,654 populated locations, of which 116,945 fall within the boundary of my calibrated rural region. These 116,945 populated rural locations are then used in a counterfactual analysis in which they serve as alternative candidates for the RUA branches opened after the reform. For each location, I also calculate the set of geospatial statistics mentioned above. The ideal case would be to use the list of locations identified by the Pakistan Population Census. Unfortunately, such data is not available. Therefore, I resort to the "NGA Gazetteer". Nevertheless, the locations from the "NGA Gazetteer" trace the populated settlements of Pakistan quite accurately as we can visually inspected in Figure 3.1, which overlays a sample of these locations identified by the "NGA Gazetteer" (blue dots) on a satellite image of the corresponding region.



Note: The blue dots represent a sample of populated locations identified by the "NGA Gazetteer". Overlaying these dots on a satellite image of the corresponding region, the settlements features of the satellite image are very accurately located by the blue dots.

Figure 3.1: Populated Locations Identified by "NGA"

I also compile from various sources a district-year panel whose variables of interest include the following. The "rural population share" of a district is obtained from the 1998 population census to measure the degree of urbanization of a district. The "number of RUA branches per million persons" is aggregated from the branch data to measure the bank network's rural outreach in a district.[4] From the SBP's Agricultural Credit and Microfinance Department, I collect a set of agricultural credit disbursement measures for the period between 2007 and 2012 with annual frequency. The SBP categorizes the agricultural credit into crop and non-crop credit.[5] For each category, the "number of borrowers" and the "amount of disbursement" are reported. I standardize these variables to per capital term by the district-specific population. Finally, I use the "Multi-dimensional Poverty Index (MPI)" from UNDP (2016), which comprises three dimensions, namely education, health and standards of living, and are further segregated into 15 indicators. The MPI is a weighted average of these 15 indicators and its value ranges between 0 and 1. A higher MPI indicates more sever poverty. The MPI data covers the period between 2004 and 2012 and is available only at biannual frequency because the household survey used to

[4]This is similar to the "social banking" measure in Burgess and Pande (2005).

[5]Non-crop items include livestock, poultry, dairy farming, fishery and forestry.

estimate the MPI is conducted biannually.[6] This district-year panel is separated into pre- and post-2007 sample and summarized in Table 3.2.

Table 3.1: Summary Statistics of Bank Branches and Locations

	Obs.	Mean	St. Dev.
Bank Branches from SBP			
Is a RUA Branch	11,293	0.22	0.42
RUA Branches			
Rural Population in 5km Radius	2,517	44,116	40,278
Urban Population in 5km Radius	2,517	68,708	191,023
Distance to district capital (km)	2,517	22	15
Non-RUA Branches			
Rural Population in 5km Radius	8,776	44,695	55,558
Urban Population in 5km Radius	8,776	652,857	633,041
Distance to district capital (km)	8,776	10	13
Rural Locations from "NGA Gazetteer"			
Rural Population in 5km Radius	116,945	19,022	26,011
Urban Population in 5km Radius	116,945	8,089	37,744
Distance to district capital (km)	116,945	32	22

Note: The branch-level sample consists of 11,293 bank branches that have ever existed between 2000 and 2012, excluding those located in Azad Jammu and Kashmir. The location-level sample consists of 116,945 rural populated locations that are identified in "NGA Gazetteer". The population is counted based on the 1998 population census. The urban and rural decomposition is based on the author's calibration, whose detailed process in explained in Appendix 3.B. According to the 1998 population census, the size of rural and urban population is 89 and 43 million, respectively.

[6] UNDP (2016) calculate the MPI using data from "Pakistan Social and Living Standard Survey (PSLM) - District-Level Survey" which is conducted biannually with the first wave initiated in 2004.

Table 3.2: Summary Statistics of District-level Variables

	2000-2006			***2007-2012***		
	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.
Rural Population Share	854	0.77	0.17	734	0.77	0.17
RUA Branch per Million Persons	854	8.49	7.54	734	10.46	8.38
Poverty (Biannually from 2004 to 2012)						
MPI	244	0.37	0.14	366	0.32	0.15
Agricultural Credit Disbursement (by credit category) in:						
Rupees per capita						
Aggregate				734	319	418
Crop				734	248	323
Non-crop				734	70	221
Borrowers per million persons						
Aggregate				734	2,295	3,517
Crop				734	2,027	3,239
Non-crop				734	268	477

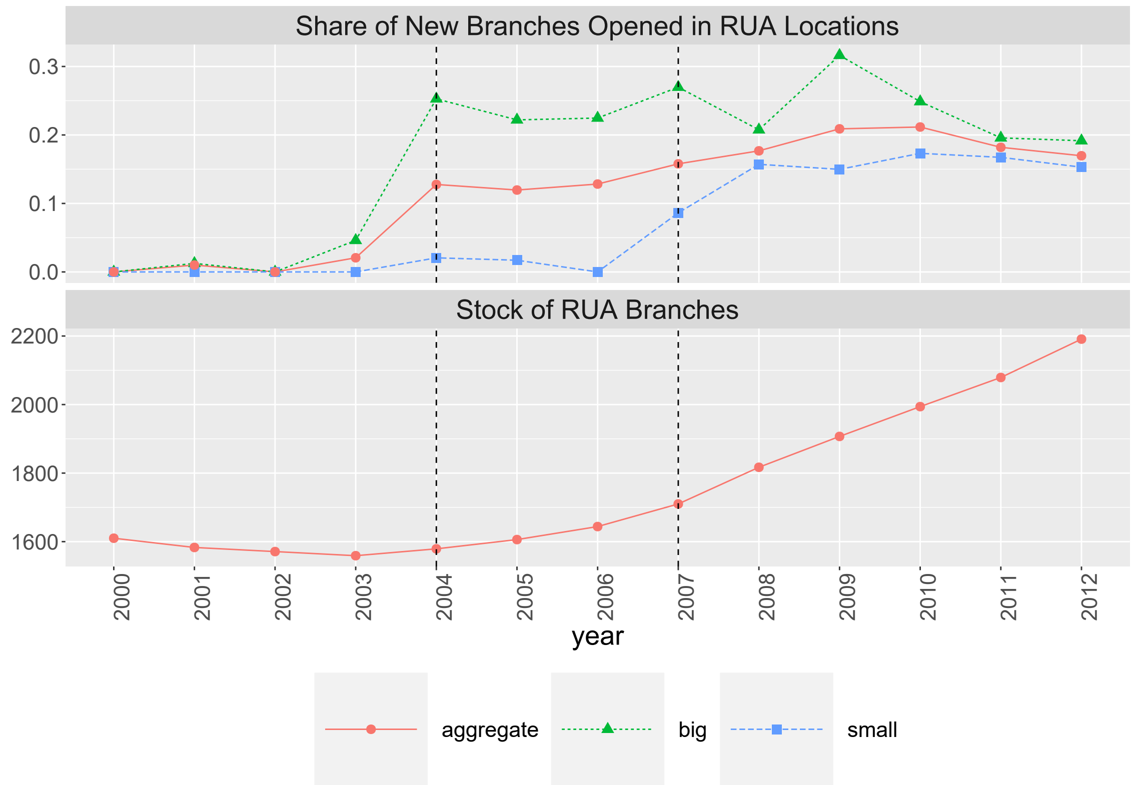
Note: An observation is a district. Except for MPI, data is collected on an annual basis. This panel consists of 122 districts, excluding those located in Azad Jammu and Kashmir. Non-crop items include livestock, poultry, dairy farming, fishery and forestry.

3.4 Evidence on Policy Compliance

I start by providing evidence that the banks have complied with the reform and significantly increased the size of their RUA branch network. The most direct evidence comes from the changes in the share of new branches opened in RUA locations along the policy timeline as shown in the upper panel of Figure 3.2. New branches refer to the branches that were opened in the calendar year of consideration. The two vertical dashed lines refer to the years since which the two reform phases have been in effect. The two dashed lines with markers trace out the cases for big and small banks respectively. As it is shown, the two groups of banks have clearly complied with the quota since 2004 and 2007 by discontinuously increasing the share of branches opened in RUA locations to 20 percent. The small banks did not meet the quota instantly in 2007 when they also became subjected to the reform. The reason, as I suspect, is that some small banks were given a grace period in adjusting their annual branch expansion plan on a case by case basis, a measure that is not uncommon to SBP in fostering small companies.[7] Nevertheless, the small banks have caught up with the quota since 2008. While it is true that multiple confounders, e.g., increase of economic demands in rural areas, can also cause expansion of RUA network, it is unlikely that these factors can explain the two discontinuous changes to exactly 20 percent in 2004 and 2007 respectively. In comparison, this share is rather low at around 2 percent before 2004. In other words, there was almost no new branch opened in RUA locations before the reform. For this fact, it is safe to claim that the policy reform drove the expansion of the RUA network. This also suggests that the reform has distorted the banks' optimal geographical allocation of new branches. The solid line traces out the case when aggregating the two bank groups. Again, a clear discontinuous change is present in 2004 and slowly converges to 20 percent after 2007. The discontinuity in 2007 is less visible because the share of small banks is relative small.

The change in the spatial composition of new branches has also affected the stock of RUA branches. This can be observed in the bottom panel of Figure 3.2. The discontinuous increase in the share of new RUA branches in 2004 and 2007 is reflected as faster growth in the stock of RUA branches starting in 2003 and 2006, i.e., the two "kinks" in 2003 and 2006 in the bottom panel. This provides a ground for my identification strategy. Intuitively, if the RUA branches are effective in delivering the policy effects, we should expect changes in the growth of the outcome variables in 2003 and 2006, a pattern that is consistent with the bottom panel of Figure 3.2.

[7]For example, at the beginning of the banking sector privatization in 1995, the opening of new branches by the new banks was eased (State Bank of Pakistan, 2000). In recent years, SBP has offered one year of grace period to loans borrowed by new companies (IH&SMEFD Circular Letter No. 09 of 2014).



Note: "big" refers to the group of banks whose branch network was no smaller than 100 branches in 2003. "small" is the complementary group. "aggregate" refers to the case including all banks.

Figure 3.2: Branches Opened in RUA Locations

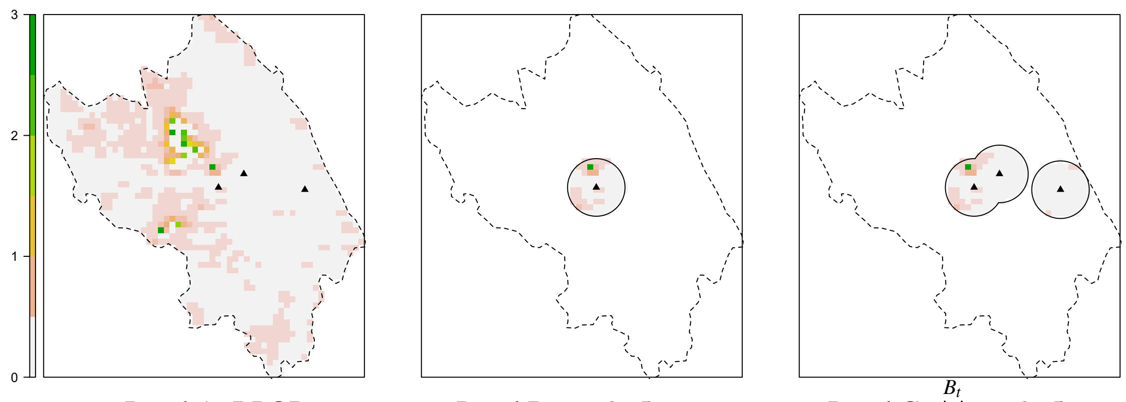
3.5 Effects on the Rural Outreach

Empirical Approach

Although the number of RUA branches has expanded massively after the reform, the number of branches does not solely determine the extent of financial access. The location of this RUA network also plays an important role. A literature has established that a closer geographical proximity to a bank branch significantly increases the take-up of bank services (Petersen and Rajan 1994, Degryse and Ongena 2005, Agarwal and Hauswald 2010, Brevoort et al. 2012). This is still true in the US today, where information technology has significantly reduced the cost imposed by geographical barrier (Nguyen, 2019). To account for the influence of the branch locations, I use an extensive margin measure of access to finance. This measure is termed as "Branch Network's Rural Outreach (BNRO)" and is defined as follow,

$$BNRO_t = (\bigcup_{b=1}^{B_t} RPOP5KM_{bt})/RPOP$$

where $RPOP5KM_{bt}$ refers to the set of rural population residing within 5 kilometers radius of the branch b of year t (banked rural population). The 5 kilometers threshold is officially used by the SBP to capture a branch's coverage extent. $RPOP$ is the total rural population in 1998, which is a constant equal to 89 million. B_t denotes to the branch stock in year t. $BNRO_t$ is calculated by taking the union of $RPOP5KM_{bt}$ over all branches that are operating in year t and it is normalized by the size of total rural population in 1998. Intuitively, $BNRO_t$ captures the share of banked rural population in year t in the total rural population. Notice that this measure assumes zero population migration across years and is based on the distribution of rural population which is decomposed from the total population distribution in 1998. A detailed decomposition procedure is provided in Appendix 3.B. The definition of $BNRO$ is visualized in Figure 3.3. $BNRO_t$ avoids double counting the set of population that are covered by multiple branches. In comparison, the intensive margin measure, i.e., branches per capita in a state as used by other studies would not capture the geographical variation within the state (i.e., in Burgess and Pande 2005, Kochar 2011, Young 2017).



Panel A: RPOP   Panel B: $RPOP5KM_{bt}$ Panel C: $\bigcup_{b=1}^{B_t} RPOP5KM_{bt}$

Note: The dashed line in all panels circles out the border of a sample region. Panel A shows the rural population density distribution at 1km by 1km resolution. The unit of the population is in 1000 persons. The sum of the grids' values gives the size of rural population. The triangles represent the locations of branches. Panel B represents the set of rural population residing within 5km radius of a branch. Panel C displays the union of the rural population sets associated with all three branches. *BNRO* is obtained by dividing the size of rural population included in Panel C by the size of the total rural population circled out in Panel A.

Figure 3.3: Conceptualization of "Branch Network Rural Outreach (*BNRO*)"

To test whether $BNRO_t$ displays a similar growth pattern as in the bottom panel of Figure 3.2, i.e., whether $BNRO_t$ grows faster after 2003 and 2006, I estimate the following simple linear trend break model,

$$\begin{aligned} BNRO_t = c + \beta_0 * (t - 2000) + \beta_1 * (t - 2003) * D_t^{2003} + \beta_2 * (t - 2006) * D_t^{2006} \\ + \lambda_1 * D_t^{2003} + \lambda_2 * D_t^{2003} + u_t \end{aligned} \tag{3.1}$$

where c is the constant term. t denotes the year. D_t^{2003} is a dummy variable that equals 1 if year t is 2003 or beyond, and 0 otherwise. The post-2006 dummy variable, D_t^{2006}, is similarly defined. u_t is the error term. Intuitively, model (3.1) breaks the time series of $BNRO_t$ into three sections at year 2003 and 2006 and fits a line to each section, allowing for intercept changes at the breaking points. Precisely, λ_1 and λ_2 estimate the intercept changes in 2003 and 2006. The slopes of the fitted lines for the three sections, i.e., [2000, 2003), [2003, 2006) and [2006, 2012], are β_0, $\beta_0 + \beta_1$ and $\beta_0 + \beta_1 + \beta_2$, respectively. Therefore, my coefficients of interests are β_1 and β_2, which capture the changes of the growth rate of $BNRO_t$ (relative to the pre-2003 period) at the year point 2003 and 2006, respectively.

Results

The results are presented in Table 3.3 where the second and third row show the estimates of β_1 and β_2 respectively. First, to confirm the existence of the kinks in the pattern of the bottom panel in Figure 3.2, I estimate model (3.1) with the left-hand side variable replaced by the number of RUA branches in year t and report the estimation in column (1). Before 2003, the stock of RUA branches decreased by around 19 branches per year. Consistent with the visual inspection of Figure 3.2, there two significant trend breaks in 2003 and 2006, respectively, suggesting that the growth of the number of RUA branches accelerated by 43 and 68 branches per year from 2003 and 2006, respectively, compared to the pre-reform period. The acceleration accumulated to a total increase of 1176 branches by the year of 2012, a value that is 73% of the total RUA branches in 2000.[8]

Column (2) shows the estimate for $BNRO_t$. The point estimate of β_0 is only nominal at 0.08 branches per million persons per year, suggesting that the network's rural population

[8]There are 9 years posterior to the reform from 2004 to 2012, with 3 years in the first phase and 6 in the second. 1176 = 43*3 + 68*6. The total number of RUA branches in 2000 is 1610.

coverage, i.e., the fraction of banked rural population in the total rural population, stayed almost stagnant during the pre-2003 period. The size of RUA network even saw a slight shrinkage at 0.06 branches per million persons per year after 2003 when the number of RUA branches started to grow. Finally, I detect a statistically significant trend break in 2006. Between 2006 and 2012, the rural population coverage grew faster than pre-2003 period by 0.2 percentage points per year. However, it is particularly noteworthy that the magnitude of this increase in growth is also nominal. According to this estimate, the total improvement in the rural outreach only adds up to 1.2 percentage points by 2012, a value that is only 2% of the value in 2000 to be precise.[9] In Figure 3.5, the solid line with circular marker traces the values of $BNRO_t$ between 2000 and 2012, depicting an overall flat growth trend of $BNRO_t$ with a slight acceleration in 2006, a pattern that is consistent with the estimates in column (2) of Table 3.3. Overall, there is a lack of growth in branch network's rural population outreach. In the appendix 3.A, I discuss the robustness of these findings.

Table 3.3: Effect on Branch Network Rural Outreach (BNRO)

Dependent Variable:	Number of RUA Branches	BNRO (%)		
		Actual	Counterfactual.1	Counterfactual.2
	(1)	(2)	(3)	(4)
β_0: [2000-2003) trend	-19.50^{**}	0.08^{***}	0.08^{***}	0.08^{***}
	(6.54)	(0.02)	(0.02)	(0.02)
β_1: [2003-2006) trend	43.00^{***}	-0.06^{**}	1.08^{***}	1.12^{***}
*Post-2003 dummy	(9.25)	(0.02)	(0.07)	(0.06)
β_2: [2006-2012] trend	67.79^{***}	0.20^{***}	0.64^{***}	0.86^{***}
*Post-2006 dummy	(6.77)	(0.02)	(0.12)	(0.12)
λ_1: Intercept change in 2003	8.83	-0.004	0.06	0.05
	(16.47)	(0.05)	(0.11)	(0.09)
λ_2: Intercept change in 2006	3.81	-0.01	0.33	0.75
	(15.48)	(0.07)	(0.46)	(0.43)
Constant	$1,607.50^{***}$	55.46^{***}	55.46^{***}	55.46^{***}
	(8.45)	(0.03)	(0.03)	(0.03)
Observations	13	13	13	13
Adjusted R^2	1.00	0.98	0.99	1.00

Note: Standard errors are clustered by year and reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

[9] 1.2 percentage points = 0.2 percentage points per year * 6 years. $BNRO_{2000}$ = 55.7 percentage points

Taken together, an economically significant improvement in the size of RUA branch network (76%) with a stagnant growth in the branch network rural outreach (2%) suggest that the locations chosen by the banks to fulfil the quota were not the most remote. This is a valid concern, since, as argued previously, the policy reform was costly to the banks because it distorted their optimal geographical allocation of branches. The banks thus have incentives to reduce the cost. Given the lack of rigor in the definition of RUA locations, a convenient way for the banks to reduce the cost imposed by the policy was to choose less remote locations for the new branches reserved by the quota, assuming that opening a branch in a less remote location is less costly to a bank. The banks' reluctance to expand their networks to remote locations are also documented in numerous studies (Morgan et al. 2016, Brown et al. 2016, Qi et al. 2019, Young 2017).

To test this hypothesis, I use a sample that contains all new RUA branches opened between 2000 and 2012. Two measures of "remoteness" are calculated for each new branch. The first measure is whether the RUA branch is "Opened in a Banked Location ($BANKED_{dt}$)". Precisely, $BANKED_{dt}$ is a binary variable that equals 1 if the RUA branch of district d that was newly opened in year t is located in a location, within 5 kilometers of which, at least one other bank branch exists up to year $t-1$. The second measure is the RUA branch's "Distance to the District Capital ($DIST_{dt}$)".

I am interested in tracing the average "remoteness" of the new RUA branches along the policy timeline. However year-by-year comparison is not permissible due to the low yearly incidence of RUA branch openings before the reform. Therefore, each new branch is grouped into one of the three period-groups based on the year of opening, namely [2000, 2004), [2004, 2007) and [2007, 2012] which represent the periods of pre-reform, reform-phase-1 and reform-phase-2. The averages of the group [2004, 2007) and [2007, 2012] are then compared against the pre-reform group using the following model,

$$y_{dt} = c + \mu_1 * D_t^{[2004,2007)} + \mu_2 * D_t^{[2007,2012]} + u_{dt} \tag{3.2}$$

where y_{dt} is one of the two "remoteness" measures of district d in year t. $D_t^{[2004,2007)}$ is a dummy variable that equals 1 if the year t is in period [2004, 2007) and 0 otherwise. $D_t^{[2007,2012]}$ follows a similar definition for period [2007, 2012]. u_{dt} is the error term. Intuitively, the constant term gives the estimate of the mean of y for the pre-reform period [2000, 2004). μ_1 captures the difference in average between [2004, 2007) and the pre-reform period. Finally, μ_2 captures the difference in average between [2007, 2012] and the pre-reform period.

The results are presented in Table 3.4. Confirming my hypothesis, the RUA locations the banks chose after the reform are less remote than those chosen before the reform. Precisely, column (1) shows that 40% of the RUA branches opened were in banked locations between 2000 and 2004. During the phase-1 and phase-2 reform, this fraction witnessed a 50 and 53 percentage points increase, accumulating to an impressive total of 90 and 93%, respectively. This suggests that almost all RUA branches opened after the reform were in banked locations. Similarly, in terms of distance to district capitals, column (2) suggests that RUA branches opened during phase-1 and phase-2 period were 7 and 10 kilometers closer to the district capitals, compared to those opened during the pre-reform period. These results are consistent with the literature on commercial banks' branching strategy and support the claim that the reform has failed to improve the branch network's rural outreach.

Table 3.4: Remoteness of the RUA Branches Opened Between 2000 and 2012

Dependent Variable:	Opened in a Banked Location	Distance to District Capital (km)
	(1)	(2)
Constant	0.400^{*}	29.642^{***}
	(0.221)	(2.680)
$D_t^{[2004,2007)}$	0.501^{**}	-7.170^{**}
	(0.227)	(3.439)
$D_t^{[2007,2012)}$	0.527^{**}	-9.754^{***}
	(0.221)	(2.858)
Observations	646	646
Adjusted R^2	0.027	0.003

Note: The unit of observation is a RUA branch. The sample consists of all RUA branches that were opened between 2000 and 2012. This analysis cannot be extended before 2000 because the records of branch openings are not complete before 2000. Standard errors clustered by district are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Counterfactual Analysis

Overall, the reform has limited success in expanding the outreach to the rural population ($BNRO_t$) because the chosen locations for the RUA branch network are on the fringes of urban centers where a branch network already existed (as implied by Table 3.4). In this section, I address the question on how the rural outreach would appear had the banks instead chosen the RUA locations based on "rural market potential (RMP)". I construct two counterfactual RUA networks based on two measures of RMP as benchmarks to further judge the economic significance of the actual expansion.

The counterfactual RUA network is constructed by reallocating the 681 new RUA branches that have been opened between 2004 and 2012 (i.e., after the reform). The pool of alternative locations contains 116,945 populated rural locations from the "NGA Gazetteer". The reallocation is done dynamically on an annual basis. To reallocate the new RUA branches opened in year t, I rank the alternative locations by their rural market potential of year t-1 and choose the alternative locations based on this ranking. The locations' rankings are then updated and forms the basis for the reallocation of year t+1. This counterfactual thus tells us where and when new RUA branches would have been opened if the banks were maximizing rural market potential. Two measures of rural market potential are used to construct two separate counterfactual networks. They are defined as follow.

The first measure, $RMP^A_{r,t}$, captures the size of rural population a branch can serve at the location r and year t, had a branch entered that location,

$$RMP^A_{r,t} = \frac{RPOP5KM_r}{Branch5KM_{r,t-1}+1}$$

where $RPOP5KM_r$ is the size of rural population (of year 1998) residing within 5 kilometers radius of the location r. $Branch5KM_{r,t-1}$ refers to the number of branches operating within 5 kilometers radius of the location r in year t-1. An additional value 1 is added to the denominator to capture the scenario of "if a branch entered the location r".

The second measure, $RMP^B_{r,t}$, captures the size of unbanked rural population a branch can serve at the location r and year t, had the branch been opened there,

$$RMP^B_{r,t} = RPOP5KM_r * Unbanked5KM_{r,t-1}$$

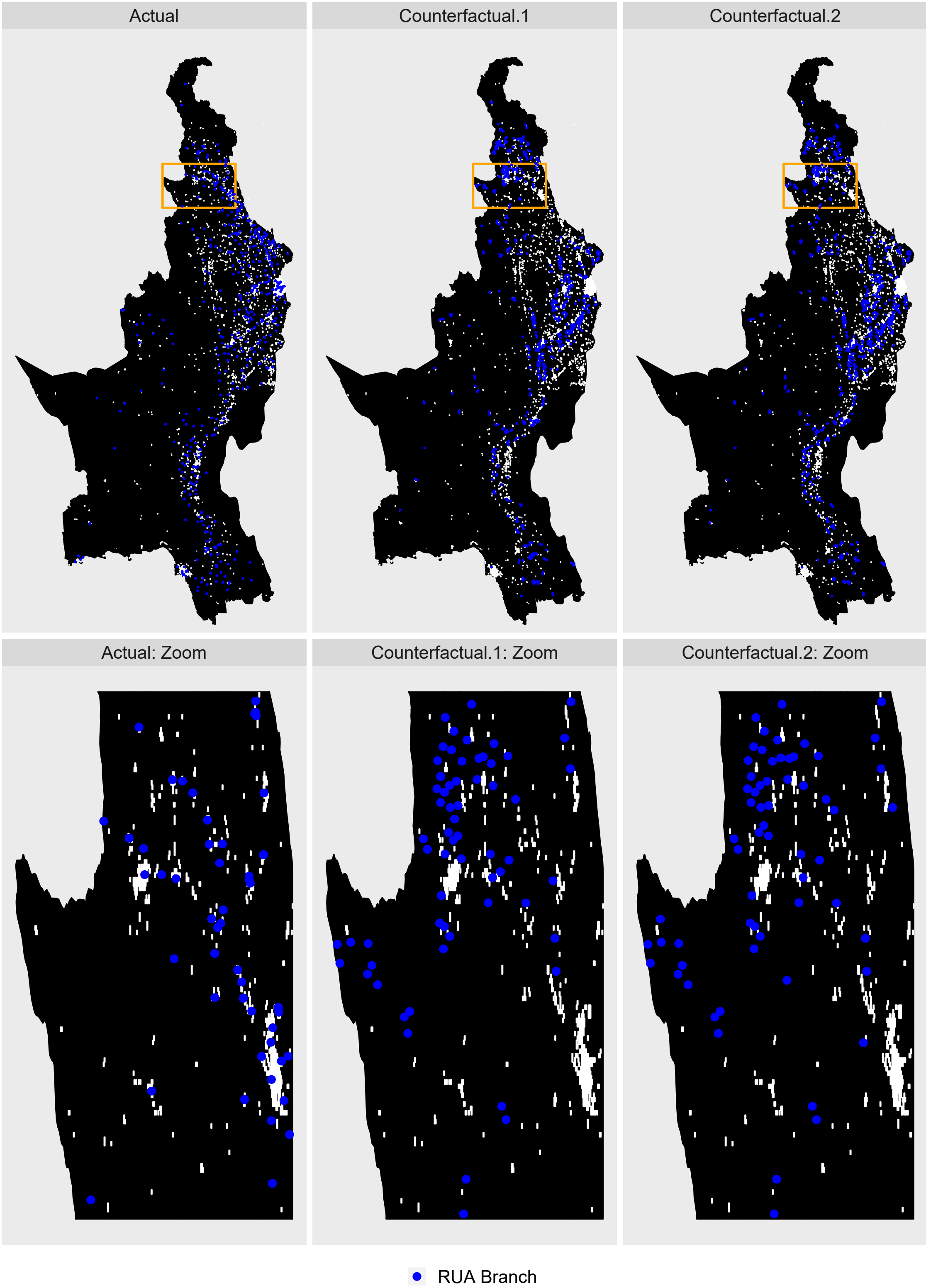
where $Unbanked5KM_{r,t-1}$ is a binary variable that equals 1 if there is no branch operating within 5 kilometers radius of the location r in year t-1. It is noteworthy that $RMP^B_{r,t}$ is equal to censoring $RMP^A_{r,t}$ by $Unbanked5KM_{r,t-1}$, i.e., $RMP^B_{r,t}$ prioritizes the unbanked locations, therefore, its corresponding counterfactual network will reach out to more remote areas than $RMP^A_{r,t}$.

The actual and the two counterfactual networks of the RUA branches opened between 2004 and 2012 are presented in Figure 3.4. The white and black regions in the base map represent the calibrated urban and rural areas respectively. The blue dots represent the locations of the branches. The actual case (i.e., left panel) shows that the new RUA branches heavily trace the urban centers while in the two counterfactual cases, these branches mostly fill the gaps between the urban centers. The starkest contrast is visible along the corridor of Peshawar-Islamabad-Rawalpindi-Gujranwala-Lahore (the northeast corner of the map), a region that contains only 6% of the rural population in Pakistan while this percentage is 23% for urban population. To make the comparison more visible, the bottom panel zooms in to the region near Peshawar. This observation shows, from another perspective, the banks reluctance to enter more rural areas. In Figure 3.5, I plot the rural outreach (i.e., $BNRO_t$) of these two counterfactual (solid lines with square and triangular markers) and the actual (solid line with circular markers) networks between 2000 and 2012. The contrast is stark. First of all, the period between 2000 and 2003 are excluded from my counterfactual simulation because the reform was initiated in 2004. As a result, the growth of rural outreach in the three cases are indistinguishable until 2003. Second, two kinks can be clearly observed in the year 2003 and 2006 in both counterfactual cases, which, in comparison, are visually absent for the actual case. This is confirmed by the estimation of model (3.1) using the simulated data of the two counterfactual cases. The results are reported in column (3) and (4) of Table 3.3.

Column (3) indicates that, under the first counterfactual network, $BNRO_t$ witnesses a growth acceleration by 1.08 and 0.64 percentage points per year in phase-1 and phase-2 reform period, respectively, compared to the pre-reform period. The acceleration is even higher under the second counterfactual network as shown in column (4) because the second counterfactual network prioritizes the unbanked locations. In comparison, the growth of $BNRO_t$ under the actual network is only nominal (column 2), at 0.2 percentage point per year after 2006, a value that is only a third of the estimate under the first counterfactual network (column 3), or a quarter under the second counterfactual case (column 4). These differences are also statistically significant at 1% level.[10]

More importantly, these estimates suggest that the improvement in the share of banked rural population in the total rural population by 2012 under the two counterfactual cases

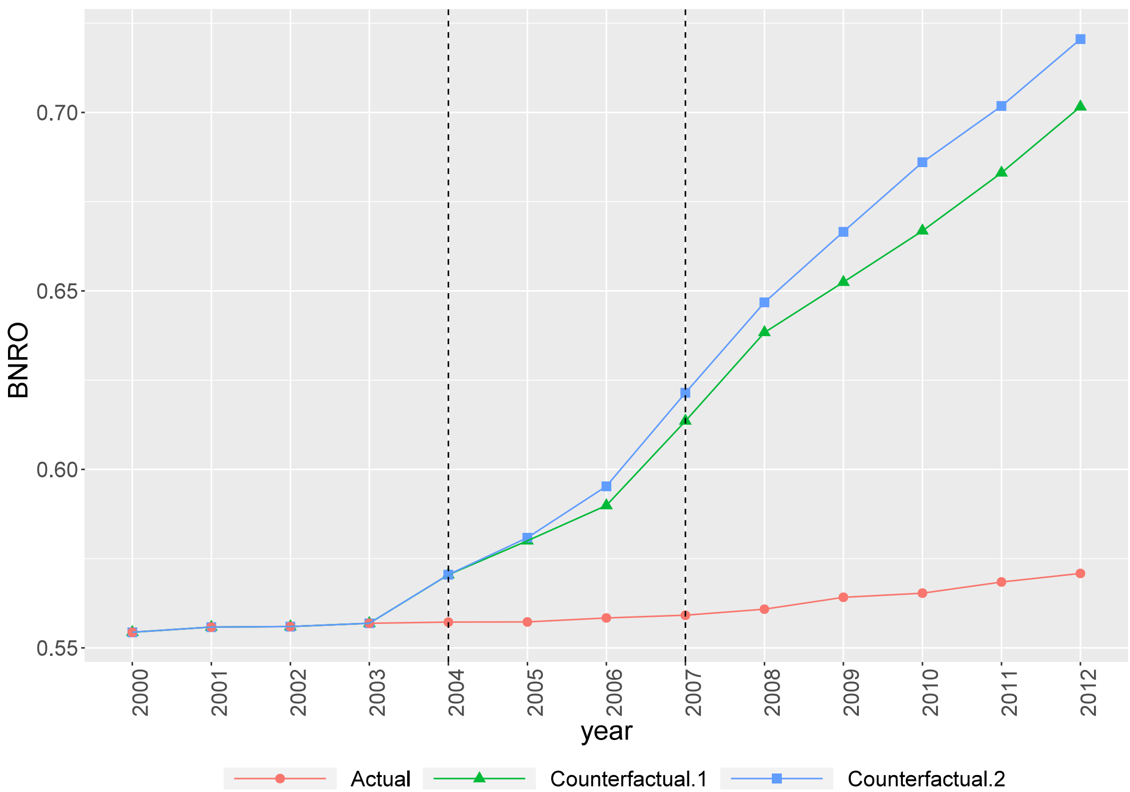
[10]A test for the null hypothesis of "$\beta_2^{Counterfactual.1} - \beta_2^{Actual} = 0$" shows a p-value of 0.002. And the p-value is $7.16 * 10^{-5}$ for the null hypothesis "$\beta_2^{Counterfactual.2} - \beta_2^{Actual} = 0$"



Note: The "blue" dots refer to the RUA branches. The "black" region refers to the rural areas identified by calibrating the rural population share to fit the one reported in the census. The calibration process is provided in Appendix 3.B. The "white" region refer to the urban areas. "Counterfactual.1" and "Counterfactual.2" present the counterfactual network based on the rural market potential measures of $RMP^{A}_{r,t}$ and $RMP^{B}_{r,t}$, respectively.

Figure 3.4: RUA Branches Opened Between 2004 and 2012

are ten times larger than that under the actual case. More precisely, the share of banked rural population reaches to 70 and 72 percent in the first and second counterfactual network, respectively, an improvement of 15 and 17 percentage points from the year 2000. In comparison, the rural outreach under the actual network has mainly stagnated, with a nominal 1.6 percentage points improvement since 2000. This stark comparison strongly demonstrates the reform's failure in terms of rural outreach enhancement.



Note: "Counterfactual.1" and "Counterfactual.2" present the counterfactual RUA network for the year 2012 based on maximizing the rural market potential measures of $RMP^{A}_{r,t}$ and $RMP^{B}_{r,t}$, respectively.

Figure 3.5: Branch Network's Rural Outreach

3.6 Effects on Agricultural Credit Disbursement and Poverty

Empirical Approach

The answer to the question on whether government interventions of financial access expansion is effective in reducing poverty is inconclusive in the literature. On the one hand, Burgess and Pande (2005) and Young (2017) have found consistent evidence that better financial development reduces poverty by relaxing the credit and saving constraints faced by the poor. On the other hand, La Porta et al. (2002) and Sapienza (2004) argue that market frictions and political considerations, especially in low income countries, distort the allocation of financial resources, preventing the poor from benefiting from financial development. This argument is supported by Kochar (2011) who shows that consumption inequality sharpened in response to the rural expansion of bank network in India.

Using the simple trend break model, i.e., model (3.1), to identify the policy effects on credit disbursement and poverty, as compared to population coverage, is more challenging because cross-sectional and temporal variations in economic factors (e.g., interest rate) could also affect the trend pattern that could potentially appear in credit disbursement and poverty. For example, a faster growth of credit disbursement after the reform could also be explained by a ceteris paribus decrease of interest rate. Moreover, the model does not account for cross-sectional variation. Therefore, it is not immediately clear that the locations where credit disbursement grew are the locations that benefited from having more RUA branches after the reform. Therefore, to reduce these biases, I employ a difference-in-differences trend break model, exploring the temporal and cross-sectional variations induced by the policy.

The underlying assumption is that more remote locations offer less profitable business opportunities for banks and therefore, in the absence of the reform, would attract less new branches and witness a smaller growth in the branch stock. The previous finding suggests that banks opened even less branches in more remote locations after the reform, suggesting that if the reform had any bite, its rollout should have sharpened the correlation between more remoteness of a location and smaller number of RUA branches. To examine this possibility I run the following difference-in-differences regression model using a district-year panel dataset. I conduct this analysis using district-level observations because credit and poverty data are available only at the district level.

$$Y_{d,t} = \sum_{\tau=2002}^{2014} \gamma_\tau * Y_{d,2000} * D_t^\tau + \sum_{\tau=2002}^{2014} \beta_\tau * RuralPopShare_d * D_t^\tau + \eta_t + \sigma_d + u_{dt} \quad (3.3)$$

where $Y_{d,t}$ is the outcome variable of district d and year t. D_t^τ are year dummies that take value 1 if year t equals τ. $Y_{d,2000}$ is the outcome Y's value in year 2000 (baseline value) and it is interacted with the year dummies to control for the time-trend specific to the baseline value. η_t and σ_d are the year and district fixed effects. $RuralPopShare_d$ is a proxy of district d's remoteness, measured by the rural population share of the district d in year 1998. This variable enters the regression interacted with year dummies, with β_τ denoting the year-specific coefficients. The difference between $\beta_{\tau+1}$ and β_τ tells us how a district's initial rural population share affected RUA branch growth between years τ and $\tau+1$.

I am interested in testing whether the growth of β_τ changes after 2007 when all banks became subjected to the reform. This is achieved by estimating a difference-in-differences trend break model following Burgess and Pande (2005),

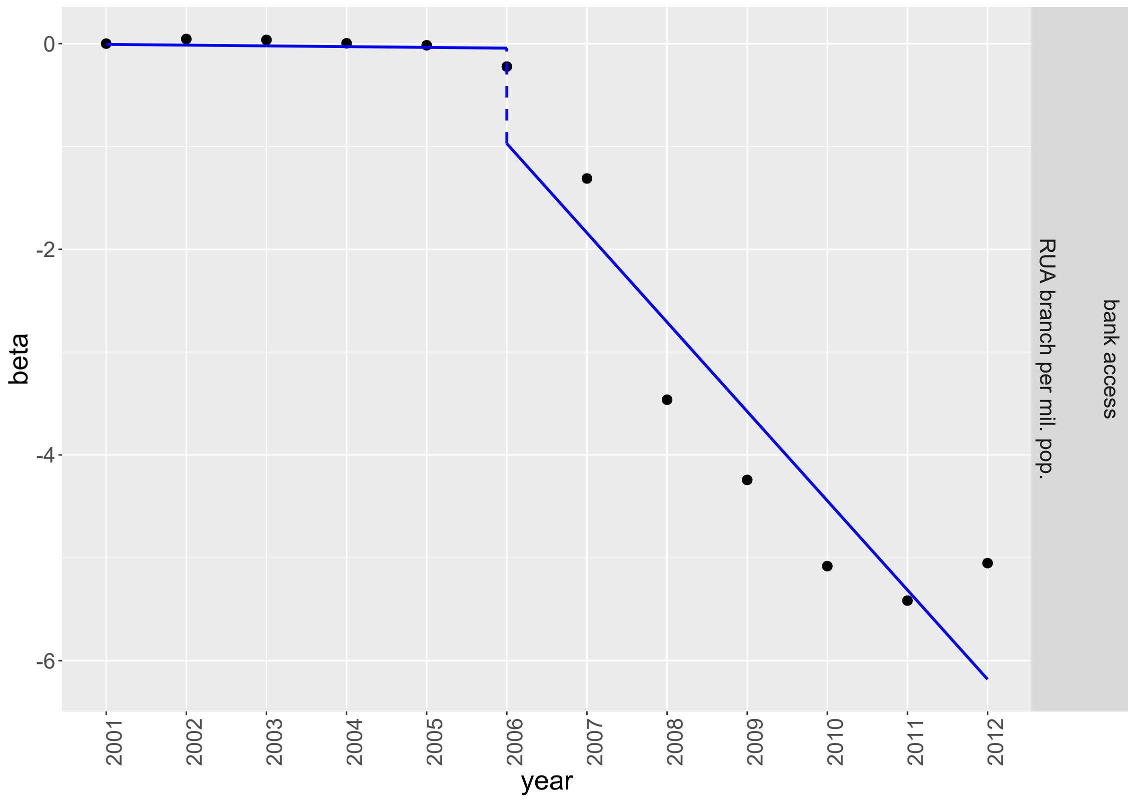
$$\begin{aligned}
Y_{d,t} &= \beta_0 * RuralPopShare_d * (t-2000) \\
&\quad + \beta_1 * RuralPopShare_d * (t-2006) * D_t^{2006} \\
&\quad + \lambda_1 * RuralPopShare_d * D_t^{2006} + \sum_{\tau=2002}^{2012} \gamma_\tau * D_t^\tau * Y_{d,2000} + \eta_t + \sigma_d + u_{dt}
\end{aligned} \tag{3.4}$$

Intuitively, model (3.4) breaks the series of β_τ estimated from model (3.3) into two sections at year 2006 and fits a line to each of the two sections, allowing for intercept changes. Precisely, the slopes of the fitted lines for the periods, [2001, 2006] and (2006, 2012] are, respectively, β_0 and $\beta_0+\beta_1$. I am interested in testing whether the slope changes, i.e., whether β_1 is significantly different from zero. Finally, λ_1 measures the intercept changes in this relationship in 2006. Ideally, I would like to test the trend break in year 2003 which is associated with the imposition of the reform's first phase. However, credit and poverty data are available only from 2004.

Results

I start by presenting the results on the cumulative number of RUA branches per million person. The dots in Figure 3.6 graph the estimates of β_τ coefficients from model (3.3) (the reference year is 2001). As it is shown, up until 2006, the trend stayed mostly flat, indicating that the banks did not change their preference for less rural districts for opening RUA branches. A trend break (i.e., a "kink") appeared in 2006 when the reform was rolled out to all banks. Precisely, the β_τ coefficients decrease with time after 2006, i.e., more rural districts (i.e., districts with larger rural population share) witness smaller growth of number of branches in RUA locations. In other words, the banks preferred less rural districts to a greater extent than they did before the reform.

The existence of the trend break is confirmed by estimating model (3.4). As shown in column (1) of Table 3.5, between 2001 and 2006, a rural district, e.g., “Lower Dir” whose rural population share equals almost 1, compared to an urban district, e.g., “Karachi” whose rural population share equals almost 0, experienced less growth in number of RUA branches per million persons by 0.01 annually. The slope of the trend decreases significantly after 2006 by 0.86 (i.e., the slope after 2006 is -0.87), a magnitude that is equal to 11% of the average stock of RUA branch at baseline. This implies that by the year of 2012 the gap between “Lower Dir” and “Karachi” in terms of RUA branch network size would be widened by 5.61 branches per million population, a value that is 66% of the average RUA branch network size in 2000. The blue lines in Figure 3.6 are fitted based on these estimates. These results confirm that the majority of the RUA branch openings after the reform was captured by the less rural districts.



Note: The dots refer to the point estimates of the betas from model (3.3). The blue lines are fitted based on the estimate of the “difference-in-differences trend break model”, i.e., model (3.4). The numbers shows the estimated line slopes, i.e., β_0 and $\beta_0 + \beta_1$ with their F-tests’ p-values in brackets. All regressions cluster the standard errors at district level.

Figure 3.6: Rural Population Share and RUA Branch Expansion

Table 3.5: Effects on Financial Inclusion and Poverty

Dependent Variable:	Outreach	Agricultural Credit						Poverty
		Aggregate		Crop		Non-crop		
	RUA branches per mil. pop.	Borrowers per mil. pop	Rupees per capita	Borrowers per mil. pop	Rupees per capita	Borrowers per mil. pop	Rupees per capita	MPI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
λ_1	-0.93^{**}							
	(0.47)							
β_0	-0.01							0.016
	(0.07)							(0.013)
β_1	-0.86^{**}	119.31	-57.96	188.01	23.02	-68.71	-80.99^{*}	-0.017
	(0.36)	(281.96)	(54.24)	(263.29)	(32.05)	(57.16)	(42.09)	(0.013)
Joint F-test	5.96							0.026
	[0.02]							[0.872]
Mean in 2007		1807	267	1685	227	121	41	
Observations	1,588	734	734	734	734	734	734	610
Adjusted R^2	0.98	0.67	0.86	0.64	0.86	0.54	0.82	0.92

Note: Standard errors clustered by district are reported in parentheses. p-values are in square brackets. Joint F-test is the joint significance test for the two β coefficients in the second and third rows.

As highlighted by Burgess and Pande (2005) and Young (2017), the rural branch expansion in India was successful in reducing poverty because the rural branches simultaneously provided more credit to and mobilized more deposit from the rural sector. In rural Pakistan, the production activities are predominantly agricultural and the credit markets, as in other developing countries, are characterized by the co-existence of formal and informal lenders. While the main source of credit of rural households in Pakistan is from informal market, e.g., friends and relatives, formal credit explains one-third of rural credit (Faruqee et al., 1996). These formal loans are mainly provided by banks which are mandated to disburse a targeted amount of credit to agricultural sector on an annual basis ("agricultural credit"). Despite its limited role, past studies have clearly highlighted the importance of formal credit in Pakistan. Zuberi (1989) finds that 70 percent of total institutional credit is used for the purchase of seed and fertilizer, and concludes that most of the increases in agricultural output can be explained by changes in the amount of seed and fertilizer expenditure. Malik et al. (1991) attempt to provide evidence for the role of institutional credit in agricultural production. Similarly, their results show that institutional credit is an important determinant of fertilizer and seed expenditure. In comparison, informal loans are mainly short-term consumption loans, and arguably, it has no impacts for rural growth which requires long-term productive investment (Khandker and Faruqee, 2003). In light of the aforementioned literature, I test whether the reform has caused expansion in agricultural credit disbursement.[11] The underlying assumption is that, if the RUA branches are successful in delivering agricultural credit, then a larger number of RUA branches correlate to a larger size of agricultural credit disbursement in a district. Therefore, the evolution of the correlations between district's remoteness and size of agricultural credit disbursement along the years follows the same pattern as Figure 3.6. With credit disbursement data available only from 2007, I restrict my analysis to estimating the slope of the post-2007 trend of the series of correlations between district's rural population share and credit disbursement in a district. Therefore, the results on the credit disbursement should be taken with a grain of salt. The slope of the trend is estimated using a modified version of model (3.4), which is specified as follow.

$$Y_{d,t} = \beta_1 * RuralPopShare_d * (t-2007) + \sum_{\tau=2007}^{2012} \gamma_\tau * D_t^\tau * Y_{d,2007} + \eta_t + \sigma_d + u_{dt} \quad (3.5)$$

where $Y_{d,t}$ denotes the one of the agricultural credit disbursement variables in district d and year t. β_1 is the slope of our interest. The hypothesis is that, if the RUA branches had succeeded in delivering credit to the rural sector after the reform, we should expect to see a slower growth of credit disbursement in a more than less rural district, as a more rural district received less RUA branch openings after the reform. This would be reflected by a negative β_1.

Two types of agricultural credit are separately reported, namely, crop and non-crop credit. Non-crop items include livestock, poultry, dairy farming, fishery and forestry. For each credit category, two disbursement variables are collected and analyzed, namely, the number of borrowers (borrowers per million people) and the total disbursed credit amount (rupee per capita). Crop-farming is the main driver for the growth of agricultural sector in Pakistan, accounting for around 89% of the total agricultural GDP (Rehman et al., 2017). Its importance is also reflected in its major share of agricultural credit

[11] Data on deposit is not compiled at district level by SBP, therefore I could not evaluate the impact on it.

disbursement. Precisely, crop credit makes up 74% and 94% in loan volume and number of borrowers, respectively, of the total agricultural credit (State Bank of Pakistan, 2007). In comparison, non-crop farming plays only a minimal role in contributing to agricultural sector growth.

Applying model (3.5) to the credit variables, the results reported in column (2) to (7) of Table 3.5 and visualized in Figure 3.7. First of all, the point estimates of the aggregate number of borrowers (column 2) is positive, contradicting to the hypothesis. It may suggest that the number of borrowers grew faster in a more, compared to a less, rural district even though the less rural district captures the majority of RUA branch openings. However, the point estimate is imprecisely estimated with a large standard error equal to 230% of the size of point estimate. It is unlikely that ceteris paribus increasing the sample size to the level of the branch outreach variable (column 1) would reduce its standard error to a level that is small enough to show statistical significance. The point estimate of the aggregate amount of agricultural credit disbursement (column 3), although statistically insignificant, is indeed negative, indicating that the amount of agricultural credit grows slower in a more than less rural district by Rs.58 per capita per year, a value that is 22% of the average credit amount in 2007. This is thus large in proportion to average credit amount, but it is only 0.7% of Pakistan's poverty line in 2007.[12] Reducing the agricultural credit by Rs.58 per capita per year (column 6) in a more, relative to a less, rural district is unlikely to change the gap in economic welfare between these two districts on an annual basis. In comparison, India's policy has caused an expansion of rural credit volume in per capita per annum term by approximately 7% of India's poverty line at the time according to Burgess and Pande (2005), an expansion that is 10 times the size of Pakistan's.[13]

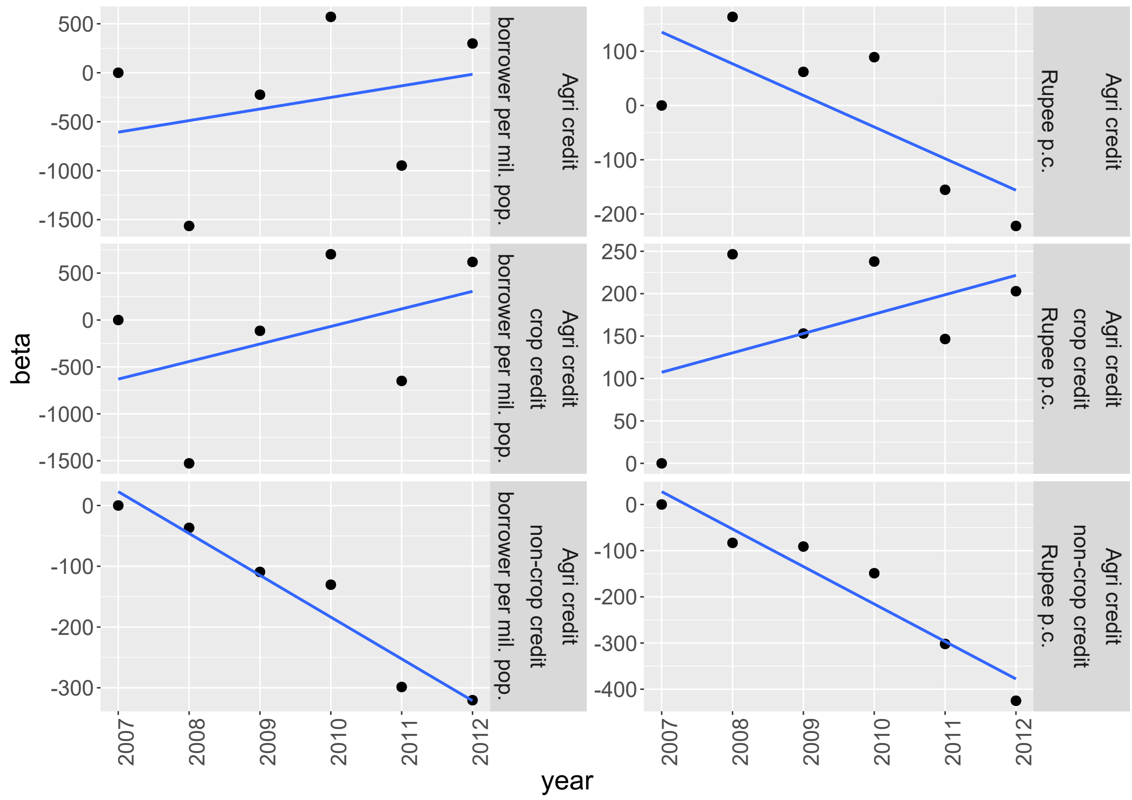
Disaggregating the credit variables by credit types, the results are presented in column (4) to (7). For crop credit, which makes up the majority of agricultural credit, both the number of borrowers (column 4) and the amount (column 5) display upward trends, indicating that the crop credit disbursement grows faster in a more than less rural district. This observation would imply that the policy benefits the more rural districts in terms

[12]Pakistan's poverty line in 2007 was estimated to be 7,650 rupees per person per year(Ministry of Finance, 2009).

[13]The estimate was taken from row and column 2 of table 1 in Burgess and Pande (2005). According to Kochar (2011) and Burgess and Pande (2005), at the beginning of India's policy, the annual average of rural credit was Rs.24,000 million, or 1.5 percent of the total credit disbursement and population in India was approximately 400 million. And the poverty line was Rs.49 per person per month, which equals Rs.588 per annum.

of crop credit disbursement even though they witness smaller growth of bank branches. However, the point estimates are imprecise, especially of the number of borrower, and small in magnitude. Therefore, we should not over interpret them. On the other hand, the trends of the non-crop credit variables are more precisely estimated. Both the number of borrowers (column 6) and the amount (column 7) of non-crop credit grew slower in a more than less rural district, a decreasing trend that is consistent with the branch expansion pattern (column 1). However, the magnitudes of the point estimates are also small. Although there is no comparable estimate from similar studies, it is highly doubted that a less of only 69 borrowers per million population per year (column 6) in a more, as compared to a less, rural district can generate meaningful difference in welfare between these two districts on an annual basis. Similarly, the magnitude of the point estimate for the amount of non-crop credit is very small compared to the poverty line and thus unlikely to generate additional disposable income that is large enough to change the life of the average person. The estimates of credit variables are visualized in Figure 3.7, with each panel presenting the effect on an outcome. The line is fitted with a slope of β_1 estimated from model (3.5) and the dots are the β_τ coefficient estimated from model (3.3).

Overall, the analysis in this part suggests that there is a lack of expansion of agricultural credit in terms of number of borrowers. Although there is evidence of expansion of the amount of non-crop credit in less rural districts, a pattern that is consistent with the branch expansion pattern, the magnitudes of the point estimates are rather small and it is doubtful that it can pass down a transformative effect on poverty.

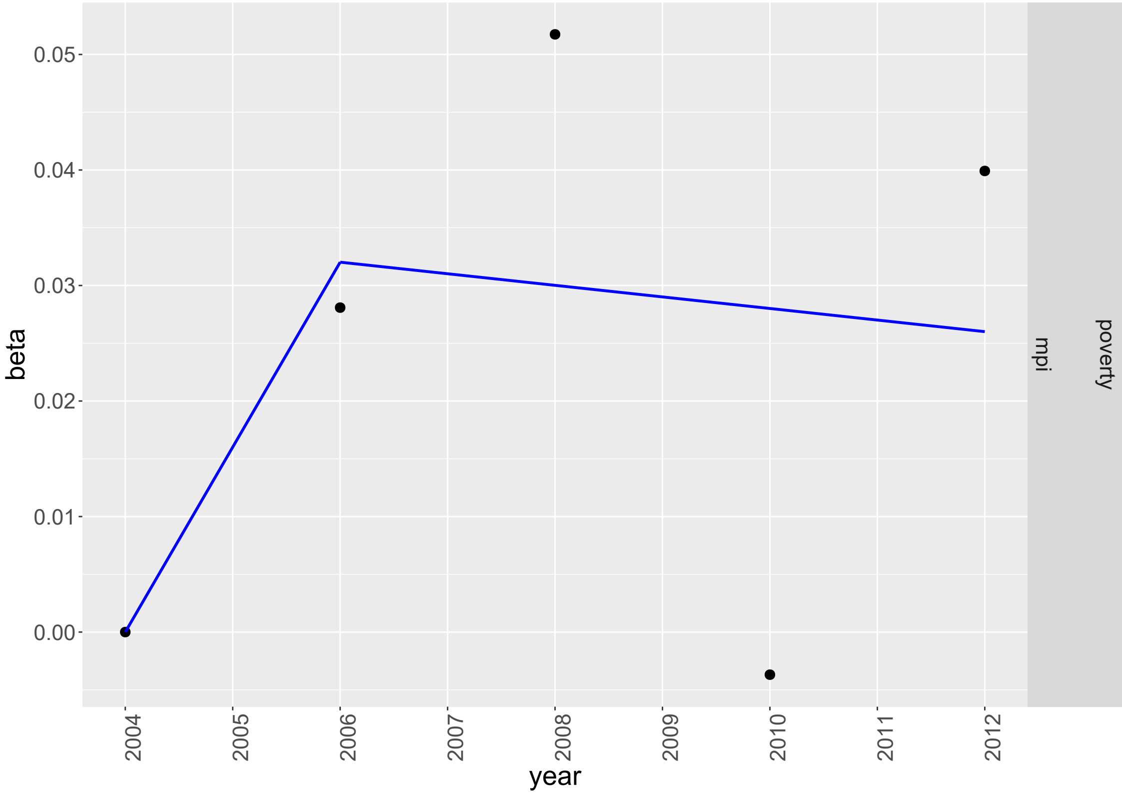


Note: The dots refer to the point estimates of the β_τ from model (3.3). The blue lines are fitted with a slope equal to the estimate of β_1 of model (3.5).

Figure 3.7: Rural Population Share and Agricultural Credit

Finally, I test whether there is policy impact on poverty, i.e., whether the imposition of the reform causes a trend break in the correlations between rural population share and poverty in a district. This is achieved by estimating model (3.4) with its left-hand-side variable replaced by "Multidimensional Poverty Index (MPI_{dt})", a poverty measurement with a range between 0 and 1. A higher value of MPI indicates severer poverty. I am interested in testing the whether β_1 is significantly different from zero.

The results are reported in column (8) of Table 3.5. Due to the lack of data before 2004, I use 2004 as the reference year and restrict the intercept change at the kink to be zero (i.e., $\lambda_1 = 0$). The insignificant β_0 and $\beta_0 + \beta_1$ suggest that the poverty in a more than less rural district evolved indistinguishably in both before and after the reform, suggesting that there is a lack of impact on poverty. Although the negative point estimates of β_1 indicates that the policy benefits the more rural districts by slowing their poverty aggravation, its magnitude is rather small as compared to the average MPI of 2004 (0.372) and therefore lacks economic significance. The estimate in column (8) is visualized in Figure 3.8. The slopes of the two line sections are the estimates of β_0 and $\beta_0 + \beta_1$ in model (3.4) and the dots are the β_τ coefficients estimated from model (3.3).



Note: The dots refer to the point estimates of the β_τ from model (3.3). The blue lines are fitted based on the estimates of model (3.4), restricting the intercept change at the kink to be zero. The slopes of the line sections of [2004, 2006) and [2006, 2012] correspond to β_0 and $\beta_0 + \beta_0$, respectively. All regressions cluster the standard errors at the district level.

Figure 3.8: Rural Population Share and Poverty

3.7 Conclusion

The main contribution of this paper is to test whether rural branch expansion was successful in enhancing financial access and reducing poverty in Pakistan. The widespread use of these programs, the mixed opinions on them and the lack of previous evaluation make this an issue of considerable interest. I show that the rural expansion in Pakistan has failed in achieving its objectives.

In 2004, Pakistan initiated a reform on its “Branch Expansion Policy” which coersed banks with a branch network size no less than 100 branches to open 20% of new branches in “Rural and Underserved Areas (RUA)”. And in 2007, this reform was rolled out to all banks regardless of their branch network sizes. I find that banks have complied with the reform by discontinuously increasing the share of new branches opened in RUA locations to exactly 20% when they became subjected to the reform. This resulted a faster growth of the stock of RUA branches in 2004 and a further acceleration in 2007. By 2012, the stock of branches in RUA locations has expanded to 2191 from 1610 branches in 2000. However, I find that the RUA branches opened after the reform were in locations that are closer to urban centers and are more likely to have another branch within 5 kilometers radius than those opened before the reform. The reform, therefore, failed to reach out to unbanked rural population, and the banked rural population share has stagnated at 57 percent. I conduct a counterfactual analysis and reveal that the fraction of banked rural population could have reached to 70 to 72 percent by 2012 had the banks chosen locations with higher rural market potential (i.e., higher rural population per branch) for those branches reserved by the quota.

Consistent with this pattern, on an aggregate level, I find that the number of RUA branches in a more (e.g., "Lower Dir" whose rural population share equals almost 1) and less rural district (e.g., "Karachi" whose rural population share equals almost 0) grew indistinguishably until 2007 when the growth rate in a more rural district was surpassed by the less rural one by 0.86 branches per million persons per year. This accounts for an accumulative difference of 5.16 RUA branches per million persons between a more and less rural district by 2012, a value that is 66% of the average RUA network size in 2000.

Although the more rural districts witnessed a smaller growth in the number of RUA branches, I do not find that they experienced a meaningfully different pattern of agricultural credit expansion from the less rural districts. The absence of agricultural credit expansion is also reflected in the previous finding that the rural population covered by the branch network has stayed stagnant even after the reform. Finally, I do not find that the poverty aggravated distinctively among districts with different degree of urbanization, suggesting a lack of impact on poverty reduction. This finding is consistent with Burgess and Pande (2005) who shows that credit disbursement expansion plays an important role in attacking rural poverty.

Taken together, my findings have shown that the rural expansion policy has failed to enhance the banking sector's rural outreach and reduce poverty. They demonstrate the need for more carefully designed government interventions for improving financial access than Pakistan's arguably naïve bank branch expansion policy. Although it is of great interest to investigate the rural branch expansion policy's effects on other sectors, especially the distortion in the profitability of the banking sector itself, severe data restrictions at the moment has post great challenges. Nevertheless, answering these questions remains an important task for future research.

Appendix

3.A Robustness of the Trend Break Model

Using the simple trend break model (3.1) to identify the effect on $BNRO_t$ relies on the change in growth pattern of $BNRO_t$. Although it is descriptive, it is capable of capturing the policy's effects on $BNRO_t$ because the confounders that affect the growth pattern are fairly limited. A valid confounder should only affect through the change in the size or location of the RUA network or the population distribution. These potential confounders include the closing and shifting of branches and population migration, however, they are unlikely to pose major threat to my identification strategy. To start with, in order to reduce the cost posed by the policy distortion, it can be expected that the banks substitute the already existed branches located in more remote locations with the new branches to be open in less remote locations. However, it has to be noticed that the regulation for branch closing and shifting remained strict after the reform. To close a branch, the bank has to make sure the closing of a branch would not leave the location unbanked. Between 2000 and 2012, only 66 branches have been closed (in comparison, 690 RUA branches have been opened during the same period) but more importantly, none of these closings happened in unbanked locations, suggesting that the bias of branch closing is limited. In terms of shifting, the banks are only permitted to shift their branch within the locality boundary (i.e., it is not allowed to shift a branch to another village/town/city). Given that my analysis is at location level (i.e., I do not consider variation within a village/town/city), the shifting would not affect my result to a meaningful extent.

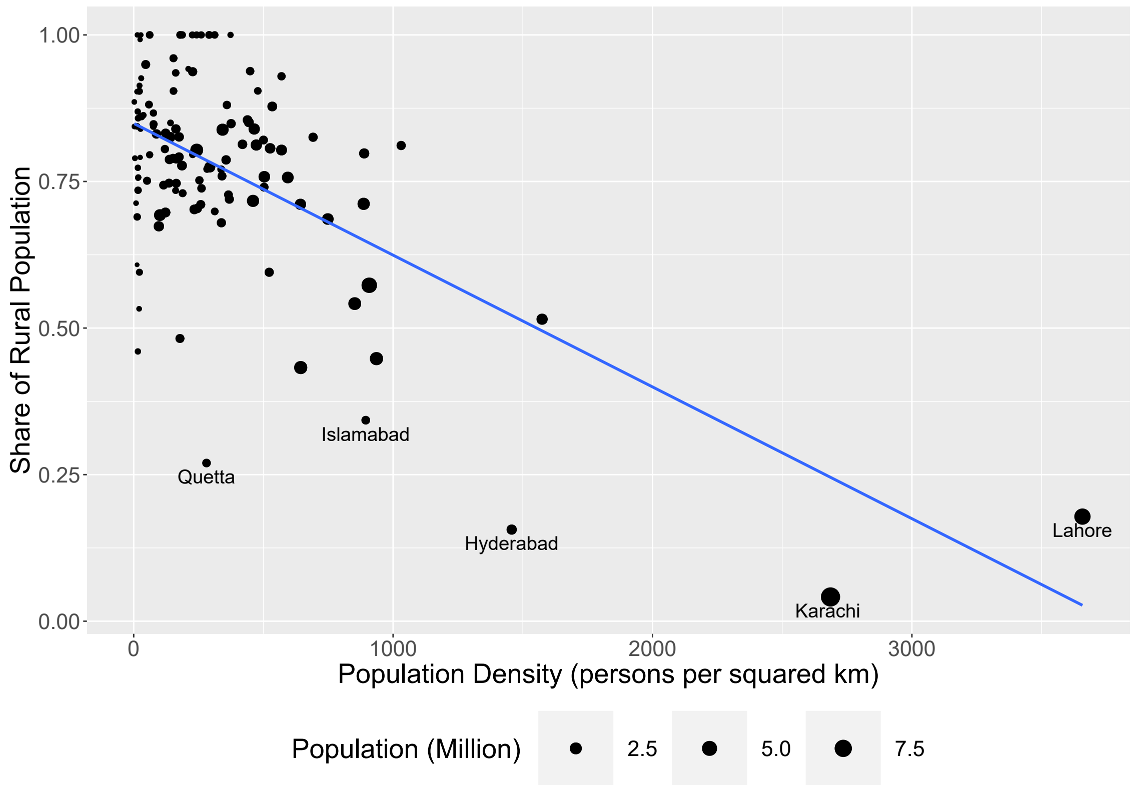
Regarding migration, my result would be underestimated if there were major migrations to RUA locations after the reform. An ideal approach to account for its influence would be to use yearly population raster data, which, unfortunately is not available as the two most recent censuses were conducted in 1998 and 2017 respectively. Nevertheless, I find that the fraction of population residing in the rural areas (calibrated using 1998 census) has stayed relatively constant, if not decreasing, from 1998 to 2017, meaning that the growth acceleration of $BNRO_t$ in 2003 and 2006 estimated in column (2) of Table 3.3 is not underestimated. If there is any bias, it is overestimated, which lends additional support my claim, suggesting that the policy failure is even more severe when accounting for population migration.

3.B Mapping Out the Rural Population

This section explains the calibration process for mapping out the rural population, which forms the basis for the calculation of branch network's rural outreach, *BNRO*.

Figure 3.10 plots the population density at 1 kilometer by 1 kilometer resolution estimated by the "Global Human Settlement (GHS)" database based on the 1998 Pakistan Population Census. The darker the color, the more densely populated. However, this map does tell us the borders between rural and urban regions. The task of this section is to draw these borders.

According to the 1998 census, urban regions refer to localities that are either metropolitan corporation, municipal corporation, municipal committee or cantonment at the time of the census. Unfortunately, a complete list of these urban localities and their boundaries are not available to the public. Nevertheless, it is well documented that the urban localities are much more densely populated than the rural (Ali, 2013). This can also be confirmed by Figure 3.9, where I collect the district-level statistics from the 1998 population census and plot each district's population density against its rural population share. Each dot represents an individual district and the size of the dot indicates the population size. A larger population density is clearly correlated with a smaller rural population share. In another word, a larger population density predicts a larger urban population share. Therefore, I resort to approximating the urban-rural border within a district by searching for a grid-level population density threshold which results in a district-aggregate rural population share that is closest to the value reported in the population census.



Note: The data are collected from the districts statistics of the 1998 population census. Each dot represents an individual district and the size of the dot indicates the population size in that district. The line is fitted by running the simple OLS model: $y_d = c + \beta * x_d + u_d$. The estimated slope, $\hat{\beta} = -0.0002$ with a standard error of 0.00003, indicates that increasing the population density by 1000 predicts an decreasing in the rural population share by 0.2. Some cities are labeled for reference.

Figure 3.9: Correlation Between Urban Population Share and Population Density

Calibrartion Procedure

The population raster data is composed by 1 kilometer by 1 kilometer sqaure grids. A grid is defined as an rural grid if its population density is smaller than the optimal threshold. Precisely,

$$rural_{id} = \begin{cases} 1, \text{if } pop_{id} < cutoff_d^* \\ 0, \text{Otherwise} \end{cases}$$

where pop_{id} is the population size in cell i of district d. $cutoff_d^*$ is the optimal population density threshold calibrated to fit the rural population share of district d according to the census ($RSCensus_d$). The $cutoff_d^*$ is obtained by solving the following minimization problem,

$$cutoff_d^* = \arg\min_x L_d(x)$$

where the loss function of district d, i.e., $L_d(x)$, is defined as,

$$L_d(x) = (RSCensus_d - \sum_{i=1}^{N_d} pop_{id} * D\{pop_{id} < x\} / \sum_{i=1}^{N_d} pop_{id})^2$$

where D{.} is a Boolean function that equals 1 if the expression within the braces is true. Intuitively, $L_d(x)$ measures the degree of difference between the census reported and calibrated (with a threshold of x) rural share of district d. Notice that although pop_{id} is real number, the N_d are finite numbers (i.e., the number of grids in a district is finite on the map). Therefore there is no guarantee of zero loss at the optimum.

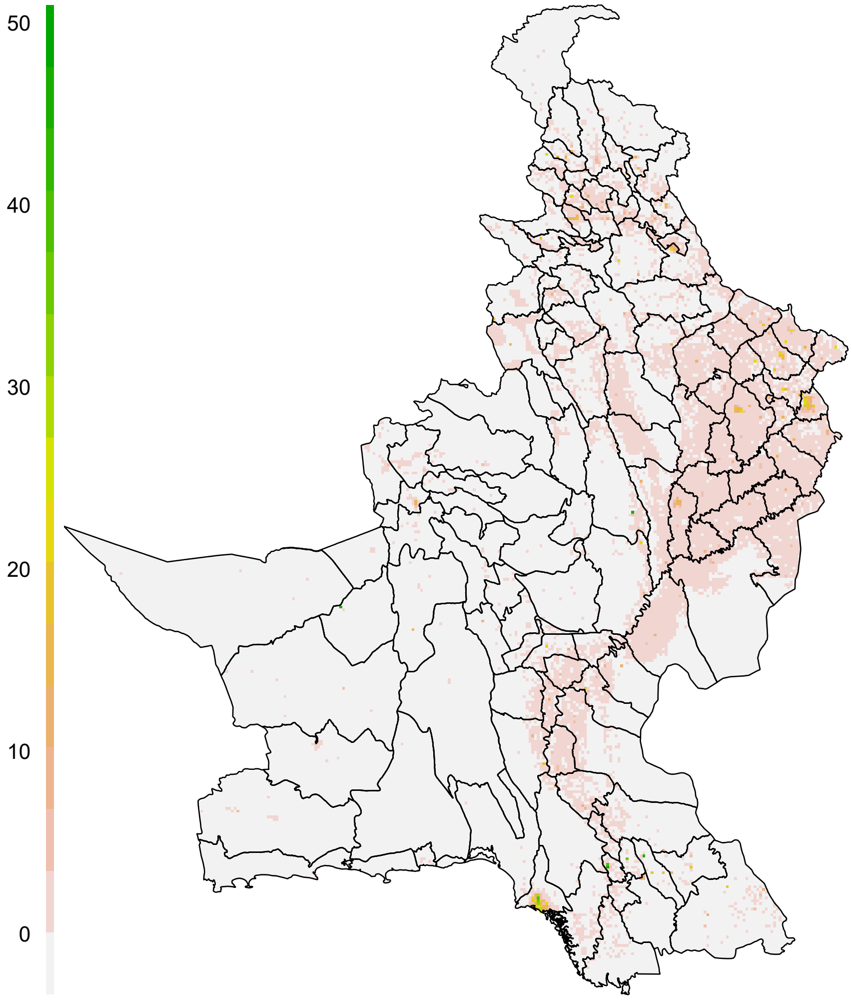
Calibration Results

The calibration outcome is presented in Table 3.6, which summarizes the loss value at the optimal thresholds. The average loss value at the optimal thresholds is 0.0004, suggesting that the calibrated rural share differ by 0.02 (i.e., $\sqrt{0.0004}$) on average from the one reported in the census. However, notice the distribution of the optimal loss values is highly right-skewed, this indicates that most of the districts have almost perfect fit (loss value equals to zero and one tenth the average at the 50th and 75th quantile, respectively).

Table 3.6: Calibrated Result: Loss at the Optimal Threshold

	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
L_d^*	131	0.0004	0.002	0.000	0.00000	0.00000	0.00004	0.017
$cutoff_d^*$	131	11,841	15,018	0	2,720	6,785	15,829	99,506

Under this calibration, the rural grids are identified. The calibrated aggregate rural population share (61.9%) fits the census (64.5%) quite well. In the following figure, I plot the $rural_{id}$ (i.e., the rural cells) which forms the base for identifying the rural locations in the following section.



Note: The base map consists of squre grid, with each grid represents an area of 1 km^2. The color of the grids indicates their population density. The unit is 1000 persons per km^2. The black lines represent district borders.

Figure 3.10: Population Distribution in 1998

Note: The areas in black indicate the rural region identified by the calibration process. The areas in white refer to urban region.

Figure 3.11: Urban-Rural Decomposition

CHAPTER 4

School Grants and Education Outcomes: The Impacts of the Non-Salary Budget Reform in Punjab Pakistan

**Joint work with Kafeel Sarwar*

Abstract Whether increasing school funding is effective in improving education quantity and quality is critical to the education policymakers. In this paper, we analyze the impacts of a decentralized school grant program in Punjab Pakistan, the "Non-salary Budget (NSB)" reform. Under this reform, each public school received about ten times the amount of grant from government prior to the reform. To identify the effects of the reform, we explore the spatial and temporal variation of the policy rollout. We find that the reform has significantly improved the school infrastructures condition. However, we do not find any discernible effects on student enrollment, attendance rate or test score in the short to medium run.

4.1 Introduction

Education has long been viewed as an important determinant of economic well-being. From the theoretical perspective, education promotes economic growth by improving human capital, increasing innovative capacity of the economy and facilitating diffusion of knowledge that is needed to implement the innovative technologies (Hanushek and Wößmann, 2010). The empirical literature has also established a positive correlation between better education (quantity and quality) and better socioeconomic outcomes, e.g., health (Strauss and Thomas 1995, Schultz 1997, Schultz 2002). Sen (2001) has regarded education itself as an intrinsic good. The development policymakers also show their interest in education, and they know how much education is crucial for the process of economic development. For example, the United Nation's "Sustainable Development Goals" call for quality education for all by 2030.[1]

These arguments have provided justifications for widespread education interventions implemented in the hope of improving education quantity (i.e., student enrollment and attendance) and quality (i.e., learning outcomes like test score). Two broad categories of interventions based on, namely human capital and physical input have been brought to test in the literature. Human capital based intervention aims to improve human capital in the education sector through pedagogy innovation (Duflo et al. 2011, Chen et al. 2017, Naik et al. 2020), teacher's accountability (Duflo et al. 2011, Duflo et al. 2012, Muralidharan and Sundararaman 2011, Mbiti 2016), students' tracking (Duflo et al. 2011) and school governance (Banerjee et al. 2010, Pradhan and de Ree 2014). This method has generally yielded positive effects on students' test scores and attendance. On the contrary, physical inputs based method seeks to improve education quantity and quality by easing the constraints on physical resources. These resources could be textbooks (Glewwe et al. 2009, Frölich and Michaelowa 2011), flipchart (Glewwe et al. 2004), radio (Pridmore and Jere, 2011), computer assistance (Banerjee et al. 2007, Cristia et al. 2017, Kremer et al. 2013), library (Borkum et al. 2012), school buildings (Newman et al. 2002), sanitation (Adukia 2017), etc. However, a consensus has emerged from these studies that investing in physical inputs does not have any discernible impact.[2]

[1]The United Nations Sustainable Development Goal 4 aims at "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". For more details on targets and facts see: https://www.un.org/sustainabledevelopment/education/

[2]Multiple review papers show that pedagogical interventions (Kremer et al. 2013, Conn 2014), the

A specific constraint that may limit the effectiveness of physical inputs on improving education outcomes is the centralized provision of these inputs. Four issues arise due to the centralized provision of inputs. First, there are often mismatch between the demand at the local level and the provision determined at the national level regarding the school resources. Decision-makers at the local level have a better understanding of the schools' needs than those at the national level. Therefore, they are in a better position to identify schools' deficiencies and to use the resources efficiently (Carneiro et al. 2015, Hanushek and Woessmann 2011). For example, in Kenya, Glewwe et al. (2009) do not find any effect of the textbook on students' test scores because the textbooks provided to the schools are too complicated for the students and teachers. Second, centralized provision is more prone to delivery failure. The information asymmetry between the central decision-maker and the local implementation administrator appears to be more severe than in a decentralized scenario. In Sierra Leone, Sabarwal et al. (2014) show that the centralized provision of textbooks does not have any impact on students' test scores. The reason is that most of the books have been kept in the schools' storage room, failing to reach the students. Third, political pressure to initiate visible education policies may also lead education systems to invest in less effective inputs (Mbiti, 2016). For example, a politician may choose to give each classroom a computer not because it is an effective way to improve education outcomes. Instead, it is due to the fact that it is more visible to show the existence of computers as an achievement to the voters. Finally, complementarities among multiple inputs may limit the effectiveness of a single input (Mbiti et al., 2019). For example, giving a tablet to each school without training the teachers to incorporate the tablet into their teaching would not be likely to affect the students' learning.

One potential way to avoid these issues and improve efficiency is to provide unconditional cash grants and devolves the financial decision-making to the school councils, i.e., decentralized school grants.[3] It seeks to increase efficiency by making financial decisions more transparent to communities, reducing corruption and incentivising localised investment in high quality teachers and materials (Carr-Hill et al., 2018). However, it

computer uses and technology (McEwan 2015), and learning material (Krishnaratne et al. 2013) are the effective ways to improve the students' learning outcomes. Above mentioned reviews generally believe that input-based policies are mostly unsuccessful in increasing students' achievements.

[3]The school council is a school-based management committee that consists of teachers, parents and notable local members. Providing decentralized school grants has emerged as a part of relatively recent practices for educational decentralization. These practices focus on devolving the decision making concerning the curricula, finance, management and teachers (Bruns et al. 2011, Slater 2013)

is debated that, decentralized school grant does not necessarily bring efficiency gains because the school councils do not necessarily possess sufficient capacity in making complex decisions at the local level. Although the allocation of school grants is associated most strongly with the devolution of financial decision-making (e.g., decisions about how resources should be allocated within a school), depending on the nature of the needs identified at school level, the allocation of school grants might also touch on managerial decision-making (e.g., the recruitment and monitoring of teacher performance) or educational decision-making (e.g., decisions related to improving the articulation of a school's curriculum) (Carr-Hill et al., 2018). Therefore, when implemented in contexts with a lack of decision-making capacity, decentralized school grants may result in inefficient and ineffective use of resources (Hanushek and Woessmann, 2011).

In this study, we address two questions. First, does providing decentralized school grants improve the education quantity and quality? Second, does this decentralized grant have any impact on human capital (measured by teacher's attendance rate) and school's physical infrastructure? Precisely, we study the impacts of the "non-salary budget (NSB)" reform in Punjab-Pakistan. Each public school in Punjab provides education for 1st to 5st graders. Before NSB reform, public schools in Punjab received Rs. 20,000 (approximately 200 USD in 2013) per annum, and an annual addition of Rs. 20,000 if the schools offer education beyond the 5st grade.[4] This fixed grant has proven to be extremely insufficient even for the basic school maintenance. A rough estimate of the annual recurrent expenditure of a school is Rs. 70,000.[5] The underfunding for the non-salary expenditure has persistently hindered public schools' ability to attract and retain students.

Against this backdrop, the Punjab government implemented the NSB reform in 2013. The NSB reform alters the funding allocation rule from a fixed to a need-based rule. The need-based formula includes indicators such as school level, student enrollment, furniture deficiency and building condition. Under this need-based rule, the average annual non-salary budget per school amounts to Rs. 220,000, which is about ten times the amount received under the erstwhile fixed allocation rule. The NSB reform was rolled out in three phases to cover the whole province. In 2013, nine districts were selected for phase one. In 2014, additional nine districts were added in phase two. Since 2015, the NSB reform

[4]Primary: grade 1-5. Middle: grade 6-8, High: grade 9-10 and Secondary High: grade 11-12.

[5]This estimation is based on informal interviews with several school headteachers and members of teacher unions across Punjab.

was extended to the whole Punjab province. The reform's staggered roll-out schedule has created temporal and geographical variations in policy treatment which we exploit to estimate the reform's impacts.

We combine data from various sources. First, we collect schools' financial accounts and infrastructure conditions from the "Annual School Census (ASC)" administered by the school and education department of Punjab (SED). The ASC is also used by the officials to calculate the school's NSB entitlements under the new rule. Second, we collect the test scores of grade five students, the terminal grade of primary school, from the Punjab Examination Commission (PEC), which is the organization for conducting standardized exams for students of Punjab province. Third, we use English reading and mathematical ability scores of children of age 5 to 16 from the "Annual Status of Education Report Pakistan (ASER)" as an alternative measure of learning outcomes. Fourth, we measure the enrollment rate by the share of children of age between 5 and 16 who are currently enrolled in a school using the "Multiple Indicator Cluster Survey Punjab (MICS-Punjab)", a UNICEF-administered household survey. Finally, we obtain student's and teacher's attendance from Punjab Monitoring and Implementation Unit (PMIU), a government department that administrates monthly attendance surveys at the school level.

We report four sets of results. First, regarding school's financial account, the NSB reform has significantly increased school's annual income and expenditure by Rs. 180,000 and Rs. 135,000, respectively. Second, we find that the reform has significantly improved the conditions of school infrastructures by 0.045 standard deviations. This effect is driven by the improvements in the existing infrastructures, e.g., the fraction of functional toilets, complete boundary wall and safe school building. In terms of human capital, we find that teacher's attendance rate has increased slightly by 0.9 percentage points, a value that is only 1% of the average at the baseline. Third, we do not find any effect on education quantity as measured by student's attendance rate and the enrollment rate among children of age 5 to 16. Fourth, in terms of education quality, we do not find any impacts on students' numeracy and literacy test scores and ability scores.

This paper contributes to the debate about the effectiveness of decentralized school grants on improving the education outcomes in developing countries. The empirical evidence is still limited and mostly based on small scale RCTs. Nevertheless, this limited empirical literature has suggested that decentralized school grants' impacts on education outcomes are only moderate and its effectiveness is highly contextual. Some lessons can be drawn from these studies. These include the following.

First, local decision-makers should be supported to build sufficient capacity in understanding school-based management and making informed spending decisions that are likely to improve education outcomes. Blimpo et al. (2015) conduct an RCT in the Gambia to study the complementarity between school grant and training for the school management committee. They find strong complementarity between the two treatments. Providing schools with grants alone does not have any impacts. Studying a school grant program that encourages parental participation in school management in Nigeria, Beasley and Huillery (2017) show that although the program is capable of improving parental participation, it fails to delivery impact on students' educational outcomes because the parents lack the capacity to make effective decisions in school management.

Second, the effectiveness may be hampered by behavioral responses by the stakeholders. Das et al. (2013) study the school block grant programs in India and Zambia using experimental designs and find that the school grants improve the students' test scores only if the grants are unanticipated to the parents. If anticipated, the grants crowd out parental investment in their children's education. Competition from the education market also affects schools' decisions. Andrabi et al. (2018) use RCT to study how market competition affects the effectiveness of school grants among private schools in Punjab Pakistan. They find that while villages with all schools given the block grant (i.e., high saturated villages) see improvement in student's test scores, villages with only one school given the grant (i.e., low saturated villages) see improvement in enrollment. They explain that, in order to increase profit, private schools in low saturated villages act as monopolists (on the residual demand from the untreated schools) and bring in additional students by increasing infrastructures investments, while schools in high saturated villages would trigger the price war if they do the same because of the lack of monopolistic power. Therefore, schools in the high saturated village resort to improving test scores (in order to raise tuition and profit).

Third, spending on improving teaching human capital seems to be promising. Carneiro et al. (2019) shows that a school grant program in Senegal improves the test scores of the second graders, especially in schools that spent on human resources training, rather than on the acquisition of school materials (e.g., textbooks). Consistent with this finding, the impacts of the Nigeria program as studied by Beasley and Huillery (2017) is mainly captured by schools where teachers have benefited from the spending.

To this strand of literature, our study's contribution is three-fold. First, we add the evidence to the limited and recent empirical literature on the effectiveness of decentralized school grants on improving education outcomes in the context of Punjab-Pakistan's "NSB" reform, which, to the best of our knowledge, has not been evaluated erstwhile. Second, rather than in the context of an experimental setting, we study a real large-scale policy, which is more likely to capture the real-world effect. Third, this study sheds light on Punjab-Pakistan's progress on achieving the "Sustainable Devlopment Goal (SDG)" which set multiple education-related targets: equal access, gender equity, equitable quality education, conducive learning environments and adequacy of trained teachers, all of which are related to our results.

A major limitation of this paper is the short-run nature of the findings because our identification strategy relies on the variation in treatment timing across the three reform phases which differ only by one to two years. Long-run analysis is impossible as the policy has covered the whole province within a time span of two years. Therefore, it is possible that different results would emerged had the reform allowed longer duration of each phase, especially for education quality which takes longer time to improve than school infrastructure. However, the results presented here are still useful. The richness of the data on school inputs gives some insights into the decisions of spending.

The rest of the study is organized as follows: section 4.2 introduces the institutional background and the NSB reform. Section 4.3 describes the data. Section 4.4 discusses the identification strategy. Section 4.5 presents the main results. In section 4.6, we conduct some robustness checks. Finally, section 4.7 concludes.

4.2 Institutional Background and the Non-salary Budget Reform

Insufficient funding for education has been a critical problem for achieving the goal of providing quality education for all in Pakistan. Twenty-five million children are still out of school in Pakistan, of which nine million are in Punjab, Pakistan's largest province (Alif Ailaan, 2014a). Education spending in Pakistan is stagnant at 2 percent of GDP in the past two decades.[6] This low level of education funding is falling short of the international benchmark, 4%. More severely, the share of the non-salary budget in the total education budget is only 12%. In Punjab, this share stands at a mere 3%, while the salary and developmental education budget make up 86% and 11%, respectively (I-SAPS, 2015).

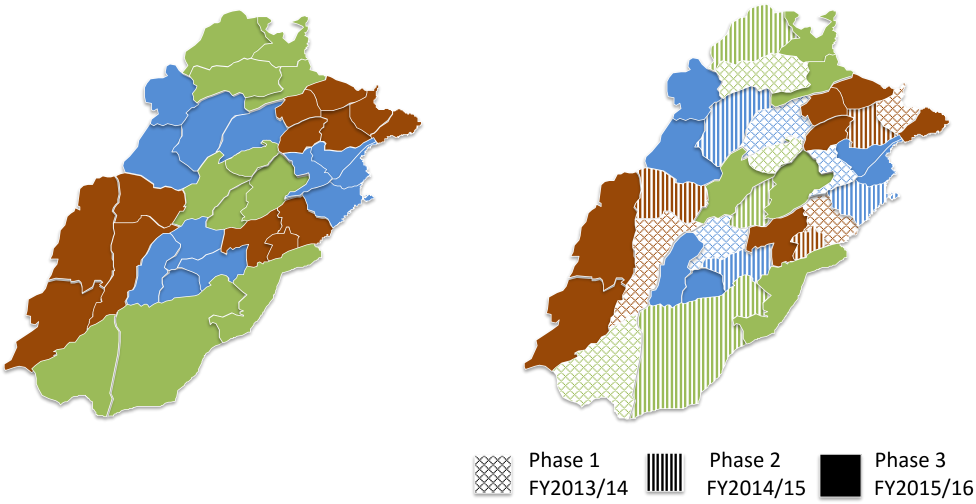
Public schools' non-salary expenditure is funded by two sources, provincial government and parents. Prior to the NSB reform, an average public school received a fixed grant of Rs. 20,000 or 40,000 per annum from government depending on the school level, and approximately Rs. 24,000 per annum from parental contribution (Rs. 20 per student per month). The funding for the non-salary expenditure is extremely insufficient for any school to maintain its facilities. A rough estimate of annual recurrent expenditure is Rs. 70,000 (excluding building maintenance). According to our interviews with the headmasters and teachers of the public schools, some schools manage to cover the deficit in NSB by selling trees from the schoolyard and asking for donations in cash or in kind from local villagers. In some cases, the headmasters or teachers use their personal relations to approach the member of the Provincial and National assembly to seek discretional funds for schools. Occassionally, teachers advance the expenditure from their own pockets, and reimburse themselves from the next round of funding. There are even cases where schools impose fine on the students, if they are late from school or not wearing school uniform.

Against this backdrop, the school education department, with the support of the World Bank, had designed and implemented a need-based non-salary expenditure funding policy, i.e., the non-salary budget reform, starting in the financial year (FY) 2013/14. Under this reform, the annual amount of NSB is no longer fixed but based on a set of school-need indicators, which include school level, student enrollment, furniture deficiency and infrastructures condition. An average public school receives an annual grant of Rs. 220,000 which is paid in 4 installments. School councils are made accountable and responsible for planning, managing and administrating the NSB grant. The grant can be used to cover a wide arrange of non-salary expenditures, e.g., construction, stationery, furniture, teaching materials, provision of electricity and water and sports goods.

The reform was rolled out in three phases and planned to cover the whole province in three years. Figure 4.1 visualizes the geographical coverage of this staggered roll-out

[6]This figure is extracted from various waves of the Pakistan Economic Survey by the Ministry of Finance of Pakistan. http://www.finance.gov.pk/survey_archieve.html

schedule. The province of Punjab is divided into nine divisions, with each division consisting of 3 to 6 districts. In total, there are 36 districts in Punjab. The selection of districts into different phases is stratified at the division level. In the first phase in FY2013/14, one district per division was selected (i.e., nine districts in total). In the second phase in FY2014/15, an additional district per division was selected (i.e., 18 districts in treatment including phase one districts). Finally, in the third phase in FY2015/16, the program was rolled out to all 36 districts.[7]



Note: The left panel shows the map of Punjab-Pakistan which consists of 36 districts whose boundaries are drawn in white. These 36 districts make up nine divisions that are separated by colors. The right panel shows the "NSB" reform's coverage across districts and phases.

Figure 4.1: Non-salary Budget Reform's Rollout

The selection of districts into different phases was determined by the reform committee that was led by three members. While the exact selection rules for pilot districts were not documented, the committee disclosed to us that the rules were not systematic and the selection was rather based on overall impressions of the districts' ability of information updating, as the committee relied on this ability to monitor the reform's implementation and progress. The information was delivered by the districts' education officers (EO). A typical EO is made responsible for several public schools. Their responsibilities include conducting school surveys, maintaining administrative records of schools, implementing education initiatives in the field and ensuring delivery of quality education. We show the

[7]Nine districts of phase 1: Chakwal, Chiniot, Khanewal, Muzaffargarh, Nankana Sahib, Okara, Rahim Yar Khan, Sargodha and Sialkot. Additional nine districts of phase 2: Attock, Bahawalpur, Gujranwala, Kasur, Khushab, Layyah, Pakpattan, Toba Tek Singh and Vehari

robustness of our main results to the addition of a proxy for the EO's effort in the section of robustness check.

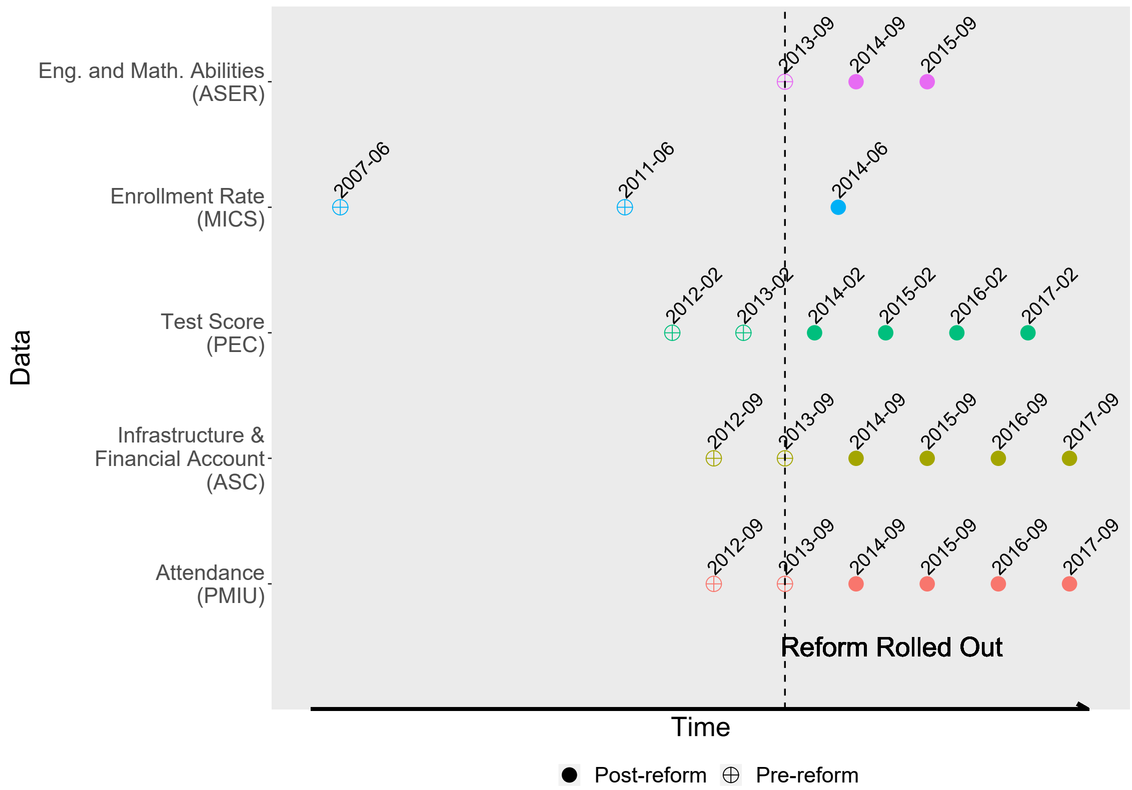
4.3 Data

We use both administrative and household survey data for our analysis. We compile a school-year panel using administrative data from three sources. First, we collect the information on school infrastructure, facilities and financial accounts from the "Annual School Census (ASC)", which is conducted in September annually by the School and Education Department (SED). The ASC is also used by the reform committee to calculate the schools' NSB entitlement. Secondly, we obtain the teacher's and student's attendance rate from the Program Monitoring and Implementation Unit (PMIU). In order to collect this information, PMIU hires approximately 900 Monitoring and Evaluation Assistants (MEAs) to administer the monthly survey in all public schools in Punjab. The MEAs are assigned to school clusters in such a way that they are able to visit at least 4 schools per day. The MEAs' assignment is rotated every month, which prevents MEAs from forming personal relationships with the school staff of a particular area. We aggregate the monthly data to annual level. Lastly, we obtain the standardized test scores of the fifth-graders who need to pass the test in order to advance to middle school. The test scores are obtained from the Punjab Examination Commission (PEC), SED's autonomous body that administers the standardized exam. The exam is held annually in February and consists of five subjects, namely Urdu, English, Islamic study, Math and Science. Each subject requires 40 points (out of 100 points) to pass. We categorize these five subjects into two categories: numeracy and literacy, where numeracy consists of Math and Science, and literacy contains the other three subjects. We calculate the category-wise average score and standardize this average score to a z-score, year by year.

In addition, we compile a district-year panel from two independent household surveys, namely, Multiple Indicator Cluster Survey (MICS-Punjab) and the Annual Status of Education Report (ASER-Pakistan). MICS-Punjab is part of the global MICS programme that was developed by UNICEF in the 1990s as an international multi-purpose household survey programme to support countries in collecting internationally comparable data on a wide range of indicators on the situation of children and women. MICS-Punjab is not available on an annual basis and we only use the three most recent survey waves (i.e., 2007, 2011 and 2014) which are cross-sectional household surveys. However, all survey waves provide a district identifier for each sample household that allows us to build a district-year panel. For each household, MICS-Punjab records basic information of all household members, including age and whether the member is currently enrolled in a school. Following the UNICEF's criteria, we calculate the enrollment rate of a district as the share of children of age between 5 and 16 that are currently enrolled in a school. To measure the children's learning outcomes from another perspective, we use the ASER-Pakistan children surveys of the wave of 2013, 2014 and 2015. ASER-Pakistan is conducted in September when the new school year starts. It aims to provide reliable estimates on the schooling status of children of age between 5 and 16 years residing in Pakistan. Similar to MICS-Punjab, ASER-Pakistan is a cross-sectional household survey with district identifiers that allow us to compile the information to a district-panel. ASER-Pakistan grades children's English and mathematical abilities by five levels.[8] For our analysis, we rescale the scores as follows. "English Ability" is graded as 1 if the child can't read, 2 if can recognize letters/words or 3 if can read sentences. Similarly, a child's "Mathematical Ability" is graded as 1 if s/he has zero knowledge, 2 if can recognize numbers or 3 if can perform mathematical operations. Under this scale, the higher the score, the higher the abilities to read English or perform mathematics.

The data structure is presented in Figure 4.2. Our school-year panel consists of six years of data on an annual frequency. The reform started to roll out in September 2013. However, according to a third-party evaluation report on the reform implementation, the grants did not arrive at the schools until the last quarter of 2013 due to the systematic technical issues at the provincial and district governments (Cambridge Education, 2014). Therefore, we regard the first two years as the pre-reform. Our district-year panel is relatively shorter with three years of data from ASER and MICS. The timing of the reform roll-out determines that we have two years of pre-reform data for enrollment rate and one for children's English reading and mathematical abilities.

[8]English: "Can't read", "Recognize capital letters", "Recognize small letters", "Can read words" and "Can read sentences". Mathematics: "Zero knowledge", "Recognize numbers 1-9", "Recognize 10-99", "Can perform Subtraction" and "Can perform division".



Note: "English and Math Abilities" and "Enrollment Rate" are district-level variables aggregated from household surveys of ASER and MICS. "Test Score", "Infrastructures & Financial Account" and "Attendance" are school-level variables collected from administrative data.

Figure 4.2: Data Structure

The summary statistics of the pre-reform data is presented in Table 4.1. There are 48,310 public schools in our sample, 47% of which are boy's schools. All public schools offer education from grade 1 to 5, which is the primary-level education. Among all public schools, 16% offer additional grade 6 to 8, i.e., middle-level education. Furthermore, 13% offer teaching at grade 9 to 10, i.e., high-level. In comparison, only 1% public schools offer additional grade 11 to 12 teaching, which is the secondary-high-level education in Punjab-Pakistan. Regarding the financial account, an average public school receives approximately Rs. 32,000 from government and spent Rs. 40,000 in total per year. It should be noted that the public schools also receive funds from parental contributions and NGOs whose records are not available for most of the schools and therefore not included in our analysis. Absence of basic facilities like drinking water, electricity, main gate or toilet is rare. However, the conditions of the school buildings and boundary wall are relatively poor, with only 59% of the schools reporting their building conditions to be safe and 87% reporting their boundary walls to be complete. To measure the overall status of the infrastructures of a school, we calculate an infrastructures index, which takes the simple average of the z-scores of all infrastructures variables aforementioned. Intuitively, it measures how different a school's infrastructures status is from the average status of all schools. Similarly, we calculate a sports facilities index to capture the overall sports supplies sufficiency in a school. These supplies include cricket, football, hockey, badminton and table tennis. Regarding attendance, student's attendance rate in different levels of education are quite high. 85% of the registered students in grade 1 to 5 were present during a MEA's visit. This value is 88%, 87% and 84% for grade 6-8, 9-10 and 11-12, respectively. Teacher's attendance rate in an average public school stands at around 87%. In terms of the scores of the grade-5 test, which a student needs to pass to advance to the middle-level education (i.e., grade 6-8), the average raw score of all numeracy subjects is 41 points which is on the verge of passing. Raw literacy score appears to be higher, reaching to 50 points. In our analysis, we normalize the scores using the same method in calculating the infrastructures index. Score normalization is a common practice in assessing the impacts on test scores. The normalized score allows for comparison across samples, as well as with the results from other studies (Glewwe and Muralidharan, 2016).

Table 4.1: Summary Statistics: Pre-reform

	Mean	Std.dev	Min	Max	N
School Cross-sectional					
Male	0.470	0.500	0	1	48,310
Primary	0.690	0.460	0	1	48,310
Provide education for:					
Grade 6-8 (Middle)	0.160	0.370	0	1	48,310
Grade 9-10 (High)	0.130	0.330	0	1	48,310
Grade 11-12 (Secondary High)	0.010	0.110	0	1	48,310
School Panel					
School grant received (10,000 Rs.)	3.250	7.490	0	218	82,550
Expenses (10,000 Rs.)	4.130	8.190	0	235	82,550
Building condition is safe	0.590	0.490	0	1	82,550
Boudary wall is complete	0.870	0.340	0	1	75,097
Drink water exists	0.980	0.150	0	1	82,550
Electricity exists	0.820	0.380	0	1	82,550
Main gate exists	0.920	0.280	0	1	82,550
Toilet exists	0.998	0.048	0	1	82,550
Share of functional toilets	0.950	0.180	0	1	82,550
Infrastructures index	0.035	0.390	-3.320	0.450	82,550
Sports Facilities index	0.003	0.600	-0.320	4.460	82,550
Teacher attendance rate	0.870	0.100	0.060	1	96,574
Student attendance rate					
Grade 1-5	0.850	0.090	0.040	1	91,678
Grade 6-8	0.880	0.070	0.080	1	26,447
Grade 9-10	0.870	0.090	0.020	1	11,993
Grade 11-12	0.840	0.150	0.210	1	1,025
Grade 5 Standardized Test Score					
Numeracy (Raw)	40.949	12.912	0.000	90.638	86,989
Literacy (Raw)	50.044	11.900	0.000	86.167	86,989
Numeracy (Normalized)	0.040	0.990	-3.220	3.800	86,989
Literacy (Normalized)	0.050	0.980	-3.840	3.170	86,989
District Panel					
Math. Ability Score	2.389	0.156	1.978	2.634	36
English Ability Score	2.244	0.162	1.812	2.543	36
Enrollment Rate	0.708	0.121	0.390	0.895	71

Note: Numeracy subjects: Mathematics and Science. Literacy subjects: English, Urdu, Islamic religious study. Infrastructures index is the average of the z-scores of the following variables: Building condition is safe, Boundary wall is complete, Drink water exists, Electricity exists, Main gate exists, Toilet exists and Share of functional toilets. sports facilities index is the average of the z-scores of five dummy variables for the existence of the following five respective sports facilities, i.e., cricket, football, hockey, badminton and table tennis.

As revealed by the reform committee, the selection into early phases is based on the districts' information update ability, rather than purposefully on the districts' improvement potentials. In order to assure that there is not targeting on the observables, we conduct a balance test using the baseline data (first year of each variable). Specifically, we run the following regression model,

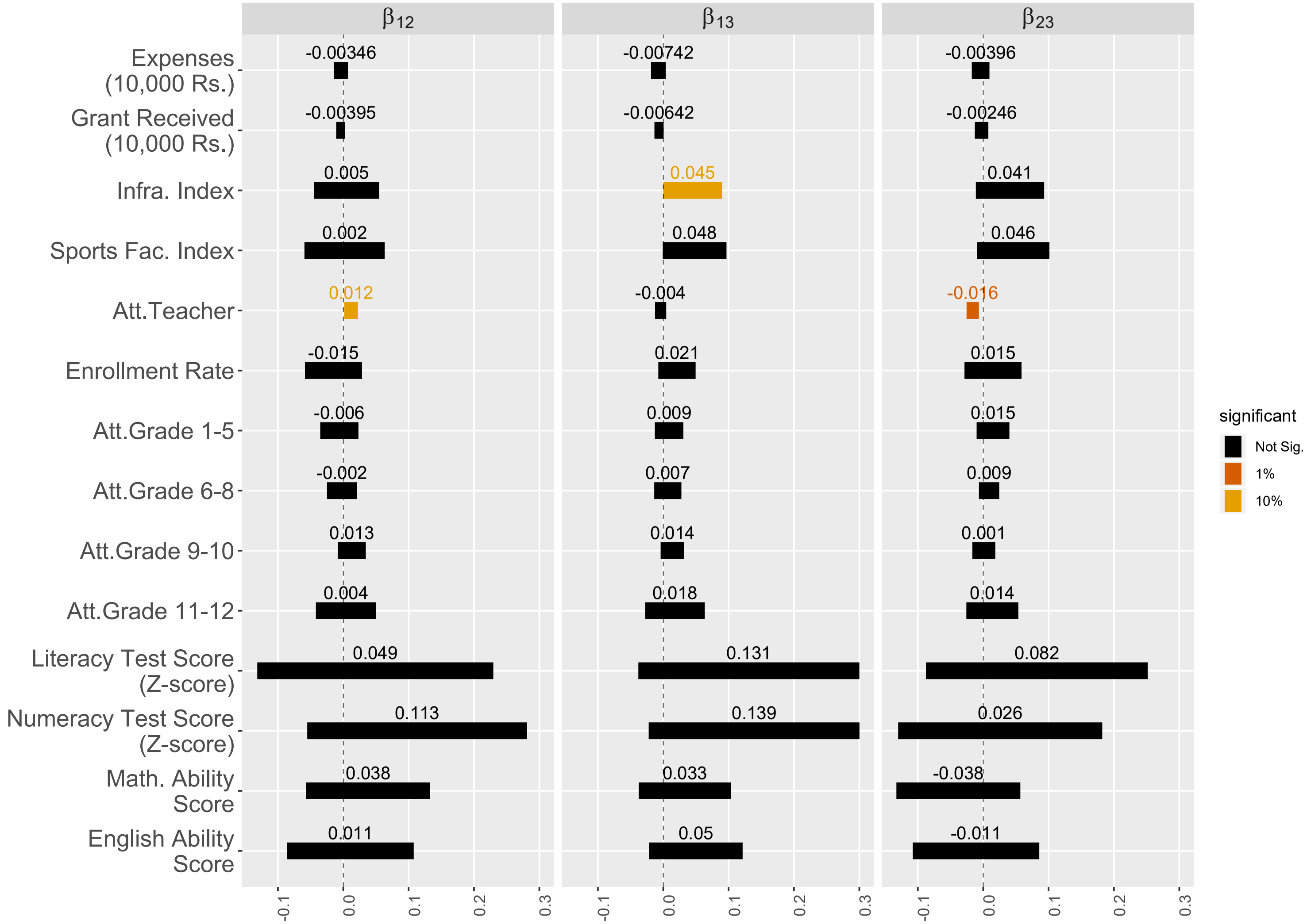
$$\begin{aligned} y_{dr} &= \beta_{13} * PHASE1_d + \beta_{23} * PHASE2_d + \eta_r + u_{dr} \\ \beta_{12} &= \beta_{13} - \beta_{23} \end{aligned} \tag{4.1}$$

where y_{dr} is the observable in district d and division r at baseline. $PHASE1_d$ and $PHASE2_d$ are respectively the dummy variables for the first and second phase assignments of district d. Since the selection is stratified at the division level, we add the division fixed effects, η_r. Therefore, β_{13} captures the baseline difference in observable of the first, relative to the third, phase districts. Similarly β_{23} is the difference of the second, relative to the third, phase districts. The difference between the first and second phase (with the second phase as reference), β_{12}, is estimated by the difference between β_{13} and β_{23}.

Running model (4.1) on 14 outcome variables, we report the results of this balance test in Figure 4.3. Each bar represents an estimate of β. We plot the point estimates of β_{12}, β_{13} and β_{23} along with their 90% confidence interval in the left, middle and right panel, respectively. There are some statistically significant difference in teacher attendance rate and infrastructures index. Teacher attendance rate in phase 2 appears to be the lowest and phase 3's infrastructures condition is worse than phase 1. However, among the 42 point estimates, only 3 show up significant and 2 of them are only at the 10% level. Given the non-experimental setting in this study, we consider the observables are sufficiently balanced at the baseline, i.e., there are no obvious selection on these observables regarding the phase assignment.

4.4 Empirical Strategy

In our analysis, we focus on estimating the average intent-to-treat effect (AITT), i.e., the average effect of becoming subjected to the NSB reform. In doing so, we exploit the spatial and temporal variations induced by the policy roll-out. Specifically, we use the



Note: Numeracy subjects: Mathematics and Science. Literacy subjects: English, Urdu, Islamic religious study. Infrastructures index is the average of the z-scores of the following variables: Building condition is safe, Boundary wall is complete, Drink water exists, Electricity exists, Main gate exists, Toilet exists and Share of functional toilets. sports facilities index is the average of the z-scores of five dummy variables for the existence of the following five respective sports facilities, i.e., cricket, football, hockey, badminton and table tennis.

Figure 4.3: Balance Test at the Baseline

following econometric specification,

$$y_{sdt} = \beta * TREATED_{dt} + \alpha_s + \gamma_t + u_{sdt} \tag{4.2}$$

where y_{sdt} is the outcome of school s in district d and year t. $TREATED_{dt}$ is the dummy variable for district d entering treatment in year t. α_s and γ_t are school and year fixed effects. β thus captures the average effects of entering the NSB reform. Essentially, this is a double-differences approach with variation in treatment timing. Effectively, β is a weighted average of all possible two-group/two-period double-differences estimator in the data (Goodman-Bacon, 2018). β delivers only the short-run (one to two years) effects of the reform because the treatment timing across the three phases differs only by one (between phase-1 and 2, or phase-2 and 3) to two (between phase-1 and 3) years.

This specification controls for the time-variant factors common to all schools, i.e., year fixed effects, and the school specific time-invariant factors, i.e., school fixed effects. The confounders of $TREATED_{dt}$, if there is any, have to be school-specific time-variant. In our case, the selection of school/district to enter the early phases could bias the estimate of β if the selection criteria are based on the schools'/districts' improvement potential on the outcomes. However, there are two reasons to reassure us that this selection bias is limited. First, the reform committee revealed that the selection was based on the districts' ability of information updating. As the reform has been planned to eventually cover all districts in Punjab, the committee did not have any obvious incentive to purposefully select the better improvers into early phases. Given the capacity constraint, the committee was more concerned about a smooth implementation process. More specifically, a district was selected into early phases if the district had a group of education officers (EOs) who are more responsive to the reform committee. Therefore, the selection bias can be reduced by controlling for the district's EOs' information update ability. We discuss the addition of this control in our robustness section. Second, the balanced outcomes across phases at baseline (Figure 4.3) reassures that the committee did not purposefully select on the districts' improvement potentials, at least on the observables.

The critical assumption in our identification strategy is that the outcomes in different phases would have evolved parallelly had the reform been absent, i.e., "the common trends" assumption. The data structure allows us to conduct a placebo test for the validity of this assumption. We demonstrate the placebo test in detail in the robustness section.

4.5 Results

4.5.1 School's Financial Account

If the policy is expected to have impacts on education outcomes, the grants need be received and spent by the schools first. We begin by presenting the results on schools' financial accounts. As suggestive evidence, we group the schools by the reform phase assignments and plot the average amount of grant received and spent in the last 12 months by phases against the years in Figure 4.4. As shown in the upper panel of the figure, an average public school received around Rs. 30,000 before the reform. However, the amount of school grants received by the phase-1 schools in the last 12 months has seen a discontinuous increase to Rs. 210,000 in September 2014, exactly 12 months after the phase-1 reform rolled out in September 2013. Similarly, phase-2 and phase-3 schools saw a similar discontinuous increase in 2015 and 2016 respectively, exactly one year after they became subjected to the reform. The three staggered discontinuities in the three phases mimic the staggered roll-out schedule of the reform. We notice that the average amount of the receipt dropped slightly after schools entered the treatment. There are two reasons contributing to this drop. First, as we mentioned, the PMIU recalculated the need-based school grant entitlements on an annual basis. The drop could reflect the decrease in school needs due to the improvement during the previous year. Second, there are delay issues in grant disbursement since 2016 (Khattak 2015, DAWN 2018). Although this delay is not yet fully understood to us at this moment, it has been accounted for by the year fixed effects in model (4.2) because the decreasing trend is common to all phases. Regarding spending, the lower panel also shows a similar staggered pattern, suggesting that the schools responsively increased their expenditure. Overall, this is a strong evidence that the need-based grants have been received and spent by the schools.

Note: The results are calculated using school-year panel data. Schools are grouped into three phases based on the policy phase assignments of the districts where the schools are located. Then for each phase-year, we calculate the average grant received/expenses.

Figure 4.4: School Grant Received from the Government by Phase and Year

To estimate the average policy effect on the schools' financial accounts, we run model (4.2) and the results are presented in Table 4.2. When an average public school becomes subjected to the NSB reform, it receives Rs. 180,000 more from the government, which is 6 times the amounts before the reform (column 1), and spends Rs. 135,000 more, a value that is three times the average before the reform (column 2).

Table 4.2: Policy Effects on School's Financial Account

	Grant Received (10,000 Rs.)	Expenses (10,000 Rs.)
	(1)	(2)
TREATED	18.18***	13.59***
	(1.85)	(1.35)
Observations	278,779	278,779

Note: The dataset is a school-year panel. Standard errors are clustered on the district level. The analysis covers all 36 districts in Punjab-Pakistan. *p<0.1; **p<0.05; ***p<0.01

4.5.2 School Inputs

Seeing that the schools have increased their spending, we are interested in understanding whether the spending has improved the schools' conditions. The rich data from the ASC provides us with an opportunity to test a large quantity of outcomes. However, it is considered problematic to do so for two reasons. First, we will encounter the issue of multiple hypotheses testing where there is a danger of overinterpreting any single significant result on a specific infrastructure, especially when there are many possible outcomes without a single possible causal pathway (Banerjee et al., 2015a). Second, due to the fungibility of the school grant, heterogeneity in grant uses can make some effects hard to detect (Banerjee et al., 2015c). Therefore we aggregate the infrastructures outcomes into an infrastructures index as a way to capture the effect on overall infrastructures condition and reduce the danger of overinterpreting. The results are presented in Table 4.3. As shown in column (1), the overall infrastructures condition indeed witnessed an economically and statistically significant improvement by 0.045 standard deviations. When looking at the impacts on the components of the infrastructures index, we see an interesting pattern. As shown in table 4.4, we find that the NSB does not avail more facilities, i.e., drinking water, electricity, main gate or toilet (column 1 to 4), but improves the condition of the existing infrastructures (column 5 to 7). Specifically, the share of functional toilets saw an increase of 1.5 percentage points. The building condition of the schools is 7.4 percentage points more likely to be safe, an improvement that is 13% relative to the pre-reform level. The boundary wall is 1.5 percentage points more likely to be complete, an improvement amounts to 1.7% of the average before the reform. Similarly, we aggregate the availabilities of various sports facilities into a sports facilities index whose result is shown in column (2) of table 4.3. We find that policy effect on the overall availablity of sports facilities is nonexistent in that the point estimate is both statistically and economically insignificant.

Of course, there are other items, e.g., textbooks, stationery and teaching materials quantity and quality, on which the schools may have spent the grants. However, we could not include them in our analysis due to the lack of data. Nevertheless, we project the improvement on these items to be limited as the improvement on the existing infrastructures involved construction works that cost a majority of the grant. In fact, Cambridge Education (2014), a review on the expenditures by public schools in Punjab, reports that in the school year of 2013/14, only 2% of the 400 sample schools ever spent the budget on learning and teaching materials. In comparison, 65% of schools reported ever spending the budget on school buildings and other infrastructures. Overall, these results are consistent with Kremer (2003) who points out that providing larger grants to schools in Kenya led school committees to shift spending toward construction from textbooks.

While the reform has sharply increased the school grants for non-salary items, teachers can still indirectly benefit from the non-salary spending, e.g., from the improvement of the classroom condition. Compared to the school infrastructures and facilities, the information on human capital is quite limited to us. The only variable available to us to approximate the human capital quality is the teacher attendance rate. In fact, teacher's absenteeism is a major concern in Punjab's education reform. Punjab has been criticized by its large share of ghost teachers and multi-grade teaching problems. It is necessary to keep teacher in the school before teachers can exert effort in improve teaching. Column (3) of table 4.3 shows that the NSB indeed improves the teacher attendance rate by 0.9 percentage points. However, this magnitude is rather nominal as it makes up only 1% of the average teacher attendance rate before the reform, or one hour per month from another perspective.[9] The reform may also affect teachers' effort on improving their pedagogy, which is in fact a promising way for education outcome improvement (Glewwe and Kremer 2006, Glewwe and Muralidharan 2016). This is especially relevant in the context of Punjab-Pakistan as the education system is characterized by its low effort of teaching (Naviwala, 2015). Unfortunately, the lack of data on teachers' effort prevents us from investigating further along this direction.

Table 4.3: Program Impacts on School Inputs

	Infrastructures Index	Sports Facilities Index	Teacher Attendance Rate
	(1)	(2)	(3)
TREATED	0.045***	-0.007	0.009**
	(0.01)	(0.01)	(0.004)
Observations	261,517	278,779	289,564

Note: Standard errors are clustered on the district level. The analysis covers all 36 districts in Punjab-Pakistan. *p<0.1; **p<0.05; ***p<0.01

[9] Assuming teachers teach 6 hours per school day and 20 school days per month, an increase of 0.9 percentage points in teacher attendance rate is equal to 0.009*6*20 = 1.08 hours per month.

Table 4.4: Program Impacts on Components of the Infrastructures Index

	drink water exists	electricity exists	main gate exists	toilet exists	share of funct. toilets	building cond. is safe	boundary wall is complete
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TREATED	-0.001	0.001	0.008	-0.001	0.015***	0.074***	0.015**
	(0.003)	(0.012)	(0.005)	(0.003)	(0.004)	(0.015)	(0.007)
Observations	278,779	278,779	278,779	278,779	274,098	278,779	261,517

Note: Standard errors are clustered on the district level. The analysis covers all 36 districts in Punjab-Pakistan. *p<0.1; **p<0.05; ***p<0.01

4.5.3 Education Outcomes

Ultimately, we are interested in the effects on educational outcomes. Specifically, we examine the effects on education quantity, as measured by enrollment rate and student attendance rate, and education quality, as measured by the students' test score and reading/mathematical abilities. The results are presented in table 4.5.

In terms of education quantity, the NSB reform does not improve school enrollment (column 1). In fact, the effect is quite precisely estimated to be zero in that the point estimate -0.009 is very small and even the upper bound of its 95% confidence interval is only 0.018, a value that is 2% of the average enrollment rate before the reform. Regarding student's school attendance, the point estimates in column (2) to (5) suggest that the NSB reform increases the attendance rate of students from various school levels. However, the magnitudes are only nominal which are approximately 0.7% of the average attendance rate before the reform, suggesting that the impact on student attendance rate is nonexistent.

Regarding education quality, we do not find any impacts on children's mathematic and English reading abilities (column 6 to 7). While the point estimates are negative, suggesting that the NSB reform may even decrease the mathematic and English reading scores of children of age between 5 and 16 by 0.026 and 0.021 points, respectively, they are both statistically and economically insignificant (column 6 and 7). In comparison, the pre-reform mathematic and English reading scores for an average child are 2.4 and 2.2 points. Similarly, the effects are lacking for the fifth graders' test scores (column 8 to 9).

Overall, despite the significant increase in spending and the subsequent improvement in school conditions, we do not find policy impacts on education quantity and quality. This is largely consistent with the literature, which shows that improving physical inputs alone does not have meaningful impacts on education outcomes (see the review by Glewwe and Muralidharan 2016). One should keep in mind that our results are relatively short-term, and the lack of more detailed measures of human resources prevents us from a more comprehensive investigation into the human resources channel. Therefore, we cannot rule out the possibility of improvements in education outcomes through higher teacher effort or better school management in the long term.

Table 4.5: Program Impacts on Education Outcomes

	Enrollment	Student Attendance				Ability		Test Score	
	Currently Enr. (Age 5-16) (%)	Primary (Gr.1-5)	Middle (Gr.6-8)	High (Gr.9-10)	H.Sec (Gr.11-12)	Math. (Age 5-16)	Eng.	Num. (Gr.5)	Lit.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TREATED	-0.009	0.006	0.006	0.005	0.002	-0.026	-0.021	-0.087	-0.059
	(0.014)	(0.005)	(0.005)	(0.008)	(0.010)	(0.042)	(0.044)	(0.07)	(0.07)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
School FE	N	Y	Y	Y	Y	N	N	Y	Y
District FE	Y	N	N	N	N	Y	Y	N	N
Data Unit	Dist.	Sch.	Sch.	Sch.	Sch.	Dist.	Dist.	Sch.	Sch.
Observations	107	276,894	84,262	38,423	3,457	107	107	269,526	269,526

Note: Standard errors are clustered on the district level. The analysis covers all 36 districts in Punjab-Pakistan. *p<0.1; **p<0.05; ***p<0.01

4.5.4 Interpretation

Taken together, we find that the NSB reform has brought significant improvement in the conditions of school infrastructures and modest improvement in teacher's attendance. However, there is no meaningful impact on education quantity and quality. In light of Beasley and Huillery (2017), the lack of impacts on education outcomes is likely to be associated with the lack of capacity among the school council members who are held accountable for managing the grant. In fact, various reports on the education sector of Pakistan have consistently shown the lack of capacity of the school councils. A report on Punjab's education sector procurements by the Institute of Social and Policy Sciences (I-SAPS) shows that, school council members from 64% of the public schools in Punjab are not familiar with the procurement rules and procedures for the school grant that they are responsible for, and 91% and 78% of the members report that they need training for financial management and school improvement planning, respectively (I-SAPS, 2014).

Although it is important to ensure a safe learning environment, it is not an effective way to improve enrollment and learning (Newman et al. 2002, Kremer 2003). And yet, we see that school infrastructures are the only outcomes that shows meaningful improvement. We argue that the reasons are twofold. First, the significant improvement in school infrastructure conditions reflects the spending priority of the public schools. Public schools in Pakistan are often criticized for their poor infrastructure conditions and school security (Naviwala, 2015). This criticism reached to a peak after school massacre in Peshawar district in 2014, where seven terrorists entered a public school by scaling the boundary wall and killed 132 schoolchildren (Sajjad et al., 2015). Improving the school buildings' conditions is a response to this widespread criticism. Second, although it has been realized in the recent years that a better school does not equal to better education, improving school infrastructures is still the focus of the education policymakers in Pakistan because it is regarded as a low hanging fruit for fixing the education crisis (Naviwala, 2019). And this perception of the policymakers is conveniently adopted and followed by the schools, especially by those with a school council that lacks capacity. The political connection of the public school teachers may also facilitate the transfer of this perception. In fact, many public school teachers have political patronage, and it is not uncommon that they play a dominant role when making school spending decisions (I-SAPS, 2014).

As it is agreed on in the literature, improving pedagogy is a promising way to improve education outcomes (Glewwe and Muralidharan 2016, Glewwe and Kremer 2006). While we find that the NSB reform has increased the teacher attendance rate by 0.9 percentage point, we doubt that this indicates teachers exert more effort to improve pedagogy. Actually, the magnitude of the impact is only nominal as it equals only 1 hour per month. If we take into account the issue of multigrade teaching, which is very common in Punjab, it is unlikely to see a sizable impact on student's learning.[10] Besides, it is not clear whether this additional hour per month was dedicated entirely to teaching as the government often assigns public school teachers to non-teaching duties, e.g., helping with anti-polio and anti-dengue drives, elections and administering government exams and surveys (Andrabi et al., 2008). Alif Ailaan (2014b), a research report based on in-depth qualitative

[10]43 percent of public schools in Pakistan are multigrade (ASER, 2018)

interviews with 1,250 teachers from more than 600 schools in Pakistan, estimates that government teachers spend a quarter of the academic year on non-teaching activities. In fact, public school teachers in Pakistan are notoriously unmotivated in teaching and it has been associated with an understanding that a child who goes to a public school will not have a chance in life, in terms of a job or being perceived as an equal (Naviwala, 2019). Holding teachers accountable is crucial but not easy in Pakistan. As it is mentioned, public teachers in Pakistan are unionized and are government employees who often have political patronage. There is no direct accountability relationship between teachers and parents who pay fees. Therefore, parents and school council do not have the power to fire a teacher for not delivering quality teaching. In fact, even a minister for education could not fire a teacher for not showing up to school (Naviwala, 2019). Furthermore, many teachers also offer private tutoring after school. It has become a norm for students in Pakistan to go to after-school tutors. Teachers have an economic incentive to compel students to sign up for tuition in order to learn material, or teachers may simply have less energy for teaching in school if they are primarily concerned with their tuition classes (Alif Ailaan, 2014b). Without a sufficient capacity to manage school grant and hold teachers accountable, school councils are unlikely to help improving education outcomes by just receiving a larger amount of grant.

4.6 Robustness Checks

4.6.1 Placebo Test

One compelling way to assess the "common trends" assumption underlying our identification strategy is to compare the temporal changes before the reform across different phases, i.e., a "placebo" test. We use two pre-reform years of the panel to conduct this placebo test. Specifically, our approach is based on the following specification:

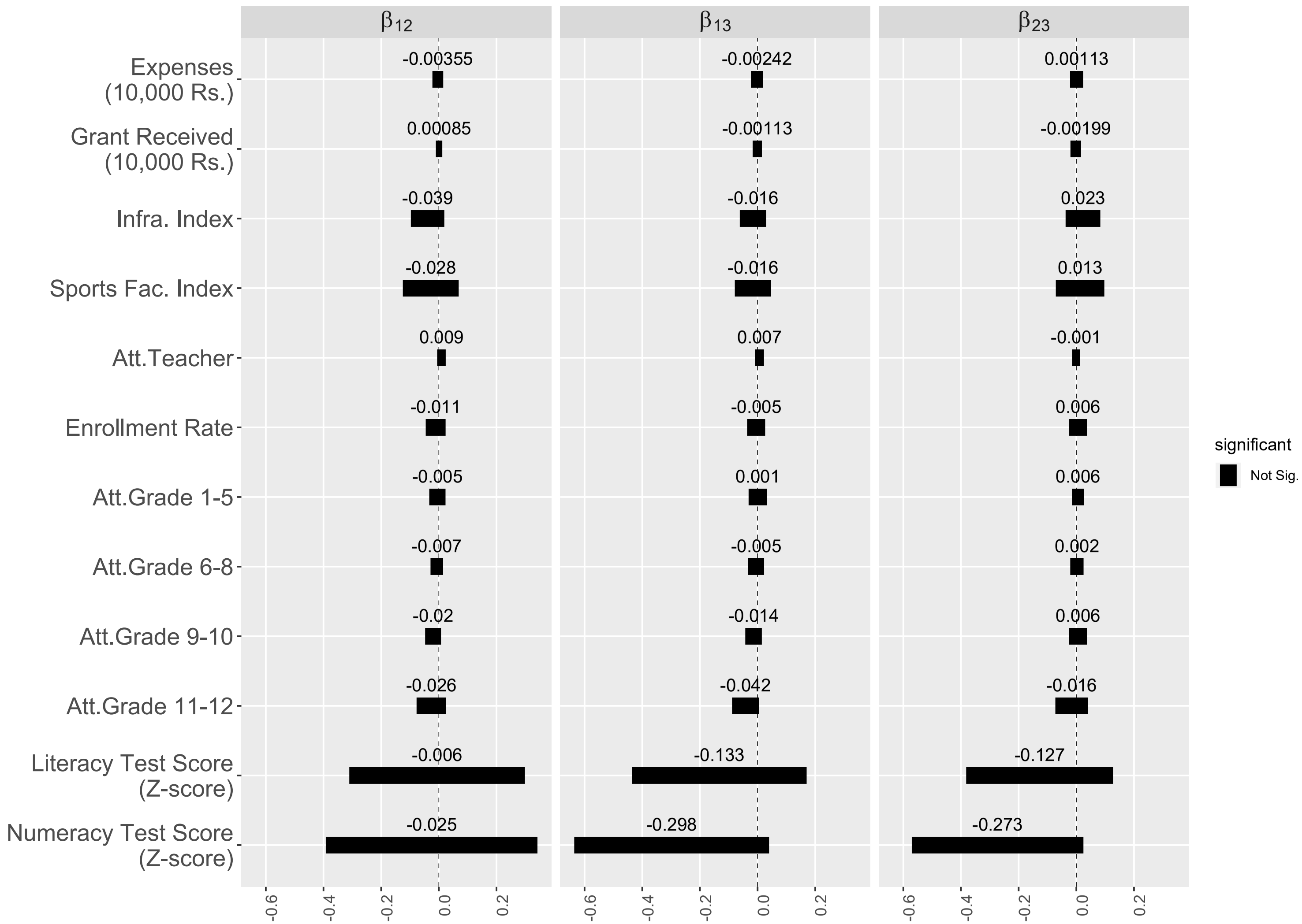
$$\begin{aligned} y_{sdt} &= \beta_{13} * PHASE1_d * PPOST_t + \beta_{23} * PHASE2_d * PPOST_t \\ &\quad + \alpha_s + \gamma_t + u_{sdt} \\ \beta_{12} &= \beta_{13} - \beta_{23} \end{aligned} \tag{4.3}$$

where $PPOST_t$ is a pseudo-post dummy variable which identifies the latter of the two pre-reform years. Other variables' definitions remain unchanged. Therefore, β_{13} effectively captures the difference in the trend of phase 1 relative to phase 3. Similar definition is applied to β_{23} and β_{12}.

We conduct this placebo test on all outcome variables of our main analysis. We could not test on the English reading and mathematical ability scores because we only have one year of pre-reform data on these two variables. The point estimates and their corresponding 90% confidence intervals are presented in Figure 4.5, with β_{12}, β_{13} and β_{23} being plotted separately in different panels. Overall, we do not find any statistically significant difference in the trends across phases before the reform, i.e., the common trends assumption is satisfied. Furthermore, the point estimates are very precisely estimated as the confidence intervals are very narrow. This assures us that the zero point estimates are not a power issue.

4.6.2 Control for District's Effort on Information Updating

According to the reform committee, a district was selected into an early phase if the education officers (EOs) in that district were more responding to the education department in terms of information updating. This could be a confounder to our estimate of the reform's impacts if the district's effort on information updating is correlated with the improvement in school outcomes. The main channel for the EOs to collect the information of the schools is to visit the school in person. To control for this potential confounder, we obtain from the PMIU a school-level cross-sectional dataset which records the number of visits by EOs during the school year 2012/13, the school year just before the rollout of the NSB reform. We then calculate the yearly visit frequency for each school and aggregate to a district average. Then, we rank the districts by frequency of school visits in a descending order, division by division, i.e., the most visited district in a division is ranked at the first position of that division. We regard the rank of visit frequency as the proxy for the district's effort on information updating. The descriptive statistics of the visits by EOs are summarized by districts and reported in Table 4.6.



Note: standard errors are clustered at district level. Test score is standardized to mean of zero and sd. of 1 in that year. Infrastructures Index is an average of z-scores of eletricity, drinking-water, toilet, boundary-wall, playground, main-gate, satisfied-building-condition, boundary-wall-completeness dummies)

Figure 4.5: Placebo Test

Table 4.6: Education Officers' School Visits by Districts

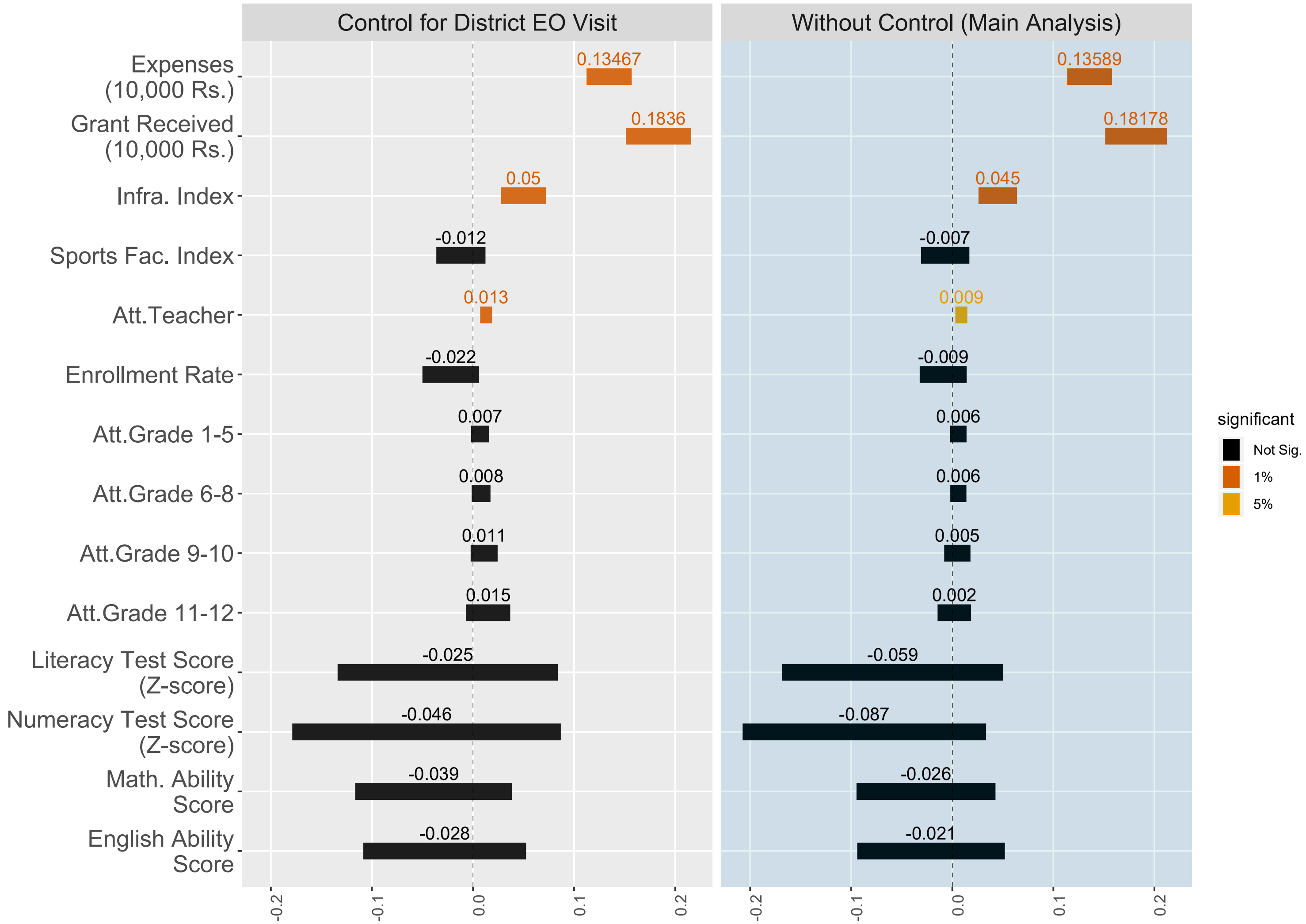
District	Nr. Visits	Ranking	Phase	District	Nr. Visits	Ranking	Phase
Bahawalpur Division				*Lahore Division*			
Bahawalpur	7.09	1	2	Nankana sahib	11.75	1	1
Rahimyar khan	5.8	2	1	Kasur	7.32	2	2
Bahawalnagar	5.6	3	3	Sheikhupura	6.51	3	3
Dera Ghazi Khan Division				Lahore	6.06	4	3
Muzaffargarh	8.86	1	1	*Multan Division*			
Dera ghazi khan	8.77	2	3	Khanewal	9.28	1	1
Rajanpur	5.99	3	3	Multan	8.86	2	3
Layyah	5.45	4	2	Vehari	6.22	3	2
				Lodhran	3.84	4	3
Faisalabad Division				*Rawalpindi Division*			
Chiniot	11.37	1	1	Chakwal	7.19	1	1
Toba tek singh	10.14	2	2	Rawalpindi	6.64	2	3
Jhang	9.8	3	3	Jehlum	5.17	3	3
Faisalabad	6.17	4	3	Attock	3.99	4	2
Gujranwala Division				*Sargodha Division*			
Hafizabad	9.74	1	3	Sargodha	13.37	1	1
Mandi bahauddin	9.07	2	3	Khushab	9.01	2	2
Gujrat	8.43	3	3	Mianwali	7.22	3	3
Sialkot	6.83	4	1	Bhakkar	5.57	4	3
Narowal	5.92	5	3	*Sahiwal Division*			
Gujranwala	5.52	6	2	Okara	8.47	1	1
				Sahiwal	7.13	2	3
				Pakpattan	5.64	3	2

Note: Nr.Visits is the average number of visits by EOs per year per school. Ranking is the ranking of Nr.Visits within a division.

We then interact the ranking of visits by EOs with the year dummies and add them to the specification (4.2). That is, we estimate the following specification,

$$y_{sdt} = \beta * TREATED_{dt} + \theta_{\tau} * Rank_d * D_t^{\tau} + \alpha_s + \gamma_t + u_{sdt} \tag{4.4}$$

where $Rank_d$ is the within-division rank of the school visit by EOs in district d in year 2012. D_t^{τ} are year dummies for year τ with τ = 2013, 2014, ..., 2017. The other variables' definitions remain unchanged. The point estimates of the effects, β, and their 90% confidence intervals are plotted in the left panel of Figure 4.6. For comparison, the results of the main analysis are plotted in the right panel. As we can see, controlling for the district's effort on information updating does not change the main results. In fact, the point estimates are extremely similar between the two specifications. In this regard, we conclude that the selection on district's effort on information updating does not bias our estimates of the policy impacts.



Note: Standard errors are clustered at the district level. In the left panel, we add district d's rank on the number of school visits by EOs at baseline, interacted with the year dummy variables as a way to control for the EOs' effort

Figure 4.6: Control for Education Officers' (EOs) Effort

4.7 Conclusion

Whether providing decentralized grants to schools is effective in improving education quantity and quality remains critical for education policymakers. While decentralized school grant has gathered popularity because it is arguably capable of reducing the mismatch between centralized provision and local needs, rigorous evaluations are still limited. We contribute to this limited literature by evaluating a large-scale policy, the "Non-salary Budget (NSB)" reform in Punjab, Pakistan, which increased the funding for non-salary expenditure by almost 10 times and made the school councils accountable for grant management. We find that the NSB reform has significantly increased the income and expenditure of the public schools in Punjab, resulting in a sizable improvement in the school infrastructures condition. Although we find that the reform has also improved teacher attendance rate, the magnitude of the effect is only nominal. Finally, we don't find any impact on enrollment rate, student attendance rate or test scores. We interpret the lack of effects on education outcomes as a result of insufficient capacity among the school council members. In fact, a report on the procurement in the education sector of Punjab, I-SAPS (2014), has revealed that the majority of school council members do not understand the procedure of school procurement. More importantly, the school councils have limited power to hold teachers accountable for delivering quality education, as teachers in public schools of Pakistan have political patronage. Simply increasing the funding to schools does not change the lack of accountability relationship between teachers and school councils. Overall, our findings are consistent with Mbiti et al. (2019) and Beasley and Huillery (2017) who show that the effectiveness of decentralized school grants depends on the capacity of the local planners who are responsible for grant management.

Bibliography

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93(1):113–132.

Adukia, A. (2017). Sanitation and Education. *American Economic Journal: Applied Economics*, 9(2):23–59.

Agarwal, S. and Hauswald, R. (2010). Distance and Private Information in Lending. *Review of Financial Studies*, 23(7):2757–2788.

Ali, R. (2013). Estimating Urbanization. *Urban Gazettee*, (December):1–10.

Alif Ailaan (2014a). 25 Million Broken Promises. Technical report, Alif Ailaan.

Alif Ailaan (2014b). The Voice of Teachers: Learning from teachers across Pakistan. *Islamabad: Alif Ailaan*.

Andrabi, T., Das, J., Khwaja, A. I., Ozyurt, S., and Singh, N. (2018). Upping the Ante : The Equilibrium Effects of Unconditional Grants to Private Schools. *Policy Research Working Paper No. 8563. World Bank.*

Andrabi, T., Das, J., Khwaja, A. I., Vishwanath, T., and Zajonc, T. (2008). Pakistan Learning and Educational Achievement in Punjab Schools (LEAPS) - Instights to inform the education policy debate. *Washington, DC: World Bank*.

Angelucci, M., Karlan, D., and Zinman, J. (2015). Microcredit impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco. *American Economic Journal: Applied Economics*, 7(1):151–182.

Angrist, J. D. and Krueger, A. B. (1999). Empirical Strategies in Labor Economics. In *Handbook of Labor Economics*, volume 3, pages 1277–1366.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics*. Princeton University Press.

ASER (2018). Annual Status of Education Report (ASER) Pakistan. Technical report, South Asian Forum for Education Development.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., and Harmgart, H. (2015). The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics*, 7(1):90–122.

Augsburg, B., Haas, R. D., Harmgart, H., and Meghir, C. (2015). The Impacts of Microcredit: Evidence from Bosnia and Herzegovina. 7:183–203.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015a). The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Pariente, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015b). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):12 60799–1–1260799–16.

Banerjee, A., Karlan, D., and Zinman, J. (2015c). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1):1–21.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., and Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, 2(1):1–30.

Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, 122(3):1235–1264.

Banerjee, A. V. and Duflo, E. (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives*, 21(1):141–167.

Banerjee, A. V., Imbert, C., and Mathew, S. (2016). E-Governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India. *SSRN Electronic Journal*.

Beasley, E. and Huillery, E. (2017). Willing but Unable? : Short-Term Experimental Evidence on Parent Empowerment and School Quality. *Policy Research Working Paper;No. 8125. World Bank*.

Bekaert, G., Harvey, C. R., and Lundblad, C. (2005). Does financial liberalization spur growth? *Journal of Financial Economics*, 77(1):3–55.

Blimpo, M. P., Evans, D. K., and Lahire, N. (2015). Parental Human Capital and Effective School Management : Evidence from The Gambia. *Policy Research Working Paper;No. 7238. World Bank*.

Boonperm, J., Haughton, J., and Khandker, S. R. (2013). Does the Village Fund matter in Thailand? Evaluating the impact on incomes and spending. *Journal of Asian Economics*, 25:3–16.

Boonperm, J., Haughton, J., Khandker, S. R., and Rukumnuaykit, P. (2012). Appraising the Thailand Village Fund. *World Bank Policy Research Working Paper*, 5998(March).

Borkum, E., He, F., and Linden, L. (2012). The Effects of School Libraries on Language Skills: Evidence from a Randomized Controlled Trial in India. *National Bureau of Economic Research*.

Brevoort, K. P., Wolken, J. D., and Holmes, J. A. (2012). Distance Still Matters: The Information Revolution in Small Business Lending and the Persistent Role of Location, 1993-2003. *SSRN Electronic Journal*, pages 1993–2003.

Brown, M., Guin, B., and Kirschenmann, K. (2016). Microfinance Banks and Financial Inclusion. *Review of Finance*, 20(3):907–946.

Bruns, B., Filmer, D., and Patrinos, H. A. (2011). Making Schools Work : New Evidence on Accountability Reforms. *Human Development Perspectives. World Bank*.

Burgess, R. and Pande, R. (2003). Do Rural Banks Matter? Evidence from the indian Social Banking Experiment. *LSE Research Online Documents on Economics 2244, London School of Economics and Political Science*.

Burgess, R. and Pande, R. (2005). Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. *American Economic Review*, 95(3):780–795.

Cambridge Education (2014). Non-Salary Budget Expenditure Review. Technical report.

Card, D. and Krueger, A. B. (1994). Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–793.

Carletti, E., Senbet, L. W., Cull, R., Allen, F., Qian, J. Q., and Valenzuela, P. (2018). Improving Access to Banking: Evidence from Kenya. *SSRN Electronic Journal*.

Carneiro, P., Koussihouèdé, O., Lahire, N., Meghir, C., and Mommaerts, C. (2019). School Grants and Education Quality: Experimental Evidence from Senegal. *Economica*, 87(345):28–51.

Carneiro, P. M., Koussihouudd, O., Lahire, N., Meghir, C., and Mommaerts, C. (2015). Decentralizing Education Resources: School Grants in Senegal. *SSRN Electronic Journal*.

Carr-Hill, R., Rolleston, C., Schendel, R., and Waddington, H. (2018). The effectiveness of school-based decision making in improving educational outcomes: a systematic review. *Journal of Development Effectiveness*.

Chen, F., Lui, A. M., and Martinelli, S. M. (2017). A systematic review of the effectiveness of flipped classrooms in medical education. *Medical Education*, 51(6):585–597.

Cole, S., Giné, X., and Vickery, J. (2017). How does risk management influence production decisions? Evidence from a field experiment. *Review of Financial Studies*, 30(6):1935–1970.

Conn, K. M. (2014). *Identifying Effective Education Interventions in Sub-Saharan Africa: A meta-analysis of rigorous impact evaluations*. PhD thesis, Columbia University.

Crépon, B., Devoto, F., Duflo, E., and Pariente, W. (2015). Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Expirement in Morocco. *American Economic Journal: Applied Economics*, 7(1):123–150.

Cristia, J., Ibarrarán, P., Cueto, S., Santiago, A., and Severín, E. (2017). Technology and Child Development: Evidence from the One Laptop per Child Program. *American Economic Journal: Applied Economics*, 9(3):295–320.

Cuong, N. V. (2008). Is a governmental micro-credit program for the poor really propoor? Evidence from Vietnam. *Developing Economies*, 46(2):151–187.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., and Sundararaman, V. (2013). School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics*, 5(2):29–57.

DAWN (2018). Government schools in Punjab yet to get two instalments of non-salary funds.

de Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to Capital in Microenterprises: Evidence from a Field Experiment. *Quarterly Journal of Economics*, 123(4):1329–1372.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *Journal of Finance*, 60(1):231–266.

Demirgüç-Kunt, A., Klapper, L., and Singer, D. (2017). Financial Inclusion and Inclusive Growth : A Review of Recent Empirical Evidence. *Policy Research Working Paper;No. 8040. World Bank*.

Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S., and Hess, J. (2018). Global Findex Database 2017 : Measuring Financial Inclusion and the Fintech Revolution. *Washington, DC: World Bank*.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5):1739–1774.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4):1241–1278.

Faruqee, R., Qureshi, S., and Nabi, I. (1996). Rural Finance for Growth and Poverty Alleviation. *Policy Research Working Paper Series 1593, orld Bank*.

Fink, G., Jack, B. K., and Masiye, F. (2014). Seasonal Credit Constraints and Agricultural Labor Supply: Evidence from Zambia. *NBER Working Paper Series*.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, 103(6):1176–1209.

Frölich, M. and Michaelowa, K. (2011). Peer effects and textbooks in African primary education. *Labour Economics*, 18(4):474–486.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., and Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice, Second Edition*. The World Bank.

Glewwe, P. and Kremer, M. (2006). Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries.

Glewwe, P., Kremer, M., and Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, 1(1):112–135.

Glewwe, P., Kremer, M., Moulin, S., and Zitzewitz, E. (2004). Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya. *Journal of development Economics*, 74(1):251–268.

Glewwe, P. and Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier.

Goodman-Bacon, A. (2018). Difference-in-Differences with Variation in Treatment Timing. *National Bureau of Economic Research.*

Hanushek, E. A. and Woessmann, L. (2011). Overview of the symposium on performance pay for teachers. *Economics of Education Review*, 30(3):391–393.

Hanushek, E. A. and Wößmann, L. (2010). *Education and Economic Growth*.

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396):945.

I-SAPS (2014). Education Sector Procurements in Punjab A Horizontal Accountability Perspective. Technical report, Institute of Social and Policy Sciences.

I-SAPS (2015). *Chronic underfunding of non-salary expenses of schools: What can policymakers do about it?* Institute of Social and Policy Sciences.

Imbens, G. W. (2014). Instrumental variables: An econometrician's perspective. *Statistical Science*, 29(3):323–358.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Imbens, G. W. and Rubin, D. B. (2010). Rubin Causal Model. In *Microeconometrics*, pages 229–241. Palgrave Macmillan UK, London.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jack, W. and Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, 104(1):183–223.

Jayaratne, J. and Strahan, P. E. (1996). The Finance-Growth Nexus: Evidence from Bank Branch Deregulation. *The Quarterly Journal of Economics*, 111(3):639–670.

Kaboski, J. P. and Townsend, R. M. (2012). The impact of credit on village economies. *American Economic Journal: Applied Economics*, 4(2):98–133.

Karlan, D. and Morduch, J. (2010). Chapter 71 – Access to Finance. In *Handbook of Development Economics*, volume 5, pages 4703–4784.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014a). Agricultural Decisions after Relaxing Credit and Risk Constraints *. *The Quarterly Journal of Economics*, 129(2):597–652.

Karlan, D., Ratan, A. L., and Zinman, J. (2014b). Savings by and for the Poor: A research review and agenda. *Review of Income and Wealth*, 60(1):36–78.

Karlan, D. and Zinman, J. (2010). Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial Studies*, 23(1):433–464.

Karlan, D. and Zinman, J. (2011). Microcredit in theory and practice: using randomized credit scoring for impact evaluation. *Science*, 332(June):1278–1284.

Karlan, D. S. and Zinman, J. (2008). Credit Elasticities in Less-Developed Economies: Implications for Microfinance. *American Economic Review*, 98(3):1040–1068.

Khandker, S. R. and Faruqee, R. R. (2003). The impact of farm credit in Pakistan. *Agricultural Economics*, 28(3):197–213.

Khandker, S. R., Koolwal, G. B., and Samad, H. A. (2010). Handbook on impact evaluation: quantitative methods and practices. *World Bank*.

Khattak, K. (2015). Technical objection delays City schools non- salary funds.

King, R. G. and Levine, R. (1993). Finance and Growth: Schumpeter Might Be Right. *The Quarterly Journal of Economics*, 108(3):717–737.

Kochar, A. (2011). The distributive consequences of social banking: A microempirical analysis of the indian experience. *Economic Development and Cultural Change*, 59(2):251–280.

Kremer, M. (2003). Randomized evaluations of educational programs in developing countries: Some lessons. *American Economic Review*, 93(2):102–106.

Kremer, M., Brannen, C., and Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, 340(6130):297–300.

Krishnaratne, S., White, H., and Others (2013). Quality education for all children? What works in education in developing countries. Technical report, International Initiative for Impact Evaluation (3ie).

La Porta, R., Lopez-De-Silanes, F., and Shleifer, A. (2002). Government Ownership of Banks. *The Journal of Finance*, 57(1):265–301.

Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355.

Levine, R. and Zervos, S. (1998). Stock Markets, Banks, and Economic Growth. *American Economic Review*, 88(3):537–558.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.

Malik, S. J., Mushtaq, M., and Gill, M. A. (1991). The role of institutional credit in the agricultural development of Pakistan. *Pakistan Development Review*, 30(4):1039–1048.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–437.

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., and Rajani, R. (2019). Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania*. *The Quarterly Journal of Economics*, 134(3):1627–1673.

Mbiti, I. M. (2016). The Need for Accountability in Education in Developing Countries. *Journal of Economic Perspectives*, 30(3):109–132.

McEwan, P. J. (2015). Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments. *Review of Educational Research*, 85(3):353–394.

Menkhoff, L. and Rungruxsirivorn, O. (2011). Do Village Funds Improve Access to Finance ? Evidence from Thailand. *World Development*, 39(1):110–122.

Meyer, B. D. (1995). Natural and Quasi-Experiments in Economics. *Journal of Business & Economic Statistics*, 13(2):151.

Ministry of Finance, G. o. P. (2009). Pakistan Economic Survey. Technical report.

Morduch, J. (1999). The Microfinance Promise. *Journal of Economic Literature*, 37(4):1569–1614.

Morgan, D., Pinkovskiy, M., and Yang, B. (2016). Banking deserts, branch closings, and soft information. Technical report, Federal Reserve Bank of New York, Liberty Street Economics.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2016). Building state capacity: Evidence from biometric smartcards in Indi. *American Economic Review*, 106(10):2895–2929.

Muralidharan, K. and Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1):39–77.

Naik, G., Chitre, C., Bhalla, M., and Rajan, J. (2020). Impact of use of technology on student learning outcomes: Evidence from a large-scale experiment in India. *World Development*, 127:104736.

Naviwala, N. (2015). Pakistan's education crisis: The real story. *Wilson Center Asia Pogram*.

Naviwala, N. (2019). Why Can't Pakistani Children Read? *Wilson Center Asia Pogram*.

Nelson, R. and Phelps, E. (1966). Investment in Humans, Technology Diffusion and Economic Growth. *The American Economic Review*, 56(1/2):69–75.

Newman, J., Pradhan, M., Rawlings, L. B., Ridder, G., Coa, R., and Evia, J. L. (2002). An impact evaluation of education, health, and water supply investments by the Bolivian Social Investment Fund. *The World Bank Economic Review*, 16(2):241–274.

Nguyen, H. L. Q. (2019). Are credit markets still local? Evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

Petersen, M. A. and Rajan, R. G. (1994). The Benefits of Lending Relationships: Evidence from Small Business Data. *The Journal of Finance*, 49(1):3.

Pitt, M. M. and Khandker, S. R. (1998). The impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy*, 106(5):958–996.

Pradhan, M. P. and de Ree, J. (2014). District Governance and Student Learning in Indonesia. *Asian Development Bank Economics Working Paper Series No. 397*.

Pridmore, P. and Jere, C. (2011). Disrupting patterns of educational inequality and disadvantage in Malawi. *Compare: A Journal of Comparative and International Education*, 41(4):513–531.

Pungprawat, K. (2012). Thaksin and Budget Allocation : A Study in Political Compromise. *Asian and African Area Studies*, 11(2):166–182.

Qi, S., De Haas, R., Ongena, S. R. G., and Straetmans, S. (2019). Move a Little Closer? Information Sharing and the Spatial Clustering of Bank Branches. *Swiss Finance Institute Research Paper No. 17-74*.

Rajan, R. G. and Zingales, L. (1998). Financial Dependence and Growth. *American Economic Review*, 88(3):559–586.

Rehman, A., Jingdong, L., Chandio, A. A., and Hussain, I. (2017). Livestock production and population census in Pakistan: Determining their relationship with agricultural GDP using econometric analysis. *Information Processing in Agriculture*, 4(2):168–177.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.

Rosenzweig, M. R. and Wolpin, K. I. (2000). Natural “Natural Experiments” in Economics. *Journal of Economic Literature*, 38(4):827–874.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701.

Rubin, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, 2(1):1.

Sabarwal, S., Evans, D. K., and Marshak, A. (2014). *The permanent input hypothesis: the case of textbooks and (no) student learning in Sierra Leone*. The World Bank.

Sajjad, S., Kazmi, N., and Ali, A. (2015). Impacts of " School Terrorism " in the Milieu of Peshawar Incident, Pakistan's Black Day, December 16th, 2014. *American Journal of Social Science Research*, 1(4):196–205.

Sapienza, P. (2004). The effects of government ownership on bank lending. *Journal of Financial Economics*, 72(2):357–384.

Schultz, T. P. (1997). Demand for children in low income countries. In *Handbook of Population and Family Economics*, pages 349–430.

Schultz, T. P. (2002). Why Governments Should Invest More to Educate Girls. *World Development*, 30(2):207–225.

Sen, A. (2001). *Development as Freedom*. Oxford Paperbacks.

Skinner, J. and Staiger, D. (2005). Technology Adoption From Hybrid Corn to Beta Blockers. *NBER Working Papers 11251*.

Slater, L. (2013). Building high-performing and improving education systems. *CfBT Education Trust*.

State Bank of Pakistan (2000). Financial Sector Assessment 1990-2000. Technical report.

State Bank of Pakistan (2001). SBP Task Force Report on Crop Loan Insurance Framework. Technical report.

State Bank of Pakistan (2004). Report of the Committee on Rural Finance. Technical report.

State Bank of Pakistan (2007). Consolidated District Wise Data. Technical report.

Strauss, J. and Thomas, D. (1995). Chapter 34 Human resources: Empirical modeling of household and family decisions. *Handbook of Development Economics*, 3:1883–2023.

Tangpianpant, P. G. (2010). Thaksin Populism and Beyond.

Tarozzi, A., Desai, J., and Johnson, K. (2015). The Impacts of Microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics*, 7(1):54–89.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309–317.

Townsend, R. M. (2013). Townsend Thai Project Household Annual Resurvey 1998-2011. *hdl:1902.1/20818, Harvard Dataverse, V1 Cite Dataset*.

UNDP (2016). Multidimensional Poverty in Pakistan.

UNESCO (2017). More Than One-Half of Children and Adolescents Are Not Learning Worldwide. *UNESCO Institute for Statistics*.

UNESCO (2018). One in five children, adolescents and youth is out of school. Technical report.

United Nations (2015). Transforming our world: the 2030 Agenda for Sustainable Development. United Nations Sustainable knowledge platform. *Sustainable Development Goals*.

World Bank (2012). *Impact Evaluations : Relevance and Effectiveness*. Washington, DC: World Bank.

World Bank (2018). UFA2020 Overview: Universal Financial Access by 2020.

Young, N. (2017). Banking and Growth: Evidence From a Regression Discontinuity Analysis. *EBRD Working Paper No. 207*.

Zuberi, H. A. (1989). Production Function, Institutional Credit and Agricultural Development in Pakistan. *The Pakistan Development Review*, 28(1):43–55.