Migration, Liquidity Constraints, and Income Generation: Evidence from Randomized Credit Access in China

Shu Cai†

December 21, 2015

Abstract

With full labor mobility, microcredit may finance production inputs that increase household labor demand and reduce migration. However, when migration is constrained by liquidity, access to credit may increase both migration and household business activity, yielding greater potential income gains. This study empirically examines the impacts of credit access on migration using data from a randomized control trial for a village banking intervention in poor villages in rural China. Consistent with theoretical predictions, I find that the program increases migration by members of treated households, in particular for households in villages with lower average assets and facing higher migration costs.

---

* I deeply appreciate the supervision and support of Albert Park. Part of the research was done while I was visiting the University of Michigan, Ann Arbor, which I thank for the hospitality. I thank Manuela Angelucci, Charlie Brown, Jing Cai, Joseph Kaboski, David Lam, David McKenzie, Imran Rasul, Jeffery Smith, Sujata Visaria, and Dean Yang for some very helpful comments and discussions, as well as participants at the Health, History, Demography, and Development (H2D2) seminar and Labor seminar at the University of Michigan, Royal Economic Society inaugural symposium of junior researchers at the University of Manchester, and the seminar at Georgetown Center for Economic Research at Georgetown University. All remaining errors are my own.

† Department of Economics, the Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. scaiaa@ust.hk. Most recent version is available here http://www.caishu.org/research.html.
1 Introduction

In spite of large wage differentials between less- and more-developed areas, throughout the world a large number of potential migrants are deterred from migration by a variety of barriers (McKenzie and Yang, 2010; Munshi and Rosenzweig, 2014). Rural-urban income gaps are particularly notable in China due in part to institutions that segment rural and urban populations even as rapid industrialization increases the demand for labor in urban areas (Park, 2008). Although restrictions associated with China’s residential permit (hukou) system have been relaxed substantially for job-related migration, many rural people, especially those in poor households, remain in rural areas (World Bank, 2009; Meng, 2014; Tombe and Zhu, 2015).

One explanation for the lack of migration by the poor is that rural-urban migration involves substantial up-front costs which the poor are unable to finance. These costs may be especially high for individuals without social networks in destination locations. At the same time, there are other plausible explanations for lack of migration by the poor. Assessing the actual importance of financing costs thus is an empirical question.

To date, credible empirical evidence on the impact of financing constraints on migration remains somewhat limited. Some nonexperimental studies have shown that high costs of migration are associated with a lower propensity to migrate (McKenzie and Rapoport, 2007, 2010; Guriev and Vakulenko, 2015). One experimental study finds that exogenous entitlement to a guaranteed income stream from cash transfers increased migration likelihood from Mexico to the U.S. (Angelucci, 2015). In this context, it may be difficult to distinguish between the impact of greater wealth and relaxing liquidity constraints. Randomized control trials (RCT) evaluating credit interventions have not focused on migration outcomes or do not find significant impacts on wage employment. These RCTs also generally have failed to find strong evidence of positive income effects.

To my knowledge, this is the first experimental study to directly test theoretical predictions about the impact of loan access on migration and to find strong empirical evidence that credit access significantly increases migration by the poor in a developing country setting. A separate paper evaluating income effects of the same Chinese program finds that it increased both self-employment income and wage employment income, and that impacts on total household income were large and statistically significant (Cai et al., 2015). An evaluation study of credit expansion in Thailand based on a natural experiment finds similar results (Kaboski and Townsend, 2012). These findings contrast with programs in India and Morocco which find that microcredit increases self-employment

\[\footnote{Such costs include transportation costs to migration destinations, cost of time and living expenses during job search, and pre-job training. When searching for a job in cities, costs of food and accommodations may be quite expensive. All such costs must be paid before any wages are received.}

\[\footnote{Other possible reasons for lack of migration by the poor include lack of human capital (poor health, low education) and lack of sufficient household labor (given needs for subsistence farming). Other barriers to migration examined in the literature include imperfect information (Shrestha and Yang, 2014), friction in job matching (Beam, McKenzie, and Yang, 2014; Franklin, 2014), risk aversion (Bryan, Chowdhury, and Mobarak, 2014), and pre-existing informal insurance networks (Munshi and Rosenzweig, 2014). See also evidence on multiple barriers in McKenzie and Rapoport (2010) and Munshi (2011).}

\[\footnote{Bryan, Chowdhury, and Mobarak (2014) focus on migration and run experiments that provide households with conditional credit (on migration) and unconditional credit, but find that migration is more responsive to incentives than to liquidity.}
income but decreases wage employment income (Banerjee et al., 2005; Crepon et al., 2015). These varying results suggest that the existence of labor market frictions and off-farm employment opportunities may condition the impacts of microcredit programs on labor allocation and income.

In China, as in many other developing countries, formal financial institutions often exclude the rural poor from gaining access to loans. Rather, most families rely upon informal loans from relatives and friends, mainly for risk-coping, or consumption purposes. The most important economic activities in poor, rural areas is agriculture (cropping and livestock) and out-migration to work in industrial and service sector jobs in cities. Both of these activities may be constrained by lack of liquidity to buy inputs or pay up-front migration costs. Migration costs may be particularly high in China because of the spatial distribution of labor and economic activity. Industrial activity is concentrated on the coast, with many rural migrants coming from the country’s vast interior.

In this study, I first present a model to compare different predictions of the microcredit program impacts for households with and without up-front migration cost. When there is no labor market friction (zero migration cost), access to credit may increase self-employment production by relaxing the liquidity constraint on inputs, and employment (migration) decreases when labor returns to self-employment activity increase. On the contrary, if off-farm opportunities are available but there are non-negligible up-front migration costs constrained by liquidity, access to credit may increase migration and relaxing liquidity constraints may increase income from both self-employment and wage employment. The study empirically tests the theoretical predictions by analyzing data from a randomized control trial designed to evaluate a nationwide village banking program in China. Starting in 2006, the Chinese national government financed the establishment of village banks in designated poor villages to provide household loans to support income-generating activities. The “banks” are credit funds that do not accept deposits. Villages formed committees to manage the funds distributed to them. A randomized control trial was implemented from August 2010, to evaluate the impacts of the village banks on household outcomes. Detailed information collected in a baseline survey and a follow-up survey two years later form a unique data set used in the analysis for this study.

In the empirical analysis, by taking advantage of the randomness of village treatment status, I find that releasing liquidity constraints on average increases migration for households in treatment villages. To more directly test the theoretical predictions, I further explore the heterogeneity in treatment impacts for households with different amounts of baseline assets. Since unobservable determinants of migration (e.g. entrepreneurship) may confound the estimated heterogeneous program impacts on migration, I instrument households baseline assets with village-level rainfall shocks. The instrumental variable (IV) estimation results show that the program increases migration for low asset households, with lesser impacts as household assets increase. These results are in line with the theoretical prediction that exogenously increasing credit access affects migration more for households that are more likely to be liquidity constrained. The results of the IV estimation are robust to taking account of potential

---

4 Other studies using randomized control trial found either both employed and self-employed income are positive or both of them are negative, but none of these effects are statistically significant (Augsburg et al., 2015; Tarozzi et al., 2015; Attanasio et al., 2015).
channels of recent rainfall shocks on migration other than via assets, such as through local wages and baseline migration network size. Examination of various margins of impacts on migration suggests that heterogeneous impacts with respect to baseline assets are mostly driven by the extensive margin (whether households have any migrants) rather than the intensive margin (migration duration of migrants), consistent with the importance of up-front migration costs in limiting migration.

To investigate the role of labor market frictions in explaining the program impacts on income generation, I show that income impacts of the program are similar to the migration impacts. Treatment effects on wage income and total income are significantly greater for low asset households than high asset households. Separately evaluating program impacts in counties of high and low migration prevalence, I find that the positive impacts on migration for low assets households are mainly driven by households in counties of low migration prevalence (higher migration cost). The microcredit programs increase income from both self-employment and wage employment for these households while the effects are much more muted for households in villages with high previous migration prevalence.

These findings provide suggestive evidence that the combination of labor market frictions and liquidity constraints makes it difficult for the poor to escape poverty, and that in such environments increasing credit access may have large effects on incomes and poverty reduction. This is consistent with Ghatak's (2015) observation that “no single friction is sufficient to trap individuals in poverty”; rather it requires multiple market failures to create poverty traps. Moreover, the significant impacts of credit access on migration and incomes in China, but not in other countries, may serve as an example of how context matters for the impact of programs (Allcott, 2012; Pritchett and Sandefur, 2015; Vivalt, 2015), justifying concerns about the external validity of field experiments (Deaton, 2012).

The rest of the paper proceeds as follows. In Section 2, I compare predictions from models with and without labor market frictions. Section 3 introduces the village bank program, experimental design, and data. Section 4 introduces the data and measurements. Section 5 presents the empirical specification, estimation strategy, and identification assumptions. Section 6 examines the impacts of exogenous credit access on migration, as well as discussion of robustness tests. In Section 7, I examine how migration cost affects the impacts of the microcredit program, and the impact of the program on incomes. The last section concludes.

Frictions may arise from imperfect credit markets, labor markets, insurance markets (Jing Cai, 2010), or other financial markets such as constraints on savings (Brune et al., 2015; Dupas and Robinson, 2009). Studies that examine single market frictions find little evidence of poverty traps (Kraay and McKenzie, 2014).
2 The Model

This section develops a model of household profit maximization under liquidity constraints. I start with the case in which there is no up-front migration cost, but only liquidity constraints on working capital inputs for household production. I then turn to the case in which both costs of migration and working capital are constrained by liquidity.

Model without Labor Market Friction

Consider a household maximizing its total profit from self-employment (i.e., household agricultural production) and wage employment by purchasing production inputs and allocating labor to the two economic activities. The profit maximization problem of the household is:

\[
\max_{m,k} f(k, n - m) + mw - pk
\]

subject to the constraint \(b(A) - pk \geq 0\).

Here, \(k\) is production inputs, \(n\) is the number of total laborers in the household, \(m\) is the number of laborers who engage in wage employment (or migrate), \(p\) is the market price for production inputs, \(w\) is the wage, \(b\) is the total liquidity available to the household, and \(A\) is the amount of assets of the household or village. I assume \(b\) increases with \(A\), that is \(b'(A) > 0\); \(w, c, p,\) and \(A\) are exogenously given; and \(w, c,\) and \(p\) are independent of \(A\).

The factors for agricultural production include inputs \((k)\) and labor \((l)\) where we assume all labor not allocated to wage employment is used in self-employment activity \((l = n - m)\). The production function \(y = f(k, l)\) is assumed to be second order differentiable and strictly monotonic with diminishing marginal returns, and factors of production are assumed to be complementary. The first order conditions of the maximization problem are

\[
\begin{align*}
\text{If I model migration choice as a household utility maximization problem, the results are qualitatively similar (proofs available upon request). The main complication in a utility maximizing framework is that higher incomes increase the demand for leisure, mitigating predicted employment and income effects.}
\end{align*}
\]

\[
\begin{align*}
\text{Since land is typically equally divided over households in the same villages according to number of laborers, the average land size per capita is typically equalized within village for all households. Therefore employment in agriculture are rare in rural China. In our data, expenditure on hiring labor force accounts for a trivial part in total agriculture inputs at baseline. Within village wage employment other than agricultural work is also unpopular. In our data, only 5% of wage employment were in home village in baseline year.}
\end{align*}
\]

\[
\begin{align*}
\text{It can also be considered as the likelihood of any household member migrated when the total labor force in the household (n) is normalized to be 1.}
\end{align*}
\]

\[
\begin{align*}
\text{Village assets are likely to influence borrowing ability by influencing the total supply of informal financing available, which comes mainly from friends and relatives living in the village. In the model, the theoretical impacts of greater household or village assets is the same.}
\end{align*}
\]

\[
\begin{align*}
\text{This assumption reflects imperfect credit market due to asymmetric information, which causes the amount of available liquidity to depend upon the amount of assets that can be used as collateral.}
\end{align*}
\]

\[
\begin{align*}
\text{In the study sites, self-employment activity is mainly crop farming or animal husbandry. Production inputs includes seeds, chemical fertilizer, feed, animal inoculations, etc. For ease of exposition, we assume no hired labor for agricultural production, which is consistent with field observation. Meanwhile, we abstract from fixed capital that may affect agricultural productivity, such as land and machines. Allowing investment in fixed capital, which is part of assets and so affects credit availability, does not alter the predictions of the model. A version of the model in which we include fixed capital is available upon request.}
\end{align*}
\]

\[
\begin{align*}
\text{For short, we define } f_1 = \frac{\partial f}{\partial k}, f_2 = \frac{\partial f}{\partial \ell}, f_{11} = (\frac{\partial^2 f}{\partial k^2}), f_{22} = (\frac{\partial^2 f}{\partial \ell^2}), f_{12} = f_{21} = (\frac{\partial^2 f}{\partial k \partial \ell}). \text{ Strict monotonicity implies } f_1 > 0, f_2 > 0. \text{ Diminishing marginal productivity implies } f_{11} < 0 \text{ and } f_{22} < 0. \text{ Complementarity between inputs implies } f_{12} = f_{21} > 0. \text{ This assumption is consistent with the fact that household loans used for cropping or animal husbandry are spent}
\end{align*}
\]
$f_1(k, n - m) = (1 + \lambda)p$ and $f_2(k, n - m) = w$, where $\lambda$ is the shadow price of liquidity constraint, $\lambda \geq 0$.

When the liquidity constraint doesn’t bind, $\lambda = 0$ and the first order conditions become $f_1(k, n - m) = p$ and $f_2(k, n - m) = w$. In this case, the marginal product of working capital equals its cost, and the marginal product of labor in self-employment activity equals the market wage. The unconstrained optimum can be solved from the above two conditions: $k^* = \phi_1(w, p, n)$, and $m^* = \phi_2(w, p, n)$. The optimal choices of inputs and labor for household production are independent of $A$. At the same time, the unconstrained optimum can be used to determine a threshold value for assets $A^*$, where $b(A^*) = pk^*$. We then have the following Lemma:\textsuperscript{13}

**Lemma 1** When $A > A^*$, $\lambda = 0$ and the liquidity constraint doesn’t bind; when $A < A^*$, $\lambda > 0$ and the liquidity constraint is binding.

I define constrained optima when the liquidity constraint binds to be $k^{**}$ and $m^{**}$. The constraint $b(A) = pk$ implies $k^{**}$ increases with $A$. From the condition $f_2(k, l) = w - c$, the constrained optimum $l^{**}$ also increases with $A$. Therefore, $m^{**}$ decrease with $A$. We thus have the following proposition:\textsuperscript{14}

**Proposition 1** When $A > A^*$, optimal $m$ and $k$ are independent of $A$; when $A < A^*$, optimal $k$ increases with $A$, and optimal $m$ decreases with $A$.

Figure 1 illustrates the constrained optima $m^{**}$ and $k^{**}$ when only $k$ is constrained by liquidity. For low asset households, there is an underuse of inputs for household production and excess migration relative to the unconstrained optima. For high asset households, since they are not liquidity constrained, optimal $k$ and $m$ are fixed and independent of $A$.

To examine the impact of a microcredit program, I introduce an exogenous increase in available liquidity by an amount $a$, and assume that $a$ is independent of $A$.\textsuperscript{15} The liquidity constraint is now $b(A) + a - pk \geq 0$, where $a > 0$. Since $a$ can be considered to be a simple transformation of $A$, we have the following corollary to Proposition 1:

**Corollary 1** By exogenously increasing liquidity, optimal $k$ increases and optimal $m$ decreases for households if $A < A^*$; optimal $k$ and $m$ are unchanged for households with assets $A > A^*$.

Figure 2 shows that, by exogenously increasing available liquidity, the curves for the constrained optimum shift to the left, indicating that less $A$ is needed to achieve the same amount of liquidity. This leads to an increase in inputs and a decrease in migration.

---

\textsuperscript{13}Mostly on fertilizer or feed, which requires labor inputs.

\textsuperscript{14}We list the propositions and corollary in the text, and leave the proof in the appendix.

\textsuperscript{15}The existence for solution is proved in the appendix.

\textsuperscript{15}Equal amount of fund were given to each treatment villages. The available funds available for each households in the village are therefore determined by the inverse of number of households in village. We didn’t find significant relationship between village bank fund available per households and average log household asset in village. Meanwhile, we found no evidence that the borrowing ability for household of different assets increase differently as results of being in the treatment villages. We measure the borrowing ability from the question that “how much money can you borrow when you have some emergency need for money".
Model with Labor Market Friction

Next, I consider the case in which households face liquidity constraints for both inputs and migration cost. The profit maximization problem is now:

$$\max_{m,k} f(k, n - m) + m(w - c) - pk$$

subject to the constraint $b(A) - pk - mc \geq 0$.

The unconstrained optima $k^* = \phi_1(w, p, c, n)$ and $m^* = \phi_2(w, p, c, n)$ as before are independent of $A$. From the liquidity constraint equation $b(A) = cm^* + pk^*$, we can derive a threshold value $A^{**}$ for which the following Lemma holds:

**Lemma 2** When $A > A^{**}$, $\lambda = 0$ and the liquidity constraint doesn’t bind; when $A < A^{**}$, $\lambda > 0$ and the liquidity constraint binds.

The constrained optima $k^{**}$ and $m^{**}$ satisfy the first order conditions $f_1(k^{**}, n - m^{**}) = (1 + \lambda)p$ and $f_2(k^{**}, n - m^{**}) = w - (1 + \lambda)c$. This means that $f_1(k, l)/p = (w - f_2(k, l))/c = 1 + \lambda$, which means that the marginal return to additional liquidity spent on inputs is equal to the marginal returns to additional liquidity spent on migration costs. Assuming that $|f_{ii}| \gg |f_{ji}|$, it can be shown that $(\partial m^{**})/(\partial k^{**}) > 0$. Given that the liquidity constraint is binding, or $b(A) = pk + cm$, it follows that $(\partial k^{**})/\partial A > 0$ and $(\partial m^{**})/\partial A > 0$, when $A < A^{**}$. We thus have the following proposition and corollary when there are up-front migration costs:

**Proposition 2** When $A > A^{**}$, optimal $k$ and $m$ are independent of $A$; when $A < A^{**}$, optimal $k^{**}$ and $m^{**}$ both increase with $A$.

**Corollary 2** By exogenously increasing liquidity, both $m$ and $k$ increase for households with assets $A < A^*$; while $m$ and $k$ are unchanged for households with assets $A > A^*$.

Figure 3 shows the relationship between optimal $k$ and $m$, and $A$ when households face up-front migration costs. Different from Figure 2, $m$ increases with $A$ for low asset households. Figure 4 illustrates the heterogeneous treatment impacts on migration and working capital for households with different amounts of assets. As shown, the curves for the constrained optima shift leftward if there is an exogenous increase in liquidity, while the horizontal lines for the unconstrained optima are unchanged. As a result, both migration and agricultural inputs increase after the program for low asset households, while nothing changes for high asset households.

I also prove that if the cost of migration is sufficiently high, access to credit will have a greater impact on migration for higher migration cost $c$.\textsuperscript{16}

\textsuperscript{16}In the appendix, I show that this is true if migration cost $c$ meets the condition $c > p\sqrt{f_{22}/f_{11}}$, and $|f_{ii}| \gg |f_{ji}|$. 

6
3 Village Bank Program, Experimental Design and Loan Characteristics

China’s formal financial institutions are dominated by state owned banks, the branches of which are mostly in cities. Rural credit cooperatives are supervised by the People’s Bank of China and are the main financial institutions that meet the credit needs of rural areas. However, many farmers lack access to rationed credit because of lack of collateral (exacerbated by the fact that land is not privately owned), high transactions costs, and perceived riskiness of projects. Initiated by the national government, the village bank program aimed to reduce poverty by providing loans to households in poor rural areas.\(^{17}\) By the end of 2011, the total funds distributed nationwide through the program reached 3.3 billion yuan (US$ 533 million), with around 150,000 yuan (US$24,000) allocated to each village. Remaining funds came from membership fees charged of households, which amounted to an average of 42,000 yuan (US$6,800) per village. The village banks are supervised by the county government Poverty Alleviation Office (PAO), and managed by self-elected committees. The committees decide on the loan terms (amount, duration, interest rate) and on who receives the loans. Although it is requested that poor households be given priority in borrowing from the village bank, there are no official restrictions on participation or on who can receive loans.\(^{18}\) Starting from the year 2006, the program began on a trial basis in 100 villages in 14 provinces. In 2007, 270 more villages were added, and by the end of 2011, the program had been scaled up nationally, reaching 16,300 villages in 28 provinces.

An experimental project was conducted with the support of China’s Office of the Leading Group for Poverty Reduction to evaluate the impacts of the village bank program in five provinces starting from August 2010. The provinces were chosen to be geographically representative of China: two are in western China (Gansu, Sichuan), two are in central China (Henan, Hunan), and one is in eastern China (Shandong). Each province recommended two counties to serve as study sites. Figure 5 shows the geographic location of the selected provinces and counties. In each of the ten counties, county officials recommended five officially designated villages to be eligible for the RCT intervention. To ensure that treatment villages were selected randomly, the research team selected the treatment villages, randomly choosing three of the five eligible villages for treatment, with the remaining two villages serving as control villages.

Figure 6 describes the timeline of the experiment and the surveys that collected data from households in both treatment and control villages. The program was introduced in treatment villages starting in September 2010, with most treatment villages beginning to lend out funds from January to June, 2011.\(^{19}\) The baseline wave of the survey was conducted in August 2010, just before initiation of the program. The survey covered 30 randomly selected households in each village. It collected detailed information on migration days, migration destination,

\(^{17}\) The village bank program differs from previous anti-poverty programs in China in two aspects. Firstly, it targets households, while most previous anti-poverty programs target villages or counties. Secondly, learning lessons from previous failure of anti-poverty programs, the village bank program aims to build up sustainable self-organized credit service in poor rural areas.

\(^{18}\) For most treatment villages, the participation fee is 200 yuan. It is usually waived for poor households.

\(^{19}\) Villages in one county in Hunan Province were an exception as they did not start lending until June 2012.
and wage earnings for each migrant in the household during calendar year 2009, as well as detailed information on household income and expenditure during the same period. It also asked about the value of durables, fixed assets, and housing at the time of survey. A village questionnaire also asked questions about village population, geographic characteristics (mountain, hill, or plain), funds from upper government for various village projects in the past year, and other village characteristics. After two years (in July 2012), the research team re-interviewed all of the households, collecting information similar to that collected in the baseline wave. For treatment villages, the survey also collected information about program implementation (in the village questionnaire) and loans borrowed from the village bank (in the household questionnaire). Households in control villages were not told anything about the village bank program, and by the second wave, no control villages had begun implementing the village bank program.

The loans are mainly provided to support household income-generating activities. The typical loan size is 5000 yuan (about US$800) with a typical duration of one year. The average annual interest rate charged by the village banks was 9.8%, compared to an average loan interest rate of 12.1% for loans from Rural Credit Cooperatives. Most banks lend to groups of households with joint liability, with five to seven households per group. Loans do not require collateral and are repaid in one installment. The average take-up rate of the village bank loans is 28% among the village population, and the repayment rate is 98%.

4 Data and Measurements

Data

As described earlier, 30 households in each village were randomly selected to be surveyed. In total, the baseline survey included 1500 households in 50 villages in 10 counties in 5 provinces. 1351 of the households were successfully re-interviewed in the follow-up survey. The attrition rate of the households thus is 9.9%. In the analysis, all villages in one county were dropped due to late implementation in program villages. Eventually, 1234 households in 45 villages and 9 counties were included in the analysis.

In addition to the household and village survey data, I use matched village-level daily rainfall data from weather station across China during the years 2005 to 2009. I also use data from China’s 2000 census to construct previous migration prevalence for each surveyed county.\textsuperscript{20}

Measurements

Next, I describe several key measurements used in the analysis:

Migration. Migration is measured by the share of migrants among laborers (those aged 16 to 65) in the household, i.e. $m/n$ in the model. A person is a migrant if he or she who worked outside a certain geographic

\textsuperscript{20}More information of these two data sets are documented in the appendix.
scope (home village, home township, home county, home province) for at least ten days during the past year.\footnote{In the analysis, we also replace this definition by migration over six months during the past year.}

**Assets.** Two measurements are used for assets. The first is baseline household assets, which are defined as the sum of the values of durables, fixed assets and housing at the time of the baseline survey, plus the values of crop and animal inventories at the end of 2009.\footnote{For implementation of the survey, information on financial assets such as deposits in financial institutions and cash in hand are not asked in the questionnaire.} The second is baseline village average assets, which is the average of household assets in the village at baseline.

**Income.** Income is the sum of self-employment income (including net income from crop farming, animal husbandry, and small businesses) and wage employment income. Income from public and private transfers are excluded.

**Village Level Rainfall Shock.** The rainfall shock is defined as the deviation of annual rainfall during the non-harvest season in 2009 from the mean for the years 2005 to 2008. The village level rainfall shock in 2009 is normalized by dividing it by the standard deviation of rainfall shocks across all villages.

**Attrition**

I focus on discussion of household attrition here and left the discussion of individual attrition in the appendix. Household attrition is that households surveyed at baseline which didn’t show up in the second wave of the survey.\footnote{Another kind of attrition is individual attrition. That is, even though households were surveyed in both year, its members may be attritted from the roster of house member in the second wave of the survey. Some new members may be added to the households as well. Both will change the composition of the household. We documents individual attrition and leave detailed discussion in the appendix.} The survey did not collect any specific information on the reasons for household attrition. According to field supervisors, it most likely occurs due to temporary absence of all household members during the survey period (for instance being in the hospital because of illness or visiting relatives) or household migration. Even though the attrition rate is relative low, only 10\%, it is worthwhile to assess whether the attrition pattern differs between treatment and control villages.

The first column of Table A1 shows the estimation results for a probit model of household attrition, including the impact of village program treatment and a battery of household and village baseline characteristics. There is no statistically significant impact of treatment on the likelihood of absence during the second wave of the survey, conditioning on other characteristics. The results suggest that lower asset households and households with a greater share of labor who are migrants or with more children at baseline are also suffer from greater attrition. Being in villages with a higher rainfall shock is associated with a lower probability of household attrition. In the second column, an interaction term between village treatment status and household’s baseline assets is included in order to examine if there is any treatment-control difference in attrition by level of household assets. The coefficient of the cross product of treatment and household assets is not statistically significant. That is, for households in treatment villages, the likelihood of attrition decreases with households’ asset in the same
way as in control villages. The last row of the table reports the p-values of a joint $F$-test for the hypothesis that the coefficients of treatment and cross products of treatment and assets are equal to 0. The result suggests no significant treatment-control difference in attrition even allowing for possible heterogeneity by asset level.

**Sample Characteristics and Balance Check**

Table 1 presents the summary statistics of baseline characteristics in the control group (columns 1 and 2) and the results of balance tests between treatment and control groups (columns 3 and 4). Baseline characteristics of villages and households are described in Panels A and B. As shown, more than half of the villages are located in mountainous areas. The average village population is about 1000 persons. 29% of household labor migrated outside of their home village at baseline, and just less than half of these migrants migrated outside of their home province. 54% of the households received interest free loans from informal sources (relatives, friends, etc.) during the period January 2009 to July 2010, while only 12% of households borrowed from financial institutions. This suggests that the formal financial market is not well-developed in these areas, and credit demand is mostly met by informal loans. The last two columns show that the treatment and control group are balanced for nearly all of the household characteristics at baseline, including households assets, income, likelihood for borrowing from various sources, and village characteristics, including rainfall shock, and public expenditure. The share of migrants outside their home village is a little lower for households in treatment villages than in control villages, with the difference being marginally significant (significant at the 10% level). For migration outside other geographic scopes, there is no significant difference between treatment and control groups. All in all, the balance check confirms the randomness of the treatment assignment. The results suggest that there is no statistically significant difference between surveyed households in treatment and control villages.

Table 2 compares the baseline share of migrants and its changes over time for households in treatment and control villages, by households baseline assets. In 2009, the share of labor that migrated to their home township, home county (not in home township), home province (not in home county), and to other provinces are 3%, 6.5%, 6.6%, and 12.3%, respectively, among low asset households in treatment villages. There are no statistically significant treatment-control differences for the low asset group, except that the share of migrants going to another province is less for households in treatment villages. By looking at the change in share of migrants between the years 2011 and 2009, I find that low asset households in treatment villages migrated less to their home township, and more to other provinces, compared to control villages. These difference are statistically significant. For high asset households, I didn’t detect any notable difference between households in treatment and control villages. Overall, these results suggest that the village bank program increases migration (and to more distant locations) only for low assets households, which is consistent with the prediction of the model that migration is affected only for liquidity constrained households.
5 Empirical Estimation

First, I estimate the average intention-to-treat effect of the village bank program on migration by estimating the following equation:

\[ \Delta m_{ijt} = \alpha_0 + \alpha_1 \Delta T_{jt} + \eta X_{ij0} + C + \Delta \varepsilon_{ijt}, \]  

(1)

where \( i \) indexes households, \( j \) indexes village, \( t \) is time period, \( m_{ijt} \) is the share of labor that are migrants, and \( T_{jt} \) is a dummy variable indicating whether the village bank program exists in village \( j \) at time \( t \). Time invariant determinants of migration, such as preferences and geographic location thus are differenced out. To capture determinants of trends in migration, I control for observable characteristics of households and villages at baseline \((X_{ij0})\), where time subscript 0 indicates the baseline period, as well as county fixed effects \((C)\). One of the control variables is migration share at baseline \((m_{ij0})\), so equation (1) can also be interpreted as estimating the determinants of migration after the program, controlling for initial migration share and other initial period characteristics.

The error term \( \Delta \varepsilon_{ijt} \) may be correlated within villages. When the treatment assignment is random, the change in share of migrants in control villages is expected to be identical to the change in share of migrants in treatment villages in the absence of village bank program. The parameter \( \alpha_1 \) thus identifies the average change in share of migrants for households in treatment villages as a result of the village bank program.

The theory in Section 2 predicts that treatment impacts on migration are likely to be heterogeneous for households and villages with different asset levels. To test whether this is the case, we empirically estimate the following equation:

\[ \Delta m_{ijt} = \beta_0 + \beta_1 \Delta T_{jt} + \beta_2 \Delta T_{jt}(A_{ij} - A_{(q)}) + \beta_3(A_{ij} - A_{(q)}) + \theta X_{ij0} + C + \Delta \varepsilon_{ijt} \]  

(2)

where \( A_{ij} \) is the log assets of household \( i \) in village \( j \) at baseline, and \( A_{(q)} \) is the value of log assets at the \( q' \)th percentile of the distribution of log assets at baseline in the full sample. Thus, parameter \( \beta_1 \) measures the treatment effect on change in migration for households with assets at the \( q' \)th percentile of the distribution of baseline assets. To focus on the impacts of the program on poor households, I specify \( q = 10 \), so that \( \beta_1 \) captures the treatment effect for households with baseline assets at the 10th percentile of the asset distribution. \( \beta_2 \) quantifies heterogeneity in the treatment effect with respect to the amount of baseline assets. The theory predicts that with no labor market friction, \( \beta_1 < 0 \) and \( \beta_2 > 0 \), and with up-front migration costs, \( \beta_1 > 0 \) and \( \beta_2 < 0 \). In other words, when there are up-front migration costs, migration of low asset households increases as a result of the program, and the impact of the program on migration decreases with greater household or village assets.

The ordinary least squares estimate (OLS) of \( \beta_2 \) in equation (2) is biased if the the assumption that \( \text{cov}(\Delta T_{jt}A_{ij}, \Delta \varepsilon_{ijt}) = 0 \) is invalid. That is, baseline household assets may be correlated with unobservables that influence the magnitude of treatment effects. For instance, if asset levels are correlated unobserved entrepreneurship, and entrepreneurship affects the extent to which treated households are likely to migrate in response to the
program, then the estimated heterogeneous treatment effects with respect to asset levels could be reflecting the effects of unobserved attributes correlated with assets.

To obtain a more consistent estimate of $\beta_2$ in equation (2), I propose using instrumental variables. Specifically, I use recent rainfall shocks measured at the village level as an instrument for baseline assets. Such shocks reflect random variation in weather, so are unlikely to be correlated with unobserved household attributes. To capture the nonlinear relationship between rainfall shock and assets, I model assets to be a quadratic formula of the shock. This suggests the following first stage equations:

$$A_{ij} = \gamma_0 + \gamma_1 R_j + \gamma_2 R_j^2 + \gamma_3 R_j \Delta T_{jt} + \gamma_4 R_j^2 \Delta T_{jt} + \gamma_5 \Delta T_{jt} + \mu X_{ij0} + C + \zeta_{ij},$$  
(3)

$$A_{ij} \Delta T_{jt} = \delta_0 + \delta_1 R_j + \delta_2 R_j^2 + \delta_3 R_j \Delta T_{jt} + \delta_4 R_j^2 \Delta T_{jt} + \delta_5 \Delta T_{jt} + \pi X_{ij0} + C + \omega_{ij}.$$  
(4)

where $R_j$ is the rainfall shock in village $j$ in the baseline year. In the second stage, the instruments used for the endogenous variables $A_{ij}$ and $A_{ij} \Delta T_{jt}$ include $R_j$, $R_j^2$, $R_j \Delta T_{jt}$, and $R_j^2 \Delta T_{jt}$. For valid IV estimation, the coefficients of the instrumental variables in equation (3) and (4) should be jointly significantly different from 0. Meanwhile, the exclusion condition requires that the rainfall shock doesn’t affect change in migration via channels other than assets, conditioning on $X_{ij0}$ and $C$ ($R_j \perp \Delta \epsilon_{ijt} | X_{ij0}, C$). Under these identifying assumptions, the two stage least squares estimation for parameters $\beta_2$ in equation (2) is unbiased.

6 Is Migration Constrained by Liquidity?

Table 3 reports the average intention-to-treat impacts on migration of the village bank program using ordinary least squares (OLS) to estimate equation (1). According to the estimated treatment effect ($\alpha_1$), the share of labor migrating outside the home township increases by five percentage points in response to the program. This impact is about 20% ($= 5/25.2$) of the mean share of such migration at baseline in control villages. The estimated impacts on migration of different distances (outside of home village, home county, and home province) are consistently positive, and the magnitudes vary in ways that seem reasonable. There is a greater impact on migration outside the home township than outside the home village, which suggests that migration to nearby areas is less likely to be liquidity constrained than migration to locations that are further away.

The positive and significant average treatment impacts of greater credit on migration provides prima facie evidence that liquidity constraints pose a significant barrier to migration for households in China’s poor villages. However, it provides limited information on mechanisms or on which types of households benefit from the program. To provide stronger tests of the theoretical predictions presented earlier, I examine how program impacts vary with the level of household and village assets, and with the cost of migration.

Columns (1) to (4) in Table 4 report the OLS estimates for equation (2), where $q$ is set at the 10th percentile. Thus, the estimated $\beta_1$ is the program impact on households with assets at the 10th percentile of assets. The impacts are positive but only marginally significant for migration outside home township. The estimates of the
interaction term, however, are close to zero and not statistically significant. Of course, the OLS estimates could be biased if household assets are correlated with unobserved factors that also influence the migration response to the program.

Next, I report the preferred results from IV estimation using baseline rainfall shocks as instruments. Columns (1) and (2) in Table A2 report the first stage results based on equations (3) and (4). They show that baseline log household assets have an inverse U-shaped relationship with rainfall shocks, suggesting that too much rainfall or too little rainfall can serve as negative shocks to assets, likely via lower crop yields. The $F$-values for the joint significance tests of the instrumental variables are reported at the bottom of the table, and show that there is not a problem of weak instruments.

The second-stage results for the parameters of interest are reported in columns (5) to (8) of Table 4. The program has a positive and significant impact on migration outside the home township for households at the 10th percentile of the asset distribution. The magnitude of the effect is larger than that suggested by the OLS estimates, suggesting that households with low assets may have unobservables that reduce the propensity to migrate when loan access is increased. The significant, negative coefficient on the interaction term indicates that the treatment effect decreases with assets, suggesting that the estimated average treatment effects are driven mostly by the response of low asset households. This finding is consistent with the theoretical predictions (see Corollary 2). The difference with the OLS results suggests that unobservables correlated with assets reduce the propensity to migrate in response to the program more for poor households than for rich households.

Because rainfall shocks are village-level variables, the IV estimates are identified from variation across villages in household asset levels. Thus, the identified impacts could reflect the impact of average assets in the village, not just household assets. To explore this further, I re-estimate OLS and IV specifications for equation (2) replacing log household baseline assets with the village mean of log household assets (Table 5). The 2SLS estimates are very similar using village mean assets or household assets, which perhaps should not be surprising since in both cases identification comes from differences in village-level rainfall shocks. However, I find that whereas the OLS results using household assets show little evidence of smaller impacts for richer households (insignificant interaction term between treatment and log of household assets), the OLS estimation using village mean assets finds a significant negative coefficient on the interaction term, which is very similar in magnitude to the IV estimate. There are several possible explanations for this finding. First, it could reflect the fact that village level assets are less endogenous to unobserved household factors that are correlated with assets and influence responsiveness of migration to greater credit access. Second, it could be due to classical measurement error in log household assets. Third, it could be that village assets are more important than household assets in determining credit access, if those with more assets lend to those with fewer assets (which could generate strong spillovers from providing

---

24There are natural grounds for taking villages as a whole unit. As put by Townsend (1994), “Village economies satisfy the explicit or implicit conditions of general equilibrium modeling, namely that individuals in the entire community can arrange their institutions and allocations in such a way as to achieve a Pareto optimum.”
credit to some households in each village). Fourth, there could be village spillovers in migration response, if migration by the poor (who take program loans) leads to more migration by the rich.

One concern about the finding that program impacts on migration decreases with average assets in the village is that this could be due to differences in treatment intensity rather than differences in liquidity. This would occur if poorer villages are on average have more available village bank funds per household than the rich villages. Figure 7 plots the total available village bank funds provided by the central government and the average available funds per household against village average assets. Consistent with our earlier description of the program, most villages received 150,000 yuan from the national government, with a few villages not obtaining the full disbursement by the time of the follow-up survey or receiving additional funds from provincial or county governments. Thus, most variation in village assets per household are due to differences in the number of households per village. As seen in the Figure 7, in fact there is no significant correlation between average household assets and the funds available per household in each village.

Table 6 examines the heterogeneous program impacts on migration using different definitions of migration. The results are consistent with the existence of significant up-front migration costs. First, I use share of migration person-days (rather than share of laborers who migrate) to allow for impacts on the intensive margins of time allocation per individual. The results are very similar to the base case using share of household laborers who migrate, suggesting that there is no additional response on the most intensive margin. Second, I examine the response to program on the most extensive margin for migration, namely whether any household member migrates. The estimates are similar but stronger using this specification, consistent with the explanation that up-front migration costs are greater for households with no migrants. Third, by using share of laborers who migrated for over 6 months during the past year (rather than 10 days), I focus on the response of long term migration. The stronger difference in program response between rich and poor households is consistent with longer-term migration involving higher up-front costs.

**Robustness**

Table 7 reports the results of several robustness checks related to identification using rainfall shocks as an instrumental variable. One might be concerned that rainfall shocks could influence migration through channels other than household assets, such as by reducing local labor demand. Panel A examines if impacts on the local market wage can explain why shocks to assets impact migration. To do this, I check whether the estimated impacts of assets interacted with treatment are robust to controlling explicitly for changes in the local market wage (as reported in village questionnaires) and its interaction with village treatment status. The results show that migration does decrease when the local market wage increases. However, the response of migration to the program increases with change in the local market wage. Given that rainfall shocks are expected to reduce local demand and reduce local market wages, this leads to the estimates of the treatment effect on wages and interaction
with assets to be even larger in magnitude than in the baseline results, and still statistically significant.

Panel B uses lagged rainfall (in 2008) rather than recent rainfall to construct the instruments. Similar to the above logic, one might be concerned that recent rainfall shocks reduce agricultural demand that is not captured in the local village wage, and so push households to migrate. Such push factors affect migration by impacting relative returns rather than incomes and asset levels. Lagged rainfall shocks should be relatively uncorrelated with agricultural productivity shocks during the period covered by the baseline survey. The results suggest that using lagged rainfall shocks does not significantly alter the magnitude of the parameters of interest.

Another concern for the validity of IV estimation is that rainfall shocks (both recent and lagged) may affect the size of village migration networks which affect the cost of individual migration by households in the village (Munshi, 2003). In Panel C, I test the robustness of the results when I control for the village level share of migrants outside the home county at baseline. The results show that household migration does increase with migration network size at baseline. However, controlling for the baseline migration network size doesn’t alter much the estimates of treatment impacts.

### 7 Migration Cost and Program Impacts

The previous section established the importance of liquidity constraints for migration. In this section, we examine how program response differs with migration cost, and examine how the responses to the program increase incomes. Previous studies have found evidence that migration networks have reduced migration costs and increased the likelihood of subsequent migration (Massey et al., 1994; Gathmann, 2004; McKenzie and Rapoport, 2007, 2010; Munshi, 2011). Historical migration networks may reduce up-front migration costs in several ways: (1) assist new migrants to find accommodations or temporary lodging; (2) provide new migrants with loans for temporary needs at the migration destination; and (3) provide information on job opportunities that reduces job search costs.

To ensure our measure of migration networks is as exogenous as possible, we measure the intensity of established migration networks by county-level share of workers who migrate outside of the township in 2000 (10 years before the baseline survey) using Chinese census data. The sample is divided into low- and high-migration prevalence counties. Figure 8 displays the relationship between migration rates in 2000 and in the baseline wave in 2010, revealing a clear positive correlation.

Table 8 repeats the IV estimation of equation (2) for separate sub-samples of households in low- and high-migration counties. The results show that, in low migration prevalence counties, the results are qualitatively similar to the results for the full sample, but the magnitude of the effects (and statistical significance) is greater. On the other hand, in high-migration counties the program does not have a significant impact on migration by low-asset households, nor are there significant differences in program response between rich and poor households. Overall, the results suggest that the impacts of the credit program are most pronounced for households with low

---

25 The results also are robust to controlling for the rainfall shock in 2009. Results are available upon request.
assets in counties with higher migration costs. These are exactly the households for whom liquidity constraints on migration are expected to be most binding.

Next, I examine the implications of the migration impacts of the program for income generation. Table 9 reports the treatment impacts on income. As shown in Panel A, the average intention-to-treat impacts on both income of self-employment and wage employment activities are positive. But only the impact on employment income is significant. On average, the total income of households in treatment villages increased much more than for households in control villages. Panel B examines the heterogeneous treatment impacts for households with different baseline assets. The IV estimates suggest that the positive program impacts are mostly driven by poor households. That is, for poor households, both self-employment income and wage employment income increased as being in the program villages, while the program impacts on income decrease with households’ baseline assets. As shown in Table 10, in low-migration counties, employment income increased significantly for low-asset households (villages), and the impact decreases with the amount of household assets. In counties of high migration prevalence, self-employment income increases for low asset households in treatment village, but employment income doesn’t change much. These income effects are consistent with the results on migration in Table 8. In other words, households that face potential frictions in both credit and labor markets benefit the most from relaxing liquidity constraints.

8 Conclusions

Analyzing data from a randomized control trial that exogenously increased credit access for households in poor villages in China, I empirically examine the role of liquidity constraints on migration. I present, for the first time, robust evidence that migration increases with greater credit access using an RCT. I also show that migration increases for low-asset households as a result of the program. Further, I find that the program increases migration and total incomes more in low-migration areas where migration costs are expected to be higher. These results are in line with the predictions of a theoretical model of household profit-maximizing decisions. They suggest that when credit can help overcome frictions in both credit and labor markets, the income benefits may be more pronounced.
References


working paper.


Maitra, Pushkar, Sandip Mitra, Dilip Mookherjee, Alberto Motta, Sujata Visaria. 2015. “Financing Smallhold-


1. Proof of Lemma 1 and Proposition 1

Define

\[ F_1(k, m, \lambda; p, w, n, A) = f_1(k, n - m) - (1 + \lambda)p = 0 \]
\[ F_2(k, m, \lambda; p, w, n, A) = f_2(k, n - m) - w = 0 \]
\[ F_3(k, m, \lambda; p, w, n, A) = b(A) - pk = 0 \]

\( F_1, F_2, \) and \( F_3 \) are continuous and differentiable. The Jacobian matrix \( J_1 \) of the equation system can be written as

\[
J_1 = \begin{bmatrix}
\frac{\partial F_1}{\partial k} & \frac{\partial F_1}{\partial m} & \frac{\partial F_1}{\partial \lambda} \\
\frac{\partial F_2}{\partial k} & \frac{\partial F_2}{\partial m} & \frac{\partial F_2}{\partial \lambda} \\
\frac{\partial F_3}{\partial k} & \frac{\partial F_3}{\partial m} & \frac{\partial F_3}{\partial \lambda}
\end{bmatrix} = \begin{bmatrix}
f_{11} & -f_{12} & -p \\
f_{21} & -f_{22} & 0 \\
-p & 0 & 0
\end{bmatrix}
\]

\( |J_1| = p^2 f_{22} < 0 \). According to the implicit function theorem, there exists functions \( \varphi_1, \varphi_2 \) and \( \varphi_3 \) which are continuous and differentiable, such that the following conditions are satisfied:

\[ k^{**} = \varphi_1(p, w, n, A) \]
\[ m^{**} = \varphi_2(p, w, n, A) \]
\[ \lambda^{**} = \varphi_3(p, w, n, A) \]

Similarly, we have for all \((k, m, \lambda) \in N_\delta(k^{**}, m^{**}, \lambda^{**})\), \( J_1 \) has full rank. Furthermore,

\[
\begin{bmatrix}
\frac{\partial \varphi_1}{\partial A} \\
\frac{\partial \varphi_2}{\partial A} \\
\frac{\partial \varphi_3}{\partial A}
\end{bmatrix} = J_1^{-1} \begin{bmatrix}
\frac{\partial F_1}{\partial A} \\
\frac{\partial F_2}{\partial A} \\
\frac{\partial F_3}{\partial A}
\end{bmatrix}
\]

\[
= \frac{1}{|J_1|} \begin{bmatrix}
0 & 0 & -p f_{22} \\
0 & 0 & -p f_{12} \\
-p f_{22} & p f_{12} & -(f_{11} f_{22} - f_{12} f_{21})
\end{bmatrix} \begin{bmatrix}
0 \\
0 \\
-b'(A)
\end{bmatrix}
\]

Under the assumption \(-f_{ii} \gg f_{ij}\), we have

\[ \frac{\partial k^{**}}{\partial A} = \frac{\partial \varphi_1}{\partial A} = \frac{1}{|J_1|} p f_{22} b'(A) > 0 \]
\[ \frac{\partial m^{**}}{\partial A} = \frac{\partial \varphi_2}{\partial A} = \frac{1}{|J_1|} p f_{12} b'(A) < 0 \]
\[ \frac{\partial \lambda^{**}}{\partial A} = \frac{\partial \varphi_3}{\partial A} = \frac{1}{|J_1|} (f_{11} f_{22} - f_{12} f_{21}) b'(A) < 0 \]

From the last equation, we have \( \lambda > 0 \) or the liquidity constraint binds for \( A < A^* \). From the first two equations, we have \( k^{**} \) increases with \( A \), while \( m^{**} \) decreases with \( A \) at the constrained optimum.
2. Proof of Corollary 1

The proof is similar to the proof of above. The differences are

\[ F_3(k, m, \lambda; p, w, n, A, a) = b(A) + a - pk = 0 \]

and

\[ k^{**} = \varphi_1(p, w, n, A, a) \]
\[ m^{**} = \varphi_2(p, w, n, A, a) \]
\[ \lambda^{**} = \varphi_3(p, w, n, A, a) \]

The Jacobian matrix of the equation system is still \( J_1 \). Similarly, we have

\[
\begin{bmatrix}
\frac{\partial \varphi_1}{\partial a} & \frac{\partial \varphi_2}{\partial a} & \frac{\partial \varphi_3}{\partial a}
\end{bmatrix} = J_1^{-1} \begin{bmatrix}
-\frac{\partial F_1}{\partial a} \\
-\frac{\partial F_2}{\partial a} \\
-\frac{\partial F_3}{\partial a}
\end{bmatrix}
\]

\[
= \frac{1}{|J_1|} \begin{bmatrix}
0 & 0 & -pf_{22} \\
0 & -p^2 & -pf_{12} \\
-pf_{22} & pf_{12} & -(f_{11}f_{22} - f_{12}f_{21})
\end{bmatrix} \begin{bmatrix}
0 \\
0 \\
-1
\end{bmatrix}
\]

Given \(-f_{ii} \gg f_{ij}\), we have

\[
\frac{\partial k^{**}}{\partial a} = \frac{\partial \varphi_1}{\partial a} = \frac{1}{|J_1|} pf_{22} > 0
\]
\[
\frac{\partial m^{**}}{\partial a} = \frac{\partial \varphi_2}{\partial a} = \frac{1}{|J_1|} pf_{12} < 0
\]

That is, when liquidity constraint binds, input of working capital in self-employment increases, while migration decreases, as results of increasing in credit access. However, when liquidity constraint doesn’t bind, increase in liquidity will not affect migration and inputs in working capital as shown in the text.

3. Proof of Lemma 2 and Proposition 2

When \(c \neq 0\), define

\[ F_1(k, m, \lambda; p, w, n, c, A) = f_1(k, n - m) - (1 + \lambda)p = 0 \]
\[ F_2(k, m, \lambda; p, w, n, c, A) = f_2(k, n - m) - w + (1 + \lambda)c = 0 \]
\[ F_3(k, m, \lambda; p, w, n, c, A) = b(A) - pk - cm = 0 \]

\(F_1, F_2, \text{ and } F_3\) are continuous and differentiable. The Jacobian matrix \(J_2\) of the equation system is

\[
J_2 = \begin{bmatrix}
f_{11} & -f_{12} & -p \\
f_{21} & -f_{22} & c \\
-p & -c & 0
\end{bmatrix}
\]
\[|J_2| = c(cf_{11} + pf_{12}) + p(pf_{22} + cf_{21}).\] By assuming \(-f_{ii} \gg f_{ij}\), we have \(|J_2| < 0\). According to the implicit function theorem, there exists functions \(\psi_1, \psi_2\) and \(\psi_3\) which are continuous and differentiable, such that the following conditions are satisfied:

\[k^{**} = \psi_1(p, w, n, c, A)\]
\[m^{**} = \psi_2(p, w, n, c, A)\]
\[\lambda^{**} = \psi_3(p, w, n, c, A)\]

Similarly, we have for all \((k, m, \lambda) \in N_\delta(k^{**}, m^{**}, \lambda^{**})\), \(J_2\) has full rank. Furthermore,

\[
\begin{bmatrix}
\frac{\partial \psi_1}{\partial A} \\
\frac{\partial \psi_2}{\partial A} \\
\frac{\partial \psi_3}{\partial A}
\end{bmatrix} = J_2^{-1} \begin{bmatrix}
-\frac{\partial F_1}{\partial A} \\
-\frac{\partial F_2}{\partial A} \\
-\frac{\partial F_3}{\partial A}
\end{bmatrix}
\]

\[
= \frac{1}{|J_2|} \begin{bmatrix}
c^2 & -cp & -(pf_{22} + cf_{21}) \\
-cp & -p^2 & -(cf_{11} + pf_{12}) \\
-(pf_{22} + cf_{21}) & (cf_{11} + pf_{12}) & -(f_{11}f_{22} - f_{12}f_{21})
\end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -b'(A) \end{bmatrix}
\]

Therefore,

\[
\frac{\partial k^{**}}{\partial A} = \frac{\partial \psi_1}{\partial A} = \frac{1}{|J_2|}(pf_{22} + cf_{21})b'(A) > 0
\]
\[
\frac{\partial m^{**}}{\partial A} = \frac{\partial \psi_2}{\partial A} = \frac{1}{|J_2|}(cf_{11} + pf_{12})b'(A) > 0
\]
\[
\frac{\partial \lambda^{**}}{\partial A} = \frac{\partial \psi_3}{\partial A} = \frac{1}{|J_2|}(f_{11}f_{22} - f_{12}f_{21})b'(A) < 0
\]

The last equation implies, \(\lambda > 0\) or the liquidity constraint binds for \(A < A^*\). Meanwhile, both \(m^{**}\) and \(k^{**}\) increase with \(A\), when the liquidity constraint binds.

4. Proof of Corollary 2

Similar to the proof of Corollary 1, we have

\[
\begin{bmatrix}
\frac{\partial \psi_1}{\partial a} \\
\frac{\partial \psi_2}{\partial a} \\
\frac{\partial \psi_3}{\partial a}
\end{bmatrix} = J_2^{-1} \begin{bmatrix}
-\frac{\partial F_1}{\partial a} \\
-\frac{\partial F_2}{\partial a} \\
-\frac{\partial F_3}{\partial a}
\end{bmatrix}
\]

\[
= \frac{1}{|J_2|} \begin{bmatrix}
c^2 & -cp & -(pf_{22} + cf_{21}) \\
-cp & -p^2 & -(cf_{11} + pf_{12}) \\
-(pf_{22} + cf_{21}) & (cf_{11} + pf_{12}) & -(f_{11}f_{22} - f_{12}f_{21})
\end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}
\]

Therefore,

\[
\frac{\partial k^{**}}{\partial a} = \frac{\partial \psi_1}{\partial a} = \frac{1}{|J_2|}(pf_{22} + cf_{21}) > 0
\]
\[
\frac{\partial m^{**}}{\partial a} = \frac{\partial \psi_2}{\partial a} = \frac{1}{|J_2|}(cf_{11} + pf_{12}) > 0
\]

That is, when the liquidity constraint binds, the exogenous increase in credit will increase both investment in migration and inputs in working capital.
5. The Effect of $c$ on the Program Impacts on Migration $\partial m^{**}/\partial a$ when $c > 0$

From the proof of Corollary 2, we have

$$\frac{\partial m^{**}}{\partial a} = \frac{\partial \psi_2}{\partial a} = \frac{1}{|J_2|} \left( cf_{11} + pf_{12} \right) = \frac{cf_{11} + pf_{12}}{c(cf_{11} + pf_{12}) + p(pf_{22} + cf_{21})} = \frac{1}{c} + \frac{cf_{11} + pf_{12}}{p(pf_{22} + cf_{21})}$$

Intuitively, the first part represents the marginal effect of $c$ on the program impacts on migration $\partial m^{**}/\partial a$, when liquidity constraint binds. That is, $\partial m^{**}/\partial a$ decreases when $c$ increases, given bounded liquidity constraint. The second part represents the marginal effect of $c$ on the likelihood of being liquidity constrained. The higher is $c$, the more likely household is constrained by liquidity, and therefore greater program impacts on migration. By assuming $-f_{ii} \gg f_{ij}$,

$$\frac{\partial m^{**}}{\partial a} \approx \frac{1}{c} + \frac{cf_{11}}{pf_{22}}$$

For $c > p\sqrt{f_{22}/f_{11}}$, $\partial m^{**}/\partial a$ increases with $c$. That is, when $c$ is greater enough, the second impact dominants the first impact. Relaxing liquidity constraint caused by higher migration cost $c$ will leads to more increase in migration.

6. Model of Subsistence Constraint

The household’s profit maximization problem is

$$\max_{k,l} \ f(k,l) + m(w - c) - pk$$

s.t. $b(A) + a \geq pk, f(k,l) - s \geq 0$, where $s$ is the subsistence needs for survive, and households are assumed to feed themselves by self-produced agricultural output. The Lagrangian function is

$$L = f(k,l) + m(w - c) - pk + \lambda_1[b(A) + a - pk] + \lambda_2[f(k,l) - s],$$

where $\lambda_1$ is the shadow price of liquidity constraint on investment of working capital, $\lambda_2$ is the shadow price of subsistence constraint. The first order conditions are

$$\frac{\partial L}{\partial k} = f_1(1 + \lambda_2) - p(1 + \lambda_1) = 0$$
$$\frac{\partial L}{\partial l} = f_2(1 + \lambda_2) - (w - c) = 0$$
$$\frac{\partial L}{\partial \lambda_1} = b(A) + a - pk = 0$$
$$\frac{\partial L}{\partial \lambda_2} = f(k,l) - s = 0$$

When neither constraint is binding, $\lambda_1 = 0, \lambda_2 = 0$. The optimum solutions are the same as those in the benchmark model. We have $\partial k^{**}/\partial a = 0$ and $\partial m^{**}/\partial a = 0$. When both constraints are binding, $\lambda_1 > 0, \lambda_2 > 0$. Equation (A3) implies $\partial k^{**}/\partial a > 0$. Equation (A4) implies $\partial l^{**}/\partial a < 0$. Therefore, $\partial m^{**}/\partial a > 0$. These are the same as predictions from model of liquidity constrain on both migration and input in self-employment.
However, when \( c = 0 \) and both liquidity constraint and subsistence constraint bind, we also have \( \partial m^{**}/\partial a > 0 \). This is different from the benchmark model in the text, where it predicts migration decreases as a result of the village bank program for low asset households when there is no migration cost. Intuitively, the equations (A3) and (A4) determine the constrained optima \( k^{**} \) and \( l^{**} \), which are unaffected by migration cost \( c \). That is, the subsistence constraint associated with liquidity constraint on only working capital is not able to predict heterogeneous treatment impacts on migration for different migration cost.

7. Check Changes in Households Composition

6378 person were recorded as household member in the baseline survey. Among them, 90% (= 5736/6378) were in households followed in the second wave of the survey (see text for discussion on household attrition). Overall, 527 person (9% of 5736) were attrited individually in the follow-up survey. Among them, only 342 person were aged between 16 and 65. For these people, 174 were active labor (the other either occasionally participated or didn’t participate labor activity). Their attrition are potentially because of migration. In the data, we find the attrition rate of person aged between 20 and 30 years old are extraordinarily higher for unmarried women than for unmarried man (18% versus 4%). These extra attrition of unmarried women are very likely be because of marriage out. By assuming balanced attrition between men and women, 14% of the 193 unmarried women aged between 20 and 30, i.e. 27, can be considered to be attrited because of marriage out. For rest 147 person, their attrition are potentially because of migration out of family property which we can’t identify from the data. Overall, 2.6% of the total 5736 person in households surveyed in both waves, or 3.6% of the 4126 person aged between 16 and 65 in households surveyed in both waves are potentially attrited because of migration.

Among the 5689 person recorded as household member in the second wave of the survey, 5209 person shown up in both waves. 480 person (8% of 5689) were new in the second wave. Among them, 126 person (26% of 480) were new born babies. Another 122 person were either younger than 16 years old or older than 65 years old. For the rest 232 person, only 142 had migration experience and potentially were new members for reasons of return migration (not take account of marry in for young females). Overall, 2.5% of total person in the second wave, or 3.5% of the 4049 person aged between 16 and 65 who recorded in the second wave of the survey are potentially be new members of the household as return migrants.

Figure A1 shows the age distribution of attrited individuals or new household members by gender. It shows most of the new members in the second wave are new born babies. Among the labor population, the number of attrition or new member between age 20 and 30 are highly unbalanced between female and male, suggesting sufficient amount of marry related changes in household composition. All in all, the changes in households composition are unlikely driven by migration. Above analysis suggests the upper bound for migration related attrition is 3.6% among the labor population, while the upper bound for new members of returned migrants is 3.5% of the
labor population. In addition, by performing tests on hypothesis that the likelihood of migration related attrition (or being new members for reason of return migration) are equal between treatment and control groups, we cant reject the hypothesis on significant level of 10%.

8. Rainfall Data

I use the daily rainfall data from weather station across China from 2005 to 2009. It was linked to our survey data by geographic latitude and longitude. More specifically, I matched each of the surveyed village to weather station within 100 kilometers in great cycle distance. The daily precipitation of each village during year 2005-2009 were constructed by the inverse distance weighted average of all matched weather station for the village. I focus on the rainfall in non-harvest season instead of calendar year, since rainfall during harvest season may hurt agricultural production by delaying the harvest and erosion. I identify the harvest period of main agricultural plants for each of the sampled county according to my review on local agricultural season. The annual rainfall for each village is constructed by excluding the days during harvest periods. The rainfall shock for each village is the deviation of annual rainfall in year 2009 from its historical mean during 2005-2008.

9. Census Data

I use the 2000 China census data to construct historical migration prevalence on county level. Specifically, I linked the census data to the survey data by county level statistical division code. The census questionnaire asked for each sampled household the number of person migrated outside home township over six month. This is used to construct the share of migrants (outside home township) on county level for each village in the survey.
Figure 1: Optimal $m$ and $k$: only $k$ is Liquidity Constrained

Figure 2: Impacts of Microcredit: only $k$ is Liquidity Constrained
Figure 3: Optimal $m$ and $k$: both $m$ and $k$ are Liquidity Constrained

$$m^* (p, w, c, n)$$

$A^*$

$A$

Figure 4: Impacts of Microcredit: both $m$ and $k$ are Liquidity Constrained

$$m^* (p, w, c, n)$$

$A^*$

$A$

$a' > a = 0$
Figure 5: Sampling of Provinces, Counties, and Villages
Figure 6: Timeline of Experiment and Survey

- **Experiment**
  - Starting to establish village bank

- **Survey**
  - Rainfall (outsource matched data)
  - Migration 1st wave
  - Income 1st wave
  - Expenditure 1st wave
  - Grain Stock
  - Livestock
  - Durables
  - Fixed assets
  - Housing
  - Migration 2nd wave
  - Income 2nd wave
  - Expenditure 2nd wave
  - Durables
  - Fixed assets
  - Housing

- **Timeline**
  - 2009 JAN to 2010 JAN: 1st Wave 2010 AUG
  - 2011 JAN: Starting to lend loans (Treatment)
  - 2012 JAN to 2013 JAN: 2nd Wave 2012 JUL
Figure 7: Average Available Program Fund and Asset Levels in Village

Figure 8: Relationship between Previous and Current Migration Prevalence

Note: The migration prevalence is the share of labor force migrated outside home township. The village level migration prevalence are measured in our data from village questionnaire which asked directly of this question. The county level migration experience is calculated from 2000 Chinese census by dividing the total number of migrants outside home township by the total number of labor force in the same county in year 2000. Standard errors are clustered by county for the ordinary least square estimation.
Figure A1: Change in Household Composition in Panel Households

Panel A: Age Distribution of Attrited Individuals by Gender

Panel B: Age Distribution of New Household Members by Gender

Notes: The sample in Panel A includes individuals in panel households who were attrited in the second wave of the survey. The sample in Panel B includes individuals in panel households who only shown in the second wave of the survey. The dash lines group the sample into children, labor force, and the elderly according to ages of 16 and 65 years old.
### Table 1: Sample Characteristics and Balance Check

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>Control</th>
<th>S.D.</th>
<th>Treat - Control</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Village baseline characteristics</strong> (# control village=18, # treatment village=27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Geographic feature</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mountain (yes=1)</td>
<td>0.56</td>
<td>0.51</td>
<td>-0.04</td>
<td>0.813</td>
</tr>
<tr>
<td>Hills (yes=1)</td>
<td>0.33</td>
<td>0.49</td>
<td>0.04</td>
<td>0.805</td>
</tr>
<tr>
<td>Plain (yes=1)</td>
<td>0.11</td>
<td>0.32</td>
<td>0.00</td>
<td>1.000</td>
</tr>
<tr>
<td>Population</td>
<td>1028</td>
<td>482</td>
<td>188</td>
<td>0.277</td>
</tr>
<tr>
<td>Area of arable land (unit: mu)</td>
<td>1490</td>
<td>1669</td>
<td>712</td>
<td>0.269</td>
</tr>
<tr>
<td><strong>Public expenditure by upper government (thousand RMB)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telephone, broadcast, cable television</td>
<td>21.94</td>
<td>72.54</td>
<td>-13.80</td>
<td>0.349</td>
</tr>
<tr>
<td>Energy (electricity, gas, etc.)</td>
<td>48.73</td>
<td>91.80</td>
<td>-14.70</td>
<td>0.518</td>
</tr>
<tr>
<td>Drinking water</td>
<td>59.00</td>
<td>150.52</td>
<td>-21.37</td>
<td>0.556</td>
</tr>
<tr>
<td>Irrigation, water conservancy</td>
<td>5.56</td>
<td>23.57</td>
<td>111.85</td>
<td>0.243</td>
</tr>
<tr>
<td>Land improvement</td>
<td>10.89</td>
<td>34.30</td>
<td>51.33</td>
<td>0.375</td>
</tr>
<tr>
<td>Environment improvement</td>
<td>0.00</td>
<td>0.00</td>
<td>27.52</td>
<td>0.392</td>
</tr>
<tr>
<td>School</td>
<td>138.33</td>
<td>438.15</td>
<td>-129.44</td>
<td>0.131</td>
</tr>
<tr>
<td>Hospital and clean toilet</td>
<td>14.67</td>
<td>28.47</td>
<td>2.59</td>
<td>0.778</td>
</tr>
<tr>
<td>Others</td>
<td>4.40</td>
<td>18.67</td>
<td>4.47</td>
<td>0.489</td>
</tr>
<tr>
<td><strong>Average yearly rainfall 2005-2009 (unit: millimeter)</strong></td>
<td>660.33</td>
<td>242.01</td>
<td>1.42</td>
<td>0.985</td>
</tr>
<tr>
<td><strong>Rainfall shock in 2009 (unit: millimeter)</strong></td>
<td>-13.57</td>
<td>39.08</td>
<td>-2.56</td>
<td>0.843</td>
</tr>
<tr>
<td><strong>Panel B: Household baseline characteristics</strong> (# control hh=493, # treatment hh=741)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Household size</strong></td>
<td>4.29</td>
<td>1.52</td>
<td>-0.12</td>
<td>0.173</td>
</tr>
<tr>
<td><strong>Share of labor force migrated outside</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home village</td>
<td>0.29</td>
<td>0.26</td>
<td>-0.03</td>
<td>0.098*</td>
</tr>
<tr>
<td>Home township</td>
<td>0.25</td>
<td>0.25</td>
<td>-0.01</td>
<td>0.493</td>
</tr>
<tr>
<td>Home county</td>
<td>0.20</td>
<td>0.24</td>
<td>-0.02</td>
<td>0.160</td>
</tr>
<tr>
<td>Home province</td>
<td>0.12</td>
<td>0.21</td>
<td>-0.01</td>
<td>0.287</td>
</tr>
<tr>
<td><strong>log household assets</strong></td>
<td>9.79</td>
<td>1.32</td>
<td>-0.08</td>
<td>0.260</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td>9572</td>
<td>9417</td>
<td>-755</td>
<td>0.154</td>
</tr>
<tr>
<td><strong>Received loans from financial institutions (Jan 2009 - Jul 2010)</strong></td>
<td>0.12</td>
<td>0.32</td>
<td>0.03</td>
<td>0.177</td>
</tr>
<tr>
<td>Received informal loans with interest (Jan 2009 - Jul 2010)</td>
<td>0.04</td>
<td>0.20</td>
<td>0.00</td>
<td>0.702</td>
</tr>
<tr>
<td>Received informal loans without interest (Jan 2009 - Jul 2010)</td>
<td>0.54</td>
<td>0.50</td>
<td>0.00</td>
<td>0.882</td>
</tr>
</tbody>
</table>

Notes: The observation unit in Panel A is village. The observation unit in panel B is household. The sample excludes one county where the village bank program started right before the second wave of the survey in treatment villages. The village level rainfall is calculated from matched village level daily rainfall data from weather station across China from 2005 to 2009. The other variables are measured from the survey data. Column 4 reports p-values on testing the hypothesis that the difference between the treatment and the control mean is equal to 0. * p<0.1.
Table 2: (Change in) Migration by Household Baseline Assets and Village Treatment Status

<table>
<thead>
<tr>
<th>Labor share in household migrated to</th>
<th>Year 2009</th>
<th>Change between year 2011 and 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Treatment</td>
</tr>
<tr>
<td>Low asset household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To home township</td>
<td>0.030</td>
<td>0.025</td>
</tr>
<tr>
<td>To home county/outside home township</td>
<td>0.065</td>
<td>0.064</td>
</tr>
<tr>
<td>To home province/outside home county</td>
<td>0.066</td>
<td>0.055</td>
</tr>
<tr>
<td>To foreign province/outside home province</td>
<td>0.123</td>
<td>0.086</td>
</tr>
<tr>
<td>High asset household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To home township</td>
<td>0.058</td>
<td>0.033</td>
</tr>
<tr>
<td>To home county/outside home township</td>
<td>0.048</td>
<td>0.069</td>
</tr>
<tr>
<td>To home province/outside home county</td>
<td>0.079</td>
<td>0.077</td>
</tr>
<tr>
<td>To foreign province/outside home province</td>
<td>0.116</td>
<td>0.128</td>
</tr>
</tbody>
</table>

Note: Households are evenly divided into low and high asset households according to their baseline household assets. Labor force is defined as household member aged between 16 and 65 years old (inclusive). Migrants are labor force who migrated outside home village more than 10 days during year 2009 (or year 2011) for baseline (or follow-up) survey.
### Table 3: Average Treatment Impacts on Migration

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Outside home village (1)</th>
<th>Outside home township (2)</th>
<th>Outside home county (3)</th>
<th>Outside home province (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treat</td>
<td>0.02</td>
<td>0.05*</td>
<td>0.05**</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>[0.788]</td>
<td>[1.922]</td>
<td>[2.442]</td>
<td>[1.606]</td>
</tr>
<tr>
<td>Mean of baseline share of migrants in control group</td>
<td>0.295</td>
<td>0.252</td>
<td>0.196</td>
<td>0.123</td>
</tr>
<tr>
<td>Observations</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
</tr>
</tbody>
</table>

Notes: Robust *-statistics in brackets. The standard errors are clustered by village. ** * p<0.05, * p<0.1. Other control variables include county dummies, village baseline characteristics (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household baseline characteristics (number of laborers, size of arable land, share of migrants outside corresponding geographic scope).
Table 4: Heterogeneous Treatment Impacts on Migration by Baseline Household Assets

<table>
<thead>
<tr>
<th></th>
<th>OLS estimation</th>
<th>2SLS estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outside home village</td>
<td>Outside home township</td>
</tr>
<tr>
<td>Treat</td>
<td>0.03</td>
<td>0.06*</td>
</tr>
<tr>
<td></td>
<td>[1.006]</td>
<td>[2.012]</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>[-0.481]</td>
<td>[0.055]</td>
</tr>
</tbody>
</table>

F values of first stage

| (Baseline log hh asset - value at 10%) × Treat | 50.08 | 48.20 | 48.04 | 48.03 |
| Baseline log hh asset - value at 10% | 12.23 | 11.35 | 10.91 | 11.38 |

p values of overidentification test

| Observations | 0.2451 | 0.3237 | 0.1266 | 0.8398 |

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. ** p<0.05, * p<0.1. Other controls include county dummies, village (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household characteristics (number of laborers, size of arable land, having any member migrated outside corresponding geographic scope) in baseline year.
Table 5: Household Level Regression by Using Village Average Assets

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>OLS estimations</th>
<th>2SLS estimations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Share of Migrants</td>
<td>Outside home village</td>
<td>Outside home township</td>
</tr>
<tr>
<td>Treat (Baseline village average log hh asset - value at 10%) × Treat</td>
<td>0.11***</td>
<td>0.12***</td>
</tr>
<tr>
<td></td>
<td>[3.497]</td>
<td>[3.965]</td>
</tr>
<tr>
<td>(Baseline village average log hh asset - value at 10%)</td>
<td>-0.11***</td>
<td>-0.09***</td>
</tr>
<tr>
<td>Baseline village average log hh asset - value at 10%</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>[1.580]</td>
<td>[1.521]</td>
</tr>
</tbody>
</table>

*F values of first stage*

| (Baseline village average log hh asset - value at 10%) × Treat | 35.03 | 34.91 | 35.15 | 35.18 |
| Baseline village average log hh asset - value at 10% | 8.71 | 8.66 | 8.72 | 8.70 |

*p values of overidentification test*

| 0.2276 | 0.3058 | 0.0700 | 0.8378 |

Observations: 1,234

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other controls include county dummies, village (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household characteristics (number of laborers, size of arable land, having any member migrated outside corresponding geographic scope) in baseline year. The village average assets is the average of log households assets in village at baseline, minused by the value at 10% of its distribution.
Table 6: Impacts on Various Margins of Migration

<table>
<thead>
<tr>
<th>Dependent variable: Change in Migration share in person-days</th>
<th>2SLS estimation</th>
<th>Any household member migrated</th>
<th>Share of laborers migrated over 6 months during a year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Migration share in person-days</td>
<td>Outside home village (1)</td>
<td>Outside home township (2)</td>
</tr>
<tr>
<td>Treat</td>
<td>0.16***</td>
<td>0.17***</td>
<td>0.13**</td>
</tr>
<tr>
<td></td>
<td>(2.642)</td>
<td>(2.867)</td>
<td>(2.449)</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>-0.09***</td>
<td>-0.08***</td>
<td>-0.06**</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>0.08*</td>
<td>0.09*</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>[1.778]</td>
<td>[1.841]</td>
<td>[1.532]</td>
</tr>
</tbody>
</table>

F values of first stage

| (Baseline log hh asset - value at 10%) × Treat              | 48.99           | 48.45                         | 48.19                         | 48.24                        | 50.53                        | 48.29                        | 47.70                        | 47.98                        | 47.98                        | 47.98                        | 47.98                        | 47.98                        |
|                                                            | 11.91           | 11.60                         | 11.37                         | 11.37                        | 12.77                        | 11.70                        | 11.16                        | 11.38                        | 11.78                        | 11.71                        | 11.58                        | 11.56                        |
| Baseline log hh asset - value at 10%                        | 0.2970          | 0.2325                        | 0.1087                        | 0.8056                       | 0.2843                       | 0.5497                       | 0.4638                       | 0.9532                       | 0.6432                       | 0.5124                       | 0.3284                       | 0.7916                       |
| p values of overidentification test                         | 1.234           | 1.234                         | 1.234                         | 1.234                        | 1.234                        | 1.234                        | 1.234                        | 1.234                        | 1.234                        | 1.234                        | 1.234                        | 1.234                        |

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other controls include county dummies, village (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household characteristics (number of laborers, size of arable land, having any member migrated outside corresponding geographic scope) in baseline year. Migration share in person-days is defined as the share of days of all household member migrated for employed work during the year. Any household member migrated is a dummy indicating if any household member migrated outside certain geographic scope.
## Table 7: Robustness of Instrumental Variable Estimation

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Outside home village</th>
<th>Outside home township</th>
<th>Outside home county</th>
<th>Outside home province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Share of Migrants</td>
<td>2SLS estimation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Treat</td>
<td>0.23***</td>
<td>0.22***</td>
<td>0.18***</td>
<td>0.09</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>[3.528]</td>
<td>[3.653]</td>
<td>[2.664]</td>
<td>[1.331]</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>-0.14***</td>
<td>-0.11***</td>
<td>-0.08*</td>
<td>-0.04</td>
</tr>
<tr>
<td>(Change in log local wage - sample mean) × Treat</td>
<td>[2.396]</td>
<td>[2.179]</td>
<td>[1.771]</td>
<td>[0.428]</td>
</tr>
<tr>
<td>Change in log local wage - sample mean</td>
<td>0.20**</td>
<td>0.22**</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>F values of first stage</td>
<td>18.22</td>
<td>17.82</td>
<td>17.72</td>
<td>17.82</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>9.00</td>
<td>8.58</td>
<td>8.17</td>
<td>8.51</td>
</tr>
<tr>
<td>p values of overidentification test</td>
<td>0.0448</td>
<td>0.0476</td>
<td>0.0448</td>
<td>0.5053</td>
</tr>
<tr>
<td>Panel B: use rainfall shock in year 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treat</td>
<td>0.14**</td>
<td>0.22***</td>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>[2.014]</td>
<td>[2.980]</td>
<td>[1.423]</td>
<td>[0.204]</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>-0.06***</td>
<td>-0.11***</td>
<td>-0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Village migration network size at baseline</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
<td>-0.09</td>
</tr>
<tr>
<td>p values of overidentification test</td>
<td>0.3611</td>
<td>0.4664</td>
<td>0.1067</td>
<td>0.1972</td>
</tr>
<tr>
<td>Panel C: control for village level migration network size at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treat</td>
<td>0.17**</td>
<td>0.18**</td>
<td>0.18**</td>
<td>0.09</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>[2.154]</td>
<td>[2.330]</td>
<td>[2.193]</td>
<td>[1.016]</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>-0.10***</td>
<td>-0.08*</td>
<td>-0.08</td>
<td>-0.03</td>
</tr>
<tr>
<td>Village migration network size at baseline</td>
<td>0.14**</td>
<td>0.08</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>p values of overidentification test</td>
<td>0.4583</td>
<td>0.3965</td>
<td>0.1318</td>
<td>0.8829</td>
</tr>
<tr>
<td>Observations</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
</tr>
</tbody>
</table>

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other controls include county dummies, village (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household characteristics (number of laborers, size of arable land, having any member migrated outside corresponding geographic scope) in baseline year. The data of local market wage is collected in village questionnaire. The village level migration network size is the share of labor force in the village who migrated outside home county at baseline.
Table 8: Heterogeneous Impacts on Migration by Migration Prevalence (2SLS estimation)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>County level migration prevalence outside home township in 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low migration prevalence</td>
</tr>
<tr>
<td></td>
<td>Outside home village (1)</td>
</tr>
<tr>
<td><strong>Treat</strong></td>
<td>0.30***</td>
</tr>
<tr>
<td></td>
<td>[3.129]</td>
</tr>
<tr>
<td><strong>(Baseline log hh asset - value at 10%) × Treat</strong></td>
<td>-0.18***</td>
</tr>
<tr>
<td></td>
<td>[-3.417]</td>
</tr>
<tr>
<td><strong>Baseline log hh asset - value at 10%</strong></td>
<td>0.13**</td>
</tr>
<tr>
<td></td>
<td>[2.241]</td>
</tr>
</tbody>
</table>

F values of first stage regression on 
(Baseline log hh asset - value at 10%) × Treat
Baseline log hh asset - value at 10% | 24.55 | 23.33 | 22.87 | 23.10 | 48.01 | 48.60 | 45.69 | 44.50 |
| p values of overidentification test | 0.2040 | 0.2279 | 0.5215 | 0.4202 | 0.8353 | 0.7343 | 0.4982 | 0.8458 |
| Observations | 567 | 567 | 567 | 567 | 667 | 667 | 667 | 667 |

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other control variables include county dummies, village baseline characteristics (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household baseline characteristics (number of laborers, size of arable land, share of migrants outside corresponding geographic scope). County level migration prevalence is measured from matched Chinese census data in 2000.
## Table 9: Impacts on Income Generation

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Self-employed income</th>
<th>Employment income</th>
<th>Total income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in household per capita income</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td><strong>Average Treatment Impacts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treat</td>
<td>564</td>
<td>755***</td>
<td>1,319***</td>
</tr>
<tr>
<td></td>
<td>[1.349]</td>
<td>[3.186]</td>
<td>[2.703]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
</tr>
<tr>
<td><strong>Heterogeneous Treatment Impacts (IV estimations)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treat</td>
<td>5,186***</td>
<td>3,171***</td>
<td>8,553***</td>
</tr>
<tr>
<td></td>
<td>[2.823]</td>
<td>[3.889]</td>
<td>[3.767]</td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>-2,711**</td>
<td>-1,759***</td>
<td>-4,710***</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>6,403***</td>
<td>1,509**</td>
<td>8,369***</td>
</tr>
<tr>
<td></td>
<td>[4.229]</td>
<td>[2.048]</td>
<td>[4.622]</td>
</tr>
<tr>
<td><strong>F values of first stage regression on</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Baseline log hh asset - value at 10%) × Treat</td>
<td>29.50</td>
<td>32.71</td>
<td>27.75</td>
</tr>
<tr>
<td>Baseline log hh asset - value at 10%</td>
<td>7.26</td>
<td>8.09</td>
<td>6.70</td>
</tr>
<tr>
<td>p values of overidentification test</td>
<td>0.6116</td>
<td>0.2085</td>
<td>0.7498</td>
</tr>
<tr>
<td>Observations</td>
<td>1,234</td>
<td>1,234</td>
<td>1,234</td>
</tr>
</tbody>
</table>

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, **  p<0.05, * p<0.1. Other control variables include county dummies, village baseline characteristics (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household baseline characteristics (number of laborers, size of arable land, share of migrants outside corresponding geographic scope).
Table 10: Heterogeneous Impacts on Income by Migration Prevalence (2SLS estimation)

| Dependent variable: | County level migration prevalence outside home township in 2000 | Low migration prevalence | | | | High migration prevalence | | | |
|---------------------|---------------------------------------------------------------|-------------------------|----------------|----------------|----------------|-----------------|----------------|----------------|
| Change in household per capita income | | | | | | | | | |
| Self-employed income | Employment income | Total income | Self-employed income | Employment income | Total income | Self-employed income | Employment income | Total income |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Treat | 1,979 | 5,422*** | 8,200*** | 5,759*** | 320 | 6,126*** | | | |
| | [1.093] | [6.206] | [4.307] | [4.950] | [0.208] | [6.484] | | | |
| (Baseline log hh asset - value at 10%) × Treat | -1,096 | -2,188*** | -3,690*** | -3,207*** | 1,030 | -2,663 | | | |
| | [-1.200] | [-6.087] | [-4.113] | [-2.617] | [0.487] | [-1.530] | | | |
| Baseline log hh asset - value at 10% | 4,317*** | 2,060*** | 6,871*** | 4,090 | 2,182 | 5,814* | | | |
| | [4.777] | [3.503] | [5.438] | [1.477] | [0.892] | [1.732] | | | |
| F values of first stage regression on (Baseline log hh asset - value at 10%) × Treat | 16.64 | 18.93 | 16.33 | 89.53 | 74.93 | 80.62 | | | |
| Baseline log hh asset - value at 10% | 8.27 | 9.07 | 7.99 | 9.89 | 12.53 | 9.53 | | | |
| p values of overidentification test | 0.7209 | 0.0271 | 0.9072 | 0.3918 | 0.6235 | 0.6094 | | | |
| Observations | 567 | 567 | 567 | 667 | 667 | 667 | | | |

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other control variables include county dummies, village baseline characteristics (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household baseline characteristics (number of laborers, size of arable land, share of migrants outside corresponding geographic scope). County level migration prevalence is measured from matched Chinese census data in 2000.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pr (household attrited in the second wave of survey)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>[0.521]</td>
<td>[0.744]</td>
</tr>
<tr>
<td>Treatment × log household assets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1.393]</td>
<td></td>
</tr>
<tr>
<td>log household assets</td>
<td>-0.12***</td>
<td>-0.17***</td>
</tr>
<tr>
<td></td>
<td>[-2.872]</td>
<td>[-2.997]</td>
</tr>
<tr>
<td>Share of migrants outside home village</td>
<td>0.36**</td>
<td>0.35**</td>
</tr>
<tr>
<td></td>
<td>[2.163]</td>
<td>[2.100]</td>
</tr>
<tr>
<td>Number of Children in household</td>
<td>0.15***</td>
<td>0.15***</td>
</tr>
<tr>
<td></td>
<td>[3.067]</td>
<td>[3.047]</td>
</tr>
<tr>
<td>Number of labor force in household</td>
<td>-0.02</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>[-0.599]</td>
<td>[-0.613]</td>
</tr>
<tr>
<td>Number of elderly in household</td>
<td>-0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>[-0.713]</td>
<td>[-0.698]</td>
</tr>
<tr>
<td>Any loans during Jan 2009 to June 2010</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>[0.157]</td>
<td>[0.086]</td>
</tr>
<tr>
<td>log Amount of loans conditional on borrowing</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.345]</td>
<td>[-0.274]</td>
</tr>
<tr>
<td>Village rainfall shock in 2009</td>
<td>-0.17***</td>
<td>-0.18***</td>
</tr>
<tr>
<td></td>
<td>[-2.807]</td>
<td>[-2.908]</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.54***</td>
<td>-1.55***</td>
</tr>
<tr>
<td></td>
<td>[-9.032]</td>
<td>[-9.065]</td>
</tr>
<tr>
<td>Joint F-test on treatment and interaction terms (p values)</td>
<td></td>
<td>0.333</td>
</tr>
<tr>
<td>Observation</td>
<td>1500</td>
<td>1500</td>
</tr>
</tbody>
</table>

Notes: The sample is all the 1500 households surveyed at baseline. The outcome variable is a dummy indicating household was attrited in the follow-up survey. The log of household assets has been demeaned. The village rainfall shock in 2009 has been standardized by the standard deviation of rainfall shock across all villages in the regression sample. z-statistics in brackets. *** p<0.01, ** p<0.05. The last row reports the p-values of joint F-test on hypothesis that the coefficients of treatment and cross products of treatment and log household assets are equal to 0.
Table A2: Impacts of Rainfall Shock on Baseline Household Assets

<table>
<thead>
<tr>
<th>OLS Estimation</th>
<th>Baseline log household assets × Treat</th>
<th>Baseline log household assets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Treat</td>
<td>1.83***</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[7.927]</td>
<td>[0.053]</td>
</tr>
<tr>
<td>The instrumental variables (IV)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainfall shock × Treat</td>
<td>0.57***</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>[5.570]</td>
<td>[1.313]</td>
</tr>
<tr>
<td>Rainfall shock² × Treat</td>
<td>-0.12</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>[-1.577]</td>
<td>[-0.146]</td>
</tr>
<tr>
<td>Rainfall shock</td>
<td>-0.53**</td>
<td>-0.63*</td>
</tr>
<tr>
<td></td>
<td>[-2.225]</td>
<td>[-1.894]</td>
</tr>
<tr>
<td>Rainfall shock²</td>
<td>-0.25***</td>
<td>-0.56***</td>
</tr>
<tr>
<td></td>
<td>[-2.829]</td>
<td>[-4.504]</td>
</tr>
<tr>
<td>F values of joint significance tests on IVs</td>
<td>50.08</td>
<td>12.23</td>
</tr>
<tr>
<td>Observations</td>
<td>1,234</td>
<td>1,234</td>
</tr>
</tbody>
</table>

Notes: Robust t-statistics in brackets. The standard errors are clustered by village. *** p<0.01, ** p<0.05, * p<0.1. Other control variables include county dummies, village baseline characteristics (type of geographic features, population, area of arable land, funds from upper governments on various village projects) and household baseline characteristics (number of laborers, size of arable land, share of migrants outside home village). The instrument variables (IV) include rainfall shock, rainfall shock square, and their interaction with treatment. The rainfall shock has been standarized by standard deviation of rainfall across all villages in the sample. The illiquidity asset includes durables, production assets, and housing. Both grain stock and animal stock measure the values at the end of year 2009. The illiquidity assets measures the values at July 2010.