Hidden Impact? Household Saving in Response to a Poor-Area Development Project

Martin Ravallion and Shaohua Chen[1]

Development Research Group, World Bank, 1818 H Street NW, Washington DC

June 18, 2003; revised July, 2004

Abstract: The possibility that poor households may prefer to save the income gains from a development project raises concerns about how well standard evaluation methods — using data collected over relatively short periods — can capture the true welfare impacts. By the widely used difference-in-difference method, the Southwest China Poverty Reduction Project had little current impact on the proportion of people in beneficiary villages consuming less than $1/day — despite a public outlay of $400 million. However, the program had much larger impacts on incomes than consumptions. Uncertainty about the project's future impact probably made it hard for participants to infer the gain in permanent income, so they saved a high proportion of the current income gains.

Keywords: Savings, development projects, evaluation, poor areas, China

JEL: D91, H43, I32, O22

[1] The work reported in this paper would not have been possible without the excellent survey data collection done by the Rural Household Surveys Team of China's National Bureau of Statistics (NBS). The support of the World Bank's Research Committee (under RPO 681-39) is gratefully acknowledged. Helpful comments were received from Quy-Toan Do, Emanuela Galasso, Alan Piazza, Gene Tidrick and seminar participants at NBS, Beijing, the Department of Economics, MIT, Beijing University, the World Bank, University College London and the journal's editor, Thomas Piketty, and two anonymous referees. These are the views of the authors, and need not reflect those of the World Bank.

1. Introduction

It is often assumed that poor people tend to consume the current income gains from a successful development project. However, this need not be so. There is a large body of evidence consistent with the view that poor people think about the longer-term implications of their current consumption and savings choices given the uncertainties they face.[1] If the current income gains are known to be permanent, and markets work well, then the consumption gains would probably be revealed within the project cycle. However, if the income gains are seen to be transient then they may well be saved rather than consumed now. High savings from the current income gains might also arise from uncertainty about future income gains, or from positive program effects on the returns to saving, given market failures.

Such inter-temporal behavioral responses raise concerns about the validity of standard impact evaluation methods, which typically rely on data collected over a limited time — typically not much more than the disbursement period for project funding. High savings from the income gains will mean that the project's impacts on living standards are not evident within the evaluation period. By the same token, the impacts on current incomes may greatly overstate the long run welfare gains.

This paper studies savings behavior in response to a large poor-area development project in rural China. Using survey data collected for this purpose, changes in consumption and income within project villages are compared to those found in a set of comparable non-project villages. We assess impacts on mean income and consumption and on poverty incidence over a wide range of poverty lines. We find that a large share of the current income gain from the project was saved by the participants, which greatly attenuated the measured impacts on current living standards.

The following section describes the setting and program. In section 3 we turn to our data, while section 4 outlines our method for identifying impacts on income and consumption. Section 5 suggests some theoretical arguments as to why the income gains from a project might not be evident in current living standards. Section 6 then presents our empirical results and discusses their implications. Section 7 concludes.

2. Poor-area programs in rural China

It is widely acknowledged that many inland rural areas have been lagging in China's overall success against poverty over the last two decades (Ravallion and Chen, 2004). Wide geographic disparities in living standards have emerged, notably between the coast and remote resource-deficient inland areas (see, for example, Jian et al., 1996; Khan and Riskin, 1998; World Bank, 1992,1997). Partly in response to this problem, anti-poverty policies in China have emphasized poor-area development (World Bank, 1992, 1997). Local infrastructure is improved and credit is provided for private (farm and non-farm) investments.

There is evidence that these programs have been reaching poor rural areas within rural southwest China. Using survey and administrative data for 1985-90, Jalan and Ravallion (1998) show that the counties chosen tended to be poorer — by a wide range of criteria — than those not picked.[2] At the same time, there are signs of unconditional (absolute and relative) divergence over time between the counties covered by the program and those not (Ravallion and Jalan, 1996). In the five years after these programs began, average consumption growth rates in the counties covered in this region of China were actually lower than growth rates in the areas not covered (Jalan and Ravallion, 1998).

However, a bias in the impact estimates for such poor-area programs can be expected if one simply compares growth rates in areas targeted by the program and those not, given that

whether or not an area is targeted depends on differences in local characteristics that are also likely to influence growth prospects (Ravallion, 1998). On controlling for geographic heterogeneity in a micro consumption growth model, Jalan and Ravallion (1998) find that households living in areas targeted by the program had higher consumption growth than one would have expected. The gains from the program were enough to prevent absolute decline. But they were not enough to reverse the underlying divergent tendencies in the rural economy. Significant impacts on average incomes from China's poor-area development programs were also found by Park et al (2002), using income growth regressions on county data over all of China. However, Park et al. found a diminished impact from the programs in the 1990s relative to the 1980s.

A substantial increase in external aid for poor-area development in China began in 1995 with the World Bank's Southwest Poverty Reduction Project (SWPRP). This aimed to reduce poverty by augmenting the private and (local) public capital stock of farm-households in poor areas. The program was targeted to poor villages within 35 designated "national poor" counties in Guangxi, Guizhou and Yunan. The SWPRP involved an investment of about $US400 million over 1995-2001 from both a World Bank loan and counterpart funding from the Central and Provincial Governments. (Local governments were not required to contribute financially.) As in other development projects financed by the Bank, there were numerous appraisal and supervision missions by Bank staff and consultants, and these missions often probed quite deeply into the project's local operations, including numerous visits to participating poor counties and villages. Both authors participated in some of these missions and worked closely with staff of the National Bureau of Statistics (NBS) on the design of the survey data collection done for the purpose of evaluating SWPRP.

The SWPRP comprised a range of income-generating activities including methods for raising grain yields, animal husbandry, and reforestation. There was also a component for off-farm employment, including voluntary rural labor mobility and support for township-village enterprises. The SWPRP also included local social services and rural infrastructure initiatives, including tuition assistance to children from poor families, upgrading village school and health clinics, the construction of rural roads and piped water supply systems. Table 1 gives the breakdown of total project investment by category. In common with other development projects, the SWPRP provided the capital and technical assistance, but it did not provide insurance, and many of the project activities are likely to entail non-negligible income risk. The income gains will depend on a number of contingencies, including the vagaries of the weather (given the evident importance of agriculture in the breakdown in Table 1), uncertain demand for the new products and risks associated with out migration.

The selection of sub-projects aimed to take account of local conditions and the expressed preferences of participants and local stakeholders. How much participation by the poor there was in practice is a moot point. We discussed this with participants, and with the sociologist responsible for assessing the extent of beneficiary participation during supervision missions; it was clear that the record was mixed, varying from village to village, and county to county.

Whether in fact the resources transferred to participants actually financed the identified project is also unclear. To some degree, all external aid is fungible. Yes, it could be verified in supervision that the proposed sub-project was actually completed. But one cannot rule out the possibility that it would have been done otherwise. Participants and local leaders would naturally have put forward the best development option they saw, even if it was something they planned to do anyway with the resources already available. Then there is some other (infra-marginal)

expenditure that was really being financed by the aid. Similarly, there is no way of ruling out the possibility that non-project villages benefited by a re-assignment of public spending by local authorities, thus lowering the differential impact of program participation.

3. Data

A baseline survey in 1995 was followed by five annual surveys over 1996-2000. All surveys were done by the Rural Household Survey (RHS) team of NBS. The sample size for the annual surveys was 2000 households spanning 20 project counties and 200 villages. (Notice that our sampled non-project villages also come from project counties; we return to this feature of the design in the next section.) It was originally intended to have 100 villages in each of the project and non-project townships within the project counties. However, the assignment of project villages had not been finalized at the time the samples of villages were drawn, and it turned out that 13 of the originally sampled non-project villages did in fact get the project.[3] So we end up with 113 project villages and 87 non-project villages in the same counties. 10 randomly sampled households were interviewed in each village (project and non-project).

There is a comparability problem between the 1995 survey and the subsequent surveys. Because of delays in the statistics office obtaining the locations of project villages, the first survey in December 1995 had little choice but to use a one-time interview method, asking for recall over the full year. The use of this long recall period is likely to lead to underestimation of income and consumption, though this is of less concern for the village-level characteristics to be used for matching. The subsequent surveys use the daily diary method over the full year and collect much more accurate income and consumption data. As a consequence, the rates of income and consumption growth are very likely to be overestimated using 1995 as the baseline. It is unclear how this would affect the difference-in-difference estimates. An option is to rely

mainly on the 1996 survey as the baseline, though this is not free of contamination by the project; 16% of the program's total disbursement on projects at household level had been made by the middle of 1996, and 23% had been made by the end of 1996.

The surveys from 1996 onwards were closely modeled on NBS's Rural Household Survey, which is described in detail in Chen and Ravallion (1996). This is a good quality budget and income survey, notable in the care that goes into reducing both sampling and non-sampling errors. Sampled households maintain a daily record on all transactions plus log books on production. Local interviewing assistants (resident in the sampled village, or nearby) visit each household at two-three weekly intervals. Inconsistencies found at the local (county-level) NBS office are checked with the respondents. The sample frame is all registered agricultural households.

The consumption expenditure aggregate we use is what is referred to as "living expenditures" in the RHS. This comprises cash spending on all goods and services (both durables and non-durables) and imputed values of in-kind spending, which is mainly consumption from household production (farming, forestry, animal husbandry, handicrafts etc.). Consumption of own-farm output is valued at local selling prices. Living expenditures exclude expenditures on production inputs, which are accounted for separately in estimating net income from own-production activities. It also excludes transfer payments (cash or imputed values of transfers to relatives living in urban areas, interest and insurance payments, fines, transaction costs in acquiring assets or changing land-usage), though these only account for a small share of total spending (3.7% over the whole sample in 1996). The income aggregate includes cash income from all sources and imputed values for in-kind income consistent with the methods used for consumption. Note that the income aggregate includes remittances received from family

members who migrate, including those supported by the SWPRP. (The migrant workers were not themselves tracked.)

In principle, savings can take a variety of forms in this setting, including higher money balances and investment in production activities. The survey was not designed to allow a complete independent accounting of all forms of saving. Some data were collected on assets and liabilities though their reliability is questionable. Simply subtracting consumption from income is clearly the best means of estimating savings from these data and given the seemingly high quality of the consumption and income data we expect that aggregate saving measured this way will be reasonably accurate.

4. Identification strategy

The standard difference-in-difference (DD) method compares changes in measured outcomes between a treatment group and a comparison group of non-participants. In this context, we point to two potentially important sources of bias in this method. The first relates to a possible source of interference between the treatment and comparison groups. Our "non-participant" villages did not receive the program but are located in a county that did. From our field work and discussions with NBS and project staff, we came to the conclusion that the physical distances involved would not mean that geographic proximity is a source of contamination. However, sharing a common local government could be a more serious problem. Since all project counties are automatically amongst China's nationally-designated "poor counties" they are covered by the Government's own national poor-area program. This is needed to assure that the comparison of income and consumption gains between project and non-project villages can reveal the impact of the Bank's program on top of the government's program. However, this is not as clean an identification strategy as it might seem at first glance.

The fact that the project and non-project villages come from the same counties covered under other programs could generate a downward bias in our estimated impacts. This will happen if SWPRP displaced other programs in the project villages, to the benefit of the non-project villages in national poor-counties.

There is a second source of bias. As already noted, DD will give a biased impact estimate if the subsequent outcome changes are a function of initial conditions that also influenced the assignment of the sample between the two groups. This is known to be a serious concern in this context, based on past research on poor area programs in the same region of rural China (Jalan and Ravallion, 1998). To heal address this problem we use a flexible, largely non-parametric, method of controlling for initial heterogeneity, based on the propensity-score matching (PSM) method of Rosenbaum and Rubin (1983).[4] Single-difference PSM gives unbiased impact estimates as long as one controls fully for the factors that jointly influence project placement and outcomes, leaving no selection bias due to latent heterogeneity. While we implement PSM on a rich data set of village attributes, selection bias cannot be ruled out.

To outline the matching method in more formal terms, let D_i be a dummy variable taking the value unity for any participating village and zero for nonparticipants. Let $P(X_i) = \Pr(D_i = 1 | X_i)$ denote the propensity score, giving the probability of participation for unit *i* conditional on a vector X_i of pre-exposure control variables. Rosenbaum and Rubin (1983) prove that if the D_i's are independent over all *i*, and outcomes are independent of participation given X_i (i.e. unobserved differences do not influence whether or not *i* participates) then outcomes are also independent of participation given $P(X_i)$, just as they would be if participation was assigned randomly. PSM uses $P(X_i)$ to select comparison subjects for each of those treated. In effect, the Rosenbaum-Rubin result establishes that if no selection bias

remains when controlling for X_i then no bias remains when controlling solely for $P(X_i)$. We follow common practice in the matching literature of using a parametric binary response model to estimate the propensity score for each observation in the participant and the comparison-group samples. The comparisons are then constrained to assure common support, i.e., that project and non-project villages share sufficiently similar values of their propensity scores. We let Ψ denote the set of treatment observations within the region of common support.

Some treatment villages will have to be dropped for lack of sufficiently similar comparators. Given that this is more likely at high propensity scores, there must be a strong presumption that Ψ is not a random sub-sample of the original sample of treatment units. This points to a potential trade off between two possible sources of bias in the resulting impact estimates. On the one hand, there is the aforementioned need to assure comparability in terms of initial characteristics, to reduce bias in the difference-in-difference estimator, given that growth rates could well depend on initial conditions. This speaks to the importance of common support. On the other hand, imposing common support creates a new possibility of sampling bias in inferences about impact on the population of treated villages, to the extent that we lose treatment villages in achieving common support.[5] Recognizing this trade-off, we also present our estimates only eliminating non-participating villages that are outside the propensity-score range found for treatment villages, while retaining the original sample of treatment villages. For comparison purposes, we also present estimates without matching.

To test the possibility that the true impact is being hidden by inter-temporal behavioral responses through savings behavior, we shall assess impacts on the time profiles of both incomes and consumptions. We measure savings by difference between the two. We can write the

outcome measures for income (Y_{it}) and consumption (C_{it}) of the *i*'th treatment household ($D_i = 1$) at date *t* as:

$$(Y_{it} \mid D_i = 1) = Y_{it}^{*} + G_{it}^{Y} + \varepsilon_{it}^{Y} \quad (i = 1,..,n; t = 0,..,T) \tag{1.1}$$

$$(C_{it} \mid D_i = 1) = C_{it}^{*} + G_{it}^{C} + \varepsilon_{it}^{C} \quad (i = 1,..,n; t = 0,..,T) \tag{1.2}$$

where Y_{it}^{*} and C_{it}^{*} are the counter-factual income and consumption for treatment household *i* if the program had not existed, G_{it}^{Y} and G_{it}^{C} are the corresponding gains attributable to the project and ε_{it}^{Y} and ε_{it}^{C} are zero-mean innovation error terms uncorrelated with program participation, to allow for measurement error in Y_{it} and C_{it}.

Indicators of the counter-factual are available from a comparison group and are given by $\hat{Y}_{it}^{*}$ and $\hat{C}_{it}^{*}$. We recognize that these may be noisy indicators due to miss-matching (selection bias) arising from latent heterogeneity. We make the standard assumption that the selection bias is separable and time invariant, and so it is swept away by taking differences over time. Then, on taking the expectation over all participants in the region of common support, the mean differences-in-differences for income and consumption are:[6]

$$E[(Y_{it} - \hat{Y}_{it}^{*}) - (Y_{i0} - \hat{Y}_{i0}^{*}) \mid D_i = 1, i \in \Psi] = E(G_{it}^{Y} - G_{i0}^{Y} \mid D_i = 1, i \in \Psi) \tag{2.1}$$

$$E[(C_{it} - \hat{C}_{it}^{*}) - (C_{i0} - \hat{C}_{i0}^{*}) \mid D_i = 1, i \in \Psi] = E(G_{it}^{C} - G_{i0}^{C} \mid D_i = 1, i \in \Psi) \tag{2.2}$$

When period 0 is a genuine baseline prior to the intervention (and not in any way contaminated by the program assignment) we have $G_{i0}^{Y} = G_{i0}^{C} = 0$. Then DD estimates the mean current gains in consumption and income for program participants (often referred to as the "treatment effect on

the treated" in the evaluation literature). We will consider the implications for our results of the possibility that $G_{i0}^{Y} = G_{i0}^{C} \neq 0$.

The above exposition has focused on impacts on mean income and consumption. Given the explicit aims of the SWPRP, we also want to know the impacts on poverty. For this purpose, we can reinterpret Y_{it} as an indicator that takes the value 1 if household *i* has an income per person below the poverty line at date *t* and 0 otherwise, and similarly for C_{it}. The means of these indicators then give the incidence of poverty in terms of income and consumption respectively. We allow a wide range of poverty lines.

5. Saving out of the income gains from a development project

By separately estimating the income and consumption gains, the above formulation of the evaluation problem allows for saving out of the current income gains. Before turning to the empirical results it is of interest to ask: why might the income gains be saved?

To begin with a simple benchmark model, consider Friedman's (1957) Permanent Income Hypothesis (PIH). This assumes that consumption is directly proportional to permanent income (the annuity value of life-time wealth). In our case, permanent income has a counter-factual component (in the absence of the program) and a component due to the program (which is zero in the absence of the program). Let G_{it}^{YP} denote the contribution of the program to permanent income while G_{it}^{YT} is a transient component, such that the full impact on income can be written as:

$$G_{it}^{Y} = G_{it}^{YP} + G_{it}^{YT} \qquad (3)$$

The counter-factual is independent of participation and we assume that this is also true of any measurement error or transient consumption.

We focus initially on the special case in which there is no saving from permanent income. Thus we have the following model for consumption with and without the program:

$$(C_{it} \mid D_i = 1) = Y_{it}^{*P} + G_{it}^{YP} + \nu_{it} \tag{4.1}$$

$$(C_{it} \mid D_i = 0) = Y_{it}^{*P} + \nu_{it} \tag{4.2}$$

in which Y_{it}^{*P} is the counterfactual permanent income and we allow for a zero-mean innovation error term in measured consumption, ν_{it}. Comparing (4.1) with (1.2) it is plain that $G_{it}^{YP} = G_{it}^{C} + C_{it}^{*} + \varepsilon_{it}^{C} - Y_{it}^{*P} - \nu_{it} = G_{it}^{C}$ since $(C_{it} \mid D_i = 0) = Y_{it}^{*P} + \nu_{it} = C_{it}^{*} + \varepsilon_{it}^{C}$. Thus, the current consumption gain attributable to the program identifies the impact on permanent income and the extent of saving from the income gain ($G_{it}^{Y} - G_{it}^{C}$) measures the program's transient income gain for participants.

This benchmark model offers a clear explanation for saving from the program's income gains: the saved income gains are perceived to be transient by participants. However, this model makes a number of strong assumptions, most notably that permanent income is entirely consumed, there are no constraints on borrowing and there are no transaction costs or sources of lumpiness in consumption.[7] As the following discussion will illustrate, more general models suggest other reasons why the current income gains from a development project might be saved.

One reason is uncertainty about how much of the income gain is in fact permanent. There may be concerns about the longer-term sustainability of the income gains due to the project, such as arising from uncertainty about future output prices. Participants may then save as a hedge against the income uncertainty. This will happen (even without borrowing constraints) if the marginal utility of consumption is a convex function of consumption. By Jensen's inequality, a mean-preserving increase in uncertainty about future incomes will then increase the marginal

utility of future consumption, leading to higher savings (Gersovitz, 1988). There is evidence of such precautionary saving in the same setting as the SWPRP (Jalan and Ravallion, 2001). Against this effect of uncertainty, a positive impact on mean income will tend to reduce precautionary savings. The outcome is ambiguous on theoretical grounds.

Introducing borrowing constraints into the PIH can also generate savings from permanent income gains. The PIH assumes perfect credit and risk markets, which does not appear to be realistic.[8] Assume instead that households can save but not borrow. The anticipation of future borrowing constraints when negative income shocks are experienced may then lead program participants to save from an increase in permanent income, as a contribution to their buffer stock. Nonconvexities in consumption could also distort the empirical relationship between the permanent income gains and changes in current consumption. The nonconvexity can stem from lumpiness in certain expenditures (consumer durables or certain production inputs), given borrowing constraints. Small income gains will be saved to overcome the constraint.

There is another way that market failures can lead participants to save the income gains. Suppose that the project's investments raise the marginal product of private capital — that the program inputs are cooperant with private capital in production — and that private capital is geographically immobile, so that the marginal product of capital is equalized with a local rate of interest, that varies geographically. (This model is outlined in more formal terms in Jalan and Ravallion, 2002, who find supportive evidence for this region of rural China.) Under these conditions, the program can induce higher saving through its effect on the marginal product of private capital. This can happen even if there is no concern about longer-term sustainability.

All these modifications to the PIH will tend to create lags between the program's initial income gains and the impacts on consumption. Higher living standards might not then be

evident until after SWPRP's completion. By tracking annual income and consumption gains over time we can look for signs of lagged impacts on consumption.

The political economy of a local development project might also generate low impacts on living standards despite the income gains. This will happen if the direct income gains are somehow expropriated by higher-level (county or provincial) authorities and diverted to other uses, possibly benefiting non-project villages. Recall that our consumption aggregates exclude transfer payments. We will check if transfers responded positively to the project, consistent with some form of expropriation. The dynamics of income and consumption impacts will also offer clues as to the plausibility of this political economy explanation. If the local income gains were being siphoned off by a higher level of government then one would expect to see little sign of lagged consumption gains after an income gain due to the project. An expropriation model would also lead one to expect declining income gains, through disincentive effects of the expropriation. We will look for these features in the income profile over time of consumption and income gains attributed to SWPRP.

6. Results

Table 2 gives sample means by year, including the baseline year 1995. Project villages started worse off on average than non-project villages, in terms of both income and consumption. The growth rates in both mean consumption and mean income are higher in the project villages, though the difference is greater for income. Indeed, relative to 1995, mean income grew by 47% in project villages, versus 27% in non-project villages; for consumption, the five-year growth rates are 19% and 21% for project and non-project villages respectively. Using 1996 as the "baseline," mean income grew by 27% in the project villages, versus 6% in non-project villages; for consumption, the four-year growth rates are 12% and 8% for project and non-project

villages respectively. By the end of the period, the project villages had caught up in mean income, but not consumption.

The results of Table 2 indicate unusually high growth rates between 1995 and 1996. In that one year, mean income grew by 16% in the project villages and 19% in the non-project villages. However, as we noted in section 3, there is a comparability problem between 1995 and the subsequent surveys, and the problem is likely to lead to an over-estimation of the growth rates in the first year. It is unclear on *a priori* grounds how this would affect the difference-in-difference estimates of impact. The proportionate impact appears to be slightly higher in the non-project villages, and the simple (un-matched) DD estimates indicate a negative impact on incomes and consumptions in the first year of the project. (The DD estimates for 1996 implied by Table 2 are –49.53 Yuan per capita for income and –46.90 for consumption.) This seems implausible, though we cannot explain why the survey comparability problem would have had greater impact on the recorded income and consumption gains in non-project villages. Nor do these problems with the 1995 survey round wash out by the end of the study period. The DD estimate for 2000 indicates an income impact of 147.45 Yuan but a negative impact on consumption of -24.59 Yuan.

In the light of these concerns about using the survey round for 1995 as the baseline, the rest of this analysis we will follow the option discussed in section 3 of treating 1996 as the baseline.

6.1 Matching methods

To estimate the propensity scores, the sampled project and non-project villages are pooled and we run a probit regression for the village assignment to these two groups. We

include as explanatory variables virtually all the village level variables for 1995 that could be constructed from the data set. Table 3 gives the results.

We find a number of significant covariates of program participation. SWPRP villages tend to be in more mountainous remote areas, are less likely to have electricity, less likely to have a school in the village or nearby, though more likely to have a health clinic within the village relative to nearby. (Remote villages may well be more likely to have a very basic health clinic, to compensate for the inaccessibility to more comprehensive township facilities.) The project villages also tend to have higher populations, with lower mean income and more land per capita. The latter characteristic probably reflects lower population density and lower land quality in the project villages. In most respects, the results of Table 3 suggest that the project villages tend to be poorer than other villages within the project counties.

Figure 1 gives the frequency distribution of the propensity scores based on Table 3 for project and non-project villages. It can be seen that there are regions of non-overlapping support. We consider two methods of redefining the treatment and comparison groups to better balance the observable characteristics in the baseline. In the first, we keep the treatment group intact but we trim the comparison group to assure that they are all within the region of common support; we refer to this as the "trimmed comparison group". By this method, 16 non-project villages are dropped (see Figure 1). In the second method, comparisons are only permitted if the absolute difference in propensity scores is within pre-determined caliper bounds; we call this "caliper-bound matching." Project and non-project villages outside the caliper bounds are discarded. This method clearly gives the closest matching of treatment and control villages, but it can do so at a cost to sample size and representativeness. We set the tolerance levels for the caliper at 0.01. The choice of this tolerance is somewhat arbitrary. However, we felt that too

many villages were lost when the tolerance went much below 0.01. If one was relying on single difference matching then one would probably want closer matches than our 0.01 absolute difference in scores. However, here we can exploit the fact that we have multiple observations to "difference-out" any (time-invariant) errors due to miss-matching. With 0.01 tolerance level, we end up with only 63 of the original sample of project villages to be matched with 34 non-project villages.

6.2 Impact estimates for mean income and consumption

Table 4 gives the unmatched DD estimates of mean impacts based on the original samples of treatment and comparison villages. We give the annual impacts, the two-year moving average of the annual impact and the cumulative impacts. Table 5 gives the results when we better balance the observable characteristics of the two groups using the two matching methods described above.

Let us focus first on the results for the final year of the study period, 2000. While we find sizeable income gains over time in the project villages, this is not evident for the counter-factual comparison group. Taking account of both the changes over time and the differences between the treatment and comparison villages, the estimated double difference for 2000 indicates an income gain attributable to the project of around 17-21% of initial mean income (depending on the matching method).[9] However, we find little or no impact on consumption; indeed, we cannot reasonably reject the null hypothesis that the consumption impact over the whole period is zero. The vast bulk of the income gain in 2000 was saved.

Recall that we are measuring consumption by "living expenditures" in the RHS. So our definition of "savings" implicitly includes transfer payments. One can question whether some of these transfer payments should be included as savings. However, transfer payments do not

account for the high savings out of the project's income gains. Indeed, mean transfer payment actually fell slightly in the project villages over 1996-2000, and we found that the DD estimate was negative though not significantly so.

As noted in section 2, there are likely to have been impacts in 1996. On the assumption that these gains would have initially impacted on incomes rather than consumptions, we will have underestimated the true income impact and underestimated the extent of saving from the current income gains. As we will see below, the inter-temporal pattern of income and consumption impacts within the evaluation period offers support for this conclusion.

We have seen that the results for 2000 suggest that virtually all of the aggregate income gain was saved. Let us now turn to the results for the three intervening years, 1997-99, as also given in Tables 4 and 5. We will focus on the results using the trimmed comparison group, noting any marked differences with the results for unmatched DD and for caliper-bound matching.

Mean income was higher in all years due to the project and significantly so in all years except 1999. The gains were lower in the second and third years than the first and last. Despite the large income gain in the first year, there was negligible impact on consumption in that year. Appreciably higher consumption only emerged in the second year (1998). The relatively low income gain in 1999 was followed by a lower impact on consumption in 2000. By the end of the study period, 50% of the cumulative income gain attributed to the project had been saved. Caliper-bound matching gives an even higher savings rate, of 58%.

While one should be cautious with only four annual observations, there is a pattern in Tables 4 and 5 that is suggestive of lagged consumption impacts from the project's income gains. By this interpretation the high consumption in 1998 reflected in part the income gains in the

previous year, on top of those in 1998. The high income gains in 2000 may then be expected to be reflected in higher future consumption, beyond the study period. Neither the signs of lagged consumption impacts nor the fact that the highest income gains were in the last year are supportive of the existence of some hidden form of expropriation of the project's income gains.

Comparing the three evaluation methods, the most notable difference is that caliper-bound matching tends to give lower impact estimates than the other two methods. This is not consistent with the expectation discussed in section 2 that the relatively poorer villages targeted by such a program would tend to have intrinsically lower growth prospects; if anything we find the opposite, though the difference is small. However, it should be recalled that our comparison villages were chosen from the same (poor) counties as the project villages. The bias in unmatched comparisons might well only emerge when making comparisons across project and non-project counties, given that there can be large inter-county differences in initial conditions relevant to growth prospects (Jalan and Ravallion, 1998).

6.3 Impacts on poverty

Given that the project's main aim was poverty reduction, it is also of interest to calculate the impact on poverty incidence in the final year of the study period. We use probably the most common measure of absolute poverty in developing countries, namely the proportion of the population living in households with consumption per person below the international poverty line of $1.08 per day at 1993 Purchasing Power Parity (Chen and Ravallion, 2001); this is equivalent to 808 Yuan per year per person at 1995 prices.

Table 6 gives the results. We find reductions in the incidence of poverty due to the program, though the magnitude varies by matching method. The biggest difference is not between the unmatched DD and matched DD, but rather between the two methods of matching.

The unmatched DD estimate and that using the trimmed comparison group indicate that the poverty rate by the end of the study period had fallen by 5-6 percentage points due to the project. However, using tighter matching according to caliper-bounds, we find no impact on poverty.

Widely used poverty lines for China are lower than $1 a day. To test robustness to the choice of poverty line, Figure 2 gives our estimate of impact over the whole distribution. The figure gives the difference between the empirical cumulative distribution function of consumption for the treatment villages and the counter-factual comparison group. (The results are similar for unmatched DD as for the trimmed comparison group, so we only give results for matched DD to make the figure easier to read.) For caliper-bound matching we find that the negligible poverty impact for the $1/day line is not robust to the choice of poverty line, with more sizable impacts emerging amongst the poorest and least poor in the project villages. (The impacts become statistically significant at about 6 percentage points.) The lack of robustness in the caliper-bound estimates may well reflect the smaller sample sizes.

To see the impact of this high savings rate on the poverty measures, we re-calculated the DD estimates using incomes. For the unmatched DD and the matched DD using the timed comparison group, the impacts on income poverty were 11.5% points (t-ratio = -4.03) and 11.3% (t = -3.65) respectively (instead of 5.0 and 6.3% for consumption poverty). The impact is greater using caliper-bound matching; instead of the very small 0.6 percentage point impact on consumption poverty using the $/day line by the caliper-bound matching, we find that the income poverty rate fell by 15.7% points (t = -4.41). Figure 3 gives the impacts on income poverty over the whole distribution. Comparing Figures 2 and 3 it is evident that the largest divergence between the income and consumption impacts tends to be in the middle of the distribution.

6.4 Rates of return

Our estimated income gains from the SWPRP can be interpreted as the output returns from the project's investments within the disbursement period. From the project documents we calculated the total investment (by the Bank and Government) by the end of the project was 1120 Yuan per person per year in 1995 prices, averaged over the population of project villages.[10] Using the annual average income gains over the project disbursement period from Tables 5 and 6, the implied rate of return is 9.8% without matching and 9.2% and 8.6% for outer-support matching and caliper-bound matching respectively. These could be underestimated, to the extent that the Bank's program displaced other programs in the project villages, to the benefit of the non-project villages. (Recalling that project and non-project villages come from national-poor areas covered under other poor-area programs.)

Note that the fact that the project and comparison villages were drawn from the same national-poor counties covered by the Government's pre-exiting programs means that these rates of return should be interpreted as incremental returns from the Bank's program on top of the Government's programs. Jalan and Ravallion (1998) estimated an average rate of return of 12% for the Government's poor area development program in the same region of China over 1985-90. Using different methods, Park et al., (2002) also estimate a rate of return to the Government's national poor-area program of 12% in the period 1992-95.[11] So our results suggest that the compound rate of return from the SWPRP and the Government's own program is 22-23%.

6.4 Variability in returns

The income gains from the program are clearly more variable over time than other income sources. From Table 2, the range of annual mean incomes is about 24% of the overall mean in the project villages while the range of the project's income impacts is over 100% of the

mean impact.[12] And this difference appears to be reflected in the savings rates. The baseline data indicate that an average of 16-17% of income was saved in the project villages (Table 2). By contrast, we find that the average saving rate from the income gains during the life of the project was 50%.

With such variability in the income gains from the project, one can conjecture that project participants would have had a hard time inferring the project's impact on permanent income. This is consistent with the argument that the high saving rate out of the income gains that we find reflects transience or uncertainty in the project's income gains. Furthermore, none of the other possible explanations for high saving from the project's income gains appear to be as plausible in the light of our empirical findings. Explanations that posit that the project increased the returns to saving (to overcome borrowing constraints) would appear to have a hard time explaining the variability over time that we find in the savings rate from the project's income gains. The facts that the high aggregate savings rate is not attributable to measured transfer payments, and that income gains do not fall over time, are not supportive of the expropriation model discussed in section 5.

The variability in returns has implications for the design of evaluations. It is common for evaluation designs to only have one follow-up survey near the end of the project. Such designs can clearly be deceptive. Suppose for example that the design had relied on only two surveys, one in 1996 (just after the project began) and one in 1999 (just before it finished). This evaluation design would have considerably underestimated the average annual income gain from the project, and overestimated the consumption gain, given the time path of the underlying income gains. Or suppose that one only knew the income gains in the last year (as given in

Table 5). One would then conclude that the rate of return was around 18% rather than 9-10% based on the mean annual income gains.

7. Conclusions

We have studied the impacts on both income and consumption of a rural development project in China over the bulk of its disbursement cycle. On comparing income changes in project villages with those in matched non-project villages, we find that the project resulted in an average income gain over five years of around 10% of baseline mean income, representing an average return on the project's disbursements of about 9-10%, on top of the impact of the Government's pre-existing assistance to poor areas.

However, we find that half of the cumulative income gain was saved, so that the project's impact is far less evident in participants' consumptions. Indeed, on comparing the final year of the study period with the first, we find little or no impact on mean consumption or on consumption poverty using an international "$/day" poverty line, though the poverty impact depends critically on the poverty line used; there are indications of significant impacts on consumption poverty for lower poverty lines.

We find large year-to-year differences in impact. For example, the estimated income gain in the final year was 23% of baseline income (an 18% return on the project's total disbursement) and virtually all of this was saved. The impact variability was primarily due to variability in the annual returns to the program's investments rather than the level of that investment.

Our results reject the commonly held view that poor people tend to rapidly consume the income gains from a development project. Indeed, we find a high saving rate. When interpreted in terms of the Permanent Income Hypothesis, our results imply that participants felt that a large share of the income gains was likely to be transient. Uncertainty about future incomes and future

borrowing possibilities can also lead to high saving out of the income gains from such a program. The considerable variability that we find in the programs' income returns suggests that participants would have had a hard time assessing the program's impact on permanent income.

Finding that even poor participants choose to save a large share of the current income gains from external aid has an important implication for assessments of the efficacy of anti-poverty programs, given their finite time horizons and that it is common to study poverty impacts within a relatively short period of time — often no more than the period of the disbursement cycle. A large share of the impact on peoples' living standards may occur beyond the life of the project. This does not necessarily mean that credible evaluations will need to track welfare impacts over much longer periods than is typically the case, raising concerns about feasibility. But it does suggest that evaluations need to look carefully at impacts on partial intermediate indicators of longer-term impacts — such as incomes in our case — even when good measures of the welfare objective — consumption in our case — are available within the project cycle. The choice of such indicators will need to be informed by an understanding of participants' behavioral responses to the program.

Table 1: Composition of spending under SWPRP

	% of total investment
Education	8.60
Health	5.37
Labor mobility	9.74
Rural infrastructure	17.24
Agriculture	43.05
Rural enterprise development	11.52
Institution building	1.69
Project and poverty monitoring	2.78
Total	100.00

Table 2: Mean household income and consumption per capita by year

		Project villages		Non-project villages	
		Mean	Std. dev.	Mean	Std. dev.
1995	Income	854.56	458.95	967.76	528.92
	Consumption	789.24	388.34	844.87	425.08
1996	Income	992.74	713.47	1155.47	603.45
	Consumption	841.13	468.63	943.66	444.38
1997	Income	1084.86	658.14	1148.86	628.80
	Consumption	874.72	441.08	954.57	512.99
1998	Income	1108.91	603.27	1189.28	680.96
	Consumption	937.01	541.27	951.11	497.81
1999	Income	1182.23	681.62	1285.25	807.03
	Consumption	1002.91	658.89	1050.27	591.22
2000	Income	1259.47	913.70	1225.22	669.92
	Consumption	943.09	579.15	1023.31	696.10

Note: Household-size weighted means in Yuan per year at 1995 prices using Provincial Rural CPI. Sample sizes: 1130 households in project villages and 870 households in non-project villages (10 households per village in both cases).

Table 3: Probit regression of village participation in the SWPRP

	Coefficient	Z score
Village on the plains	*Reference category*	
Hills	4.6023	2.651
Mountainous	2.6301	1.616
Whether village has electricity	-0.8272	-1.722
... telephones	-0.1088	-0.248
... road passing through it	0.4085	0.971
... radio transmitters	0.4683	0.972
Whether village can receive TV transmission	0.2141	0.531
Located <5 km from the nearest market	0.3084	0.364
...5 -10 km from the nearest market	-0.3476	-0.406
...10 –20 km from the nearest market	1.1554	1.167
…> 20 km	*Reference category*	
# of days in a cycle during which the market assembles	-0.0888	-0.662
County town within 5 km	*Reference category*	
Distance from village to county town is 5-10 km	1.1096	1.230
...10-20 km	-0.6387	-0.842
...>20 km	-0.4168	-0.596
Township=village	*Reference category*	
Distance from village to township is within 5 km	0.5466	0.609
...5 –10 km	0.7836	0.877
...10-20 km	-1.0477	-1.141
Main mode of transportation used by the villager: bicycle	-0.5539	-1.026
...bus	-0.1329	-0.415
...other automobile	0.6948	1.440
…walking	*Reference category*	
Nearest train station is within 5 km	-0.1729	-0.192
...5-10 km	1.1186	1.137
...10-20 km	0.4978	0.429
…>20 km	*Reference category*	
Nearest bus station is within 5 km	-0.0173	-0.050
...5-10 km	0.2013	0.432
...10-20 km	0.3736	0.718
…> 20 km	*Reference category*	
Whether village has a day-care center	0.5773	0.848
Elementary school is in village	*Reference category*	
Nearest elementary school is within 5 km	0.0520	0.128
...5-10 km	0.5050	0.900
Middle school is in village	*Reference category*	
Nearest middle school is within 5 km	0.8846	1.871
...5-10 km	-0.0652	-0.142
...10-20 km	1.6566	2.416
...>20 km	1.3317	1.847

Medical clinic in village	*Reference category*	
Nearest medical clinic is within 5 km	-1.0271	-2.322
...5-10 km	-0.2405	-0.518
...10-20 km	-0.8605	-1.290
...>20 km	-0.5790	-0.581
Total population of the village	0.0004	2.097
Elevated land (mu)	-0.0016	-2.653
Forest land (mu)	0.0000	-1.160
# of people work in TVE over # of labor.	0.0845	1.135
Whether village has TVE	-0.4689	-1.027
Output of grain per capita (kg/person)	0.0019	1.732
Net income per capita	-0.0033	-3.349
(End of year) # of pigs per person	0.7031	1.274
(End of year) # of cows per person	0.3248	0.267
(End of year) # of sheep, goat per person	0.6432	1.034
(End of year) # of poultry per person	0.4133	2.608
(End of year) # of honey bee per person	-5.1474	-1.765
Workforce per capita	0.0463	1.506
Average household size	-0.0785	-0.992
Share of workforce female	-0.1132	-1.875
Cultivated land per capita (mu).	1.3591	2.685
Grassland per capita (mu)	2.5915	1.926
Guangxi	1.4329	2.198
Guizhou	1.1390	1.656
Yunnan	*Reference category*	
Intercept	-4.2891	-1.649
Pseudo-R^2	0.3130	

Note: The village is the unit of observation (n=200) and all explanatory variables are pre-intervention (1995).

Table 4: Difference-in-difference estimates for impacts on mean income and consumption

	(1) Gain in project villages	(2) Gain in comparison villages	Difference-in-difference (1)-(2)
1997			
Income	92.12	-6.61	98.72 (3.07)
Consumption	33.59	10.91	22.68 (1.07)
Saving	58.53	-17.51	76.04 (2.34)
1998			
Income	116.17	33.81	82.36 (2.63)
Consumption	95.88	7.45	88.43 (3.77)
Saving	20.29	26.36	-6.07 (-0.18)
1999			
Income	189.48	129.78	59.70 (1.65)
Consumption	161.77	106.61	55.16 (1.93)
Saving	27.71	23.17	4.54 (0.13)
2000			
Income	266.73	69.76	197.97 (5.14)
Consumption	101.96	79.65	22.31 (0.81)
Saving	164.77	-9.89	174.66 (4.49)

Note: Household-size weighted means in Yuan at 1995 prices with all 113 sampled project villages compared to 87 sampled non-project villages. T-ratios for the null hypothesis that DD=0 in parentheses.

Table 5: Matched difference-in-difference estimates

	(1) Gain in project villages	(2) Gain in comparison villages	Difference-in-difference (1)-(2)
Trimmed comparison group (113 villages matched with 71 comparison villages)			
1997			
Income	92.12	-9.02	101.14 (2.90)
Consumption	33.59	17.16	16.44 (0.71)
Saving	58.53	-26.18	84.70 (2.43)
1998			
Income	116.17	46.29	69.88 (2.06)
Consumption	95.88	7.90	87.98 (3.50)
Saving	20.29	38.39	-18.10 (-0.51)
1999			
Income	189.48	146.95	42.53 (1.09)
Consumption	161.77	84.83	76.94 (2.55)
Saving	27.71	62.12	-34.41 (-0.92)
2000			
Income	266.73	69.11	197.62 (4.77)
Consumption	101.96	78.47	23.49 (0.80)
Saving	164.77	-9.36	174.13 (4.17)
Caliper-bound matching (63 project villages matched with 34 comparison villages)			
1997			
Income	110.70	15.35	95.35 (2.37)
Consumption	47.79	30.36	17.43 (0.63)
Saving	62.91	-15.00	77.92 (1.92)
1998			
Income	113.47	31.68	81.79 (2.19)
Consumption	99.26	18.87	80.38 (2.86)
Saving	14.22	12.81	1.41 (0.03)
1999			
Income	187.81	179.49	8.32 (0.16)
Consumption	148.52	93.95	54.57 (1.61)
Saving	39.29	85.54	-46.25 (-0.88)
2000			
Income	178.66	-22.36	201.02 (4.55)
Consumption	85.60	75.94	9.66 (0.27)
Saving	93.06	-98.30	191.36 (4.21)

Note: Household-size weighted means in Yuan at 1995 prices. T-ratios for the null hypothesis that DD=0 in parentheses.

Table 6: Impacts of SWPRP on poverty incidence in 2000

1996 poverty incidence (*H*) in project villages (%)	(1) Change in *H* in project villages	(2) Change in *H* in comparison villages	Double difference (1)-(2)
No matching (113 project villages compared to 87 non-project villages)			
57.86	-6.66	-1.63	-5.03 (-1.75)
Trimmed comparison group (113 villages matched with 71 comparison villages)			
57.86	-6.66	-0.33	-6.33 (-2.07)
Caliper-bound matching (63 project villages matched with 34 comparison villages)			
59.72	-4.00	-3.39	-0.61 (-0.17)

Note: Poverty line =808 Yuan per year per person (1995) prices, equivalent to $1.08 per day at 1993 consumption PPP. 1130 sampled households in project villages; 870 in non-project villages. T-ratios for the null hypothesis that DD=0 in parentheses.

Figure 1: Histograms of the propensity scores

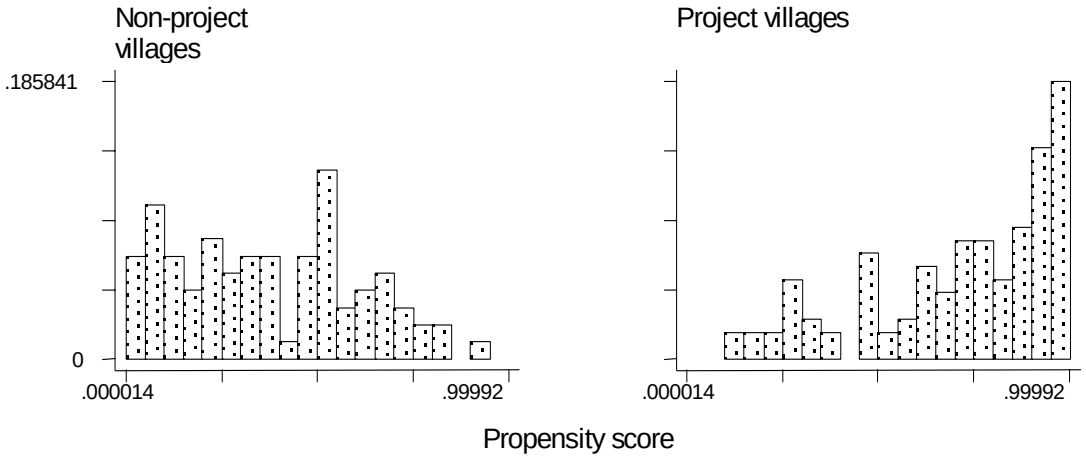


Figure 2: Impacts on consumption poverty

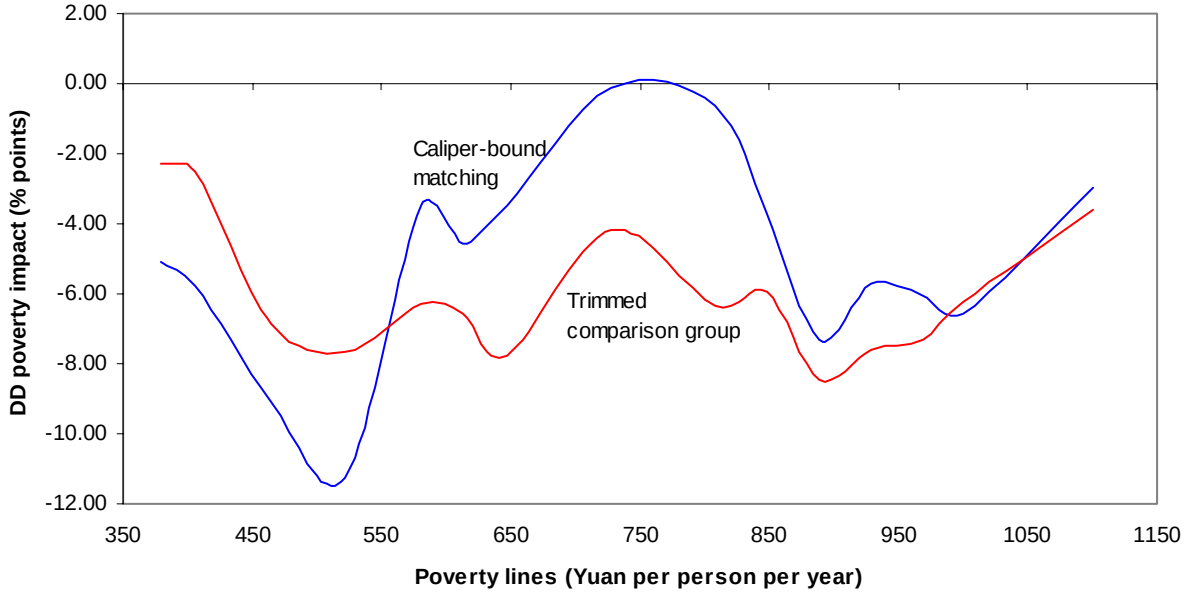
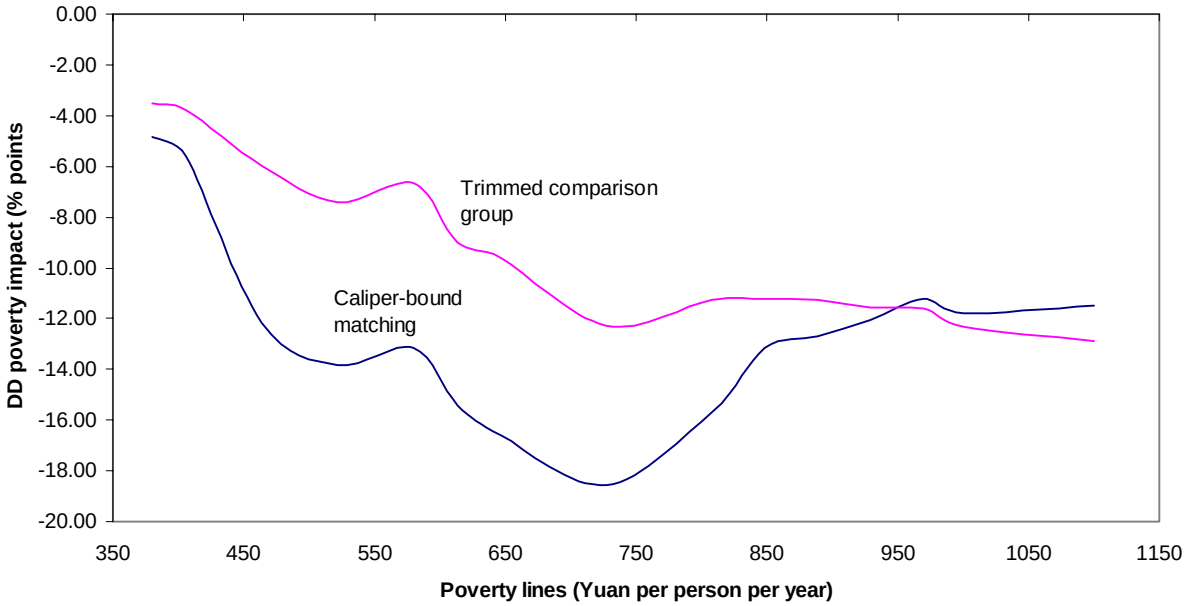


Figure 3: Impacts on income poverty



References

Besley, Timothy, 1995, "Savings, credit and insurance", in Jere Behrman and T.N. Srinivasan (eds) *Handbook of Development Economics Volume 3,* Amsterdam: North-Holland.

Chen, Jian, and Belton M. Fleisher, 1996, "Regional Income Inequality and Economic Growth in China." *Journal of Comparative Economics,* 22, 141-164.

Chen, Shaohua and Martin Ravallion, 1996, "Data in Transition: Assessing Rural Living Standards in Southern China," *China Economic Review*, 7, 23-56.

_____ and _____, 2001, "How Did the World's Poor fare in the 1990s?", *Review of Income and Wealth,* 47(3), 283-300.

Deaton, Angus, 1992, *Understanding Consumption*, Oxford: Oxford University Press.

Friedman, Milton, 1957, *A Theory of the Consumption Function*, Princeton N.J., Princeton University Press.

Gersovitz, Mark, 1988, "Savings and development," in H. Chenery and T.N. Srinivasan (eds) *Handbook of Development Economics Volume 1*, Amsterdam: North Holland.

Heckman, J., H. Ichimura, and P. Todd, 1997, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies* 64(4): 605-654.

Heckman, J., H. Ichimura, J. Smith, and P. Todd, 1998, "Characterizing Selection Bias using Experimental Data," *Econometrica*, 66: 1017-1099.

Jian, Tianlun, Jeffrey Sachs and Andrew Warner, 1996, "Trends in Regional Inequality in China," *China Economic Review*, 7(1), 1-21.

Jalan, Jyotsna and Martin Ravallion, 1998, "Are There Dynamic Gains from a Poor-Area Development Program?" *Journal of Public Economics*, 67, 65-85.

_____ and _____, 1999, “Are the Poor Less Well Insured? Evidence on Vulnerability to Income Risk in Rural China,” *Journal of Development Economics*, 58(1), 61-82.

_____ and _____, 2001, “Behavioral Responses to Risk in Rural China,” *Journal of Development Economics*, 66, 23-49.

_____ and _____, 2002, “Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China,” *Journal of Applied Econometrics*, 7(4), 329-346.

Khan, Azizur Rahman and Carl Riskin, 1998, “Income Inequality in China: Composition, Distribution and Growth of Household Income, 1988 to 1995.” *The China Quarterly*, 154, 221-253.

Leading Group, 1988, *Outlines of Economic Development in China's Poor Areas*, Office of the Leading Group of Economic Development in Poor Areas Under the State Council, Agricultural Publishing House, Beijing.

National Bureau of Statistics (NBS), 2000, *The Poverty Monitoring Report of Rural China 2000*, Beijing: China Statistics Press.

Park, Albert, Sangui Wang and Guobao Wu, 2002, “Regional Poverty Targeting in China,” *Journal of Public Economics*, 86(1), 123-153.

Ravallion, Martin, 1998, “Poor Areas,” in David Giles and Aman Ullah, eds., *The Handbook of Applied Economic Statistics*. New York: Marcel Dekkar.

Ravallion, Martin and Shaohua Chen, 2004, “China’s (Uneven) Progress Against Poverty,” Policy Research Working Paper, World Bank, Washington DC.

Ravallion, Martin and Jyotsna Jalan, 1996, “Growth Divergence due to Spatial Externalities,”

Economics Letters, 53(2), 227-232.

Rosenbaum, Paul R., and Donald B. Rubin, 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41-55.

World Bank, 1992, *China: Strategies for Reducing Poverty*, Washington DC.: World Bank.

_____, 1997, *China 2020: Sharing Rising Incomes*, Washington DC.: World Bank.

Notes

[1] For reviews of the theory and evidence see Deaton (1992) and Besley (1995).

[2] Though this is not to say that targeting was perfect. Using a county-level panel data set for all of China for the period 1981–1995, Park et al., (2002) find signs that political factors have affected targeting and that leakage to non-poor counties has increased over time while coverage has improved.

[3] There were no obvious systematic differences between the 13 villages re-assigned and the original 100. A probit regression for whether a non-project village was in the original sample or the extra sample of 13 villages (with essentially the same set of regressors as in the main regressions used for estimating the propensity scores) had very low explanatory power and almost nothing was significant.

[4] An alternative approach to controlling for heterogeneity is to run a linear regression of the outcome indicators on a dummy variable for program placement allowing for the same observable covariates entering as linear controls. This method requires essentially arbitrary assumptions about the functional form of the regression for outcomes and the properties of its error term. PSM does not require these assumptions. Indeed, there is no outcome regression as such.

[5] This is a known problem in the evaluation literature. See the discussion of non-overlapping support bias in Heckman et al. (1997, 1998).

[6] Noting that, by assumption, the differenced error terms $\varepsilon_{it}^{Y} - \varepsilon_{i0}^{Y}$ and $\varepsilon_{it}^{C} - \varepsilon_{i0}^{C}$ have zero expected value amongst participants. Equations 3.1 and 3.2 also implicitly average over the distributions of the control variables used in matching.

[7] As originally formulated, the PIH also assumes that labor supply is exogenous and that preferences are homothetic.

[8] Jalan and Ravallion (1998) provide evidence for this region of China that rural households are not well insured against income shocks, and that this insurance failure is more severe for the asset-poor.

[9] Note that the baseline means differ for caliper-bound matching, given the change in the number of project villages used for the analysis. The 1996 mean income for the 63 project villages used for the caliper-bound matching is 968.75 Yuan.

[10] This was calculated from the project documents using the cumulative total project investments (deflated to 1995 prices) normalized by the number of households in participating villages. The project documents only give the number of households covered by the project. To obtain the per capita disbursements we used mean household size from the sample survey of project villages.

[11] Park et al., (2002) used regional growth regressions, estimated at county level. (The Jalan and Ravallion, 1998, method was described earlier in this paper.)

[12] This calculation is for the series excluding 1995. From Table 2, the range of mean incomes in the project villages is 266.73 (= 1259.47-992.74) Yuan in the project villages while overall mean income is 1125.64 Yuan. The range in project impacts from Table 5 is 138.27 (=197.97-59.70) while mean impact is 109.69.