

Are the poor benefiting from China's land conservation program?¹

EMI UCHIDA

Department of Environmental and Natural Resource Economics, University of Rhode Island, 219 Coastal Institute, 1 Greenhouse Road, Kingston, RI 02881, USA. Tel: +1-401-874-9196. Email: emi@uri.edu

JINTAO XU

Center for Chinese Agricultural Policy, Chinese Academy of Sciences, 5th Floor, Institute of Geographical Sciences and Natural Resources Research, Datun Road, Anwai, Beijing, China 100101.

ZHIGANG XU

Center for Chinese Agricultural Policy, Chinese Academy of Sciences, 5th Floor, Institute of Geographical Sciences and Natural Resources Research, Datun Road, Anwai, Beijing, China 100101.

SCOTT ROZELLE

Helen F. Farnsworth Senior Fellow, Freeman Spogli Institute for International Studies, Stanford University, Encina Hall East, Stanford, CA 94305, USA.

ABSTRACT. This paper studies the impact of the largest conservation set-aside program in the developing world: China's Grain for Green program, on poverty alleviation in rural areas. Based on a large-scale survey, we find that the program was implemented mostly in the areas of China that are fairly poor. In addition, we show that income from livestock activities and some types of asset holdings of participants have increased significantly more than those of non-participants (due to program effects). Only weak evidence is provided demonstrating that participating households have begun to shift their labor into the off-farm sectors. Overall, the results suggest that the program has been moderately successful in achieving its poverty alleviation objectives.

¹ The authors are grateful for the assistance during the fieldwork part of this study provided by Yazhen Gong and other members of the Center for Chinese Agricultural Policy, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences. An earlier version of the paper was presented at UC Davis and the 2004 AAEA Annual Meeting in Denver, Colorado. We thank Keiji Otsuka, Doug Miller, Aaron Smith, Bryan Lohmar, participants of the seminars, and two anonymous reviewers for helpful comments and suggestions. The authors acknowledge the support from Ford Foundation. Ditto National Science Foundation of China (70024001) for financial assistance. RI AES Contribution No. 5105

1. Introduction

In recent years there has been an increasing interest in using incentive mechanisms, known as Payments for Environmental Services (PES) to finance conservation in developing countries (Landell-Mills and Porras, 2002; Pagiola and Platais, 2005). Some developing countries have introduced PES programs to encourage land owners to enhance certain environmental services, such as biodiversity conservation or carbon sequestration. In contrast to many of the programs that failed, at least in part because they lacked incentives to generate environmental services, households in an increasing number of countries, such as Costa Rica, Columbia, and Ecuador, are participating in PES schemes. Governments in other parts of the world also have begun designing similar programs (Pagiola, 2002; Pagiola and Platais, 2005).

Aside from the effect on the environment, one of the most critical questions about the use of PES programs in developing countries is the impact that the programs have on the poor. There are two contrasting views. Proponents of PES argue that although the PES approach was conceptualized as a mechanism primarily to improve natural resource management, it could also serve as a tool for poverty alleviation (Pagiola and Platais, 2005). A PES program can offer a means to increase the income of the rural poor and reduce risk by diversifying income sources, mainly through its monetary payments. Furthermore, depending on the environmental service demanded, the land owners that provide such environmental services could be predominantly poor. For example, a PES program may target farmers that live in the upper reaches of a river basin and pay them to plant trees to reduce sedimentation in a watershed. To the extent that the upstream farmers are poor, a well-designed program could contribute to the alleviation of poverty.

In contrast, others have suggested that PES programs are likely to have negligible or even adverse effects on poverty. Poor farmers participating in a PES program (particularly in a non-voluntary one) may end up worse off if they are not compensated for their opportunity costs either due to a flawed incentive mechanism design or poor program implementation. Furthermore, land in a region could be distributed in a way such that the particular types of land that are productive in generating environmental services may not be owned by the poor. Even if some appropriate land were owned by the poor, the program might be ill-targeted and systematically exclude the poor. In particular, if a program provides high compensation, it could attract both poor and rich. Given their higher inherent abilities and social connections, if participation is limited, higher-income households may have an advantage in gaining participation rights. Lastly, PES programs often are designed on the assumption that structural changes in household income-earning activities will occur. However, if farmers are unable to find alternative income sources during the program period, and are not allowed to convert their land back to cultivation when the program payments stop, they may be trapped in poverty (Landell-Mills and Porras, 2002). Surprisingly, despite the controversy, little empirical evidence exists on the impact of PES programs on poverty, or the distributional consequences of these incentive mechanisms. Whether or not the poor are

able to take advantage of the PES programs remains to be examined (Pagiola *et al.*, 2002).

To address this question, this paper uses information from a data collection effort that we designed and implemented to evaluate a large-scale PES program in China. Pushed into action by a series of devastating floods in 1998, China's government started a conservation set-aside program, frequently known as Grain for Green, in 1999. Known as one of the largest conservation set-aside programs in the world, its main objective is to increase forest cover on sloped cropland in the upper reaches of the Yangtze and Yellow River Basins to prevent soil erosion. When available in their community, households set aside all or parts of certain types of land and plant seedlings to grow trees. In return the government compensates the participants with in-kind grain, cash payments, and free seedlings. By the end of 2002, officials expanded the program to some 15 million farmers in more than 2000 counties in 25 provinces and municipalities in China (Xu *et al.*, 2004a). If the program meets its original goals, by 2010 leaders will have set aside nearly 15 million hectares of cropland, affecting the land of more than 50 million households. If implemented to its intended size, such a program would surpass in area the Conservation Reserve Program (CRP) in the United States (Xu *et al.*, 2004a).

While the environmental goals are clear, under the current program design it is ambiguous whether or not the Grain for Green program will benefit the poor. On the one hand, proponents of the program argue that the Grain for Green program could have significant positive impact on the poor. While preventing soil erosion is the primary objective, the State Forestry Administration also explicitly states that poverty alleviation and restructuring agricultural production into a more environmentally and economically sustainable set of activities are also program goals (State Forestry Administration, 2002). Moreover, the compensation level of the Grain for Green program is higher on per acre basis compared to the CRP (Uchida *et al.*, 2005). In addition, when rural households retire their cropland, it is possible that farmers could use their freed up on-farm labor and choose to allocate the labor off the farm into the wage earning or self-employed labor market. Migration is an important way out of poverty in China (Li *et al.*, 2003). In general, since urbanization and industrialization are the main conduits of economic development (e.g., Lewis, 1954; Nurkse, 1957), transition of the labor force from rural to urban and from agriculture sector to industrial sector would be desirable for China's long-run economic development. If the program could aid the process of migration, we believe it shows that Grain for Green has longer-term dynamic consequences for China's overall development.

On the other hand, it is possible that the Grain for Green program may leave its participants worse off. Potential problems include problems with incentive design, targeting, as well as implementation. Although the central government states that poverty alleviation is one of the program goals, there is no clear statement that poverty is one of the criteria in the program area selection process. Similar to previous poverty alleviation programs in China, political factors could affect targeting. Park *et al.* (2002) found that in China's poverty alleviation programs, political factors have

led to less accurate targeting and deteriorated impact of programs on rural income growth. Given that the program offers a relatively high level of compensation, it is possible that local officials strategically favor a particular group – and given their higher inherent abilities and social connections, higher-income households may have an advantage in being allowed to participate in the program. Furthermore, if the compensation is not delivered – as some official reports indicate – the welfare of participants could become worse. Even if compensation is delivered, it is possible that its level may still be lower than the household's opportunity cost (i.e., the household's foregone agricultural profit), as was found in Uchida *et al.* (2005). The program could also be another repetition of afforestation projects in the past where lack of technical assistance and insufficient attention paid to management of the planted seedlings led to low survival rate (e.g., Smil, 1993). If officials stopped payments because of low survival rate, households could be left with less income and a hillside full of withered seedlings.

To meet the overall goal of evaluating the impact of China's PES program on the poor, this paper has two specific objectives. First, we seek to assess who is participating in the Grain for Green program, and thus understand whether China's program officials are disproportionately targeting the poor or not. To do so, we will examine the determinants of participation at the household level. Second, we seek to understand the impact of the program on income, assets, and the labor allocation decisions of participating households. We utilize recent program evaluation methods in our attempt to overcome the selection bias problem that is inherent in non-experimental studies. We also examine whether or not the program impacts, if any, are different for the poorer participants.

2. The data and the Grain for Green program

The data used in this paper come from a household survey executed by the authors in 2003. The survey was commissioned by China's State Forest Administration as part of their effort to evaluate the nation's Grain for Grain program after the third year of implementation. To the best of our knowledge this is the only existing data set that includes both participating and non-participating households. A total of 359 households in three provinces (Sichuan, Shaanxi, and Gansu), six counties, 18 townships and 36 villages were interviewed. Of these, 75 per cent of the households participated in the Grain for Green program.² Summary statistics of household characteristics, income, asset holdings, and labor allocation outcomes for participating and non-participating households in 1999 (before the program) and 2002 (after the program) are presented in tables 1 and 2.

The household survey employed a sampling strategy designed to collect data on a random sample of households in the program area. From the three

² The Grain for Green program was announced in 1999 and approved in 2000. Although some households started to retire land in 1999, the earliest year that the households could have received some compensation was in 2000. We drop the 20 households who said they started to receive compensation in 1999 to reduce possible biases.

Table 1. Sample average for participating and non-participating households in the Grain for Green program (2003)

	(1) Participating households	(2) Non- participating households	(3) Difference in mean (1)–(2) (t-statistics in parentheses)
Household Size	4.84 (1.69)	4.47 (1.30)	0.38 (1.86)*
Number of Children (≤ 15 years old)	1.14 (0.98)	0.95 (0.93)	0.12 (1.57)
Number of Adults (≥ 16 years old)	3.69 (1.36)	3.51 (1.01)	0.18 (1.14)
Household head's age	47.77 (10.98)	47.69 (12.40)	−0.08 (0.06)
Household head's educational attainment	4.68 (3.33)	4.54 (3.68)	0.15 (0.34)
Female headed household (%)	7.54 (0.26)	9.30 (29.22)	−0.02 (0.52)
Household landholding (mu)	13.85 (9.35)	10.25 (8.74)	3.59 (3.13)**
Household landholding per capita (mu)	3.05 (2.17)	2.54 (2.37)	0.52 (1.86)*
Weighted average of distance from each plot to house (m) ^a	1,029.47 (880.92)	760.92 (803.56)	268.56 (2.50)*
Weighted average of distance from each plot to water (m) ^a	1,068.59 (1,886.26)	810.99 (1,401.14)	257.59 (1.16)
Weighted average of distance from each plot to road (m) ^a	852.33 (1,098.32)	814.17 (1,604.97)	38.16 (0.25)
Number of households in sample	253	86	

Notes: ^aWeighted by the size of each plot.

Standard deviations are reported in parentheses for columns (1) and (2); absolute values of t-statistics are reported in column (3). * significant at 10%; ** significant at 5%; *** significant at 1%.

Source: Authors' survey.

provinces that had been participating in the Grain for Green program since 2000, two counties in each province and three townships in each county were randomly selected. In each township, two participating villages were selected, and, within each village, ten households were randomly selected. There is at least one household participating in the program in every village. In two out of 36 villages, all of the households in our sample were participating households.

Table 2. *Sample average for participating and non-participating households in the Grain for Green program in 1999 and 2002*

	Participating households		Non-participating households		(5) Difference in mean in 1999 (1)–(3) (t-statistics in parentheses)
	(1) 1999	(2) 2002 ^a	(3) 1999	(4) 2002 ^a	
Income per capita (yuan)	1,404.41 (1,516.38)	1,675.25 (1,797.70)	1,850.41 (2,126.33)	2,026.11 (1,926.95)	–446.00 (2.11)*
Agricultural income per capita (yuan)	648.17 (807.01)	687.02 (1,018.76)	869.06 (1,392.09)	774.47 (920.37)	–220.89 (1.79)*
Off-farm work (number of adults with off-farm work in household)	3.58 (1.42)	3.63 (1.38)	3.44 (1.06)	3.45 (1.06)	0.13 (0.21)
Migration status (number of adult migrants in household)	0.35 (0.64)	0.48 (0.72)	0.23 (0.50)	0.41 (0.62)	0.12 (1.53)
Value of house (yuan)	13,659.45 (19,104.72)	13,744.45 (18,994.64)	20,066.54 (29,639.19)	19,682.79 (28,599.90)	–6,407.10 (2.31)*
Fixed productive assets (yuan)	842.80 (1,412.12)	1,619.50 (3,683.74)	948.47 (1,912.12)	1,259.68 (2,273.94)	–105.67 (0.54)
Consumer durables (yuan)	569.20 (988.09)	1,623.46 (2,645.82)	930.30 (2,391.40)	2,326.95 (3,224.18)	–361.10 (1.96)*
Livestock inventories (yuan)	384.58 (518.15)	613.86 (897.62)	414.39 (442.43)	478.65 (360.38)	–29.80 (0.48)
Total land in Grain for Green program (mu)	2.26 (5.24)	7.15 (9.39)		N.A.	
In-kind and cash subsidy from Grain for Green program per capita (yuan)	N.A.	59.25 (88.94)		N.A.	
Number of households in sample		253		86	

Notes: Standard deviations are reported in parentheses for columns (1) through (4); absolute values of t-statistics are reported in column (5). *significant at 10%; ** significant at 5%; *** significant at 1%. N.A.=Not Applicable.

^aValues in 2002 are converted to real terms using Consumer Price Index for each province.

Source: Authors' survey.

The household survey asked respondents about a number of variables from both before (1999) and after (2002), the time the program began. Enumerators collected information on the household's production activities on a plot-by-plot basis, as well as detailed information on each household's total asset holdings, its demographic make-up, and other income earning activities from both on- and off-farm activities. We acknowledge the problems inherent in recall data.³ Long-term recall data are possibly inaccurate, although the literature continues to debate the issue. Unfortunately, the government's quick decision to implement Grain for Green and lack of transparency in the details of implementation precluded the option of conducting interviews with potential participants at program onset. Concerns regarding recall bias were addressed both through the design of the survey and careful training and monitoring of the enumerators to ensure that respondents gave their best recollection of past amounts and activities. In addition, in our analyses we attempt to deal with the recall bias inherent with a retrospective survey.

China's Grain for Green program

Starting in 1999 as a pilot program, the Grain for Green program was implemented by China's government as a cropland set-aside program to increase forest cover and prevent soil erosion on sloped cropland. By the end of 2002, officials expanded the program to some 15 million farmers in more than 2000 counties in 25 provinces and municipalities in China (Xu *et al.*, 2004a). According to the program rules, each participating farmer receives three types of compensation: in-kind grain compensation, cash compensation, and free seedlings. In-kind grain and cash compensation are both given out every year upon successfully passing an inspection; free seedlings are provided in the initial year. The program also is designed so that there are only two levels of in-kind grain compensation nationwide. Differences in in-kind payments are justified by the inherent differences in regional average yields. The compensation level is 1,500 kilograms per hectare per year in the Yellow River Basin and is 2,250 kilograms per hectare per year in the Yangtze River Basin. In cash equivalent terms these are 3,150 yuan and 2,100 yuan per hectare per year, respectively.⁴ In total, the three types of compensation amount to 3,150 yuan per hectare in the middle

³ Errors occur in retrospective surveys because respondents may either completely forget events or mis-date them (Dex, 1991). If these errors are systematically correlated with explanatory variables in the regression models, they may bias the coefficients. In particular, it would be a serious problem if the bias is correlated with program participation. Although we cannot test such correlation using our data set, we do not have compelling reasons why either participating or non-participating households would systematically misreport pre-program status because of forgetting or mis-dating. The difference in average educational level between the two groups is not statistically significant and there is no reason to believe that one group is more biased than the other.

⁴ The annual average official exchange rate in 2001 was 8.28 Chinese yuan to the US dollar. The Purchasing Power Parity conversion factor in 2001 was 1.9 yuan to the dollar World Bank (2003).

and upper reaches of Yellow River for the first year of conversion, and 2,400 yuan per year from the second year on. In the upper reaches of the Yangtze River, the program pays farmers 4,200 yuan per hectare the first year and 3,450 yuan per year from the second year on. If Purchasing Power Parity is used as the conversion rate, the average first year compensation amounts to a payment that is more than ten times the average per hectare rental payment under a similar program in the US (Uchida *et al.*, 2005). Moreover, the level of compensation is not trivial relative to the earnings of the typical participating household in our sample. For example, if an average household in Sichuan Province (Yangtze River Basin) received full compensation, it would receive 340 yuan per capita, which is a sum of money equal to 24 per cent of the household's pre-program average total income per capita in 1999.⁵ Xu *et al.* (2004a) have found, however, that the actual delivery of payments often falls short (although this happens only in some villages, and not in most).⁶

Since the main objective of China's program is to restore the nation's forests and grasslands to prevent soil erosion, program designers have made the steepness of the slope one of the main criteria by which plots are selected for inclusion into the Grain for Green program. The steepness criterion means that the program in southwest China targets land with 25 degrees of slope or more for inclusion into the program. In the northwest the program targets land with 15 degrees of slope or more. Using a targeting mechanism such as slope also has the advantage that it is relatively easy to observe, which could make project monitoring more effective. Some studies have found, however, that the program's targeting has been less than optimal in some regions where some productive and low sloped plots were retired when less productive and high sloped plots were still available (Xu *et al.*, 2004a; Uchida *et al.*, 2005).⁷ Most close observers believe the Grain for Green program has been 'quasi-voluntary'; although households had the ultimate decision of whether or not to participate, some households that have plots meeting the steepness criterion were 'strongly encouraged' to participate. In fact, Xu *et al.* (2004a) finds that only 15 per cent of the participating households in the sample were consulted before program

⁵ This was calculated as follows. The average program area of households in our sample in Sichuan Province (Yangtze River Basin) in 2002 was 7.15 mu (=0.477 hectare). If a participating household received full compensation, it would amount to 1,646 yuan (using the compensation level from the second year and on). Using an average household size of 4.84 members (which it was in Sichuan in 1999), this amounted to 340 yuan per capita, which was a sum of money equal to 24 per cent of pre-program average total income per capita.

⁶ Full compensation may not reach participating farmers for several reasons. Village officials may deduct labor cost to plant trees on the farmer's program plots. They may deduct back-taxes owed by the farmer. In other cases, some plots are retired before being fully certified under the program (Zuo, 2002; Xu *et al.*, 2004a).

⁷ From the perspective of program goals, however, it may be favorable to retire plots with no slope under certain circumstances. For example, low sloped plots that may be productive could also be effective in preventing soil from entering a river system.

implementation, and only 30 per cent had the autonomy to choose which plots to retire.

While preventing soil erosion is the primary objective, poverty alleviation has been stated as another goal of the Grain for Green program (State Forestry Administration, 2002). Many local governments consider it as an opportunity to promote the transformation of the county's local economic structure. Field interviews conducted by the authors revealed that many county, town, and village leaders as well as farmers were attracted to the program's high level of compensation.⁸ Therefore, if returns were truly high, as long as poor households are included in the project, it is reasonable to expect some positive impact of the program on poverty alleviation. In the rest of this paper we examine this issue more carefully, seeking to assess whether the poor households are disproportionately participating in the program, and what impact the program has had on the participating households.

3. Determinants of participation

While the Grain for Green program did not directly target the poor, we expect that most of the participants are relatively poor compared to China's average rural household. Forestry officials designed the program to target sites that have a high potential to minimize the effects of off-site soil erosion; the program criterion is not designed to target the poor directly. But since poor households were historically the ones that were pushed on to marginal plots in the upland regions of China's main watersheds, the program may be implicitly targeting poor households.

Strategy to explain determinants of participation

The first step in understanding the determinants of participation is to understand the nature of participating households and compare them to those that did not participate. In our sample, participating and non-participating households share certain characteristics, but differ in income level and asset holdings (tables 1 and 2).⁹ For example, households in both groups are similar in terms of the number of children and adults as well as the age and level of education of the head of household (table 1). The two

⁸ We recognize that compensation standards do not account for regional and inter-household production variability and opportunity costs. Without accounting for this, of course, the program is subject to mis-targeting in both ways – both by overpaying some farmers that would have been willing to participate at a lower compensation rate, and by making farmers participate at a rate which they believed was not sufficient to offset their opportunity cost (particularly if a household participates involuntarily.) In some villages, however, the leaders converted more cropland than the allocated program quota, hoping to be included in future rounds of the program. Some interviewed farmers were willing to convert more cropland even with half of the current compensation level. From these observations we believe that the promised compensation level generally has been attractive compared to the opportunity cost of setting aside the cropland (i.e., the forgone revenue from cultivating the land).

⁹ All values for 2002 are adjusted for inflation using consumer price index for each province.

groups differ, however, in initial income levels and asset holdings (table 2). Interestingly, the participating households have lower levels of income and assets prior to the program. In particular, total income per capita, housing value, and consumer durables in 1999 were lower for the participating households.

On average participating households had access to more land but had lower levels of other assets before the program: participating households had about 14 mu of cropland, whereas non-participating ones had 10 mu (1 mu equals 1/15th of a hectare). Housing value, often recognized as a more accurate measure of wealth than income (since it is more easily observable and measured with less error), also is lower for participating households (13,659 yuan) compared to non-participating ones (20,067 yuan).¹⁰ Participating households also had lower levels of livestock holdings in 1999 (385 yuan) compared to non-participating ones (414 yuan).¹¹ With such differences, any analysis of the effect of Grain for Green clearly must control for the differences in the initial characteristics of participating and non-participating households.

To further investigate the factors determining participation in the program, we first estimate kernel densities to trace out the income distributions of participating and non-participating households. If the poor were disproportionately selected into the program, the income distribution of the participating households would be to the left of the distribution of non-participants. We then use limited dependent variable regression methods to examine the marginal effects of each determinant on program participation.

To examine the decision to participate or not, we use a fixed-effect logit model, which can be expressed by $\Pr(P_i = 1) = \Lambda(X_i'\beta)$, where $P_i = 1$ if household i participates in the Grain for Green program and 0 otherwise and X is a matrix of the explanatory variables described above. The model includes county fixed effects to control for unobservable county characteristics.¹² We use explanatory variables corresponding to three groups of factors that may have been used by local governments as criteria for deciding if villages and households should be included in the Grain for Green program: environmental factors (slope, distance to waterway), wealth factors (income, land and asset holdings before the program; migration status before the program), and implementation costs

¹⁰ Housing value is based on self-reported values.

¹¹ We collected full data on livestock for 2002 including revenue from livestock, prices, initial and ending stocks, outflow, and death counts. For 1999 we only have data on prices, initial stock and outflow, and death counts. To approximate the livestock value in 1999 we used the prices and outflow counts, multiplied by 1.02, the correlation coefficient between ending stock value and the sum of that and the revenue from livestock in 2002.

¹² We cannot include village or township fixed effects because a few villages and townships in our sample only had participating households. Had we included village fixed effects we could control for unobservable village characteristics that are invariant across households in each village, but at the expense of dropping more than 20 households from the analyses. Given the relatively small sample size we chose not to do so.

(e.g., proximity to a public road). In addition, we also included other household characteristics.

We then use a random effect Tobit model to explain the total area of land in the program by each household, where y_i^* is the total cropland area of household i that was enrolled in the Grain for Green program. The variable y_i^* is truncated since all non-participating households have zero values and participating households have strictly positive values. The covariates in the model are the same as those in the logit model except that the maximum slope of the household's plots is not included.¹³ This model is estimated with random village-level effects.

Results

The kernel density estimates of log of income per capita for participating and non-participating households illustrate that the income distributions of the two groups before the program (1999), in fact, were quite similar (figure 1). Although the means are statistically different at the 10 per cent level and the variance of participating households was somewhat smaller than for the non-participants, it is difficult to see striking differences. Most importantly, while there is considerable heterogeneity in levels of per capita income among households, it should be noted that most if not all of the households in our sample are absolutely poor with nearly 99 percent below the US \$2 per day poverty line.

The results of the determinants of participation analyses also suggest that low income households were neither disproportionately included nor excluded from China's Grain for Green program (table 3). The estimated marginal effect of income level in 1999 was zero and insignificant (column 1). Holding all other factors constant, the poverty status of a household was not a determining factor of participation. The household's income level also did not affect how much cropland area was retired (column 2). Although on the one hand such a result would suggest that the program did not excessively target the poor, a more positive spin can also be put on it. First, the poor were not disproportionately excluded; second, since most people in the study areas were poor (average income was only 69 per cent of the average income of rural households in China), a substantial number of at least relatively poor households were included.¹⁴

The results of the logit and tobit analyses illuminate other determinants of participation. For example, households with higher-sloped plots were more likely to participate (0.083), a finding that was expected since the steepness of a plot's slope was regarded as the principal criterion of site selection for the program. The estimated coefficient on household head's age (0.008) suggests that a household with an older household head is more likely to participate, a finding consistent with studies examining the determinants in the US CRP program (e.g., Skaggs *et al.*, 1994). Finally, the estimated

¹³ This variable was not included because although it may be important for making a household eligible for the program, we believe that it will not affect the quantity of the land in the program.

¹⁴ Annual net income per capita in 1999 was 2,210 yuan for rural households and 5,854 yuan for urban households. Our sample average was 1,518 yuan.

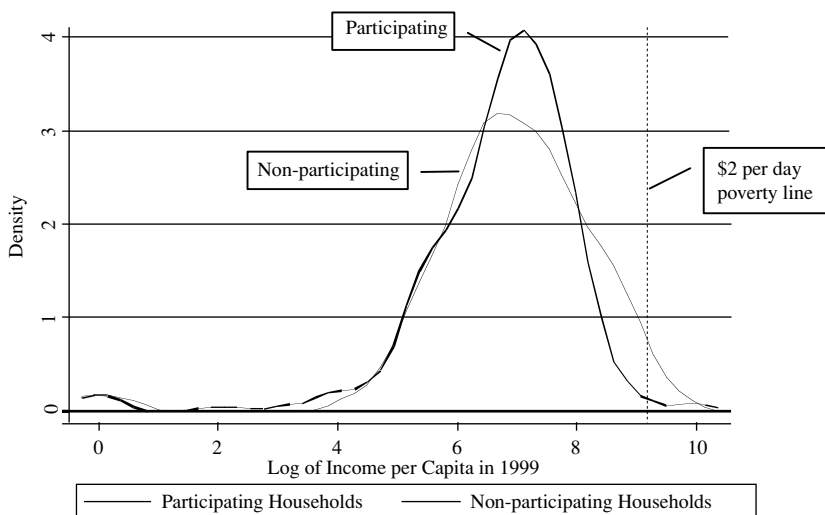


Figure 1. Kernel Density of Log of Income per Capita for Participating versus Non-participating Households in 1999.

Notes: The vertical line corresponds to the two-dollars-a-day poverty line in 1999 (US \$1,112.59 or 9212 Chinese Yuan). In logarithm the value is equivalent to 9.13. The two-dollars-a-day poverty line in 1999 was calculated based on the 1996 poverty line reported in Sala-i-Martin (2002) (US \$1064) and by adjusting for inflation using deflators reported in the World Economic Outlook 2002 (Sala-i-Martin, 2002).

Source: Authors' survey.

coefficient of the variable measuring the number of adult migrants in the household (0.099) suggests that the probability of participating in the program is 10 per cent higher for a household with an additional adult migrant (or at least there is a positive correlation between migration and participation). For households that had already sent a household member into the migrant labor force before the program, the Grain for Green program may have become an opportunity to take some cropland out of production, which could be attractive because many forestry activities result in labor savings compared to cultivation.

We also tested whether or not the slopes of the explanatory variables differed among the three provinces. In the logit model, we found that coefficients are statistically significantly different for two regressors: 'distance to waterway' and 'maximum slope'. The marginal effect of 'distance to waterway' is small but positive and statistically significant for Sichuan Province but statistically insignificant for the other two provinces. The marginal effect of maximum slope is positive and statistically significant for Sichuan Province but statistically insignificant for the other two provinces. These findings suggest that the participation decision was systematically different for Sichuan Province. Importantly, the slope criterion may have been best implemented in Sichuan. In the Tobit model, we found that coefficients are statistically significantly different for four

Table 3. Fixed-effect logit and random-effect tobit regressions results explaining determinants of households' program participation and land size in the Grain for Green program

	(1) Program participation (1=participate)	(2) Area in Grain for Green (mu)
Total agricultural revenue per capita in 1999 (yuan)	-0.000 (0.58)	-0.001 (1.02)
Income per capita in 1999 (yuan)	-0.000 (1.04)	0.000 (1.17)
Land holding per capita (mu)	0.010 (0.50)	2.150 (9.57)***
Maximum slope among household's plots	0.083 (2.88)***	
Fixed asset value in 1999 (yuan)	0.000 (0.23)	-0.000 (1.15)
House value in 1999 (yuan)	0.000 (0.91)	-0.000 (1.80)*
Livestock value in 1999 (yuan)	0.000 (0.94)	-0.001 (0.86)
Number of adult migrants in household	0.099 (1.81)*	-0.370 (0.58)
Distance from road to plots, weighted average (meter)	0.000 (1.81)*	-0.001 (1.52)
Distance from plots to house, weighted average, ln (meter)	0.113 (2.55)**	1.710 (2.91)**
Distance from plots to water, weighted average (meter)	0.000 (0.890)	0.000 (1.17)
Household size	0.029 (1.78)*	2.120 (7.84)***
Household head's age	0.008 (1.83)*	0.018 (0.26)
Household head's age squared	-0.000 (1.65)*	-0.001 (1.55)
Household head's educational attainment	-0.004 (0.45)	-0.210 (1.60)
Constant		-19.520 (3.61)***
Observations	338	338
Number of counties	6	
Number of villages		36
Prob>chi2	0.0000	0.0000
Pseudo R ²	0.1724	

Notes: Means of marginal effects are reported in column (1). Absolute value of z-statistics in parentheses. The standard errors are clustered on village. In column (1) a constant is not estimated since we include county fixed effects. *significant at 10%; ** significant at 5%; *** significant at 1%

regressors: 'land per capita', 'maximum slope', 'household size', and 'age of household head squared'. Land per capita is positive and statistically significant for all three provinces, but the magnitude of the coefficient for Shaanxi is three times that of the other two provinces; maximum slope is negative and statistically significant for Shaanxi but insignificant for the other two; household size is positive and statistically significant for Shaanxi but insignificant for the other two; finally, the household head's age squared is negative for Shaanxi but insignificant for the other two. In sum, factors

determining the extent of participation were systematically different for Shaanxi.

We have some evidence that a large portion of non-participating households wanted to participate but could not. Based on a question where we asked the non-participant respondents why they did not participate, of the 87 non-participating households in our sample, a majority (47 households) indicated that they did not participate because their sloped plots were not included in the program area.¹⁵ Only eight households clearly indicated that they could have participated but did not want to participate (mostly because either the compensation level was not high enough or they did not believe that the compensation would be delivered by the government). In sum, a majority of the households not enrolled in the program also have plots with steep slopes and appear to have wanted to participate if they had been given the opportunity.

4. Effects on income, assets, and labor allocation

We expect a direct and indirect income effect on the participants of the Grain for Green program. The direct effect on income and assets is expected to come through the grain and monetary compensation. In addition, participation in Grain for Green also may change labor allocation decisions. A land retirement program, like the Grain for Green program, has two effects on household labor: a substitution effect and an income effect. With a substitution effect a household retiring its cropland would shift its labor freed by the program into other productive activities, such as on-farm activities on the household's remaining cropland, an off-farm wage job, or the self-employment sector. But households may not shift all of their free time that arose from participating in the program into productive activities because of an income effect: a farmer may reallocate the time saved from the program into leisure. Hence, whether or not we could expect increases in off-farm labor is theoretically indeterminate.

Strategy for explaining program impacts

While the descriptive statistics show that on average income per capita and asset values for the participating households increased between 1999 and 2002 (table 2), we cannot attribute these increases solely to the Grain for Green program. It could be that other characteristics of participants contributed to the income rises, and, in fact, there is a reason to suspect that this is so. According to our data, average income and holdings of certain assets are also increasing for non-participating households. In order to obtain a more unbiased estimate of the impact of the Grain for Green program on income, asset holdings, and labor allocation we first need to hold constant the other effects in order to isolate the impact of the program on these changes.

What we are trying to measure is often called the mean impact of 'treatment on the treated', which is the average impact of the program

¹⁵ Ten households were ineligible because they did not have sloped plots suitable for the program.

among those treated

$$TT = E((Y_1 - Y_0)|X, D = 1) = E(Y_1|X, D = 1) - E(Y_0|X, D = 1), \quad (1)$$

where we denote Y_1 as the outcome conditional on treatment, Y_0 as the outcome conditional on non-participation, $D = 1$ the group of individuals who participated in the program for whom Y_1 is observed, $D = 0$ those who do not participate in the program for whom Y_0 is observed. In reality we do not observe the counterfactual mean, $E(Y_0|X, D = 1)$, or the mean outcome for the participating group had they not participated in a program.

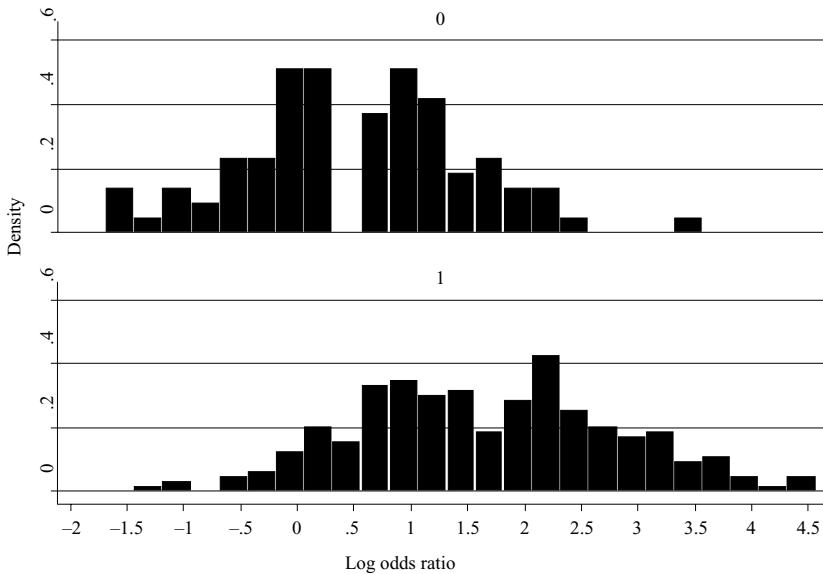
Since officials did not implement the Grain for Green program on the basis of a randomized experiment, we cannot assume that the selection bias is zero. To overcome the problem we employ three approaches: propensity score matching method (PSM), difference-in-differences (DD), and difference-in-differences matching method (DDM). PSM allows the analyst to match the treated and the controls when observable characteristics of participants and non-participants are continuous, or when the set of explanatory factors that determine participation contains many variables (Rosenbaum and Rubin, 1984). With the right data, it is possible to estimate the propensity score and compare the outcomes of participating and non-participating households that have similar propensity scores. We can obtain the mean impact of the treatment on the treated by

$$E(Y_1 - Y_0|D = 1) = E\{E(Y_1|p(Z), D = 1) - E(Y_0|p(Z), D = 0)|D = 1\}, \quad (2)$$

where $p(Z) \equiv \Pr(\text{Treated}|Z)$. By matching participating and non-participating households with different values of Z but the same values of $\Pr(D = 1|Z)$, any differences in $E(Y_0)$ between the two groups difference out when calculating the above equation. The observable covariates Z should include all of the relevant characteristics that determine participation and the outcomes of interest. In our analyses, Z includes a number of variables including total income per capita, agricultural income per capita, asset values before the program (housing values, consumer durables, productive fixed assets), labor allocation (number of migrating adult household members), and other household and plot characteristics. We also include county fixed effects to control for unobservable factors at the county level that may affect participation.¹⁶

To implement PSM successfully, the nature of the samples of participating and non-participating households must meet certain criteria and several other choices must be made. Importantly, the common support of the balancing score for participating and non-participating households should be and is, in fact, fairly wide (figure 2), so information from a large share of the households can be used in the analysis. We choose to use the nearest neighbor matching method with replacement. Following Smith and Todd (2005), we match on the log odds-ratio so that the estimates are robust to choice-based sampling, where the proportion of the treatment group in the data is over-sampled relative to their frequency in the population of persons

¹⁶ The specification is the same as the logit model in the previous section.



Graphs by HH Participation

Figure 2. *Distribution of estimated log odds-ratios for participating and non-participating households.*

Note: 0=Non-participating households, 1=Participating households,

Source: Authors' survey.

eligible for the program. We follow Dehejia and Wahba (2002) and estimate an 'adjusted' version: a weighted regression where the comparison units are weighted by the number of times that they are matched to a treated unit. This method can potentially improve the precision of the estimates (Dehejia and Wahba, 2002). Standard errors are bootstrapped using 1,000 replications. A balancing test that follows Dehejia and Wahba (1999, 2002) was satisfied for all covariates.

While PSM is increasingly being used in program evaluations, it relies on a key underlying assumption that is not always valid – that the outcomes (in our case, the conditional means) are independent of program participation conditional on the propensity score estimated from a set of observable characteristics. Formally, this can be written as

$$E(Y_0|P(Z), D = 1) = E(Y_0|P(Z), D = 0). \quad (3)$$

While we control for unobservable differences at the county level using fixed effects when estimating the propensity score, there may still be systematic, unobservable differences between the outcomes of participants and non-participants. The systematic differences could arise, for example, because households are selected into the Grain for Green program based on unmeasured household, village, or township characteristics, or because earning levels differ among labor markets in which the participating and non-participating households reside. Such differences would violate the identification conditions required for matching.

To control for such unobservable characteristics, we next employ difference-in-differences. DD compares outcomes before and after a policy change for a group affected by the change (participants in the program) to a group not affected by the change (non-participants of the program – Meyer, 1995). Let t and t' denote time periods after and before the program, respectively. The DD estimate is given by

$$DD = [E(Y_t|D = 1) - E(Y_{t'}|D = 1)] - [E(Y_t|D = 0) - E(Y_{t'}|D = 0)]. \quad (4)$$

The idea is to correct the simple difference before and after for the treatment group by subtracting the simple difference for the control group. By comparing the before–after change of treated units with the before–after change of control units, any *common* trends, which will show up in the outcomes of the control units as well as the treated units, get differenced out (Smith, 2004). The DD estimator can also get rid of the recall bias inherent in a retrospective survey to the extent that such bias is the same for both participants and non-participants. In the paper we report the unrestricted and adjusted DD estimate which relaxes the implicit restrictions in the standard DD estimator that the coefficient associated with $Y_{t'}$ (pre-program outcome) and covariates in t' (pre-program period) equals 1, respectively.

Identification of the causal effects using DD relies on the assumption that absent the policy change, the average change in $Y_t - Y_{t'}$ would have been the same for treated and the control. This is called the ‘parallel trend’ assumption, which can be expressed as

$$E(Y_{0,t}|D = 1) - E(Y_{0,t'}|D = 1) = E(Y_{0,t}|D = 0) - E(Y_{0,t'}|D = 0). \quad (5)$$

As might be expected, the effectiveness of DD in a very real sense depends on the validity of this assumption. Whether or not the assumption is valid depends on the context of the study and on how similar the control and treatment groups are. In general, the more similar are the treatment and control groups, the more convincing the DD approach. Using our village and household data we found (discussed above) that the participating and the non-participating groups were significantly different in several aspects, suggesting that the parallel trend assumption may not hold. While we control for the observable pre-program variables in the reported DD estimate, there could be other unobservable factors that may compromise the parallel trend assumption.

To better control for the differences between participating and non-participating households, we also can combine DD with a matching method and implement a difference-in-differences matching strategy. With DDM we can exploit the data on the participants in 1999 to construct the required counterfactual, instead of just using the data in 2002 (as is used in the PSM analysis). One additional advantage of DDM over DD is that the assumptions that justify DDM estimation are weaker than the assumptions necessary for DD or the conventional propensity score matching estimator (Smith and Todd, 2005). DDM requires that in the absence of treatment, the average outcomes for treated and controls would have followed parallel paths

$$\begin{aligned} &E(Y_{0,t}|P(Z), D = 1) - E(Y_{0,t'}|P(Z), D = 1) \\ &= E(Y_{0,t}|P(Z), D = 0) - E(Y_{0,t'}|P(Z), D = 0) \end{aligned} \quad (6)$$

which is weaker than the assumptions necessary for DD. DDM removes any parallel, time-variant, unobservable differences between the participating and non-participating households conditional on $P(Z)$, a clear advantage over cross-sectional propensity score matching. In fact, using outcomes from experimental data as a benchmark, Smith and Todd (2005) found that DDM performed better than DD or PSM.

In performing DDM we match by log odds-ratio and use the same nearest neighbor matching method with replacement as in PSM. We report the 'adjusted' version, where the control units are weighted by the number of times that they are matched to a treated unit. The standard errors are bootstrapped using 1,000 replications.

Results of cross-sectional propensity score matching analyses

The results of cross-sectional PSM analysis, which (unlike DD) compares the matched participating and non-participating households with similar probability of participation, reveal that Grain for Green had some positive effects on participating households (table 4, column 1). Although there is no statistically significant effect on the household's total income per capita, the PSM results suggest that the program had a significant positive effect on other agricultural income (from livestock activities), which increased by 172 yuan. In contrast, crop income dropped by 172 yuan. In addition, house value and livestock inventory values of the participating households increased by 486 yuan and 180 yuan; the estimates for these variables were statistically significant. We did not, however, find evidence of significant program effect on the employment prospects of the adults in the household.

The results that show only a marginal (or negligible) impact on income are consistent with findings in Xu *et al.* (2004b). Using the difference-in-differences analysis, the authors find that there is negative impact on cropping income and positive impact on income from subsidies (Xu *et al.*, 2004b). In contrast to our paper, however, they use a model that is restricted and unadjusted for other variables and do not examine the impacts of the conservation set-aside program on household assets or labor allocation.

Results of difference-in differences analysis

The findings from difference-in-differences analysis suggest that the program has had a significant impact on several of the income categories and several of the asset categories when comparing participating households with the non-participating ones (table 4, column 2). While crop income decreased (significantly) by 114 yuan, other non-crop agricultural income (from livestock enterprises) increased by 181 yuan, offsetting the decrease in crop income. Although the estimates are not statistically significant, the point estimates for fixed productive assets and livestock inventories were 683 yuan and 161 yuan, respectively. The DD results also did not find significant program impact on change in labor allocation into off-farm work or migration.

Table 4. *Estimated effects of the Grain for Green program on changes in income, labor allocation and asset holdings using difference-in-differences, 1999 to 2002*

Dependent variable	<i>Propensity score matching^a</i>	<i>Difference-in-differences^b</i>	<i>Difference-in-difference matching^b</i>
	Y(2002)	Y(2002) – Y(1999)	Y(2002) – Y(1999)
Income per capita (yuan)	–11.36 (0.10)	88.19 (0.61)	–11.36 (0.09)
Crop income per capita (yuan)	–172.21 (3.14)***	–114.34 (3.75)***	–167.14 (3.32)***
Other agricultural income per capita (yuan)	171.99 (2.27)**	180.56 (1.67)*	168.02 (2.36)**
Non-agricultural income per capita (yuan)	–80.38 (0.82)	–2.92 (0.04)	–29.28 (0.36)
Off-farm work (number of adults with off-farm work in household)	–0.04 (0.04)	0.0034 (0.08)	0.045 (1.62)*
Migration status (number of adult migrants in household)	–0.04 (0.04)	–0.018 (0.25)	–0.04 (0.58)
Value of house (yuan)	485.80 (2.39)**	323.13 (0.90)	521.80 (2.51)**
Fixed productive assets (yuan)	320.23 (0.89)	682.97 (1.36)	312.77 (0.80)
Livestock inventories (yuan)	180.00 (2.94)***	161.15 (1.46)	220.02 (2.74)***

Notes: Absolute values of t-statistics are in parentheses. The estimates are adjusted for inflation. * significant at 10%; ** significant at 5%; *** significant at 1%

^aThe standard errors were bootstrapped using 1,000 replications. The propensity score model is the same as the logit model reported in table 3, column 1. The balancing property was satisfied using the specification. Following Smith and Todd (2005) we match on the log odds-ratio so that the estimates are robust to choice-based sampling. The matching method used is nearest neighbor matching method (random draw version) with replacement. 252 participating households were matched to 86 non-participating households. Following Dehejia and Wahba (1999), the effects are estimated by performing weighted least squares where treatment observations are weighted as 1 and control observations are weighted by the number of times they are matched to a treatment observations (Dehejia and Wahba, 1999).

^bThe explanatory variables used in the regressions-adjusted equations are value of the outcome variable in 1999, household size, land holding per capita, education attainment of household head, age of household head, whether or not the household head is female, maximum slope among plots belonging to each household, minimum soil quality among plots belonging to each household, income per capita in 1999, fixed assets in 1999, and village fixed effects.

Results of difference-in-differences matching analyses

In findings largely consistent with the DD and PSM analyses, the DDM analysis results demonstrate that other agricultural income (from livestock), the value of livestock inventories, and housing are higher for participating households (table 4, column 3). Income from livestock, activities (and other non-cropping agricultural activities) increased by 168 yuan; the family's housing value and livestock inventories also increased by 522 yuan and 220 yuan, respectively. The number of adults with off-farm work in households also increased slightly (by 0.05 persons). Crop income, as expected from a cropland set aside program, declined by 167 yuan, also consistent with the previous results.

Although most of the signs on the estimates are consistent with the cross-sectional PSM results, the magnitude of the coefficients differs to some extent between the two methods for some outcomes of interest (e.g., crop income and livestock inventories.) As described previously, DDM differs from cross-sectional PSM in that it removes any unobservables that change in parallel between participating and non-participating households conditional on the propensity score, suggesting that the conditional independence assumption of cross-sectional PSM does not hold in these data. Instead, it appears that a substantial amount of bias remains after conditioning on the covariates included in the propensity score specification. This bias that is parallel between the two groups is differenced out by the DDM estimator.

Discussion

In summary, using DD, PSM, and DDM approaches with different models, we have found that there have been positive, although somewhat nuanced, effects on participating households. The strongest finding is that participants have increased their non-cropping income and asset base to offset the fall in cropping income. Although we have not yet distinguished between different categories of households (poor versus relatively better off) as we do below, the findings do show that households in these fairly poor areas have benefited from the implementation of the largest Payment for Environmental Services mechanism in the developing world.

While it is unclear exactly why households enjoyed increases in some assets and livestock income and not greater participation in off-farm activities despite the direct payments and labor savings provided by the program, there are several plausible explanations. First, since we use cash accounting methods to measure assets, the higher direct income effects that might be associated with participation in the program may be being offset by lower realized income from families that chose to increase their holdings of livestock. If participating households decide to build up their livestock inventories (as our data suggest) in anticipation of higher incomes in the coming years, they would be deferring income (in a cash accounting sense). In other words, if accrual accounting methods had been used, there would have been higher incomes instead of assets. In fact, during a number of trips to the field, we observed that in many areas Grain for Green is accompanied by efforts of villagers to increase their livestock enterprises. Such a shift, in fact, might be expected since most of the compensation is paid in grain that

can be used for feed, and also since in some regions, households – which have more time available, especially after the first year of the program – are allowed to plant fodder, alfalfa, and other mulches that not only provide protection for the soil during the initial growth periods of the newly planted trees, but also can be used as feed.

Households also may be using the newly freed up labor in the first years of the program, not for activities that lead immediately to greater incomes, but for investing in other parts of their asset portfolios, including the construction of their home. In several interviews, farmers told us that with the reduced need to spend time in agriculture due to the lower labor requirement of the newly planted forests after the first year, they were able to spend their time on building a new house or renovating their old one. Housing in rural China is by far the largest asset in a household's portfolio and many households use hundreds of hours or more of their labor in building their house. This is especially true in some forested areas (where the Grain for Green program is mostly implemented) since households use their labor to cut trees and saw them into planks that can be used for constructing their homes. Our results would be consistent with a scenario in which many participating farm households have chosen to use the labor that they otherwise would have devoted to the cultivation (or some other activity) to build their houses.

One of our most surprising findings is that there was not a significant increase in off-farm employment activities.¹⁷ Interviews before and during the implementation of Grain for Green frequently found that local leaders and villagers expected that setting aside an appreciable amount of their land would shift labor into the off-farm sector. It could be, however, that such expectations have not been realized, since finding an off-farm job may require more than a desire to work off the farm; connections, human capital and social capital have often proved to be the most important determinants of access to off-farm jobs for rural households. Furthermore, since the household surveys collected information on labor allocation decisions for only the first three years after the initiation of the program, it may have been too soon to detect changes.

5. Impact on the poorer participants

So far we have addressed the program's impact on *typical* participating households. As our last step in the assessment of Grain for Green on poverty we examine whether or not the impact of the program was different for the relatively poorer participants compared to the relatively better-off participating households (which as discussed are often still absolutely poor – most being below US\$2/day). Different program impacts are plausible given the fact that livestock activities, for example, require labor investment

¹⁷ These findings are different from the case studies in Xu and Cao (2002), who found that non-agricultural income (income from off-farm wage earnings and self-employed sectors) increased on average after the program in all six case study counties. These case studies examine only sample statistics of the participating households and do not compare with the non-participating households; therefore, the income changes cannot be attributed solely to the program.

(for collecting feed and for increased care), which may be more attractive for the poorer households who have fewer outside options and thus lower shadow wages compared to some of the wealthier participating households. In addition, constructing a house requires not only labor but also investment into housing materials, which should be more readily available for better-off households.

To test whether or not the program impacts depend on the level of poverty, we first divided all households into quantiles based on total income in 1999. Next, to examine the poverty effect, we estimate the following county fixed-effect regression

$$\Delta Y_i = \beta_0 + \sum_{j=1}^4 \beta_j \text{Quantile}_j * \text{Participation}_i + \delta Z_i + \gamma a_c + \varepsilon_i,$$

where ΔY_i is the change in our outcomes of interest (i.e., income, asset holdings, and labor allocation), $\text{Quantile}_j * \text{Participation}$ is an interaction term between a quantile dummy (equals 1 if the household belongs to the j th quantile) and a participation dummy, Z_i is, as before, a vector of household characteristics, and a_c is a county fixed effect to control for time-invariant unobservables at the county level. In essence, this approach jointly estimates difference-in-differences for each income quantile. We will then test the null hypothesis $\beta_1 = \beta_2 = \beta_3 = \beta_4$ using a joint F-test.

The impact of Grain for Green on poverty

The results from the regression analyses suggest that the program does not systematically favor the poor or the better-off participants (table 5). The program impact on value of fixed assets is clearly largest for participants in the richest quantile (column 8, row 4). In fact, the value of fixed assets was the only variable in which the null hypothesis of all coefficients being equal was rejected (column 8, last row). The point estimates of program impact on other variables such as crop income (column 2), other agricultural income (column 3), and livestock (column 9) also differ across quantiles. The joint F-test, however, was not rejected in those models, suggesting that the program impact does not differ across quantiles.

Overall, the results suggest that within an area that is relatively poor, there are few distributional effects, in that there are minimal statistically significant differences between the impacts on the poorer and the better-off participating households. The main reason why we do not find systematic differences in program impacts may be because the majority of households in our sample are poor, both in relative and absolute terms, and their surrounding market conditions are similar so that the ways participants are affected through program compensation and use their freed-up labor do not differ after controlling for household characteristics and county fixed effects.

Impact on future income streams

In addition to the short-term program effects on income and household assets, one added potential effect on participating households is the future income stream associated with the forest and non-timber commodities that

Table 5. *Estimated impact of the Grain for Green program and initial income quantile on changes in income, labor allocation and asset holding, 1999 to 2002*

<i>Dependent variable: change in income, labor, asset holding (2002–1999)</i>									
	<i>Total income</i>	<i>Crop income</i>	<i>Other agricultural income</i>	<i>Non-agricultural income</i>	<i>Off-farm work</i>	<i>Migration status</i>	<i>House value</i>	<i>Fixed assets</i>	<i>Livestock</i>
Quantile1*participate	111.67 (0.60)	-122.14** (-3.15)	67.23 (0.52)	34.60 (0.37)	0.003 (0.06)	0.092 (1.07)	388.23 (0.90)	798.37 (1.37)	56.12 (0.43)
Quantile2*participate	203.33 (1.14)	-98.27* (-2.59)	336.13* (2.65)	-60.06 (-0.66)	0.007 (0.13)	-0.068 (-0.79)	154.24 (0.37)	-49.18 (-0.09)	159.73 (1.24)
Quantile3*participate	128.92 (0.78)	-83.32* (-2.37)	209.12* (1.75)	6.83 (0.08)	-0.022 (-0.46)	-0.096 (-1.19)	166.32 (0.42)	-46.60 (-0.09)	316.27** (2.62)
Quantile4*participate	97.43 (0.53)	-98.88* (-2.59)	253.48* (1.96)	-84.38 (-0.90)	0.007 (0.13)	-0.029 (-0.33)	334.51 (0.79)	1767.81** (3.00)	130.94 (1.00)
Ho: All coefficients equal	Not rejected	Not rejected	Not rejected	Not rejected	Not rejected	Not rejected	Not rejected	Rejected	Not rejected

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Households were divided into quantiles based on total income in 1999 (before program) as shown in appendix table. Quantile 1 is the poorest and quantile 4 is the richest. All models control for base value of the outcome variable in 1999, household size, land holding per capita, household head's age and education, a dummy for a female household head, maximum slope among household's plots, weighted average of distance from plots to a road, water source, and house, and county fixed effects.

Table 6. *Types of trees planted in the Grain for Green program by income quantile, 1999 to 2002*

Tree type	Quantile of household income in 1999				All participants
	1st quantile	2nd quantile	3rd quantile	4th quantile	
Ecological trees	43%	51%	61%	68%	56%
Cash crop trees	50%	43%	33%	41%	41%
Grass	20%	8%	7%	14%	12%
Grass and trees	32%	14%	21%	8%	19%
Others	4%	0%	0%	2%	1%
Number of participating households in each quantile	56	65	72	59	252

Note: Each household may plant one or more types of trees. The households were grouped into quantiles based on income in 1999 using the whole sample. *Source:* Authors' survey.

are created during Grain for Green. While it is too early to precisely quantify the future income stream, our data indicate that households in general should benefit, but the expected returns vary, depending on the types of trees that were planted. The two types of trees with large shares of the newly planted forestry areas are the so-called 'ecological trees', such as Chinese fir and Japanese cedar, and 'cash crop trees', such as chestnut, walnut and peach. In our sample ecological trees and cash crop trees were planted by 56 per cent and 41 per cent of the participating households, respectively. While the cash crop trees are generally expected to generate revenue from fruit and nut production within a few years, ecological trees will not be profitable to harvest for at least a decade.

Interestingly, there are differences in types of trees planted depending on a household's income level (table 6). The households in the fourth quantile (the richest quantile) tend to plant more ecological trees (68 per cent) and less grass (14 per cent) and grass-tree intercropping mixes (8 per cent), whereas the households in the first quantile (the poorest quantile) tend to plant not only ecological trees (43 per cent) and cash crop trees (50 per cent) but also more grass (20 per cent) and grass-tree intercropping mixes (32 per cent).¹⁸

¹⁸ The distribution of tree types may reflect the suitability of tree types for different geographical conditions. For example, Gansu Province is generally flatter and is more suitable for livestock grazing than Sichuan Province, and thus more suitable to plant grass or grass/trees. In our sample, Gansu Province has disproportionately more households in the 1st quantile, where Sichuan Province has disproportionately more households in the 4th quantile (appendix table). Therefore the observation that more households in the 1st quantile plant grass

The tree types that the participating households chose have implications on the extent of income smoothing from on-farm activities and availability of labor for off-farm work. The poorer participants that plant grasses typically grow fodders to feed the livestock. This may help them smooth the income stream from on-farm activities compared to planting only ecological trees, which only become profitable after one or more decades. The ability to smooth income, however, comes at the cost of increased on-farm labor; planting and using fodder for livestock requires more on-farm labor input. The richer participants, on the other hand, tended to plant more ecological trees that only generate income years later but require less labor during the growth period. While the program impact on the long-term income stream is difficult to predict, the differences in tree types may suggest that the richer participants have the potential to achieve the structural change from on-farm to off-farm labor sooner than the poorer participants.

6. Conclusion

Our paper has examined the effect of China's PES program on the welfare of rural households. Using a number of alternative approaches, the study found that the income of participants from livestock activities and some asset holdings have clearly increased more than those of non-participants. Even though the majority of participating households may not have participated voluntarily and the actual delivery of payments in some areas have been shown to fall short of the original promised amounts (Xu *et al.*, 2004a), the program has had positive effects. Although the program did not disproportionately favor poorer population, most of the households in the targeted program regions are poor both in relative and absolute terms; since the participants are better off on average, the program is benefiting the poor. Hence, in addition to providing environmental services (most observers agree that soil erosion has been greatly reduced), China's Grain for Green program has raised the wealth of villagers who are mostly quite poor.

Somewhat surprisingly, the results do not generate strong evidence for the expected finding that participating households have changed labor allocation by shifting labor into off-farm work. Nevertheless, in the analysis using the most stringent assumptions, there is weak evidence that participating households have begun to marginally increase the allocation

or grass/trees and that Gansu has disproportionately more households in the 1st quantile suggest that the tree types may be reflecting geographical and climatic conditions. We also should note that only 36 per cent of the participating households had a choice over which trees to plant. When we repeated the same analysis for those households who had the tree type choice ($n = 89$), we still find that richest households tended to plant more ecological trees and less grass, whereas the poorest households tended to plant not only ecological and cash crop-type trees but also grass and grass/cash crop-type tree intercropping mixes. We also found that a somewhat smaller share of households chose to plant ecological trees and more chose instead to plant cash crop trees, grass, or grass/trees. There were no major differences found in the impact of the program when considering whether or not a household was consulted before the program or not.

of household labor to off-farm activities. Our empirical results suggest that so far much of the freed up labor from cultivation and program compensation, most of which has been paid as in-kind grain, have been used to build up livestock. Moreover, it may be too early in the life of the program to have seen large changes, and it is also possible that additional efforts, such as training and other human capital increasing programs, are needed to provide villagers with the tools and connections needed to help households move their activities off the farm.

While the results illustrate the potential for a PES program to benefit the participants, our analysis necessarily suggests caution in assigning more widespread implications. First, it is difficult to draw nationwide conclusions based on a small sample from three provinces. Based on the same sample, Xu *et al.* (2004a) found that many participating households were not consulted prior to the program and that the actual compensation delivered did not always meet the level that was promised *ex ante*. In other regions in China, the compensation may have been delivered as promised, which would increase the effect of the program on poverty alleviation.¹⁹ We did not find evidence that the compensation level is any higher or lower relative to the opportunity cost for households who participated voluntarily; but because our sample sites only covered a small share of the entire program's command area, it may not be true for other regions that were not included in our survey. Lastly, although survival rate of the planted seedlings was not a concern in our sample (above 70 per cent for almost all villages reported in 2002), other regions may suffer from low survival rate and fail the inspections necessary to receive the compensation.

Additional research also is needed, since three years is almost certainly too short of a time to assess the impact of a program on any more fundamental structural transformations. An effort also should be made to understand the mechanism by which Grain for Green led to higher income, asset holdings, and increased off-farm labor. We are concerned that if farmers are unable to find alternative off-farm income sources during the program period and if they are not allowed to convert the program land back into cultivation when the program period ends, they may end up being trapped in a state of poverty that was worse than before the program.

These caveats aside, China's Grain for Green program appears to be an example of a successful PES program that is being implemented in a developing country. While the program officers of the Grain for Green program do not seem to disproportionately target the poor, by far the majority of the households in the program areas are relatively poor. Given this fact, even without explicitly targeting the poor, the program clearly has had a positive impact on many of the poor participants. Such results may mean that as the world continues to invest in PES programs there will be positive implications for poverty alleviation.

¹⁹ There is no reason, however, to believe that the households in our sample were any more or less likely to be in regions that systematically underpay them for their tree planting activities.

References

- Dehejia, R.H. and S. Wahba (1999), 'Causal effects in nonexperimental studies: reevaluating the evaluation of training programs', *Journal of American Statistical Association* **94**: 1053–1062.
- Dehejia, R.H. and S. Wahba (2002), 'Propensity score-matching methods for nonexperimental causal studies', *The Review of Economics and Statistics* **84**: 151–161.
- Dex, S. (1991), 'The Reliability of Recall Data: A Literature Review', Occasional Papers of the ESRC Research Centre on Micro-Social Change No. 6, University of Essex, Colchester.
- Landell-Mills, N. and I. Porras (2002), *Silver Bullet of Fools' Gold? A Global Review of Markets for Forest Environmental Services and Their Impact on the Poor*, London: International Institute for Environment and Development (IIED).
- Lewis, W.A. (1954), 'Economic development with unlimited supplies of labour', *Manchester School* **22**: 139–191.
- Li, H., S. Rozelle, and L. Zhang (2000), 'Micro-credit programs and off-farm migration in China', *Pacific Economic Review* **9**: 209–223.
- Meyer, B.D. (1995), 'Natural and quasi-experiments in economics', *Journal of Business and Economic Statistics* **13**: 151–161.
- Nurkse, R. (1957), *Problems of Capital Formation in Underdeveloped Countries*, Oxford: Blackwell.
- Pagiola, S. (2002), 'Paying for water services in Central America: learning from Costa Rica', in S. Pagiola, J. Bishop and N. Landell-Mills (eds), *Selling Forest Environmental Services: Market-based Mechanisms for Conservation and Development*, London: Earthscan Publications Ltd, pp. 37–62.
- Pagiola, S., N. Landell-Mills, and J. Bishop (2002), 'Market-based mechanisms for forest conservation and development', in S. Pagiola, J. Bishop and N. Landell-Mills (eds), *Selling Forest Environmental Services: Market-based Mechanisms for Conservation and Development*, London: Earthscan Publications Ltd.
- Pagiola, S. and G. Platais (2005), 'Can payments for environmental services help reduce poverty? An exploration of the issues and evidence to date from Latin America', *World Development* **33**: 237–253.
- Park, A., S. Wang, and G. Wu (2002), 'Regional poverty targeting in China', *Journal of Public Economics* **86**: 123–153.
- Rosenbaum, P. and D.B. Rubin (1984), 'Estimating the effects caused by treatments: comment', *Journal of the American Statistical Association* **79**: 26–28.
- Sala-i-Martin, X. (2002), 'The World Distribution of Income (Estimated from Individual Country Distributions)', *NBER Working Paper Series* **8933**.
- Skaggs, R.K., R.E. Kirksey, and W. M. Harper (1994), 'Determinants and implications of post-CRP land use decisions', *Journal of Agricultural and Resource Economics* **19**: 299–312.
- Smil, V. (1993), *China's Environmental Crisis: An Inquiry into the Limits of National Development*, Armonk, NY: M.E. Sharpe.
- Smith, J. (2004), 'Evaluating the local economic development policies: theory and practice', in A. Nolan and G. Wong (eds), *Evaluating Local Economic and Employment Development*, Paris: OECD Publishing.
- Smith, J. and P. Todd (2005), 'Does matching overcome Lalonde's critique of nonexperimental estimators?', *Journal of Econometrics* **125**: 305–353.
- State Forestry Administration (2002), *Master Plan for the Sloping Land Conversion Program*.
- Uchida, E., J. Xu, and S. Rozelle (2005), 'Grain for Green: cost-effectiveness and sustainability of China's Conservation Set-aside Program', *Land Economics* **81**: 247–264.

- World Bank (2003), *World Development Indicators*, Washington, DC: The World Bank.
- Xu, J. and Y. Cao (2002), 'Efficiency and sustainability of converting cropland to forest and grassland in the Western region', in J. Xu, E. Katsigris, and T. A. White (eds), *Implementing the Natural Forest Protection Program and the Sloping Land Conversion Program: Lessons and Policy Recommendations*, Beijing, China: China Forestry Publishing House.
- Xu, Z., R. Tao, and J. Xu (2004b). *Sloping Land Conversion Program in West China – Analysis of Household Impact and Sustainability*, Beijing: Center for Chinese Agricultural Policy.
- Xu, Z., M.T. Bennett, R. Tao, and J. Xu (2004a), 'China's Sloping Land Conversion Programme four years on: current situation, pending issues', *International Forestry Review* 6: 317–326.
- Zuo, T. (2002), 'Implementation of the SLCP', in J. Xu, E. Katsigris, and T. A. White (eds), *Implementing the Natural Forest Protection Program and the Sloping Land Conversion Program: Lessons and Policy Recommendations*, Beijing: China Forestry Publishing House.

Appendix table: *Distribution of household income by province, 1999.*

Province	Quantile of household income in 1999				Number of participants in each province
	1st quantile	2nd quantile	3rd quantile	4th quantile	
Shaanxi	15%	25%	40%	19%	84
Gansu	35%	28%	20%	18%	97
Sichuan	13%	24%	27%	37%	71
Number of participating households in each quantile	56	65	72	59	252

Note: The households were grouped into quantiles based on income in 1999 using the whole sample (participating and non-participating households.)
Source: Authors' survey.