

Did China's Tax-for-Fee Reform Improve Farmers' Welfare in Rural China?

James Alm and Yongzheng Liu*

Abstract

China enacted a rural tax reform known as the “Tax-for-Fee Reform” (TFR) in the late 1990s, in large part driven by an attempt to address farmers’ complaints about their perception of a heavy tax burden. A crucial but unanswered question is whether this reform actually improved farmers’ welfare. This paper applies several estimation methods, including both cross-sectional propensity score matching and difference-in-difference propensity score matching methods, to village-level survey data from the Chinese Household Income Project in order to examine the causal effect of the TFR on farmers’ direct and indirect welfare. “Direct” welfare is measured by net income per capita in the village, and “indirect” welfare is proxied by changes in per capita public expenditure composition. We find no evidence that the direct welfare effects of the reform improved farmer’s net income. In contrast, the reform appears to have reduced the villages’ financing capacity, and hence to have lowered their overall fiscal expenditures, especially expenditures on local welfare programs such as education, public health, and infrastructure. These indirect effects have had significant and negative impacts on farmers’ welfare. Indeed, our analysis suggests that it is poorer villages that largely undertook these expenditure adjustments, while richer villages experienced no significant changes in these welfare relevant categories.

Keywords: Tax-for-Fee Reform; welfare; public expenditures; rural China.

JEL classifications: H7, I2, I3, O1, O5, P3.

* Tulane University and Georgia State University. Please address all correspondence to James Alm, Professor and Chair, Department of Economics, 208 Tilton Hall, Tulane University, New Orleans, LA 70118 (phone +1 504 862 8344; fax +1 504 865 5869; email jalm@tulane.edu).

1. Introduction

In the late 1990s, the government of the People's Republic of China (PRC) enacted a rural tax reform known as the "Tax-for-Fee Reform" (TFR), as a response to rural farmers' bitter complaints about what they saw as a heavy fiscal burden. The reform was intended largely to reduce this fiscal burden (as well as to improve local governance). The reform was first formally introduced on a local pilot village basis in 2000, and then was widely carried out in tens of thousands of villages across the nation.¹ This paper utilizes the dramatic policy changes introduced by the TFR, both over time and over regions, to examine its impact on farmers' welfare. In doing so, we are able to quantify the impact of public policy in promoting rural development in a developing country like China.

Although a number of papers exist on rural tax reform in China, few have empirically evaluated the impact of this reform on farmers' welfare. In large part the absence of these studies is due to the absence of detailed information in rural areas, a problem that is common in many developing countries (Dethier, 1999). Lin and Liu (2007) describe the historical evolution and performance of rural tax reform in great detail, and Yep (2004), Li (2006), Lin and Liu (2007), and Tao and Qin (2007) provide some suggestive evidence that the fiscal burden on farmers was reduced following the introduction of the TFR.² However, a reduction in farmers' fiscal burdens may not necessarily imply an increase in farmers' net income, given the complicated process by which income in rural areas may be affected by different aspects of the reform. For one thing, the TFR brought a direct saving to farmers via the elimination of local

¹ As a further step toward reducing farmers' burdens, all agricultural-related taxes were completely eliminated in year 2006 by the central government. However, given data availability, we examine only the impacts of the TFR, rather than the abolition of agricultural taxes implemented in the later years after our sample information was collected.

² For example, Tao and Qin (2007) calculate that the burden of taxes and fees per capita dropped by more than half from RMB 145 in 2000 to RMB 72 in 2004 for the 6 surveyed provinces in their sample.

fees, which should increase farmers' net income, even though this change is likely to be temporary and may diminish over time. Also, the resulting increases in the rate of agricultural taxation after the TFR may adversely affect agricultural production and consequently decrease farmers' income (Mushtaq et al. 2008). Further, the TFR largely reshaped the landscape of local governance in rural China, which in turn affected some village-specific factors in the determination of farmers' income (e.g., the provision of production inputs, human capital, and social capital at the village level) (Sato, 2006).

Of some importance, the TFR also affected the resources available to local governments, with possible effects on farmers' welfare via changes in the actions of these governments (Yi, 2006; Li, 2006). As emphasized by Luo et al. (2007), the TFR may have introduced greater fiscal discipline at the local level, and may thereby have also imposed constraints on public goods provision due to the structural changes of revenue sources and expenditure composition.³ Several studies have in fact suggested that local governments in rural China experienced a dramatic fall in revenues for financing basic public services provision in the post-TFR period (Zhang et al., 2004; Fork and Wong, 2005; Luo et al., 2007). In a recent study, Meng and Zhang (2011) use a two-year panel of village data from rural China to measure these impacts on local budgets and, especially, to demonstrate their effects on local governance. Due to the large disparities of fiscal capacities in their upper-level governments and the significant differences in the development of off-farm industries, these impacts are also generally known to be geographically different across the nation. While in rich (e.g. eastern) provinces the elimination of the rural fees was offset with transfers to local governments from upper-level governments

³ Especially in rural China, methods to improve the efficiency of rural public finance, including the delivery of public goods, has always been a challenge for policymakers (Zhang et al., 2004).

and/or smoothed out by the revenue collection from rural industries, this was not the case in many of the poorer (e.g. central and western) provinces (Yep, 2004; Kennedy, 2007).

All of this work suggests that the TFR may have significantly affected the welfare of farmers across several dimensions. Given the vast numbers of people in rural China, measuring these impacts is a crucial issue in evaluating this – and other – reform efforts. However, evidence on the direct and indirect impacts of the reform remains elusive, and indeed to our knowledge there is no systematic quantitative evaluation of the TFR on farmers' welfare.

In this paper we address this gap by conducting a comprehensive analysis of the welfare implications of the reform based on a detailed village survey data set. Specifically, we examine in this paper whether the TFR has reduced farmers' fiscal burden, and also whether the TFR has affected farmers' overall welfare. We apply several estimation methods, both cross-sectional propensity score matching and difference-in-difference propensity score matching methods, to village-level survey data from the Chinese Household Income Project (CHIP) in order to examine the causal effect of the TFR on farmers' direct and indirect welfare. We are able to measure both direct and indirect impacts on farmers' welfare, in ways that deal with potential selection issues. We measure the "direct" effect of the TFR on farmers' welfare by the change in net income received by farmers after the reform. We measure the "indirect" effect by changes in their receipt of benefits from village level provision of public services, using changes of the composition of public expenditures as a proxy for this indirect welfare effect of the reform. We find no evidence that the direct welfare effects of the TFR improved farmer's net income. In contrast, the TFR appears to have reduced the villages' financing capacity, and hence to have lowered their overall fiscal expenditures, especially expenditures on local welfare programs such as education, public health, and infrastructure. These indirect effects have had significant and

negative impacts on farmers' welfare. Indeed, our analysis suggests that it is poorer villages that largely undertook these expenditure adjustments, while richer villages experienced no significant changes in these welfare relevant categories. Overall, in the absence of sufficient compensation from upper-level governments either in the form of additional transfers or of realigned expenditure assignments, we conclude that the TFR seems likely to have lowered overall farmers' welfare in poor villages. This quantitative result extends and complements the more qualitative arguments of Yep (2004) and Kennedy (2007).

The remainder of this paper is organized as follows. Section 2 provides some background introduction on the tax system in rural China, including its reform via the TFR. Section 3 develops the empirical framework to evaluate the impact of the TFR on farmers' welfare, and also discusses our data. We present our results in section 4 and our conclusions in section 5.

2. The Tax System in Rural China and the Tax-for-Fee Reform

Ever since the establishment of the People's Republic of China in 1949, the tax system in rural China has been in flux, with changes driven by the overall national development strategy and by the role of agriculture in the overall economy (Wang, 2008). There have been three broad stages in the local tax system. The first stage spans the period from the establishment of the PRC until the dawning of the "open-door" reforms in 1978. During this period there was an explicit state agriculture tax imposed on farmers at a rate around 10 percent and collected mainly through a mandatory procurement system; there were also a number of other agriculture taxes imposed implicitly by increasing the prices of agricultural inputs and depressing the prices of

agricultural outputs (e.g., so-called “price scissors”).⁴ The main purposes of agricultural policies during this period were to collectivize the sector and to extract agriculture resources to support the priority industrialized sector, as driven by the choice of a heavy-industry-oriented development strategy in China’s planned economy period (Lin et al., 2003).

A second stage began in the late 1970s and extended to the early 1990s. This period was characterized by de-collectivization reforms in the agriculture sector. The Household Responsibility System (HRS) was adopted, which recognized the legal status of farmers in claiming the farming production residual after they had fulfilled the required grain quota and agricultural taxes and fees (Lin, 1992; De Braw et al., 2004). As a consequence of the HRS reform, the mandatory grain quota system was abolished and replaced by grain procurement contracts under which the state continued to tax agriculture implicitly through the price margin between the state grain sector and the market (Lin and Liu, 2007). In this stage, two major categories of fees were introduced (in addition to the implicit taxes) to offset revenue reductions in township governments and villages arising from the HRS reform: five “township-pooling funds” and three “village levies”.⁵ There were also many illegal levies in the form of fines and financial contributions for expenditure or public projects. As noted by Aubert and Li (2002), it is very difficult to estimate the value of these illegal levies, or even to list all of them.

The surge of these diverse and often illegal local charges on farmers (especially those in heavily agriculture-based areas and in poorer regions) generated rising opposition among farmers, which began in the 1990s to threaten rural social stability and even to endanger the

⁴ In addition, farmers bore the burden of some informal fees imposed both by local communes to finance local public services and by production brigades to pay for cadre compensation.

⁵ The township-pooling funds included education supplements, social help, family planning, collective transportation, and militia exercises. The village levies refer to the public accumulation fund, the public welfare fund, and administrative fees.

state's political legitimacy (Bernstein and Lu, 2000; Tao and Qin, 2007).⁶ Indeed, large-scale protest, even conflict, against local authorities was often observed during the process of taxes and fees collection (Aubert and Li, 2002; Chen, 2003).

These developments led to the third stage of agricultural policies. To accommodate the farmers' bitter complaints, the central government launched the Tax-for-Fee Reform in the late 1990s, with the broad goal of reducing the fiscal burden on farmers. The TFR had several main features. First, all existing township and village levies, including the previous five township pooling funds, the three village levies, and other kinds of informal local fees, were abolished. Second, to substitute for these reduced local charges, there was an increase in the rates of the agriculture tax and agriculture tax supplements. Third, "Case-by-case fundraising" (or *yishiyiyi*) was introduced to finance special public projects, and budgetary transfers from upper-level governments were adjusted to accommodate local needs. This reform was first formally introduced on a local pilot basis in 2000, and was then carried out nationally. By the end of 2002, 20 of the 31 provinces in China (including municipalities and autonomous regions) had commenced the TFR on a pilot basis (Tao and Qin, 2007). As a further step to reduce the tax burden on farmers, all agricultural-related taxes of the central government were completely eliminated in 2006.⁷ There is in fact some anecdotal evidence that the implementation of the TFR did in fact reduce farmers' fiscal burdens. For instance, in Anhui province a 31 percent burden reduction across the whole province was reported for the first year of the reform (Yep, 2004).

Our study is able to document the changes in the fiscal balance sheet of villages with and without reform between 1998 and 2002. In our sample, 1998 is the initial year prior to any

⁶ For a detailed analysis of the increasing fiscal burden on farmers in the 1990s, see Lin and Liu (2007).

⁷ Remember that data limitations prevent us from analyzing this later step of the overall reform process.

reform in any village. Those villages that enacted the TFR in some year before 2002 are grouped as treated villages, and other villages are grouped as control villages. As shown in Table 1, the TFR had a significant impact on the village balance sheets. In control villages (i.e., where the TFR was not in effect), village total revenues per capita rose by 8.98 percent annually between 1998 and 2002. This same measure declined by 6.26 percent annually in the same period in treated villages. Although local fees per capita had been fully eliminated in treated villages in 2002, transfers per capita in 2002 were only enough to cover 63.91 percent off the loss of local fees collected in 1998 (or 10.49 yuan per capita). Also, total expenditures per capita rose 7.48 percent more annually in control villages than in treated villages from 1998 to 2002, implying that treated villages responded to fiscal shortfalls by cutting expenditures.

It seems certain that these many changes induced by the TFR had many, potentially conflicting, effects on farmers' welfare. Any decline in fiscal burdens should have directly increased their welfare; however, any changes in local government services could well have reversed these effects through indirect effects. The next section presents our methodology for quantifying these direct and indirect effects.

Note, however, that the TFR had wide-ranging impacts on the landscape of the fiscal balance sheet in the villages, as well as on the budgets of other levels of government. As we show later, village revenues declined significantly due to the termination of informal fundraisings and the absence of sufficient transfers from upper-level governments to replace these funds. At the same time, some village expenditure assignments were assigned to higher levels of government. For example, higher-level governments (mainly county governments) were required to assume a larger responsibility for compulsory education costs in rural areas as a complementary policy to the TFR. These changes in expenditure responsibilities across levels of

government suggest that the changes in village budgets alone may not be sufficient to capture the overall changes of farmers' welfare. Detailed information on county-level government budgets is not available in our data. However, we believe that it is unlikely that a significant change in expenditure assignments across levels of government did in fact occur, at least changes that would bias our results. For example, Kennedy (2007), and Sato (2008) show that county governments did not cover all education funding losses at the village level. Also, although in principle upper-level governments were supposed to provide sufficient fiscal transfers to compensate for the loss of revenues in the villages, in practice, they generally failed to do so (Li, 2006). We discuss the ways in which changes in assignments across levels of government might affect our results in more detail later.

3. Empirical Strategy: Methods and Data

Our strategy uses village-level survey data from the Chinese Household Income Project (CHIP) to quantify changes in farmers' welfare, and applies both propensity score matching and difference-in-difference methods to these data. We first introduce the estimation methods, and we then discuss our data in detail.

3.1. Propensity Score Matching Estimation

Let TR_i be an indicator of whether the TFR is implemented in village i , defined as 1 if reform is enacted and 0 otherwise. Let Y_{i1} be the observed value of outcome variables (e.g., net income per farmer) for village i following the implementation of the reform. Also denote Y_{i0} as the observed value of the outcome variables if the TFR had not been implemented in the village. The treatment effect τ_i from the TFR for village i can be written as:

$$\tau_i = Y_{i1} - Y_{i0}. \tag{1}$$

The fundamental problem of program evaluation arises because we can only observe one of the outcomes for each village, either Y_{i1} or Y_{i0} . Assessing the impact of the TFR requires making an inference about what would be the counterfactual outcome in a non-reform state for villages where the TFR has indeed been implemented. Therefore, we focus instead on the average treatment effect of the reform on the villages, where the TFR has been in place (ATT), defined for village i as:

$$\begin{aligned} ATT_i &\equiv E[\tau_i | TR_i = 1] \\ &= E[Y_{i1} - Y_{i0} | TR_i = 1] = E[Y_{i1} | TR_i = 1] - E[Y_{i0} | TR_i = 1]. \end{aligned} \quad (2)$$

However, the counterfactual mean for the last term in equation (2), or $E[Y_{i0} | TR_i = 1]$, is not observed. Consequently, we have to choose an appropriate substitute in order to estimate the average treatment effect of the reform on the treated villages.

Now the average outcome value of control villages, or $E[Y_{i0} | TR_i = 0]$, is a valid approximation for $E[Y_{i0} | TR_i = 1]$, as long as the selection of treated and control villages was a random process under the experimental design of the TFR implementation. However, this assumption seems unlikely to hold, given that factors that determine the implementation of the reform in one village may also simultaneously determine the outcome variables of interest.

Using $E[Y_{i0} | TR_i = 0]$ would likely lead to selection bias, given the systematic difference of outcomes between treated and control villages even in the absence of the TFR.

Following the microeconomic evaluation literature (Dahejia and Wahba, 1999, 2002; Lee, 2005), we address this selection bias by using matching techniques to construct a valid counterfactual estimator for the average outcome of treated villages. The underlying logic of the matching estimator is to construct an “artificial” experimental subset of the original sample in such a way that, conditional on observed characteristics X_i of village i , the selection process of

the implementation of the TFR is random. As shown by Rubin (1977), if the outcomes of the TFR are assumed to be independent of program participation after conditioning on a set of covariates, then the average treatment effect ($\tau|TR = 1$) is equal to $(\tau|TR = 1, X)$, averaged over the distribution of $(X|TR_i = 1)$.

To implement this approach, we first find a set of comparable control villages for each treated village on the basis of similarity of observable characteristics. We then compute the difference in the outcome variables of interest and take its mean. This procedure is straightforward if there are only a few covariates. However, with an increase in the dimensions of covariates, this method become difficult to implement because of the difficulty of finding exact matches for each treated village.⁸

We therefore adopt the propensity score matching approach pioneered by Rosenbaum and Rubin (1983) in order to reduce the dimensionality of matching problem. This approach creates a summary measure of similarity in the form of a “propensity score”. To implement this, we first estimate the probability of the TFR being implemented in village i using a binary discrete choice model, or:

$$p(X_i) = Prob(TR_i = 1|X_i), \tag{3}$$

where X_i denotes observed covariates for village i that are not affected by the implementation of the TFR (or the anticipation of it).⁹ We then match each treated village with a control village on the basis of the predicted probability of implementation of the reform (e.g., the “propensity score”). The average treatment effect of the TFR on the treated villages ATT is finally obtained by computing the expected value of the difference in the outcome variable between each treated

⁸ For instance, if n dichotomous covariates are contained in X_i , then the possible number of matches will be 2^n .

⁹ To ensure this, variables should either be fixed over time or measured in the time period before the implementation of the TFR (Caliendo and Kopeinig, 2008).

village and the matched control villages. As shown by Todd (2006), a standard matching estimator can be written as:

$$\hat{\tau} |_{TR=1} = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} [Y_{i1} - \hat{E}(Y_{i0} | TR = 1, p_i)] \quad (4)$$

$$\hat{E}(Y_{i0} | TR = 1, p_i) = \sum_{j \in I_0} W(i, j) Y_{j0}, \quad (5)$$

where I_1 denotes the set of treated villages, I_0 represents the set of control villages, S_p is the region of common support, $W(i, j)$ is a weighting function, and n_1 is the numbers of villages in the set $I_1 \cap S_p$. Denote a “neighborhood” of village i as $C(P_i)$, where P_i is the propensity score of village i .¹⁰ Then the village matched to village i is that village in set A_i such that $A_i = \{j \in I_0 \mid P_j \in C(P_i)\}$ and Y_{j0} is the corresponding outcome of the matched control village for each treated village i belonging to the set $I_1 \cap S_p$.

Note that the weighting function $W(i, j)$ in equation (5) assigns the weights for the matched control village j in constructing the counterfactual for the treated village i . Depending on the different definitions of the neighborhood $C(P_i)$ and of the weighting function $W(i, j)$, several type of matching approaches have been proposed in the literature, such as nearest neighbor matching, caliper matching, and kernel matching. The “nearest neighbor matching” uses for each treated village only a single control village with the closest difference in propensity score to the treated village. “Caliper matching” is a variation of nearest neighbor matching that attempts to avoid “bad” matching by imposing a tolerance level on the maximum propensity score distance. The “kernel matching” method follows a nonparametric approach to match each treated village with a weighted average of all control villages, using weights that are inversely proportional to the distance between the propensity scores of treated and control group. The

¹⁰ Recall that the propensity score is the probability that the TFR is implemented in village i , generated from the estimation of equation (3).

nearest neighbor matching and kernel matching are the most commonly used approaches, but we use all three methods in our analysis. We implement the kernel matching estimator by using a weighted average of all villages in the control village group as a baseline estimator to construct the counterfactual outcome; we also construct nearest neighbor and caliper matching estimators to check the robustness of our results.

The propensity score matching method provides a reliable estimate for the average treatment effect of the TFR on treated villages only under some assumptions. First, the selection of the TFR must be independent of potential outcomes Y_{i1} and Y_{i0} , after conditioning on the propensity score (i.e. the “conditional independence assumption”). Second, the average treatment effect of the TFR on treated village must be computed only within the region of common support, which ensures that villages with the same pre-reform observable characteristics values have a positive probability of being assigned to the treated or the control village groups (Heckman et al., 1999).¹¹ The matching procedure can be checked to determine whether it is able to balance the distribution of the observed covariates in both treated and control villages; a lack of balance suggests either a misspecification in the model used to estimate the propensity score or a failure of the conditional independence assumption (Dahejia and Wahba, 2002; Smith and Todd, 2005). The “balancing test” compares the situation before and after matching, and checks to see whether there remain any differences in the groups after conditioning on the propensity score (Caliendo and Kopeining, 2008).¹²

We employ two different balancing tests: a standardized test of differences between the groups, and a t-test for the equality of each covariate means for both groups. The “standardized

¹¹ There are various ways of defining common support. One method used in Dehajia and Wahba (1999, 2002) is to discard the comparison units with an estimated propensity score less than the minimum or greater than the maximum estimated propensity score for treated units. We employ this method here.

¹² Differences are expected *before* matching is performed. *After* matching, the covariates should be balanced in both treated and control village.

test of differences” was developed by Rosenbaum and Rubin (1985), and has been widely used (Lechner, 1999; Sianesi, 2004; Caliendo and Kopeining, 2008). This test checks the balance by comparing the sample means of both treated and control villages as a percentage of the square root of the average variances in the corresponding sample before and after matching. The formulae for the standardized difference are:

$$SB_{before} = 100 \cdot \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\frac{V_1(X) + V_0(X)}{2}}} \quad SB_{after} = 100 \cdot \frac{\bar{X}_{1m} - \bar{X}_{0m}}{\sqrt{\frac{V_{1m}(X) + V_{0m}(X)}{2}}}, \quad (6)$$

where for each covariate \bar{X}_1 ($V_1(X)$) and \bar{X}_0 ($V_0(X)$) represent the mean (variance) for both groups before matching, and \bar{X}_{1m} ($V_{1m}(X)$) and \bar{X}_{0m} ($V_{0m}(X)$) represent the mean (variance) for both groups after matching. A sufficient reduction of the standardized difference before and after matching or a low enough standardized difference value after matching will be treated as a “good” balancing outcome. Although there is no consensus on how large a standardized difference should be defined as the threshold of identifying balancing outcome, Rosenbaum and Rubin (1985) suggest a value of 20.

We also employ the usual t-test, which performs a paired t-test between treated and control villages to check if there are significant differences in each covariate means for both groups.

3.2. Difference-in-difference Matching Estimation

An alternative estimation approach is the use of difference-in-difference (DID) matching methods, as defined in Heckman et al. (1997) and Heckman et al. (1998).¹³ Instead of matching on the basis of level as in propensity score matching (e.g., equation (4)), the DID matching approach matches on the basis of differences in outcomes (before and after treatment); that is, the treatment effect becomes:

¹³ We also conduct simple DID estimation as well.

$$\hat{\tau} |_{TR=1} = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} [(Y_{i1t} - Y_{i1t'}) - \hat{E}((Y_{i0t} - Y_{i0t'}) | TR = 1, p_i)], \quad (7)$$

where t and t' are the time period after and before the year of the reform. Since taking the difference in outcomes before and after treatment eliminates any individual-specific, time-invariant unobserved characteristics, the DID matching method has the additional advantage of avoiding the hidden bias due to unobservables between participants and nonparticipants.¹⁴ As with our propensity score matching estimation, we apply different definitions of the neighborhood and of the weighting function via our use of nearest neighbor matching, caliper matching, and kernel matching, in order to test the robustness of our DID matching estimation.

3.3. Data

Our data are based upon a survey from the Chinese Household Income Project (CHIP), which was conducted by the Institute of Economics at the Chinese Academy of Social Sciences with the assistance of the State Statistical Bureau in Beijing.¹⁵ The dataset consists of samples from both urban and rural populations in China, collected through a series of questionnaire-based interviews at the end of 2002. Ten separate datasets are created. The first four datasets (1-4) survey different living aspects (e.g. income, consumption) of individuals and households in urban areas; the last five datasets (6-10) contain similar living aspects for individuals and households in rural areas. The fifth dataset concentrates on village-level data, obtained by interviewing village leaders. Since our focus is on the effects of the TFR in which the village

¹⁴ To the extent that there are unobserved time-varying characteristics between treated and control villages, we are unable to account for these features. However, it is not clear what unobservable village characteristics could vary over time and across the two groups of villages. In addition, since the outcome variables (e.g., income, public expenditures) depend on many factors that are determined after the TFR (especially public expenditures), the selection problem of unobservables, if any, seems likely to be much weaker than the selection problem of observables. See Altonji et al. (2005) for a detailed discussion of selection on unobservables.

¹⁵ For the issue of the quality and representativeness of the CHIP data, see Gustafsson et al. (2006) for a detailed discussion. Other studies that have used this dataset include Gustafsson and Li (2002), Li and Zahniser (2002), Sicular et al. (2007), and Meng and Zhang (2011).

was the basic unit of implementation, the fifth village-level dataset in CHIP is a high quality dataset ideal for our purpose.

There are two main advantages of the village-level dataset. First, the information provided by the survey is very comprehensive. The survey contains 259 variables for 961 villages distributed across 22 provinces in China.¹⁶ These variables include nearly all aspects of the villages, such as basic geographic information, arable land, agriculture activities, collectives, enterprise, labor force, income, productivity, population, government budget, taxes, expenditures, local election results, and characteristics of village government officials. Such rich information is crucial for us to predict the determinants of the TFR implementation in our matching techniques. Second, the village-level dataset also provides information on almost all variables for periods both before and after the TFR implementation; that is, the survey uses the same questionnaire to provide data for years 1998 and 2002, asking each question for each of two years. This is significant because the conditional independence assumption requires that the variables included in the specification that predicts program participation be unaffected by the program treatment, which requires in turn that the variables should either be fixed over time or measured before participation (Sianesi, 2004; Smith and Todd, 2005). The pre-reform information for year 1998 is required in order to meet this requirement.

Our working sample in this paper is derived by imposing the restriction that only those villages have not introduced the TFR by the end of 1998 are included in the sample; several villages that implemented the reform before 1999 are excluded because we use year 1998 information as pre-treatment information. This ensures that all observations having the same

¹⁶ These 22 provinces are: Anhui, Beijing, Chongqing, Gansu, Guangdong, Guangxi, Guizhou, Hebei, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Jilin, Liaoning, Shanxi, Shandong, Shanxi, Sichuan, Xinjiang, Yunnan, and Zhejiang.

initial status in 1998 (e.g., pre-treatment status) can be grouped into either treated villages or control villages by the end of 2002. This restriction gives us 841 villages for both years.

For measuring the “direct” welfare effect of the reform on farmers, we use the annual net income per capita in the village (*netinc.pc*) as a proxy. We use various categories from collective expenditures in per capita term to reflect “indirect” welfare changes from the reform. There are nine main expenditure items under the village government collective expenditures account: collectively operated reproduction, productive service for farmers, education, public health, infrastructure, other commonweal expenditures, wage and subsidies for village and group cadres, other administrative expenditures, and other expenditures. We group the various specific expenditure items into welfare and non-welfare expenditure program in accordance with their properties and objectives. Collectively operated reproduction, productive service for farmers, education, public health, and infrastructure expenditures are grouped into “welfare” expenditure programs (*welexp1.pc*); wage and subsidies for village and group cadres and other administrative expenditures are treated as “non-welfare” expenditure programs (*nonwelexp1.pc*). Since the classification of the two remaining items – other commonweal expenditures and other expenditures – are less clear, we test for sensitivity of our basic results by adding them separately in the welfare and non-welfare programs, thereby creating two alternative measures of welfare and non-welfare expenditure programs (*welexp2.pc* and *nonwelexp2.pc*). Data on net income and public expenditures is converted into 1990 Yuan using the rural CPI published in *China Statistical Yearbook*.

In our estimation of the probability of the TFR being implemented in village i using equation (3), the choice of covariates is guided by the criterion that selected variables should influence simultaneously the TFR assignment and these various outcome variables (Dehejia and

Wahba, 1999, 2002). In the absence of previous research on the determinants of the TFR assignment, we introduce the covariates linearly, and perform the balancing tests to check whether we succeed in balancing the covariates within each stratum. If our balancing tests are successful, then we accept the specification; if not, then we add higher orders of the covariates until the balancing condition is satisfied.¹⁷

Definitions and summary statistics for the outcome and explanatory variables are shown in Tables 2 and 3. Differences in the (unadjusted) features of treated and control villages are presented in Table 4, alongside their t-values. The t-values suggest that the characteristics of treated and control villages are generally different, except in a few cases such as poverty status and net income per village before the TFR implementation. Comparison of the outcome variables indicates that treated villages exhibit deteriorating public services, as reflected in the decline of fiscal expenditures in all dimensions, while net income per village is not found to be significantly different between the two groups. However, as emphasized in our earlier discussion of methodology, these comparisons of mean differences do not account for the potential selection bias generated by other characteristics that may affect the implementation of the TFR and outcome variables simultaneously. Therefore, we now turn to the analysis of our propensity score matching and DID matching estimation results.

It should be noted that we also examine whether the TFR had different effects by region (e.g., “coastal” versus “inland” villages) and by income (e.g., “poor” versus “rich” villages); these definitions and the associated results are discussed in detail later. We have also estimated all of our various models and specifications using the levels of the direct and indirect measures

¹⁷ The twelve covariates in our basic specification are: dummy variables for whether the village is a suburb, a minority county, a poverty county, a pilot village for productive and non-productive reason, and a village election; the distance from the nearest county seat; the total population and the square of the total population of the village in 1998; the total planting area and the square of the total planning area in 1998; net income per capita in the village in 1990; and net income per capita in the village in 1998.

rather than the per capita measures; our results are largely unaffected, and so are not reported.¹⁸ Finally, we assess the possible effects of changing government responsibilities on our results, even though detailed information on these effects is not available in our CHIP data.

4. Empirical Results

4.1 Propensity Score Matching Estimation Results

We first estimate a specification that generates the propensity scores for the TFR, using a probit model to predict the probability of introducing the reform. Second, we perform a balancing test to check the success of propensity score estimation in balancing covariates between treated and control villages. Third, we generate kernel and other matching estimators to calculate the causal effect of the TFR on farmers' direct and indirect welfare.

Probit Estimation

Results of the propensity scores estimation are reported in Table 5. We estimate various specifications: all villages, coastal villages, inland villages, rich villages, and poor villages. (These classifications are discussed later.) In all cases, our probit estimation reveals the significant role of specific characteristics in the selection of villages for the reform.

Balancing and Common Support Evidence

Table 6 reports the standardized differences and regression-based balancing test results after Gaussian kernel matching. The standardized differences between treated and control villages are all less than 8 percent, and most are less than 4 percent. The bias reduction as a result of employing matching technique is substantial, as indicated in Table 6.

The balancing result based on standardized differences test is confirmed by the regression-based tests. The t-statistics reported in the last column of Table 6 demonstrate that we

¹⁸ All estimation results are available upon request.

fail to reject the hypothesis that the mean differences for all covariates between treated and control villages are equal to zero. Thus, both balancing tests suggest the effectiveness of our chosen propensity score specification in accounting for selection bias in our sample.

Since our objective is to make treated villages comparable to control villages in order to estimate the average treatment effect of the TFR on farmers' welfare, the common support condition is imposed to ensure that the matching estimation is taken in the region of common support. As a result, there are 684 (out of 841) observations in the common support region, of which 489 are treated villages and 195 are control villages. Figure 1 shows the histogram for the propensity scores before matching for both treated and control villages. This figure clearly reveals that the region of common support is ample, and in fact relatively few cases are dropped because they lie off the common support.

Kernel Matching Estimates

Having established that propensity scores are balanced and that the common support condition is justified, we conclude that the treated and the matched control villages are comparable, and we proceed to present the empirical results on the average treatment effect on the treated villages (*ATT*) from the kernel matching estimator in Table 7. These findings suggest several main results.

First, the estimates shown in the top of Table 7 indicate that the TFR had no statistically significant impact on farmers' direct welfare, where "direct" welfare is measured either via the level of net income or its log form. Note that there is weak evidence that the impact of the TFR on farmers' direct welfare may be different in rich and poor villages. The argument here is that the log form outcome variables generally exhibit higher statistical significance than the level form and that the log form puts more emphasis on an increase for poorer farmers because the

proportional effects are given more weight. Later, we explore this issue further by dividing villages into rich and poor groups based on the levels of net income per villager in year 1998.

Second, there is strong evidence of significant reductions in many (although not all) categories of public expenditures in treated villages versus matched control villages, consistent with other findings that the introduction of the TFR led to deteriorating village public expenditures in at least some expenditure categories (Zhang et al., 2004; Fork and Wong, 2005; Luo et al., 2007; Meng and Zhang, 2011). For example, total expenditures are significantly lower in treated villages than in matched control villages, and expenditures on our classifications of welfare programs are also generally lower in treated than in matched control villages. In contrast, our classifications of non-welfare expenditures show no statistically significant differences of the reform between treated and matched control villages. Overall, without sufficient compensation from upper-level governments, these results suggest that farmers' indirect welfare was reduced by the TFR, where "indirect" welfare is measured by village government expenditures on welfare-related categories.

Altogether, we find no evidence supporting the existence of a direct welfare effect of the TFR on improving farmer's net income in rural China. In contrast, the TFR appears to have reduced villages' ability to finance expenditures, thereby reducing total public expenditures and especially welfare expenditures on such services such as education, public health, and infrastructure. These latter reductions likely had a significant and negative impact on the indirect welfare of farmers.

Robustness Checks: Alternative Propensity Score Matching Estimates

How robust are these findings? We use two alternative matching estimates (e.g., caliper matching, nearest neighbor matching). As shown in Table 8, the results using either matching

method show no statistically significant direct welfare effects of the reform on our two measures of income. As for indirect welfare effects, the results from both matching methods are comparable to our earlier results in Table 7, indeed with somewhat larger, negative, and significant level estimates of the reform.

Robustness Checks: Regional Specific Effects

As further robustness checks, we examine whether the TFR had different effects in different regions. We group the full sample into “coastal” and “inland” villages,¹⁹ and we reapply the propensity score matching technique to each village group, using the relevant probit estimates in Table 5.²⁰

Similar to the findings in full sample, the results in Table 9 reveal no significant effect of the TFR on farmer’s direct welfare. However, reform had significant negative effects on all categories of village government expenditures in inland villages. Most reductions of total expenditures in inland villages occur by cutting expenditures on welfare program. As shown in Table 9, the absolute magnitudes of expenditure reductions for welfare program are much larger than the corresponding magnitudes for non-welfare program. In coastal villages, there is no noticeable change in most categories of expenditures (except one measure of expenditures on non-welfare program in log form is significantly positive). These results suggest that the reform had quite different impacts on expenditures in coastal versus inland villages. Since inland provinces are generally underdeveloped regions in China relative to coastal provinces, this finding may imply an especially serious negative effect of the TFR on poor villages. The large

¹⁹ Among the 22 provinces in our sample, we classify as coastal provinces Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shandong, Guangdong, and Guangxi; we group as inland provinces Shanxi, Jilin, Anhui, Jiangxi, Henan, Hubei, Hunan, Chongqing, Sichuan, Guizhou, Yunnan, Shanxi, Gansu, and Xinjiang.

²⁰ As before, the propensity score matching estimator for each village group is based on the specification that satisfies the balancing test and common support condition. See Table 5 for these estimation results. We do not report the results of the balancing test and the common support, but all results are available upon request.

decline of revenues after the TFR may have driven the poor villages to cut expenditures both on welfare programs and on non-welfare programs, such as wages and subsidies for village cadres.

To explore further the impact of the TFR on “poor” versus “rich” villages, we split all villages into income categories based on their net income per capita in year 1998 relative to the mean value of the full sample in 1998; that is, villages with net income per capita less than the sample mean value in year 1998 are defined as “poor” villages, while villages with greater value than sample mean are grouped as “rich” villages. Again, we reapply the propensity score matching technique to each village group, using the relevant probit estimates in Table 5.²¹

The result (Table 10) is that the impact of the TFR on farmers’ welfare is only statistically significant in poor villages, and again only significant for the indirect welfare aspects. Similarly, non-welfare expenditure programs are also reduced significantly in poor villages after the reform, but the reduced magnitude for non-welfare programs is much smaller than for welfare expenditure programs. In contrast, rich villages experienced no significant change in any relevant welfare aspects, direct or indirect.

4.2. Difference-in-difference Matching Estimation Results

Tables 11 to 14 report results from DID matching methods. As with our propensity score estimation results, we find no direct welfare effects in any of the cases, or for all villages (Tables 11 and 12), coastal and inland villages (Table 13), or rich and poor villages (Table 14)). Similar to our earlier results, we also find that inland and poor villages generally experienced significant indirect welfare losses from the reform, as indicated by reductions in most aspects of expenditures (Tables 13 and 14). Indeed, the reductions in welfare expenditures are typically much greater than the reductions in non-welfare expenditure. In contrast, there are no significant

²¹ Also as before, the propensity score matching estimator for each village group is based on the specification that satisfies the balancing test and common support condition. All results for the balancing test and the common support are available upon request.

changes in rich and coastal villages (Tables 13 and 14). These results largely confirm our earlier propensity score matching estimates.

4.3. Other Considerations

Overall, then, we find robust evidence supporting some previous arguments that the TFR likely worsened farmers' welfare in poor villages (Yep, 2004; Kennedy, 2007). To understand more fully our findings, it is important to note that the central issue of rural taxation in China before the reform was the increasingly regressive nature of rural taxes, rather than the increase in the average rural tax rate (Lin and Liu, 2007). Our findings therefore suggest that village expenditures before the TFR in poor villages were mainly financed by taxing villagers via informal taxes and fees, which in turn generated a heavy fiscal burden on (and increased resentment from) those living in these poor villages. Meanwhile, because of a higher level of industrialization in rich regions, public expenditures in rich regions were financed via taxes from the large non-agricultural tax base that were less regressive in their impact, so that the fiscal burden was not seen as excessive in these regions and also so that rich villages could continue to collect revenues from rural industries and provide local services after the TFR. However, the formalization of the tax system after the TFR ruled out the possibility of poor villages financing their expenditures through imposing informal taxes and fees on villagers, as they had done prior to the reform. In the absence of adequate post-reform transfers from upper-level governments, poor villages were forced to cut back on virtually all categories of public expenditures.²² In

²² It has often been noted that intergovernmental fiscal transfers in China play little role in equalizing the fiscal disparity among regions, largely because the distribution of its main component (tax rebates) seems mainly to protect the interests of the richer provinces that existed prior to the fundamental tax-sharing system reform in 1994 (Zhang and Martinez-Vazquez, 2003). As a result, the lack of progressivity in the distribution of central transfers ensures that the richer provinces obtain a larger proportion of central transfers, while the poorer provinces with less fiscal capacity and fewer transfers are adversely affected. Although there is a specific transfer program for rural regions with relatively lower fiscal capacities, the amounts are far from sufficient. For instance, the TFR has led to a reduction of approximately RMB 150-160 billion in agricultural taxes and fees in 2005 alone, while in the same year the central transfers were only RMB 66.4 billion (Tao and Qin, 2007). In addition, the transfer payments from upper-

contrast, because rich villages did not depend as heavily on informal taxes and fees for financing public expenditures, their levels of public expenditures did not change significantly after the reform.²³

However, one complicating factor here is the possibility that, following the TFR, upper-level governments (for the most part county governments) took over themselves sufficient expenditure responsibilities for which village budgets experienced reductions. If county governments responded in this way, then the enactment of the TFR may not necessarily imply a negative impact on farmers' welfare, despite the losses from village budgets, because reductions in village government expenditures may have been offset by increases in county government expenditures.

Detailed information on the post-TFR responses of upper-level governments is not available. However, as pointed out by the work of Yep (2004), Kennedy (2007), Luo et al. (2007), and Sato (2008), among others, such a response from upper-level governments seems unlikely, for several reasons.

First, there was no explicitly defined change in expenditure assignments from the village level to county governments after the TFR, except for expenditure on primary schools. In 2001, the Central Committee of the Communist Party and the State Council of China introduced a reform of educational finance whereby county governments were required to take over the payments of teachers' salaries from village budgets. Nevertheless, the evidence shows that county governments did not cover all education funding at the village level. Since the education reform only required that county governments pay for the salaries of full-time primary school

level governments have typically been based on complicated and opaque procedures, which largely reduce their effectiveness.

²³ As pointed out by Lin and Liu (2007), because of the uneven level of industrialization between coastal and inland regions, the large non-agricultural tax base in coastal regions has contributed to a lower dependency on agricultural taxation.

teachers, funds for paying village teachers and maintaining school facilities seem likely to have been reduced after the TFR (Yep, 2004).

Second, the central government increased its transfer payments to county governments with the hope that they would increase the remittance for township and village governments after the TFR. Unfortunately, as noted earlier, the transfer allocation process at the county level is quite complex, which significantly reduces their effectiveness.

Finally, although county budget statistics in documenting the funds allocated to each of the villages in our sample are not available, it happened that the designers of the CHIP questionnaire were explicitly interested in the possibility of changes in county expenditure on welfare-related items at the village level (Sato, 2008). The CHIP designers asked village cadres for their judgments of the changes in public funding on primary schools after the TFR, in the belief that village cadres might be able to judge the overall financial conditions (including county expenditure and any other transfer payments) of primary schools after the TFR. Our tabulation of this survey question for the treated villages in our sample reveals that only 7.9 percent of the villages reported an increase in overall public expenditure on primary schools after the TFR, while 49.0 percent reported a decrease and 43.1 percent reported no change. This evidence, while not conclusive, suggests that a majority of the treated villages did not obtain sufficient compensation from upper-level governments even for expenditure on primary education where the payment responsibility was clearly shifted after the TFR to upper-level governments; for other expenditure categories, it seems even less likely that county governments compensated for reduced village expenditures.

For all of these reasons, we believe that our results for the impact of the TFR on farmers' welfare seem unlikely to be affected significantly by any post-TFR changes in expenditure assignments between village and county governments.

5. Conclusions

We examine in this paper the impact of the TFR in China on farmers' direct and indirect welfare, where direct welfare is measured by net income per capita in the village and where indirect welfare is proxied by the composition of public expenditure. Given the non-experimental nature of the dataset we used, we pay particular attention to identify the causal effect of the reform on farmers' welfare by using both a propensity score matching approach and a difference-in-difference matching approach.

Our results show no evidence that the TFR led to any direct increase in farmers' net income, despite this suggestion from some other work (Luo et al., 2007). The exact reasons for this result are not entirely clear, but, as suggested at the beginning of the paper, the answer to this puzzle may stem from the increases in the rate of agricultural taxation after the reform, which adversely affected agricultural production and consequently decreased farmers' incomes. Meanwhile, this adverse impact on the agricultural production may also have reduced the role of agricultural taxation in supplementing the revenue loss from the elimination of legal and illegal fees due to the shrinking tax bases. It is thereby unsurprising to observe that, while the elimination of the legal and illegal fees to farmers reduced the revenues of village government and led to a reduction of public services, this elimination did not result in an increase in the net income of the farmers.²⁴ Another possible answer to the puzzle lies in the fact that the reform

²⁴ Ideally, if there were data separately available for the changes in agricultural taxes and also the changes in legal and illegal fees during the reform, we would be able to examine this hypothesis. Unfortunately, these data are not

may also have resulted in a “double burden” on farmers; that is, the farmers were not only asked to pay the new agricultural tax, but they were also required to pay some illegal fees to the local authority when these authorities ran into financial difficulties (Yep, 2004; Yi, 2006).

Our results also indicate that any possible post-reform gain to farmers from the alleviation of fiscal burden was more than offset by indirect welfare effects via reduced public services from the villages. These expenditure reductions were largely driven by the new fiscal discipline that restricted villages in raising informal taxes and fees (with no offsetting increases in government transfers). Indeed, our results clearly show a substantial post-reform decline in expenditures on welfare programs such as education, public health, and infrastructure. These negative indirect effects were especially pronounced in inland (versus coastal) villages and in poor (versus rich) villages. Coastal and rich villages experienced no significant change in all welfare relevant aspects.

Our results do not deny the possibility that the TFR had other, potentially positive, effects, such as a strengthening of governance at the local level. However, our results suggest that any such positive effects must be weighed in light of the clear negative effects on overall farmers’ welfare.

References

- Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *The Journal of Political Economy*, 113 (1), 151-184.
- Aubert, C., & Li, X. (2002). Peasant Burden: Taxes and Levies Imposed on Chinese Peasants. In *Agricultural Policies in China after WTO Accession*. Paris, France: OECD.
- Bernstein, T. P., & Lu, X. (2000). Taxation without Representation: Peasants, the Central and the Local States in Reform China. *The China Quarterly* 163, 742-763.
- Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for The Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22 (1), 31-72.

reported in the dataset we used.

- Chen, X. (2003). *A Study on China's County and Township Public Finance and Farmer Income Growth*. Taiyuan, China Shanxi Economic Press (in Chinese).
- De Brauw, A., Huang, J., & Rozelle, S. (2004). The Sequencing of Reform Policies in China's Agricultural Transition. *Economics of Transition*, 12(3), 427-465.
- Dehejia, R. H., & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association*, 94 (448), 1053-1062.
- Dehejia, R. H., & Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics*, 84 (1), 151-161.
- Dethier, J.-J. (1999). Governance and Economic Performance: A Survey. Discussion Paper on Development Policy, Number 5. Bonn, Germany: Center for Development Research (ZEF).
- Fork, A., & Wong, C. (2005). Extending Public Finance to Rural China. The World Bank International Seminar on Public Finance for Rural Areas. Washington, D.C.: The World Bank.
- Gustafsson, B., & Li, S. (2002). Income Inequality within and across Counties in Rural China 1988 and 1995. *Journal of Development Economics*, 69 (1), 179-204.
- Gustafsson, B., Li, S., & Sicular, T. (Eds.). (2006). *Inequality and Public Policy in China: Issues and Trends*. New York, NY: Cambridge University Press.
- Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing Selection Bias Using Experimental Data. *Econometrica*, 66 (5), 1017-1098.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64 (4), 605-654.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The Economics and Econometrics of Active Labor Market Programs. *Handbook of Labor Economics*, 3, 1865-2097.
- Kennedy, J. J. (2007). From the Tax-for-Fee Reform to the Abolition of Agricultural Taxes: The Impact on Township Governments in Northwest China. *The China Quarterly*, 189, 43-59.
- Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification. *Journal of Business and Economic Statistics*, 17 (1), 74-90.
- Lee, M. J. (2005) *Micro-econometrics for Policy, Program, and Treatment Effects*, Advanced Text Series in Econometrics. Oxford, UK: Oxford University Press.
- Li, H., & Zahniser, S. (2002). The Determinants of Temporary Rural-to-urban Migration in China. *Urban Studies*, 39 (12), 2219-2235.
- Li, L. C. (2006). Embedded Institutionalization: Sustaining Rural Tax Reform in China. *The Pacific Review*, 19 (1), 63 - 84.
- Lin, J. Y. (1992). Rural Reforms and Agricultural Growth in China. *The American Economic Review*, 82(1), 34-51.
- Lin, J. Y., Cai, F., & Li, Z. (2003). *The China Miracle: Development Strategy and Economic Reform*. Hong Kong: The Chinese University Press.
- Lin, J. Y., & Liu, M. (2007). Rural Informal Taxation in China: Historical Evolution and an Analytic Framework. *China & the World Economy*, 15 (3), 1-18.
- Luo, R., Zhang, L., Huang, J., & Rozelle, S. (2007). Elections, Fiscal Reform, and Public Goods Provision in Rural China. *Journal of Comparative Economics*, 35 (3), 583-611.
- Meng, X., & Zhang, L. (2011). Democratic Participation, Fiscal Reform, and Local Governance:

- Empirical Evidence on Chinese Villages. *China Economic Review*, 22 (1), 88-97.
- Mushtaq, S., Khan, S., Dawe, D., Hanjra, M. A., Hafeez, M., & Asghar, M. N. (2008). Evaluating the Impact of Tax-for-Fee Reform (Fei Gai Shui) on Water Resources and Agriculture Production in the Zhanghe Irrigation System, China. *Food Policy*, 33 (6), 576-586.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70 (1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*, 39 (1), 33-38.
- Rubin, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*, 2 (1), 1-26.
- Sato, H. (Ed.) (2006). *The Impact of Village-Specific Factors on Household Income in Rural China*. New York, NY: Cambridge University Press.
- Sato, H. (2008). Public Goods Provision and Rural Governance in China. *China: An International Journal*, 6 (2), 2008.
- Sianesi, B. (2004). An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s. *The Review of Economics and Statistics*, 86 (1), 133-155.
- Sicular, T., Ximing, Y., Gustafsson, B., & Shi, L. (2007). The Urban-rural Income Gap and Inequality in China. *Review of Income and Wealth*, 53 (1), 93-126.
- Smith, J. A., & Todd, P. E. (2005). Does Matching Overcome La Londe's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 125 (1-2), 305-353.
- Tao, R., & Qin, P. (2007). How Has Rural Tax Reform Affected Farmers and Local Governance in China? *China & the World Economy*, 15 (3), 19-32.
- Todd, P. E. (Ed.) (2006). *Matching Estimators*, Second Edition. New York, NY: Palgrave Macmillan.
- Wang, L. (2008). Tax Reform, Local Resources and Village Leader Elections in Rural China. University of Toronto Working Paper. Toronto, Canada.
- Yep, R. (2004). Can Reform Reduce Rural Tension in China? The Process, Progress and Limitations. *The China Quarterly*, 177 (1), 42-70.
- Yi, Z. (2006). The Main Problems and Countermeasures of China's Rural Financial System. *Chinese Economy*, 39 (2), 57-70.
- Zhang, X., Fan, S., Zhang, L., & Huang, J. (2004). Local Governance and Public Goods Provision in Rural China. *Journal of Public Economics*, 88 (12), 2857-2871.
- Zhang, Z., & Martinez-Vazquez, J. (2003). The System of Equalization Transfers in China, Working Paper Series 0312, International Studies Program, Andrew Young School of Policy Studies, Georgia State University. Atlanta, GA.

Table 1: Fiscal Balance Sheet of Control and Treated Villages, 1998 and 2002 (real Yuan per capita)

	Control Villages		Treated Villages	
	1998	2002	1998	2002
<i>Total revenues</i>	57.58	78.25	75.22	56.40
Main items:				
Revenues from enterprises ^a	18.24	23.89	9.87	23.75
Local fees ^b	5.73	6.32	10.49	0
“Case-by-case” fundraising ^c	–	–	–	3.26
Other revenues	13.46	11.83	6.26	8.87
Transfers from upper-level government	1.46	2.89	2.54	6.71
<i>Total expenditures</i>	44.71	69.28	33.02	41.28
Main items:				
Education	4.90	3.70	2.59	2.67
Public health	0.44	0.85	0.41	0.45
Infrastructure	7.80	15.69	5.07	9.62
Wage and subsidies for village cadres	6.95	7.87	8.18	8.40
Other administrative expenditures	3.64	6.28	3.67	4.71

Source: CHIP, 2002.

^a Revenues from enterprises include revenues collected from collective and non-collective enterprises in the village.

^b Local fees refer to the revenues collected directly from farmers in the forms of the “administrative village levy” (*cuntiliu*) and other informal fundraisings by the village.

^c Case-by-case fundraising was introduced after the reform as a tool to finance case-by-case public good provision.

Table 2: Definition of Variables

Variable	Definition
<i>Treatment Variable</i>	
Treated	Indicator variable for a village that implemented the TFR, equal to 1 if the reform is implemented, 0 otherwise
<i>Observable Variables in Probit model</i>	
Suburb	Suburb of large/middle city
dist_county	Distance from nearest county seat (km)
Minority	1 if village is in Ethnic minority area, 0 otherwise
Poverty	1 if the county (the village in) is the national/province level poverty county, 0 otherwise
Pilot	1 if village be appointed as a pilot village for productive or non-productive reason at national/province level, 0 otherwise
election98	1 if village practices the direct election by the end of 1998, 0 otherwise
pop98	Total population at the end of 1998
pop98_squ	Square of pop98
planting98	Total planting area at the end of 1998 (Mu)
planting98_squ	Square of planting98
netinc98.pc	Net income per capita in the village, 1998 (Yuan)
netinc98.pc_squ	Square of netinc98.pc
<i>Outcome Variables</i>	
netinc.pc	Net income per capita in the village, 2002 (Yuan)
totexp.pc	Total fiscal expenditure per capita in the village, 2002 (Yuan)
welexp1.pc	Total fiscal expenditure on welfare programs per capita, including collectively operated reproduction, productive service for farmers, education, public health, infrastructure, 2002 (Yuan)
welexp2.pc	Total fiscal expenditure on welfare programs per capita, including collectively operated reproduction, productive service for farmers education, public health, infrastructure and other commonweal expenditure, 2002 (Yuan)
nonwelexp1.pc	Total fiscal expenditure on non-welfare programs per capita, including wage and subsidies for village and group cadre, other administrative expenditure, 2002 (Yuan)
nonwelexp2.pc	Total fiscal expenditure on non-welfare programs per capita, including wage and subsidies for village and group cadre, other administrative expenditure and other expenditure, 2002 (Yuan)
$\Delta(\text{netinc.pc})$	Difference of net income per capita in the village between year 1998 and 2002
$\Delta(\text{totexp.pc})$	Difference of total fiscal expenditure per capita in the village between year 1998 and 2002
$\Delta(\text{welexp1.pc})$	Difference of total fiscal expenditure on welfare programs per capita in the village between year 1998 and 2002 (items are defined as in welexp1.pc)
$\Delta(\text{welexp2.pc})$	Difference of total fiscal expenditure on welfare programs per capita in the village between year 1998 and 2002 (items are defined as in welexp2.pc)
$\Delta(\text{nonwelexp1.pc})$	Difference of total fiscal expenditure on non-welfare programs per capita in the village between year 1998 and 2002 (items are defined as in nonwelexp1.pc)
$\Delta(\text{nonwelexp2.pc})$	Difference of total fiscal expenditure on non-welfare programs per capita in the village between year 1998 and 2002 (items are defined as in nonwelexp2.pc)

Table 3: Descriptive Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
treated	841	0.71	0.46	0	1
suburb	841	0.08	0.27	0	1
dist_county	835	23.88	20.27	0.50	160
minority	841	0.15	0.35	0	1
poverty	841	0.29	0.45	0	1
pilot	841	0.02	0.14	0	1
election98	841	0.78	0.41	0	1
pop98	840	1793.06	1119.49	178	8586
pop98_squ	840	4466811	6698071	31684	73700000
planting98	838	3650.24	2910.02	3.00	22370
planting98_squ	838	21800000	39700000	9	500000000
netinc98.pc	839	1082.00	634.25	128.45	7232.28
netinc98.pc_squ	839	1572509	2657996	16500.22	52300000
netinc.pc	837	1307.54	791.27	61.81	7857.52
totexp.pc	807	60.74	207.53	0.0016	3408.10
welexp1.pc	704	32.34	142.91	0.0002	2956.44
welexp2.pc	733	37.49	155.62	0.0003	3135.83
nonwelexp1.pc	803	16.69	52.48	0.0008	972.62
nonwelexp2.pc	809	26.13	89.76	0.0013	1967.91

Table 4: Mean Differences in Characteristics of Treated and Control Villages

Variable Characteristics	Control Villages	Treated Villages	t-values
<i>Observable Variables Used in Probit Model</i>			
suburb	0.101	0.066	1.772
dist_county	30.473	21.140	6.193
minority	0.352	0.059	11.875
poverty	0.300	0.283	0.489
pilot	0.045	0.012	3.001
election98	0.684	0.822	-4.434
pop98	2020.328	1698.393	3.828
pop98_squ	6360172	3678177	5.375
planting98	4310.077	3374.466	4.287
planting98_squ	31500000	17700000	4.650
netinc98.pc	1117.415	1067.306	1.042
netinc98.pc_squ	1939762	1420157	2.587
<i>Outcome Variables</i>			
netinc.pc	1362.002	1284.871	1.286
totexp.pc	85.710	50.789	2.163
welexp1.pc	46.596	26.682	1.670
welexp2.pc	53.526	31.130	1.759
nonwelexp1.pc	19.231	15.676	0.868
nonwelexp2.pc	34.758	22.680	1.731
Number of villages	247	594	

Table 5: Propensity Score for TFR Implementation–Probit Estimation Results

Variables	Full Villages	Coastal Villages	Inland Villages	Rich Villages	Poor Villages
suburb	0.128 (0.206)	0.476* (0.270)	-0.187 (0.367)	0.370 (0.253)	-0.017 (0.434)
dist_county	-0.011*** (0.003)	-0.024*** (0.006)	-0.003 (0.004)		
minority	-1.486*** (0.164)	-0.567* (0.326)	-1.722*** (0.196)	-0.969*** (0.342)	-1.474*** (0.206)
poverty	-0.071 (0.140)	-0.117 (0.259)	0.048 (0.177)	-0.404 (0.259)	0.248 (0.184)
pilot_b	-0.635* (0.374)	-0.883* (0.523)	0.213 (0.695)	-0.648 (0.443)	0.011 (0.832)
election98	0.410*** (0.132)	0.273 (0.199)	0.539*** (0.188)	0.284 (0.191)	0.747*** (0.193)
pop98	0.000*** (0.000)	-0.000 (0.000)	0.001*** (0.000)	0.000 (0.000)	0.001*** (0.000)
pop98_squ	-0.000*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)
planting98	0.000 (0.000)	-0.000 (0.000)		0.000*** (0.000)	-0.000 (0.000)
planting_sq	-0.000** (0.000)	-0.000 (0.000)		-0.000** (0.000)	-0.000 (0.000)
netinc98.pc	-0.000*** (0.000)	0.001** (0.001)	-0.000 (0.000)	0.001 (0.001)	0.003 (0.002)
netinc98.pc_squ		-0.000*** (0.000)		-0.000* (0.000)	-0.000 (0.000)
Constant	0.720*** (0.259)	0.004 (0.581)	-0.015 (0.366)	-0.735 (0.852)	-1.632** (0.776)
Observations	684	261	424	303	385

Standard errors are in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Balancing Tests from Kernel Matching Estimators

Variables	Sample Mean		Percent Bias	Percent Bias Reduction	t-test t-statistic (p-value)
	Treated	Control			
suburb	0.078	0.098	-6.9	33.0	-1.11 (0.267)
dist_county	20.008	20.526	-2.4	94.3	-0.49 (0.622)
minority	0.057	0.060	-0.7	99.0	-0.18 (0.855)
poverty	0.245	0.262	-3.9	20.8	-0.60 (0.545)
pilot	0.012	0.009	2.2	87.8	0.54 (0.593)
election98	0.810	0.798	2.7	90.5	0.46 (0.648)
pop98	1731	1722.6	0.7	96.4	0.14 (0.889)
pop98_squ	3.80E+06	3.90E+06	-1.4	95.1	-0.40 (0.692)
planting98	3426.9	3170.3	8.0	71	1.60 (0.109)
planting98_squ	1.80E+07	1.60E+07	4.1	86.3	1.10 (0.270)
netinc98.pc	1120.1	1167.4	-6.4	28.0	-1.27 (0.206)

Table 7: Impact of TFR on Farmers' Welfare–Kernel Matching Estimators

Measure	ATT	t-statistic
<i>Direct Welfare Effect</i>		
netinc.pc	-131.650	-1.100
ln(netinc.pc)	-0.028	-0.370
<i>Indirect Welfare Effect</i>		
totexp.pc	-65.377**	-2.000
welexp1.pc	-45.930***	-2.700
welexp2.pc	-51.065***	-2.640
nonwelexp1.pc	0.543	0.070
nonwelexp2.pc	-9.755	-0.610
ln(totexp.pc)	-0.516***	-2.720
ln(welexp1.pc)	-0.746***	-3.220
ln(welexp2.pc)	-0.771***	-3.420
ln(nonwelexp1.pc)	-0.104	-0.590
ln(nonwelexp2.pc)	-0.213	-1.190

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Robustness Tests of TFR on Farmers' Welfare–Alternative Matching Methods

	Nearest Neighbor Matching		Caliper Matching	
	ATT	t-statistic	ATT	t-statistic
<i>Direct Welfare Effect</i>				
netinc.pc	-108.161	-0.820	-123.182	-0.930
ln(netinc.pc)	-0.026	-0.270	-0.036	-0.380
<i>Indirect Welfare Effect</i>				
totexp.pc	-139.686**	-2.200	-150.183**	-2.340
welexp1.pc	-56.279**	-2.420	-59.344**	-2.520
welexp2.pc	-58.505**	-2.290	-62.496**	-2.420
nonwelexp1.pc	1.392	0.450	0.770	0.250
nonwelexp2.pc	-10.749	-1.000	-12.068	-1.120
ln(totexp.pc)	-0.498*	-1.930	-0.509**	-1.960
ln(welexp1.pc)	-0.801***	-2.600	-0.825***	-2.660
ln(welexp2.pc)	-0.795***	-2.620	-0.817***	-2.680
ln(nonwelexp1.pc)	-0.098	-0.410	-0.092	-0.380
ln(nonwelexp2.pc)	-0.234	-0.960	-0.228	-0.940

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Welfare Effects of TFR in Coastal and Inland Villages–Kernel Matching**Estimators**

	Coastal Villages		Inland Villages	
	ATT	t-statistic	ATT	t-statistic
<i>Direct Welfare Effect</i>				
netinc.pc	-192.335	-1.040	-98.031	-1.350
ln(netinc.pc)	-0.048	-0.500	-0.038	-0.420
<i>Indirect Welfare Effect</i>				
totexp.pc	-13.921	-0.230	-65.704***	-3.160
welexp1.pc	-12.379	-0.400	-51.444***	-2.650
welexp2.pc	-13.379	-0.380	-55.567***	-2.810
nonwelexp1.pc	15.796	1.060	-5.770***	-3.210
nonwelexp2.pc	1.874	0.060	-7.359***	-2.870
ln(totexp.pc)	0.355	1.100	-0.964***	-5.750
ln(welexp1.pc)	0.238	0.640	-1.261***	-4.970
ln(welexp2.pc)	0.223	0.600	-1.322***	-5.840
ln(nonwelexp1.pc)	0.634**	2.130	-0.575***	-3.530
ln(nonwelexp2.pc)	0.468	1.520	-0.584***	-3.810

*** p<0.01, ** p<0.05, * p<0.1

Table 10: Welfare Effects of TFR in Rich and Poor Villages–Kernel Matching Estimators

	Rich Villages		Poor Villages	
	ATT	t-statistic	ATT	t-statistic
<i>Direct Welfare Effect</i>				
netinc.pc	-113.931	-0.670	-23.455	-0.330
ln(netinc.pc)	-0.012	-0.150	0.021	0.240
<i>Indirect Welfare Effect</i>				
totexp.pc	-30.417	-0.500	-61.635***	-3.270
welexp1.pc	-24.208	-0.830	-44.491**	-2.550
welexp2.pc	-25.287	-0.740	-47.114***	-2.640
nonwelexp1.pc	10.792	0.710	-4.436**	-2.450
nonwelexp2.pc	-3.266	-0.110	-5.682**	-2.250
ln(totexp.pc)	0.058	0.190	-0.759***	-3.440
ln(welexp1.pc)	-0.156	-0.440	-0.895***	-3.160
ln(welexp2.pc)	-0.130	-0.370	-0.963***	-3.570
ln(nonwelexp1.pc)	0.308	1.150	-0.217	-0.960
ln(nonwelexp2.pc)	0.224	0.810	-0.344	-1.580

*** p<0.01, ** p<0.05, * p<0.1

Table 11: Impact of TFR on Farmers' Welfare–DID and Kernel Matching Estimators

	ATT	t-statistic
<i>Direct Welfare Effect</i>		
$\Delta(\text{netinc.pc})$	-73.425	-1.000
$\Delta\ln(\text{netinc.pc})$	-0.011	-0.250
<i>Indirect Welfare Effect</i>		
$\Delta(\text{totexp.pc})$	-28.989	-1.080
$\Delta(\text{welexp1.pc})$	-21.957*	-1.660
$\Delta(\text{welexp2.pc})$	-21.638	-1.430
$\Delta(\text{nonwelexp1.pc})$	-0.210	-0.030
$\Delta(\text{nonwelexp2.pc})$	-5.756	-0.410
$\Delta\ln(\text{totexp.pc})$	-0.108	-1.300
$\Delta\ln(\text{welexp1.pc})$	-0.176	-1.290
$\Delta\ln(\text{welexp2.pc})$	-0.175	-1.490
$\Delta\ln(\text{nonwelexp1.pc})$	-0.111*	-1.670
$\Delta\ln(\text{nonwelexp2.pc})$	-0.169***	-2.720

*** p<0.01, ** p<0.05, * p<0.1

Table 12: Robustness Tests of TFR on Farmers' Welfare–Alternative DID Matching Estimators

	DID- Nearest Neighbor Matching		DID- Caliper Matching		Simple DID Estimates	
	ATT	t-statistic	ATT	t-statistic	DID	t-statistic
	<i>Direct Welfare Effect</i>					
$\Delta(\text{netinc.pc})$	-104.884	-0.960	-106.102	-0.950	-88.428	-1.56
$\Delta\ln(\text{netinc.pc})$	-0.045	-0.940	-0.044	-0.910	0.003	0.09
<i>Indirect Welfare Effect</i>						
$\Delta(\text{totexp.pc})$	-34.490	-1.520	-35.777	-1.530	-35.36	-1.53
$\Delta(\text{welexp1.pc})$	-23.327	-1.500	-24.417	-1.530	-13.986	-1.05
$\Delta(\text{welexp2.pc})$	-23.062	-1.350	-24.084	-1.380	-15.46	-1.05
$\Delta(\text{nonwelexp1.pc})$	-0.305	-0.100	-0.352	-0.110	-6.557	-1.26
$\Delta(\text{nonwelexp2.pc})$	-8.375	-0.990	-8.665	-0.990	-20.214*	-1.80
$\Delta\ln(\text{totexp.pc})$	-0.116	-1.130	-0.128	-1.270	-0.130*	-1.88
$\Delta\ln(\text{welexp1.pc})$	-0.245*	-1.650	-0.278*	-1.880	-0.157	-1.37
$\Delta\ln(\text{welexp2.pc})$	-0.219*	-1.720	-0.245*	-1.910	-0.165*	-1.66
$\Delta\ln(\text{nonwelexp1.pc})$	-0.091	-1.130	-0.094	-1.140	-0.134**	-2.44
$\Delta\ln(\text{nonwelexp2.pc})$	-0.173**	-2.23	-0.181**	-2.32	-0.215***	-4.10

*** p<0.01, ** p<0.05, * p<0.1.

Table 13: Welfare Effects of TFR in Coastal and Inland Villages–DID and Kernel Matching Estimators

	Coastal Villages		Inland Villages	
	ATT	t-statistic	ATT	t-statistic
<i>Direct Welfare Effect</i>				
$\Delta(\text{netinc.pc})$	-110.887	-1.310	-116.585	-0.920
$\Delta\ln(\text{netinc.pc})$	-0.009	-0.200	-0.100	-1.270
<i>Indirect Welfare Effect</i>				
$\Delta(\text{totexp.pc})$	-21.126	-0.420	-20.749**	-2.350
$\Delta(\text{welexp1.pc})$	-14.728	-0.540	-14.583*	-1.840
$\Delta(\text{welexp2.pc})$	-12.813	-0.410	-14.842*	-1.850
$\Delta(\text{nonwelexp1.pc})$	5.937	0.480	-2.861***	-3.010
$\Delta(\text{nonwelexp2.pc})$	-9.037	-0.350	-4.433***	-3.560
$\Delta\ln(\text{totexp.pc})$	0.090	0.600	-0.131	-1.140
$\Delta\ln(\text{welexp1.pc})$	-0.167	-0.840	0.015	0.070
$\Delta\ln(\text{welexp2.pc})$	-0.133	-0.750	-0.023	-0.140
$\Delta\ln(\text{nonwelexp1.pc})$	0.034	0.300	-0.196**	-2.520
$\Delta\ln(\text{nonwelexp2.pc})$	-0.052	-0.460	-0.215***	-3.150

*** p<0.01, ** p<0.05, * p<0.1

Table 14: Welfare Effects of TFR in Rich and Poor Villages–DID and Kernel Matching Estimators

	Rich Villages		Poor Villages	
	ATT	t-statistic	ATT	t-statistic
<i>Direct Welfare Effect</i>				
$\Delta(\text{netinc.pc})$	-80.499	-0.600	-45.582	-0.930
$\Delta\ln(\text{netinc.pc})$	-0.014	-0.250	-0.015	-0.230
<i>Indirect Welfare Effect</i>				
$\Delta(\text{totexp.pc})$	-18.197	-0.350	-23.473***	-3.190
$\Delta(\text{welexp1.pc})$	-11.609	-0.460	-16.053**	-2.530
$\Delta(\text{welexp2.pc})$	-10.176	-0.350	-15.575**	-2.390
$\Delta(\text{nonwelexp1.pc})$	3.737	0.290	-2.523**	-2.520
$\Delta(\text{nonwelexp2.pc})$	-7.720	-0.290	-4.564***	-3.680
$\Delta\ln(\text{totexp.pc})$	0.004	0.040	-0.281**	-2.490
$\Delta\ln(\text{welexp1.pc})$	-0.196	-1.060	-0.156	-0.790
$\Delta\ln(\text{welexp2.pc})$	-0.144	-0.880	-0.290*	-1.750
$\Delta\ln(\text{nonwelexp1.pc})$	-0.076	-0.780	-0.124	-1.350
$\Delta\ln(\text{nonwelexp2.pc})$	-0.114	-1.160	-0.246***	-3.240

*** p<0.01, ** p<0.05, * p<0.1

Figure 1: Propensity Scores Histograms

