Tulane Economics Working Paper Series

Chinas Tax-for-Fee Reform and Village Inequality

James Alm
Department of Economics
Tulane University
New Orleans, LA
jalm@tulane.edu

Yongzheng Liu
Department of Economics, Andrew
Young School of Policy Studies
Georgia State University
Atlanta, GA
yliu39@gsu.edu

Working Paper 1304
February 2013

Abstract

In the late 1990s, China enacted a rural tax reform known as the "Tax-for-Fee Reform" (TFR), largely driven by a desire to address farmers' complaints about their perception of a heavy and regressive tax burden. This paper examines the impact of the TFR on inequality in rural villages in China. Our results suggest an effective role of the TFR in reducing inequality within villages. Its impact on a consumption-based measure of inequality took effect immediately; its impact on per capita household income inequality took somewhat longer. Our results also suggest that it is "rich" and/or "coastal" villages that exhibited a significant reduction of inequality from the TFR, while "poor" and/or "inland" villages experienced no significant changes in inequality from the reform.

Keywords: Tax-for-Fee Reform; inequality; rural China
JEL: H7, I2, I3, O1, O5, P3

January 2013

China's Tax-for-Fee Reform and Village Inequality

James Alm and Yongzheng Liu*

Abstract

In the late 1990s, China enacted a rural tax reform known as the "Tax-for-Fee Reform" (TFR), largely driven by a desire to address farmers' complaints about their perception of a heavy and regressive tax burden. This paper examines the impact of the TFR on inequality in rural villages in China. Our results suggest an effective role of the TFR in reducing inequality within villages. Its impact on a consumption-based measure of inequality took effect immediately; its impact on per capita household income inequality took somewhat longer. Our results also suggest that it is "rich" and/or "coastal" villages that exhibited a significant reduction of inequality from the TFR, while "poor" and/or "inland" villages experienced no significant changes in inequality from the reform.

Keywords: *Tax-for-Fee Reform; inequality; rural China.*

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

---

* Tulane University and Georgia State University. We are grateful to two anonymous referees and to the Editor for many helpful comments. Please address all correspondence to James Alm, Professor and Chair, Department of Economics, 208 Tilton Hall, Tulane University, New Orleans, LA USA 70118 (phone +1 504 862 8344; fax +1 504 865 5869; email jalm@tulane.edu).

1. Introduction

China's economic growth in the last three decades has been remarkable. This rapid growth has been accompanied by serious social and economic problems, especially concerns about inequality. Indeed, there is a large and growing literature on various aspects of inequality in China (Rozelle 1994; Khan and Riskin, 1998; Kanbur and Zhang, 1999; Gustafsson and Li, 2002; Morduch and Sicular, 2002; Benjamin et al., 2005; Kanbur and Zhang, 2005; Wan and Zhou, 2005; Shen and Yao, 2008; Xing et al., 2009; Shi et al., 2010).

In this paper we examine how a major recent tax reform, the Tax-for-Fee Reform (TFR), affected inequality in rural China, focusing especially on its impact on within-village inequality. In the late 1990s, the government of the People's Republic of China enacted the TFR, largely as a response to rural farmers' bitter complaints about what they saw as a heavy and regressive fiscal burden. The reform was intended to reduce this fiscal burden, although it also had other important effects. The TFR 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] Given the motivation for the TFR, a natural question to ask is whether the reform actually affected the distribution of income (and of consumption) at the village level. It is this issue that we examine here.

Despite the many insights from existing work on inequality in China, few studies have examined rural inequality, and these studies that have done so have largely focused on interregional inequality, as opposed to inequality within villages.[2] An investigation focusing on

---

[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 later years, after our sample information was collected.

[2] There are some notable exceptions. Of most relevance to our work, Sato et al. (2008) evaluate the redistributive impacts of rural taxation in China, with a special focus on the change of tax regressivity between 1995 and 2002. By comparing before and after tax Gini coefficients at an aggregated level, they conclude that the unfavorable redistributive impact of rural taxation has been alleviated from 1995 to 2002. Our work here differs in several

inequality within villages is important, we believe, for several reasons. There are substantial economic disparities among villages within the same county or even in the same township (Knight and Li, 1997; Perkins, 2002; Sato, 2003, 2008a), and interregional inequality analysis largely ignores these disparities. Further, rising inequality within a small and closed community is more likely to generate unrest than inequality among regions, since people's happiness and/or dissatisfaction are largely based on comparisons to other members of their village (Knight and Gunatilaka, 2010a, 2010b, 2011, 2012). Finally, although there is some work that has attributed rising inequality in China to spatial differences linked to regional factors (Tsui, 1993; Yao, 1997; Kanbur and Zhang, 1999; Lee, 2000), several recent studies have ruled out geographical differences as the most important factor for understanding the dispersion of income (Benjamin et al., 2000; Gustafsson and Li, 2002; Benjamin et al., 2005). Instead, these studies have found that most inequality in rural China is due to inequality between neighbors within a village.

In principle, the TFR could affect inequality within villages in several ways. One is that the TFR largely reshaped the landscape of local governance by introducing greater fiscal discipline at the local level, which in turn imposed constraints on public good provision at the village level, including roads investment, irrigation, compulsory education, and health expenditures (Zhang et al., 2004; Liu and Lin, 2007; Luo et al., 2007; Knight, Li, and Quheng, 2009, 2010; Wang and Zhao 2012; Wang, Zheng, and Zhao, 2012; Alm and Liu, forthcoming). Since these village-level factors are among the most important determinants of household

aspects. First, we are able to examine the impacts of the TFR at a more disaggregated (e.g., village) level. Second, we employ a difference-in-difference propensity score matching approach to identify the causal effect of tax reform on inequality, with special focus on addressing the selection bias problem in the sample. See also Shen and Yao (2008), Ha et al. (2009), and Shi et al. (2010) for studies of income inequality within villages in rural China. Also, Sato (2008a) uses a hierarchical linear model to demonstrate the importance of village specific factors (e.g., physical infrastructure, human capital and social capital), as well as public management and public policy in influencing household income determination. Park and Wang (2010) evaluate the policy effects of China's flagship poverty alleviation program, begun in 2001, and they find that the program did not increase the income or consumption of poorer households, while it did increase the income and consumption of richer households by 6.1 to 9.2 percent.

income in the rural areas of developing countries (Narayan and Pritchett, 1999; Sato, 2008a; Shen and Yao, 2008; Martinez-Bravo et al., 2012), the TFR should have some impact on inequality within villages. Relatedly, other aspects of the intergovernmental system may have been affected (e.g., expenditure assignements between governments), with results effects on inequality. A second channel is through the potential impacts of the TFR on household income sources. As emphasized by many previous studies (Khan and Riskin, 1998; Tsui, 1998; Kung and Lee, 2001; Benjamin et al., 2005), the changing structure and composition of income plays an important role in generating inequality in rural China. The reform of these rural taxes may have induced behavioral responses of individuals that change their income composition, and so income inequality within villages. Finally, and most directly, the TFR reduced the regressivity of the old tax system in rural China. Local levies and fees were well-known to be more regressive than formal state taxation (Sato et al., 2008; Lin and Liu, 2007), and the elimination of all local levies and fees in the TFR seems likely to reduce tax regressivity and contribute to a lower inequality within villages. The net impact of these various factors on village-level equality is, however, unclear and remains an open question.

We use a comprehensive and specialized dataset containing detailed information before and after the TFR at the village level to examine the impact of the TFR on inequality within villages. Our data come from the 2002 Chinese Household Income Project (CHIP). The data are based upon a sample of 961 villages located throughout China, in which 9,200 households were surveyed. These data provide information both before and after the TFR, allowing us to calculate Gini coefficients within villages, calculated separately for per capita household net income and per capita household consumption; this latter measure of inequality (based on household consumption) is usually argued to be more reliable than the former (based on household net income) due to the widespread under-reporting of net income of the respondents, and it is also more related to reflect the true permanent income of the households (Deaton 1997; Milanovic 1999). With these measures of inequality, we employ a difference-in-difference (DID) propensity score matching approach to evaluate the impacts of the TFR on inequality within villages.

We find an effective role of the TFR in reducing inequality within villages. There is some immediate impact on both Gini coefficients, as well as a lagged impact, especially for the net income measure. We also find that it is "rich" and/or "coastal" villages that exhibited a significant reduction of inequality from the TFR, while "poor" and/or "inland" villages experienced no significant changes in inequality from the reform.

In section 2 we provide a brief overview of the fiscal system in rural China. Section 3 develops our empirical methods, and section 4 describes our data. We discuss our results in section 5, and we conclude in section 6.

2. The Fiscal System in Rural China and its Evolution over Time

The multilayered and highly decentralized local fiscal system in China has been operating in a weak institutional environment since its establishment in the late 1970s. Townships and villages were assigned an important responsibility for providing local public goods in rural areas. However, legally assigned revenue sources were rarely sufficient in meeting expenditure needs. In the absence of adequate fiscal transfers from upper-level governments, a number of illegal levies and fundraisings were imposed by townships and villages to offset the revenue gaps for financing expenditures or public projects.[3]

By the 1990s, this system of diverse and often illegal charges paid by farmers (especially

[3] An exhaustive list and estimated value of these illegal levies is very difficult to establish (Aubert et al., 2002).

those in poorer agriculture-based areas) to local authorities generated rising protests from farmers, which started to threaten rural social stability and endanger the state's political legitimacy (Bernstein and Lu, 2000; Tao and Qin, 2007).[4] Indeed, large-scale protests, even conflicts, against local authorities were observed during the process of taxes and fees collection (Aubert et al., 2002; Chen, 2003). As a response to farmers' bitter complaints, the central government launched its so-called "Tax-for-Fee Reform" (TFR) in the late 1990s, largely with the objective of reducing farmers' fiscal burdens.

The TFR had several main features. First, all existing township and village levies, including five township pooling funds, three village levies, and other kinds of informal local fees, were abolished.[5] Second, agriculture tax and agriculture tax supplements rates were increased, to accompany the introduction of an additional levy for agricultural taxes as a substitute for the decline of local levies. Third, "Case-by-case Fundraising" (or *yishiyiyi*) was introduced to finance special public projects, and at the same time budgetary transfers from upper-level governments were adjusted to accommodate local needs. The TFR was first formally introduced on a local pilot basis in 2000, and was widely carried out nationally after that. By the end of 2002, 20 provinces in China had commenced reform on a pilot basis (Tao and Qin, 2007). As noted earlier, an additional reform that eliminated all agricultural-related taxes was enacted in 2006 by the central government; however, our data do not cover this more recent period.

In line with its major policy objective, the implementation of the TFR reduced farmers' fiscal burdens by a substantial amount, at least according to somewhat anecdotal information.

[4] For a detailed analysis of the rising fiscal burden on farmers in the 1990s, see Lin and Liu (2007).
[5] The township-pooling funds included education supplements, social help, family planning, collective transportation, and militia exercises. The three village levies refer to the public accumulation fund, the public welfare fund, and administrative fees.

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. In Anhui province, the national model area of rural reform, Yep (2004) reports a 31 percent burden reduction across the whole province for the first year of the reform. At the same time, the TFR also reshaped the landscape of the fiscal balance sheets in villages. Revenues declined significantly due to the termination of informal fundraisings and an absence of sufficient transfers from upper-level governments to replace these funds. Expenditure assignments remained largely unchanged, other than additional mandates from upper-level governments. In principle, upper-level governments were supposed to provide sufficient fiscal transfers to compensate for the loss of revenues in the villages and to guarantee the provision of public services; however, according to Li (2006), they generally failed to do so.

More systematic evidence is provided by Alm and Liu (forthcoming). They find that, in villages where no reform was in effect between years 1998 and 2002 (i.e., "control" villages), 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 villages where the TFR was enacted in some year after year 1998 but before 2002 (i.e., "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.

Importantly, the changing fiscal system (and local governance) in rural China raises the question as to whether the TFR affected income distribution within villages. It is this issue that we examine here.

3. Empirical Methods

Let TR_i be an indicator of whether the Tax-for-Fee Reform 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 (i.e., Gini coefficients of per capita household net income and per capita household consumption) 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 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:

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

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 microeconometric 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.[6]

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

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 village 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:

$$p(X_i) = Prob(TR_i = 1|X_i), \quad (3)$$

where X_i denotes observed covariates for village i that are not affected by the implementation of the TFR (or the anticipation of it), so that these X_i variables should either be fixed over time or be measured in the time period before the implementation of the TFR (Caliendo and Kopeinig, 2008); these variables should also control for the likelihood of the TFR implementation. We then match each treated village with a control village on the basis of the predicted probability of implementation of the TFR (i.e., 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 village and the matched control villages. As shown by Todd (2008), a standard matching estimator can be written as:

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

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

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.[7] Then the village matched to village i is that village in set A_i such that $A_i =$

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

$\{ j \in I_0 \mid P_j \in C(P_i) \}$ 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 types 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 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 (i.e., the "common support assumption") (Heckman et al., 1999).[8] 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).[9]

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 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:

$$SB_{before} = 100 \cdot \frac{\overline{X_1} - \overline{X_0}}{\sqrt{\frac{V_1(X)+V_0(X)}{2}}} \qquad SB_{after} = 100 \cdot \frac{\overline{X_{1m}} - \overline{X_{0m}}}{\sqrt{\frac{V_{1m}(X)+V_{0m}(X)}{2}}}, \tag{6}$$

where for each covariate $\overline{X_1}$ ($V_1(X)$) and $\overline{X_0}$ ($V_0(X)$) represent the mean (variance) for both groups before matching, and $\overline{X_{1m}}$ ($V_{1m}(X)$) and $\overline{X_{0m}}$ ($V_{0m}(X)$) represent the mean (variance) for both groups after matching. A sufficient reduction of the standardized difference before and

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

[9] Differences are expected before matching is performed. After matching, the covariates should be balanced in both treated and control villages.

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.

In the light of our two years panel dataset, we then employ a difference-in-difference (DID) matching estimator.[10] The argument here is that DID matching estimator has the potential advantage of allowing for any individual-specific, time-invariant unobserved characteristics between treated and control villages, which may otherwise lead to a violation of the conditional independence assumption under standard matching estimator(Heckman et al., 1997; 1998).[11] Following Heckman et al.(1997), the DID matching estimator can be expressed as:

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

where t and t' are the time period after and before the year of the TFR.

4. Data

The data used in this study are a combination of rural household-level and village-level survey data from the 2002 Chinese Household Income Project (CHIP) (Li, 2002), conducted by

[10] Note that we also use a cross-section propensity score matching estimator, as defined in equation (4). These results are the same as the results from our DID matching estimators, so we do not report them here. All results are available upon request.

[11] Even after conditioning on observed covariates, there are still various reasons that systematic differences may exist between the outcomes of treated and control villages (Smith and Todd, 2005). Such differences may arise because of the selection into treated based on unmeasured characteristics. To the extent that there are unobserved time-varying characteristics between treated and control villages, the DID matching estimators 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.

the Institute of Economics, Chinese Academy of Social Sciences, with the assistance of the State Statistical Bureau in Beijing.[12] The 2002 CHIP was drawn from a large-scale sample by National Bureau of Statistics for its annual household survey in both rural and urban areas in selected provinces, with a special focus on examining the dynamic changes in income distribution in China. The data were collected through a series of questionnaire-based interviews at the end of 2002. Ten separate datasets were created. The first four datasets (1-4) survey different living aspects of individuals and households in urban areas, such as income and consumption; the last five datasets (6-10) contain similar living aspects for individuals and households in rural areas. The fifth dataset (5) concentrates on village-level data, obtained by interviewing village leaders. Since our focus is on the effects of the TFR in which the village was the basic unit of implementation, the rural household dataset and the village-level dataset are what we use in this study.

In total, 22 provinces, representing various regions (i.e. metropolitan city, coastal, central, and western) in rural China, were selected in a manner roughly consistent with their populations for the rural survey. Sample villages were then selected directly in each province on the basis of their income levels, while households were drawn from each village (Gustafsson et al., 2008). Overall, 961 villages with 9,200 sample rural households were included in the project dataset. Among these villages, 10 households were selected from each of the 871 villages, 9 households were selected from each of 10 villages, and 5 households were selected from each of 80 villages.

The household sample contains for year 2002 household information on: demographic characteristics; income, consumption, and their detailed components; assets and liabilities; work and employment information; social network information; quality of life information; and village

[12] The 2002 CHIP is well-known as one of the most representative household-level datasets in China (Gustafsson et al., 2006). Studies that have used this dataset include Gustafsson and Li (2002), Li and Zahniser (2002), Sicular et al. (2007), Gustafsson et al. (2008), Meng and Zhang (2011), and among many others.

affairs data. At the same time, retrospective information was collected for years 1998 to 2001 for three variables: total household net income, total household consumption expenditure, and total resident numbers in the household.[13] Our analysis takes advantages of this retrospective information to calculate the Gini coefficients of per capita household net income and per capita household consumption within villages, before and after the TFR implementation.[14] It should be recognized that the small number of households (or 10) used to calculate the Gini coefficients suggest some caution regarding the representativeness of this small sample in measuring inequality within villages. However, we believe that these concerns should not be overstated. The CHIP datasets were specifically designed for the study of inequality; that is, the survey instrument was first deemed appropriate for the topic of inequality and addressing relevant theoretical concerns, and was then used to conduct interviews over the selected sample households from an even larger sample of households that the Chinese National Bureau of Statistic (NBS) has conducted. The involvement of the CHIP research team in the data collection and the cleaning process also addresses some quality and representativeness concerns; for more details, see Gustafsson et al. (2006). Finally, there is evidence that the income-related data from the village survey and the data aggregated from the household survey give similar results (Sato 2008a).[15]

---

[13] Total household net income refers to household disposable income, which consists of: wages earned by household laborers; income from household production (i.e., farming and non-farm activities); income from property; rental value of owner-occupied housing; remittance income sent back by household members who work elsewhere; cash and in-kind social benefits (including health, housing, food, and other in-kind benefits); and other miscellaneous income. Household disposable income is net of costs for household production, depreciation of productive fixed capital, and cash expenditure on taxes and fees. Total household consumption includes: all food, nondurable and durable goods consumed during the year (e.g., clothing, transportation, communication, electronic appliances, and cars); expenditures on education and health; housing expenses (e.g., imputed rents of owned dwellings, utilities); costs for household production; purchases of fixed capital for production; depreciation of productive fixed capital; interest payments on borrowing; cash expenditures on taxes and fees; and other expenditures.

[14] We use the STATA command *ineqdeco* to calculate the Gini coefficients within a village.

[15] Indeed, the village mean income for 2002 collected from the village questionnaire has a strong positive correlation with the village mean income aggregated from the household survey (r=0.809, n=951) (Sato, 2008a).

Even so, the representativeness of sample household income and consumption information cannot be automatically extended to all other variables. Other village characteristics used in the empirical analysis are derived directly from the village-level survey dataset.

In addition, there are two prominent advantages of the village-level dataset. First, the information is very comprehensive, with 259 variables in 961 villages distributed across 22 provinces in China.[16] Variables cover almost all aspects of villages, including basic geographic information, arable land, agriculture activities, collectives, enterprise, labor force, income, productivity, population, government budget, taxes, expenditures, local elections, and the characteristics of government officials. This information plays a crucial role in predicting the determinants of the TFR implementation when we employ the matching techniques. Second and more importantly, the village-level dataset has information both before and after the TFR implementation for most all variables; that is, the village survey dataset provides information for years 1998 and 2002, by asking each question for the two years.[17] This feature allows us to invoke the conditional independence assumption, which requires that variables included in the specification of program participation either be fixed over time or be measured before participation (Sianesi, 2004; Smith and Todd, 2005). The pre-TFR information for year 1998 is clearly essential to meet this requirement. In addition, the availability of information in the household–level survey data for these two years also makes the implementation of our DID matching estimator possible, which is crucial for addressing the effects of time-invariant unobserved heterogeneity.

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

[17] For the question of "net income per capita in the village", the survey elicited answers for 1990, 1998, and 2002.

Our working sample is derived by imposing the restriction that only those villages that have not introduced the TFR by the end of 1998 are included in our estimation.[18] This restriction ensures that all observations having the same initial status in 1998 (or pre-treatment, pre-TFR status) can be grouped into either treated villages or control villages by the end of 2002. Imposing this condition gives us 906 villages for both years.

Tables 1 and 2 give definitions and summary statistics on the various outcome and explanatory variables used in the analysis. Differences in the features of treated and control villages are presented in Table 3, along with the associated t-values. These t-values suggest differences between treated and control villages for most (if not all) variables. Focusing especially on the Gini coefficients, Figure 1 provides histograms for the within-village Gini coefficients for per capita household net income and per capita household consumption in 1998 and 2002. Both measures demonstrate a slightly upward shifting of inequality within villages, a trend that matches some other studies showing that income distribution was becoming more and more unequal in both rural and urban China (Benjamin et al., 2005; Shen and Yao, 2008; Cai et al., 2010). Even so, a simple comparison of the Gini coefficients for treated and control villages (Table 3) indicates no significant difference between them. However, these simple comparisons do not account for the selection bias generated by characteristics that may affect the implementation of the TFR and the outcome variables simultaneously. Therefore, we now turn to our detailed analysis of the matching estimator results.

5. Empirical Results

[18] A few villages that implemented tax reform before 1999 are excluded from the analysis because we take information for the year 1998 as pre-treatment information, which by requirement should be independent of treatment status.

We first estimate a specification that generates the propensity scores for the Tax-for-Fee Reform, using a probit model to predict the probability of introducing the TFR. Second, we perform a balancing test to check the success of propensity score estimation in balancing covariates between treated and control villages. Finally, we generate difference-in-difference (DID) kernel estimators to estimate the causal effect of the TFR on income and consumption inequality within the village.

It should be noted that we have examined other DID matching estimators (nearest neighbor matching estimators, caliper matching estimators). We have also examined whether the TFR had different effects by region (i.e., "coastal" versus "inland" villages) and by income (i.e., "poor" versus "rich" villages). These robustness checks are reported later.

Probit Estimation

Results for the probit estimation are in Table 4. Our choice of covariates is guided by the criterion that selected variables should influence simultaneously the TFR assignment and the outcome variables (Dehejia and Wahba, 1999, 2002). More specifically, our choice of covariates was largely based on discussions with officials on the "determinants" of TFR implementation. These discussions indicated that the central government assigned the authority to select pilot villages to the provincial governments without making any specific criteria on what type of villages should be included, and so each authorized provincial government had its own right on the selection of the pilot villages. The provincial governments in turn usually selected a pilot county first, and then asked the county government to select the pilot villages. There may have been some element of "random" selection by the county governments. However, in practice the provincial (and county) governments seemed to follow some implicit rules. In general, a county (or village) had a greater probability of selection: if it was poorer; if it had a history of being selected for other and previous reforms; if it had more minorities; if it had more provincial government officers; or if had more "shocks" (e.g., natural disasters). We controlled for these variables in our estimation when we had relevant measures.

Following standard practice, we then build the specification by introducing the covariates linearly or in log term, and perform the balancing tests (see below) to check whether we succeed in balancing the covariates within each stratum. If the balancing tests are satisfied, then we accept the specification; if not, then we add other covariates until the balancing conditions are satisfied. Our final specification is in Table 4. These results indicate the significant role of specific village-level characteristics in the selection of villages for the TFR.[19]

Balancing and Common Support Evidence

Table 5 reports the standardized differences and regression-based balancing test results after performing DID Gaussian kernel matching on the full sample. The standardized differences between treated and control villages are all less than 9.2 percent, and most are less than 5 percent. The bias reduction as a result of employing the matching technique is therefore quite substantial.

The balancing results are also confirmed by the regression-based tests. The t-statistics in the last column of Table 5 demonstrate that we 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 the selection bias problem.

Since our objective is to make treated villages as comparable to control villages as possible, in order to estimate the average treatment effect of the TFR on inequality within

[19] Note that we also estimated separate specifications with different combination of covariates using only "rich" villages, "poor" villages, "coastal" villages, and "inland" villages, as these variables are defined later.

villages, 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 835 (out of 906) observations in the common support region, of which 592 are treated villages and 243 are control villages. Figure 2 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 get dropped because they lie off the common support.

Difference-in-difference Kernel Matching Estimates

Having established that propensity scores are balanced and that the common support condition is justified, we conclude that the treated and matched control villages are comparable. We now present in Table 6 the empirical results from DID kernel matching estimator.

The first row of Table 6 contains the results from applying the DID kernel matching to the full sample of all villages. These results suggest that the impacts of the TFR on both measures of inequality are negative, though the effect is only statistically significant for the consumption-based measure of inequality.

However, the TFR may take some time to have any impact, given the time lag between its implementation and its resulting detectable impacts. Accordingly, we explore the possible lagged effect of the reform by dividing the treated villages into three subgroups based on the particular year that the TFR took place in the villages. Numbers ranging from 1 to 3 in the first column of Table 6 represent the post-reform periods that the particular treated villages have experienced. For example, “1” denotes a village that enacted the TFR in year 2002, so that the number of years post-TFR before our data were collected after the end of year 2002 is 1; similarly, “2” denotes a village that enacted the TFR in year 2001, so that the number of years post-TFR before our data were collected is 2; and so on.[20] DID kernel matching is then re-applied to each subgroup of treated villages and accompanying control villages.[21]

The results in Table 6 show that the impacts of the TFR in the first and second year of effective periods on the net income measure of inequality remain statistically insignificant, while its impacts on consumption-based measure of inequality are negative and significant. However, when the reform is in effect for three years or more, the results in the last row of Table 6 reveal a significantly negative impact of the TFR on both measures of inequality.

These results suggest that the TFR played some role in reducing the inequality within villages, though its full impact on inequality takes some time to be realized, especially for per capita household net income inequality. This lagged effect on income inequality is consistent with the traditional notion that economic agents take time to learn and adjust their behaviors to policy changes. Therefore, it is reasonable that the income distribution in the year of reform and its immediately following years does not change significantly. The lack of much impact in the immediately effective year of reform may also suggest that the TFR was not fully implemented by the village leaders in the year of reform. Instead, they may have gradually adopted the reform procedures. Finally, the increases in the formal agricultural tax rates in the TFR may have partially offset the elimination of illegal levies and fees in reducing income inequality, because it is usually poorer farmers who are more subject to the agricultural taxation due to their relatively heavier reliance on agricultural production. Nevertheless, in comparison to the results of no *immediate* impact of the reform on net income measure of inequality, the reform's *immediate*

[20] Among the 658 treated villages in the full sample, 518 implemented the reform in year 2002, 86 in year 2001, 50 in year 2000, and 4 in year 1999. Given the relatively small number of villages that implemented in year 1999, we group those villages that implemented the reform in years 1999 and 2000 into the same subgroup.

[21] As before, the DID matching estimators for each village group is based on the specification that satisfies the balancing test and common support condition. We do not report the results of the balancing test and the common support condition, but all results are available upon request.

impact on reducing consumption-based measure of inequality is somehow puzzling. This is so because external shocks like the TFR implementation are more likely to affect individual's disposable income than consumption, since the latter is mostly determined by permanent income and can be smoothed over time through mutual insurance or other saving actions.

Although a full investigation on this puzzle is beyond the scope of this study, we believe that one of the explanations is as follows. As argued by some recent studies (Cai et al., 2010), the immature insurance market in China suggests that the Chinese household does not typically smooth consumption over time and so does not insure fully against income shocks, implying that the Chinese household consumption is mainly determined by disposal incomes rather than permanent incomes.[22] The elimination of illegal levies and fees thereby brings an anticipation of the increments of disposal incomes, which in turn stimulates households' current consumption. Since rich households have already enjoyed a relatively higher level of consumption before the reform and the share of illegal levies and fees in their total income before the reform is relatively small, the stimulant effect of the reform can be smaller for rich households than poor households. Consumption-based measure of inequality is thereby reduced by the reform immediately.

Robustness Checks: Alternative Matching Estimates

How robust are these findings? Two alternative matching estimates, DID nearest neighbor matching and DID caliper matching outcomes, are given in Table 7. These results are comparable to the kernel matching estimates. The TFR significantly reduced both per capita household net income inequality and per capita consumption inequality within villages, with the effect on the former again only emerging in later years of the reform.

Robustness Checks: Alternative Measures of Inequality

[22] Cai et al. (2010) use this argument to explain the counterintuitive phenomenon that consumption inequality in China parallels but stays above income inequality.

As we discussed earlier, the 2002 CHIP is recognized as one of the most representative household-level datasets in China. Even so, the small sample size (i.e., 10 households) used to measure the Gini coefficient within villages may generate a downward-biased measure of inequality. As pointed out by Deltas (2003), the maximum value of the Gini coefficient is usually a function of the sample size *N*, so that the estimated Gini coefficient may be subject to small-sample bias when its calculation is based on a small sample size.

This possibility seems unlikely to have a significant effect on our results. We use exactly the same households (including subjects and numbers) to calculate the Gini coefficients for the same village before and after the TFR, and the DID matching estimator is then based on the difference between these two values. If there are any biases caused by the small sample size, it is likely that these biases will be consistent across the years, so that taking the difference of the two values should be able to cancel out these potential biases in our estimation.[23]

Even so, it is useful to try other popular measures of inequality that may vary in small-sample bias sensitivity, as a further robustness check. To do this, we re-calculate the within-village inequality based on three other popular measures of inequality: the Theil index, the mean logarithmic deviation, and the Atkinson index.[24] Results from the DID kernel matching estimator with these other measures of inequality are presented in Table 8. These alternative results are consistent with our previous results. The TFR continues to exchibit a lagged effect on

[23] Deltas (2003) addresses this small-sample bias by proposing an adjusted Gini coefficient, which is simply a multiplication by a factor *(N/(N-1))* of the original Gini coefficient. In our context, applying this same factor *(N/(N-1))* to the original Gini coefficients for the same village for both years would not affect our main results, given a constant and unchanged sample size *N*.

[24] The Atkinson index is computed with a value of “inequality aversion” equal to 0.5. We also use levels of “inequality aversion” equal to 1 and 2, with similar results. All results are available upon request.

reducing within-village income inequality, along with an instant effect on reducing consumption-based measure of inequality.[25]

As an additional robustness check to deal with the potential small-sample bias, we also reapply the DID kernel matching estimator to a sub-sample that excludes those villages that only contain 5 households in measuring the Gini coefficients.[26] The results are also consistent with our previous results, and are not reported here.

Regional-specific Effects

We also examine whether the TFR had different effects in inland and coastal provinces, given different social and economic characteristics in these regions (Park and Wang, 2010). We group the full sample into "coastal" and "inland" villages, and we reapply the DID kernel matching technique to each village group.[27] Among the 22 provinces in our sample, we classify as coastal provinces Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shangdong, Guangdong, and Guangxi; we group as inland provinces Shanxi, Jilin, Anhui, Jiangxi, Henan, Hubei, Hunan, Chongqing, Sichuan, Guizhou, Yunnan, Shananxi, Gansu, and Xingjiang.

These results are reported in the upper half of Table 9. We find no significant effect of the TFR on the net income measure of inequality in either inland or coastal villages with full subsamples. Although we would like to analyze the lagged effect of the reform in these subsamples as what we did for the full sample, we are unable to do so due to the small sample size; that is, by restricting the subsample to one particular geographic region or income category, with reform implemented in one particular year, the sample size is reduced to such an extent that

[25] In Table 8, the estimates of consumption-based measure of inequality for the first two post-reform years are statistically significant in a one-tail test at the 10 percent level.
[26] As noted in the data description, 80 (out of 961) villages contain only 5 households.
[27] As before, the propensity score matching estimator for each village group is based on the specification with the different combination of covariates that satisfies the balancing test, and common support condition. We do not report the results of the probit estimation, the balancing test, and the common support to save space, but all results are available upon request.

the balancing tests for treated and control villages are not met. In addition, we also find a negative effect on the consumption measure of inequality for coastal villages, while we find no statistically significant impact on inland villages. Since inland provinces are generally underdeveloped regions in China relative to coastal provinces, this finding may imply asymmetric effects of the TFR on poor versus rich villages.

To examine this possibility, we further split all villages into income categories based on their net income per capita in year 1998 relative to the median value of the full sample in 1998. Villages with net income per capita less than the sample median value in year 1998 are defined as "poor" villages, while villages with a greater value than the sample median are grouped as "rich" villages. Again, we reapply the DID kernel matching technique to each village group.[28]

The results in the bottom half of Table 9 largely confirm our previous finding that the TFR only has a significant impact on the consumption-based measure of inequality in rich villages. For poor villages, there are no noticeable changes in either measure of inequality.[29] In addition, the comparison of the estimates between rich villages and coastal villages reveals that the absolute magnitude of the effect in rich villages is higher than in coastal villages, which further confirms the hypothesis that the effect of the TFR depends on the villages' initial income.

These results are robust to our use of the three alternative measures of inequality, as indicated in Table 10.[30]

We suggest two explanations for these results. First, the initial pre-TFR income distribution within rich villages is more unequal than for poor villages.[31] This would lead to a

[28] The propensity score matching estimator for each village group is again based on the specification with the different combination of covariates that satisfies the balancing test and common support condition. All results are available upon request.
[29] Note that, due to small sample size, we are again unable to analyze the lagged effect of the TFR.
[30] The estimates of the consumption-based measure of inequality for coastal villages in Table 10 are statistically significant in a one-tail test at the 10 percent level.

larger asymmetric effect of the reform on consumption between rich households and poor households within rich villages than within poor villages. Consequently, the TFR leads to a larger change of consumption inequality within rich villages than within poor villages. Second, the TFR has resulted in a fall in village revenues due to its elimination of the rural fees, and this revenue loss has been especially burdensome for poor villages where the transfers in the post-reform periods from upper level governments were far from sufficient to offset the revenue loss (Alm and Liu, forthcoming). In order to meet basic public services provision and village administrative expenses, the TFR may have imposed a double burden on farmers in poor villages; that is, they were asked to pay the new agricultural taxes, and they also had to pay some illegal fees (Yep, 2004; Yi, 2006). In contrast, because rich villages did not depend as heavily on informal taxes and fees for financing public expenditures, their levels of public revenues and expenditures did not change significantly after the reform.[32] It is thereby unsurprising to observe that the TFR only had significant impacts in rich villages.

Further Discussion

Our empirical results reveal a positive role of the TFR in reducing within-village inequality in rural China. This outcome can be due to the actions of any or a combination of the potential three channels discussed earlier: imposing greater fiscal discipline at the local level (along with other intergovernmental fiscal adjustments), changing household income sources, and reducing the regressivity of the previous tax system in rural China. Which of these channels emerges as most relevant?

[31] In fact, when we calculate the mean values of the Gini coefficient of per capita household net income for coastal versus inland villages and for rich versus poor villages in year 1998, these calculations indicate that rich villages have higher income inequality than poor villages before the TFR.

[32] 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 there on agricultural taxation.

As discussed in some previous studies (Zhang et al., 2004; Luo et al., 2007; Alm and Liu, forthcoming), the TFR significantly reduced village capacity to provide public good services. However, the net impacts of this first channel cannot be identified explicitly because doing so requires information on how different income groups within villages are affected by various village-level factors, and this information is not available.

It is also possible that other changes in village governance following the TFR contributed to the TFR impact on village inequality. The TFR not only changed local taxes, but it also altered other aspects of the intergovernmental fiscal system in rural areas. Following the TFR, upper-level governments (for the most part county governments) may have either increased their transfers to villages or assumed for themselves some expenditure responsibilities of villages. In either case, the result may have been to reduce the fiscal demands on village governments, and thereby to reduce their need for regressive taxes.

It is not possible to disentangle these various factors because detailed information on the post-TFR responses of upper-level governments is not available. However, such responses from upper-level governments seem 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 expenditures 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 teachers, funds for paying village teachers and maintaining school facilities seem likely to have been reduced after the TFR (Yep, 2004).[33]

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. However, the transfer allocation process at the county level is quite complex, which significantly reduces their effectiveness.[34]

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, 2008b). 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

---

[33] For a detailed analysis of education in rural areas of China and its relationship with poverty, see Knight, Li, and Quheng (2009, 2010); see also Wang, Zheng, and Zhao (2012) for an analysis of central government education reforms and their impacts on local government fiscal behavior. Gong and Wu (2012) examine central government mandates on local governments, and they conclude that local governments were not always compliant.

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

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.

The identification of the second channel (i.e., household income sources) is also difficult. As described in the data section, the 2002 CHIP survey only provides detailed household income sources and consumption components for the year 2002, while retrospective information for household income and consumption is only provided at the total level for the year 1998. The lack of available data therefore does not allow us to properly capture this channel using our DID matching technique.

Although not decisive, we are left with the third and the most direct channel (i.e., reducing the regressivity of the previous tax system) as the likely main factor in the entire process. The elimination of all local levies and informal fees and taxation in the TFR had a direct and immediate effect that significantly reduced the regressivity of the tax system in China. This impact seems likely to have been a major contributing factor to lower inequality within villages. This last explanation echoes the more qualitative argument of Sato et al. (2008).

6. Conclusions

In this paper we examine the impacts of the Tax-for-Fee Reform on village-level inequality in rural China, using a difference-in-difference propensity score matching approach. The presence of an unbalanced distribution of covariates between treated and control villages in our sample suggest that employing proper techniques to account for selection bias is a significant issue.

Our results suggest an effective role of the TFR in reducing inequality within villages, as measured by per capita household net income and per capita household consumption, although there tends to a lagged effect of the TFR especially on income inequality, due to the time needed for economic agents to learn and to adjust their behaviors to the policy change. We also find that it is rich and/or coastal villages that tend to experience a significant reduction of inequality from the TFR reform, while poor and/or inland villages experienced no significant changes in inequality from the reform. However, although our results provide evidence that the TFR reduced inequality in at least some villages, we should caution that these results do not seem to apply to all villages. Indeed, we find no significant TFR impact in poor and/or inland villages.

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 (Lin and Liu, 2007). Our findings are consistent with the likelihood that village government expenditures before the TFR were mainly financed by taxing villagers via informal taxes and fees, which in turn generated a heavy and regressive fiscal burden on those living in these villages. The formalization of the tax system after the TFR ruled out the possibility of villages financing their expenditures through imposing informal taxes and fees on villagers, as they had done prior to the reform. The result was a decline in the use of regressive taxes and a consequent reduction in inequality.

Overall, our work contributes to a better understanding of the performance of the recent TFR. On top of the purely economic significance of the redistributive effects of the policy reform, the reduction of within-village inequality itself is also consistent with a positive impact on the happiness of the residents in rural China. Recent studies in this regard (Knight et al., 2009; Knight and Gunatilaka, 2010a, 2010b, 2011, 2012) have demonstrated that happiness in rural

China is determined both by the famers' absolute income but also by their aspirations, which in turn are governed by their income relative to a "reference group", mostly neighbors or fellow villagers inside their village.[35] In light of this, our finding that the TFR reduced within-village inequality is consistent with a positive role of the TFR in promoting happiness of farmers, since the reform put them in a more advantageous (relative) position within the village income distribution. Even so, systematic evaluation of the TFR requires incorporating all possible impacts of reform into consideration, and further policy reforms should be built on the knowledge – both positive and negative – that we have learned.

References

Alm, J., & Liu, Y. (forthcoming) Did China'sTax-for-Fee Reform improve farmers' welfare in rural areas? *Journal of Development Studies*.

Aubert, C., & Li, X. (2002) Peasant burden: Taxes and levies imposed on Chinese peasants. *Agricultural Policies in China after WTO Accession*(Paris, France: OECD).

Benjamin, D., Brandt, L., & Giles, J. (2005) The evolution of income inequality in rural China. *Economic Development and Cultural Change, 53* (4), 769-824.

Benjamin, D., Brandt, L., Glewwe, P., & Guo, L. (2000) Markets, human capital, and inequality: Evidence from rural China. William Davidson Institute Working Papers Series 298, University of Michigan (Ann Arbor, MI).

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.

Brockmann, H., Delhey, J., Welzel, C., & Yuan, H. (2009) The China puzzle: Falling happiness in a rising economy. *Journal of Happiness Studies, 10* (4), 387-405.

Cai, H., Chen, Y., & Zhou, L.-A. (2010) Income and consumption inequality in urban China: 1992-2003. *Economic Development and Cultural Change*, *58* (3), 385-413.

Caliendo, M., & Kopeinig, S. (2008) Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22* (1), 31-72.

Chen, X. (2003) *A Study on China's County and Township Public Finance and Farmer Income Growth* (Taiyuan: Shanxi Economic Press) (in Chinese).

Deaton, A. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* (Baltimore, MD: The John Hopkins University Press).

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.

[35] Brockmann et al. (2009) also attribute the decline of happiness in China for the past decade to the rising income inequality in China in the same period.

Dehejia, R. H., &Wahba, S. (2002) Propensity score matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84* (1)*,* 151-161.

Deltas, G. (2003) The Small-Sample Bias of the Gini Coefficient: Results and Implications for Empirical Research. *The Review of Economics and Statistics, 85* (1), 226-234.

Gong, T., & Wu, A. M. (2012) Central mandates in flux: Local noncompliance in China. *Publius – The Journal of Federalism*, *42* (2), 313-333.

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.) (2008) *Income Inequality and Public Policy in China* (New York, NY: Cambridge University Press).

Ha, W., Yi, J., & Zhang, J. (2009) Internal migration and income inequality in China: Evidence from village panel data. MPRA Paper 16896, University Library of Munich (Munich, Germany).

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., Orley, C. A., & David, C. (1999) The economics and econometrics of active labor market programs (Chapter 31). In: O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 3, Part 1* (Amsterdam, The Netherlands: Elsevier B. V. – North Holland Publishers, 1865-2097).

Kanbur, R., & Zhang, X. (1999) Which regional inequality? The evolution of rural-urban and inland-coastal inequality in China from 1983 to 1995. *Journal of Comparative Economics*, *27* (4), 686-701.

Kanbur, R., & Zhang, X. (2005) Fifty years of regional inequality in China: A journey through central planning, reform, and openness. *Review of Development Economics*, *9* (1), 87-106.

Kennedy, J. J. (2007) From the Tax-for-Fee Reform to the abolition of agricultural taxes: The impact on township governments in north-west China. *The China Quarterly*, 189, 43-59.

Khan, A. R., & Riskin, C. (1998) Income and inequality in China: Composition, distribution and growth of household income, 1988 to 1995. *The China Quarterly*, 154, 221-253.

Knight, J., & Gunatilaka, R. (2010a) The rural-urban divide in China: Income but not happiness? *Journal of Development Studies*, *46* (3), 506-534.

Knight, J., & Gunatilaka, R. (2010b) Great expectations? The subjective well-being of rural-urban migrants in China. *World Development, 38* (1), 113-124.

Knight, J., & Gunatilaka, R. (2011) Does economic growth raise happiness in China? *Oxford Development Studies, 39* (1), 1-24.

Knight, J., & Gunatilaka, R. (2012) Income, aspirations and the hedonic treadmill in a poor society. *Journal of Economic Behavior and Organization, 82* (1), 67-81.Knight, J., Li, S., & Quheng, D. (2009) Education and the poverty trap in rural China: Setting the trap. *Oxford Development Studies*, *37* (4), 311-332.

Knight, J., Li, S., & Quheng, D. (2010) Education and the poverty trap in rural China: Closing the trap. *Oxford Development Studies*, *38* (1), 1-24.

Knight, J., & Li, S. (1997) Cumulative causation and inequality among villages in China. *Oxford Agrarian Studies*, *25* (2), 149-172.

Knight, J., Song, L., & Gunatilaka, R. (2009) Subjective well-being and its determinants in rural China. *China Economic Review, 20* (4), 635-649.
Kung, J. K. S., & Lee, Y.-f. (2001) So what if there is income inequality? The distributive consequence of nonfarm employment in rural China. *Economic Development and Cultural Change*, *50* (1), 19-46.
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, J. (2000) Changes in the source of China's regional inequality. *China Economic Review*, *11* (3), 232-245.
Lee, M. J. (2005) *Micro-econometrics for Policy, Program, and Treatment Effects*, Advanced Text Series in Econometrics (Oxford, UK: Oxford University Press).
Li, L. C. (2006) Embedded institutionalization: Sustaining rural tax reform in China. *The Pacific Review*, *19* (1), 63-84.
Li, H., & Zahniser, S. (2002) The determinants of temporary rural-to-urban migration in China. *Urban Studies*, *39* (12), 2219-2235.
Li, S. (2002) *Chinese Household Income Project* (Ann Arbor, MI: Inter-university Consortium for Political and Social Research).
Lin, J. Y., & Liu, M. (2007) Rural informal taxation in China: Historical evolution and an analytic framework. *China and 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.
Milanovic, B. (1999) Explaining the increase in inequality during transition. *Economics of Transition, 7* (2), 299-341.
Morduch, J., & Sicular, T. (2002) Rethinking inequality decomposition, with evidence from rural China. *Economic Journal*, *112* (476), 93-109.
Narayan, D., & Pritchett, L. (1999) Cents and sociability: Household income and social capital in rural Tanzania. *Economic Development and Cultural Change, 47* (4), 871-897.
Park, A., & Wang, S. (2010) Community-based development and poverty alleviation: An evaluation of China's poor village investment program. *Journal of Public Economics*, *94* (9-10), 790-799.
Perkins, T. (2002) *Village, Market, and Well-being in a Rural Chinese Township* (London, UK: Routledge).
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.
Rozelle, S. (1994) Rural industrialization and increasing inequality: Emerging patterns in China's reforming economy. *Journal of Comparative Economics, 19* (3), 362-391.
Rubin, D. B. (1977) Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics*, *2* (1), 1-26.
Sato, H. (2003) *The Growth of Market Relations in Post-Reform Rural China. A Micro-analysis of Peasants, Migrants and Peasant Entrepreneurs* (London, UK: Routledge Curzon).

Sato, H. (2008a) The impact of village-specific factors on household income in rural China. In: H. Sato, B.A. Gustafsson, S. Li, and T. Sicular (Eds.), *Inequality and Public Policy in China* (Cambridge, UK: Cambridge University Press, 287-311).
Sato, H. (2008b) Public goods provision and rural governance in China. *China: An International Journal*, 6 (2), 281-298.
Sato, H., Li, S., & Yue, X. (2008) The redistributive impact of taxation in rural China, 1995-2002: An evaluation of rural taxation reform at the turn of the century. In: H. Sato, B.A. Gustafsson, S. Li, and T. Sicular (Eds.), *Inequality and Public Policy in China* (Cambridge, UK: Cambridge University Press, 312-336).
Shen, Y., & Yao, Y. (2008) Does grassroots democracy reduce income inequality in China? *Journal of Public Economics, 92* (10-11), 2182-2198.
Shi, X., Nuetah, J. A., &Xin, X. (2010) Household income mobility in rural China: 1989-2006. *Economic Modelling*, *27* (5), 1090-1096.
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 LaLonde'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 and the World Economy*, *15* (3), 19-32.
Todd, P. E. (2008) Matching estimators. In: S. N. Durlauf and L. E. Blume (Eds.) *The New Palgrave Dictionary of Economics* (London, UK: Palgrave Macmillan).
Tsui, K.-Y. (1993) Decomposition of China's regional inequalities. *Journal of Comparative Economics*, *17* (3), 600-627.
Tsui, K.-Y. (1998) Trends and inequalities of rural welfare in China: Evidence from rural households in Guangdong and Sichuan. *Journal of Comparative Economics*, *26* (4), 783-804.
Wan, G., & Zhou, Z. (2005) Income inequality in rural China: Regression-based decomposition using household data. *Review of Development Economics*, *9* (1), 107-120.
Wang, W., and Zhao, Z. J. (2012) Rural taxation reforms and compulsory education finance in China. *Journal of Public Budgeting, Accounting & Financial Management*, *24* (1), 136-162.
Wang, W., Zheng, X., & Zhao, Z. J. (2012) Fiscal reform and public education spending: A quasi-natural experiment of fiscal decentralization in China. *Publius – The Journal of Federalism*, *42* (2), 334-356.
Xing, L., Fan, S., Luo, X., & Zhang, X. (2009) Community poverty and inequality in western China: A tale of three villages in Guizhou province. *China Economic Review*, *20* (2), 338-349.
Yao, S. (1997) Industrialization and spatial income inequality in rural China, 1986-921. *Economics of Transition, 5* (1), 97-112.
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: Definition of Variables

Variable	Definition
Treatment Variable	
Treat	Indicator variables for a village that implemented the TFR
Observable Variables in Probit Model	
Geographic	Geographic condition, where 1 represents a plain (basin) area, 2 represents a hilly area, and 3 represents a mountainous area
Suburb	Suburb of large/middle city
Dist_county	Distance from nearest county seat (in km.)
Dist_town	Distance from nearest township government (in km.)
Minority	1 if the village is in ethnic minority area, 0 otherwise
Poverty	1 if the county of the village is the national/province level poverty county, 0 otherwise
Election98	1 if the village has direct elections by the end of 1998, 0 otherwise
Pop98	Total population at the end of 1998
Pop98_squ	Square of pop98
l_Netincome_per98	Log of net income per capita in the village, 1998 (in RMB)
l_Netincome_per90	Log of net income per capita in the village, 1990(in RMB)
l_Planting98	Log of total planting area at the end of 1998 (in Mu.)
Disaster98	1 if there is any natural disaster during the year of 1998, 0 otherwise
Officer98	1 if any county level or upper level officials originated from this village during the years 1990-1998, 0 otherwise
Outcome Variables	
Δ(Gini_Netinc)	Difference of within-village Gini coefficient of per capita household net income between year 1998 and 2002
Δ(Gini_Consum)	Difference of within-village Gini coefficient of per capita household consumption between year 1998 and 2002

Table 2: Descriptive Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Treat	935	0.716	0.451	0	1
Geographic	933	1.734	0.792	1	3
Suburb	935	0.076	0.265	0	1
Dist_county	929	23.899	20.339	0.5	160
Dist_town	927	4.868	4.858	0	40
Minority	935	0.145	0.353	0	1
Poverty	935	0.294	0.456	0	1
Election98	935	0.787	0.410	0	1
Pop98	934	1770.263	1113.077	178	8586
Pop98_squ	934	4371446	6553519	31684	7.37E+07
ln_Netincome_per98	928	6.836	0.531	4.856	8.886
ln_Netincome_per90	878	7.007	0.644	5.004	8.826
ln_Planting98	929	7.859	0.897	1.099	10.015
Disaster98	935	0.471	0.499	0	1

Officer98	935	0.545	0.498	0	1
Δ(Gini_Netinc)	906	0.031	0.081	-0.262	0.380
Δ(Gini_Consum)	906	0.042	0.111	-0.355	0.508

Table 3: Mean Differences in Characteristics of Treated and Control Villages

Variable	Control Villages	Treated Villages	t-values
Observable Variables in Probit Model			
Geographic	1.856	1.686	2.966
Suburb	0.109	0.063	2.413
Dist_county	30.905	21.118	6.773
Dist_town	6.456	4.239	6.396
Minority	0.342	0.067	11.476
Poverty	0.293	0.294	-0.037
Election98	0.688	0.827	-4.723
Pop98	2037.590	1663.813	4.683
Pop98_squ	6436719	3549047	6.198
ln_Netincome_per98	6.814	6.845	-0.816
ln_Netincome_per90	6.912	7.048	-2.894
ln_Planting98	7.995	7.805	2.932
Disaster98	0.496	0.460	0.990
Officer98	0.617	0.517	2.761
Outcome Variables			
Δ(Gini_Netinc)	0.025	0.034	-1.376
Δ(Gini_Consum)	0.037	0.044	-0.879
Number of Villages	248	658	

Table 4: Propensity Score for the TFR Implementation – Probit Estimation Results

Variable	Coefficient (Standard Error)
Dist_county	-0.011*** (0.003)
Dist_town	-0.042*** (0.011)
Suburb	-0.305 (0.188)
Minority	-1.341*** (0.152)
Poverty	0.105 (0.133)
Election98	0.529*** (0.121)
Pop98	0.000*** (0.000)
Pop98_squ	-0.000*** (0.000)
ln_Netincome_per98	-0.581*** (0.144)
ln_Netincome_per90	0.340*** (0.107)
ln_Planting98	0.020 (0.071)
Officer98	-0.110 (0.107)
Constant	2.123** (1.001)
Observations	835

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

Table 5: Balancing Tests from DID Kernel Matching Estimators

Variable	Sample Mean		Percent Bias	Percent Bias Reduction	t-statistic (p-value)
	Treated	Control			
Suburb	0.061	0.055	1.9	89.4	0.39 (0.695)
Dist_county	20.858	20.613	1.1	97.6	0.26 (0.794)
Dist_town	4.286	4.549	-5.0	87.7	-1.15 (0.249)
Minority	0.057	0.070	-3.3	95.3	-0.86 (0.389)
Poverty	0.296	0.323	-5.9	-59.1	-1.02 (0.310)
Election98	0.834	0.840	-1.4	96.4	-0.28 (0.783)
Pop98	1685.3	1657.5	2.2	93.2	0.55 (0.582)
Pop98_squ	3600000	3500000	1.9	95.3	0.69 (0.489)
ln_Netincome_per98	6.860	6.895	-6.2	-31.4	-1.07 (0.286)
ln_Netincome_per90	7.044	7.053	-1.4	93.1	-0.23 (0.819)
ln_Planting98	7.854	7.772	9.2	43.1	1.57 (0.117)
Disaster98	0.454	0.446	1.7	82.6	0.29 (0.774)
Officer98	0.514	0.471	8.7	60.2	1.48 (0.140)

Table 6: Impact of TFR on Village Inequality – DID Kernel Matching Estimators

Reform Year	Δ(Gini_Netinc)		Δ(Gini_Consum)	
	ATT	t-statistic	ATT	t-statistic
Full sample	-0.009	-1.18	-0.024	-1.86*
1 (=2002)	-0.007	-0.83	-0.019	-1.66*
2 (=2001)	-0.005	-0.41	-0.036	-1.92*
3 (=1999 or 2000)	-0.030	-2.02**	-0.040	-1.83*

Note: Numbers ranging from 1 to 3 in the first column represent the post-reform periods that the particular treated villages have experienced. Given the relatively small number of villages (only 4 observations) that implemented in year 1999, those villages that implemented the reform in years 1999 and 2000 are grouped into the same subgroup. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Impact of TFR on Village Inequality – Alternative Matching Estimators

Reform Year	Δ(Gini_Netinc)		Δ(Gini_Consum)	
	ATT	t-statistic	ATT	t-statistic
DID Nearest Neighbor Matching Estimator				
Full sample	-0.014	-1.33	-0.045	-2.66***
1 (=2002)	-0.011	-0.95	-0.034	-1.94*
2 (=2001)	-0.027	-1.48	-0.058	-2.28*
3 (=1999 or 2000)	-0.042	-2.05**	-0.044	-1.90*
DID Caliper Matching Estimator				
Full sample	-0.011	-1.10	-0.045	-2.71***
1 (=2002)	-0.009	-0.81	-0.034	-1.95*
2 (=2001)	-0.027	-1.44	-0.048	-1.91*
3 (=1999 or 2000)	-0.043	-2.13**	-0.043	-1.82*

Note: Numbers ranging from 1 to 3 in the first column represent the post-reform periods that the particular treated villages have experienced. Given the relatively small number of villages (only 4 observations) that implemented in year 1999, those villages that implemented the reform in years 1999 and 2000 are grouped into the same subgroup.
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Impact of TFR on Village Inequality – Alternative Measures of Inequality

Inequality Measures	Theil Index Δ(Netinc)	Theil Index Δ(Consum)	Mean Log Deviation Δ(Netinc)	Mean Log Deviation Δ(Consum)	Atkinson Index Δ(Netinc)	Atkinson Index Δ(Consum)
Full sample	-0.008	-0.026*	-0.015	-0.025*	-0.005	-0.012*
	(-0.95)	(-1.66)	(-1.36)	(-1.82)	(-1.31)	(-1.76)
1 (=2002)	-0.003	-0.022	-0.012	-0.021	-0.004	-0.01
	(-0.39)	(-1.56)	(-1.18)	(-1.63)	(-0.89)	(-1.60)
2 (=2001)	-0.005	-0.038	-0.013	-0.038	-0.004	-0.017
	(-0.36)	(-1.49)	(-0.86)	(-1.61)	(-0.62)	(-1.60)
3 (=1999	-0.030**	-0.043*	-0.042**	-0.042*	-0.016**	-0.019*
or 2000)	(-2.30)	(-1.66)	(-2.37)	(-1.88)	(-2.52)	(-1.77)

Note: Numbers ranging from 1 to 3 in the first column represent the post-reform periods that the particular treated villages have experienced. Given the relatively small number of villages (only 4 observations) that implemented in year 1999, those villages that implemented the reform in years 1999 and 2000 are grouped into the same subgroup. Δ(Netinc) represents the difference of within-village inequality of per capita household net income between year 1998 and 2002. Δ(Consum) represents the difference of within-village inequality of per capita household consumption between year 1998 and 2002. The Atkinson index is computed with a value of "inequality aversion" equal to 0.5. *t*-values are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Impact of TFR on Village Inequality – Coastal versus Inland Villages and Rich versus Poor Villages (DID Kernel Matching Estimators)

	Δ(Gini_Netinc) ATT	Δ(Gini_Netinc) t-statistic	Δ(Gini_Consum) ATT	Δ(Gini_Consum) t-statistic
Coastal versus Inland Villages				
Coastal	-0.016	-1.42	-0.031	-1.76*
Inland	0.013	1.31	0.001	0.04
Rich versus Poor Villages				
Rich	-0.013	-1.17	-0.038	-2.17**
Poor	-0.007	-0.51	-0.005	-0.27

** p<0.05, * p<0.1

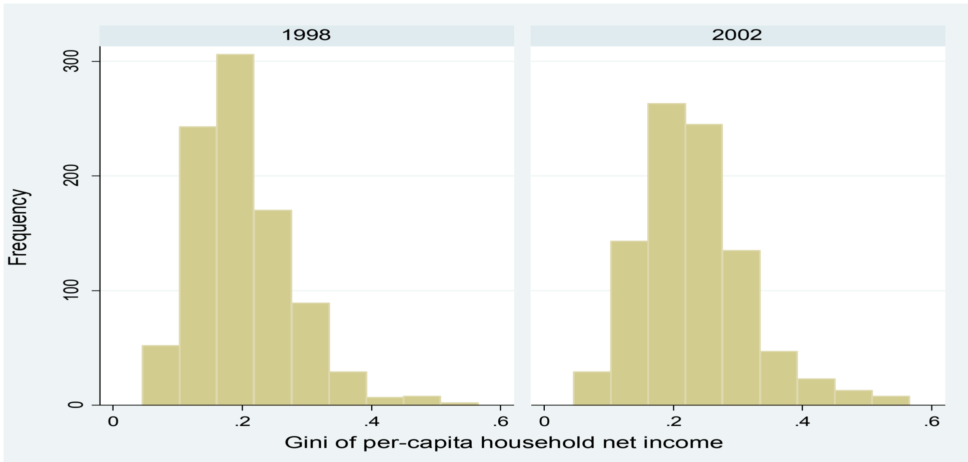
Table 10: Impact of TFR on Village Inequality – Coastal versus Inland villages and Rich versus Poor Villages (Alternative Measures of Inequality)

Inequality Measures	Theil Index Δ(Netinc)	Theil Index Δ(Consum)	Mean Log Deviation Δ(Netinc)	Mean Log Deviation Δ(Consum)	Atkinson Index Δ(Netinc)	Atkinson Index Δ(Consum)

Coastal versus Inland Villages						
Coastal	-0.01	-0.029	-0.019	-0.027	-0.006	-0.013
	(-0.83)	(-1.24)	(-1.46)	(-1.29)	(-1.19)	(-1.32)
Inland	0.012	0.004	0.011	-0.002	0.005	0.0004
	(1.09)	(0.32)	(0.69)	(-0.20)	(0.87)	(0.07)
Rich versus Poor Villages						
Rich	-0.012	-0.041*	-0.028*	-0.036*	-0.008	-0.017*
	(-0.99)	(-1.68)	(-1.87)	(-1.72)	(-1.46)	(-1.74)
Poor	-0.008	-0.008	-0.008	-0.008	-0.004	-0.004
	(-0.62)	(-0.51)	(-0.52)	(-0.50)	(-0.72)	(-0.50)

Note: Δ(Netinc) represents the difference of within-village inequality of per capita household net income between year 1998 and 2002. Δ(Consum) represents the difference of within-village inequality of per capita household consumption between year 1998 and 2002;. The Atkinson index is computed with a value of "inequality aversion" equal to 0.5. *t*-values are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1. Gini Coefficient Histograms, 1998 and 2002



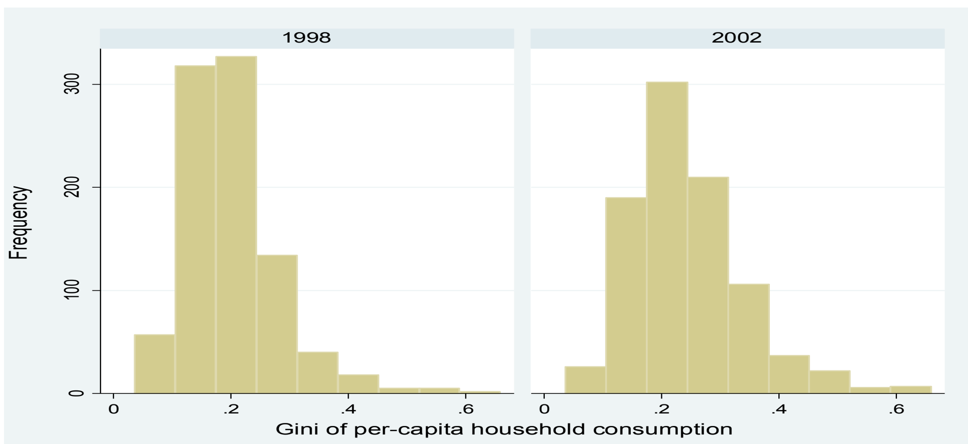


Figure 2. Propensity Score Histograms, Control versus Treated Villages