

Fiscal Decentralization, Fiscal Reform, and Economic Growth in China

Yi Ding, Alexander McQuoid, Cem Karayalcin*

July 30, 2018

Abstract

In 1994, after a period of substantial fiscal decentralization that has been credited with leading to historically unprecedented growth rates but significant fiscal decline, China introduced a new fiscal system that recentralized the collection of tax revenues. The economic and political consequences of this new Tax Sharing System (TSS) have been debated extensively in the literature, especially because of the renewed interest in fiscal federalism and its interaction with political institutions and economic outcomes. The question central to this debate has been whether the TSS constitutes a significant departure from decentralization with adverse effects on fiscal federalism or whether the recentralization of revenues under the TSS corrects for the overshooting in decentralization with beneficial economic outcomes. This paper exploits the staggered introduction of the TSS across regions and over time for econometric identification purposes and finds robust causal evidence that the TSS had a positive impact on economic outcomes.

*Ding: Tianjin Precious Metal Exchange, Dongli District, Tianjin, China (email: dingalpha@gmail.com); McQuoid: Economics Department, US Naval Academy, 589 McNair Road Annapolis, Maryland 21402-5030 (email: mcquoid@usna.edu); Cem Karayalcin, Department of Economics, Florida International University, 11200 SW 8 St, Miami, FL, 33199 (email: karayalc@fiu.edu).

1 Introduction

In 1994, China fully and formally introduced a new fiscal system under the name the “Tax Sharing System” (TSS or fenshuizi (tax assignment)). TSS reassigned taxes between the central and subnational governments (SNGs) with the intention of reducing the discretion of the latter in negotiating the sharing of revenues and ensuring a greater share for the former. This paper studies the growth consequences of the TSS by exploiting the staggered introduction of the system across regions and over time for econometric identification purposes.

The significance of TSS, which has been called a watershed event¹ for the Chinese fiscal system, lies in the question as to whether it constitutes a notable divergence from the “Fiscal Responsibility System (FRS, also known as the fenzao chifan or “eating from separate kitchens”) that preceded it and which has been credited with placing the Chinese economy on the high growth path it followed after 1980.² To put this claim in perspective note that the FRS is generally recognized as being a highly fiscally decentralized system that stands in stark contrast to the centralized system that allowed only minor amounts of local discretionary spending in the period 1949-1979.

We find that TSS treatment increased growth rates by economically meaningful amounts, with estimates that are robust to alternative specifications and controls. While methods based on endogenous proxies for fiscal federalism used by the existing literature are unstable and unreliable, our preferred approach using difference-in-difference estimation tells a consistent and compelling story: TSS reform increased economic growth, with highly significant estimates that are robust to alternative controls and inferential assumptions. In particular, we find that TSS reform resulted in per capita growth rates that were approximately 18% higher compared to the average per capita growth rates in the pre-reform (FRS) period (2.2 percentage points higher relative to a baseline per capita annual growth rate of 11.9 percentage points prior to fiscal reform).

Early theories of public finance (now commonly called “first generation theories of fiscal federalism”) had emphasized potential welfare gains from the decentralized financing of local public

¹See World Bank (2002).

²See Montinola et al. (1995) and Jin et al. (2005) among others.

goods and services in contrast to the centralized provision of a single, uniform public output in all jurisdictions.³ Such gains are viewed to flow from the relaxing of informational problems faced by central authorities who have only imperfect information with regard to local tastes and conditions. These gains, however, need to be contrasted with the efficiency cost of failure to internalize interjurisdictional externalities.

Second generation theories of fiscal federalism, on the other hand, emphasize the benefits decentralization affords from a political and institutional perspective. Thus, centralized fiscal systems are seen as leading to extortionary behavior on the part of the central government, whereas competition among decentralized governments is believed to limit the capacities of the monopolist central government, thereby fostering local economic development, affording more accountability and better outcomes for the citizenry.⁴ Not every decentralized fiscal system would, however, lead to beneficial outcomes. Soft budget constraints faced by SNGs, mismatches in responsibility and resources, and local capture can easily derail a decentralized system, giving rise to opportunities for increased local rents and corruption.⁵

Under the FRS, the second generation theorists of fiscal federalism see the Chinese central government as delegating economic policy-making and significant fiscal authority with hard budget constraints to the SNGs.⁶ Given their new policy authority and facing fiscal incentives to promote reform, a number of SNGs adopted pro-market policies and raised their own tax revenue. Contracts negotiated with the central government allowed SNGs to share 50 percent of all revenue raised up to a pre-specified revenue level and to retain all revenue that exceeded this level. As Jin et al. (2005) show, the average province faced a marginal tax retention rate of 89 percent while 68 percent of all provinces faced a marginal retention rate of 100 percent. Facing strong fiscal incentives to promote reform, many SNGs experienced high rates of economic growth. With the reforms proving successful, the central government devolved additional incentives and political

³See Oates (1972 and 2005) for a formulation of this view. See Brueckner (2005) for theoretical work that links federalism with growth.

⁴See Brennan and Buchanan (1980), Besley and Coate (2003), and Weingast (1995).

⁵See Blanchard and Shleifer (2000) and Weingast (2014).

⁶Montinola et al., (1995), Oi (1992), Shirk (1993), and Weingast (2014).

power to the SNGs, which then acted as a counterbalance to the political power of the central government.

However, the FRS also resulted in significant fiscal decline for the central government, with the share of total tax revenue transferred to the central government falling from 39 percent in 1985 to 22 percent in 1993. Overall revenue collections declined from 31 percent of GDP in 1978 to 11.2 percent in 1994.⁷ The FRS thus led to an “overshooting in decentralization” and a rethinking of the fiscal system.⁸

The crucial question that emerged with the TSS reform package that replaced the FRS in 1994 was whether this meant a recentralization to the extent of a reversal to the old system or rather a move to further the process of rationalization of market institutions. Reversal would threaten the crucial political component that supported the success of economic reforms, whereas rationalization would constitute a shift towards a formal fiscal federalism that promised to resolve the coordination problems associated with an overly weak central government without endangering the benefits that accrue from decentralization.

The debate on whether the TSS reversed the decentralization course started by the FRS is still ongoing. The literature has taken two paths to answer this question. The first focuses on the elements of the new fiscal regime introduced by the TSS and compares these with either those of the FRS or of an ideal decentralized regime. The second tries to determine empirically whether the two regimes led to differences in economic outcomes. Thus, within the former literature, Zhang (1999), for instance, views the 1994 reform as stemming budgetary decline, but leaving the actual pattern of revenue distribution and spending responsibilities between the central government and the SNGs fundamentally unchanged. Weingast (2009) perceives TSS as continuing the market-preserving federalism initiated by the FRS. However, Wong (2000) associates the TSS with significant recentralization and emphasizes the fiscal problems (e.g. unclear expenditure assignments, unfunded mandates) that continue to plague fiscal relations between the center and

⁷Ahmad et al., (2002), Dabla-Norris (2005), Bahl and Martinez-Vazquez (2006). The positive connection between fiscal centralization and fiscal revenues has been observed widely across space and time.

⁸See, for instance, Montinola et al. (1995).

the provinces.

The empirical literature that follows the second path includes Zhang and Zou (1998) who, using a province-level panel data set for the 1978–1992 period, find that fiscal decentralization is inversely associated with provincial economic growth. Lin and Liu (2000), on the other hand, examine a similar province level panel data set for the longer 1970–1993 period and find evidence that fiscal decentralization is associated positively with economic growth. Jin and Zou (2005) use a panel data set of Chinese provinces covering the FRS and TSS periods of 1979–1993 and 1994–1999. They provide evidence that provincial economic growth is negatively associated with expenditure decentralization and positively associated with revenue decentralization in the 1979–1993 FRS phase. Further, they find that in the 1994–1999 TSS phase, there is no statistically significant relationship between provincial economic growth and expenditure decentralization, while there is a highly significantly association between provincial economic growth and revenue decentralization. Finally, Ding (2008) focuses on the TSS period of 1994–2002 and finds a positive association between province level economic growth and various measures of fiscal decentralization.

Existing empirical studies of the growth consequences of TSS reform have two primary challenges, measuring fiscal decentralization and causal identification. First, the appropriate measurement of fiscal decentralization is a difficult task given the complexity of intergovernmental relations.⁹ The empirical literature typically uses the share of subnational government expenditure (or revenue) in consolidated general government expenditure (or revenue) as a proxy for the degree of fiscal decentralization. However, it was recognized very early on¹⁰ that such proxies have very limited use as they do not take into account the vertical structure of decision-making and thereby misrepresent the degree of fiscal decentralization.

As a result, empirical studies of both cross-country comparisons and descriptions of long-term trends may be seriously distorted. For instance, it is not clear whether an observed increase in the SNG share of total public expenditure (or revenue) implies fiscal decentralization in the sense of devolution of decision-making powers to the SNG. A fundamental question that arises here

⁹See Stegarescu (2005) for an insightful discussion.

¹⁰See, for instance, Oates (1972).

is the extent to which SNGs can keep local tax revenue instead of transferring it to the central government and the autonomy they may have in making spending decisions.

In what follows, we move beyond the existing literature in this respect by first showing that the two most commonly used measures of fiscal federalism, namely the SNGs' share of total public revenue and expenditure, move in opposite directions with the introduction of the TSS. This highlights the problem with using endogenous variables to proxy for the object of interest, fiscal federalism. Instead, to analyze the effects of the TSS, we exploit the quasi-natural experiment afforded by its staggered introduction, which resulted in certain regions in China starting the reform earlier than other regions. Using a difference-in-difference strategy, we identify the impact of TSS reform on economic outcomes. This approach also allows us to deal with the second question raised by the existing empirical literature, namely that of causality.

As the extant literature leaves unanswered potential omitted variable bias in its empirical strategy, it is commonly recognized that the results obtained so far can only be used to establish associations between the economic outcomes investigated and measures of fiscal policy employed, but not a causal relationship between the two. The difference-in-difference approach adopted here provides a more satisfactory method in establishing causality.

The plan of the paper is as follows. In the next section we provide a brief description of the institutional and policy background for the TSS. Section 3 presents our empirical analysis with the main results. Section 4 concludes the paper.

2 Institutional Background

The “watershed” tax reform of TSS arose in an institutional setting where the most important fiscal element was the FRS, which was introduced in 1984. This system shifted the collection of most tax revenue to local governments under a fiscal contract with the central government. A typical fiscal contract allowed a given Chinese province to share 50 percent of the revenues it raised up to a certain level with the central government. Any revenues the province raised beyond that

level were retained by the province. In most cases, however, the central government was not able to monitor tax collection at the local level and was forced to renegotiate revenue shares with local governments who hid resources from it in extra-budgetary funds.

At the same time, the main taxpayers, state-owned enterprises (SOEs), experienced substantial declines in profitability as prices adjusted to newly established markets, reducing tax revenues. With SOEs in financial difficulty, a significant share of safety net expenditures (unemployment insurance, subsidies for housing and fuel, etc.) was shifted from the SOEs to the local governments, who now faced higher local expenditures. Thus, while overall tax revenues were in decline, expenditure burdens of local governments were on the rise.

The 1994 tax reform, the TSS, was introduced mainly to improve the so-called “two ratios”—the ratio of budgetary revenue to GDP and the ratio of central budgetary revenue to total budgetary revenue. The major elements of the new system were as follows.¹¹ First, in an attempt to reduce the complexity of the tax system due to the proliferation of taxes after 1984, the number of taxes was cut from 32 to 18. Second, surtaxes were eliminated and the top marginal income tax rate for enterprises was cut from 55 percent to 33 percent, with tax rate schedules being unified to subject all enterprises to the same regime.

Third, to reduce the incentive to hide revenues from the central government through deals with enterprises, SNGs were assigned all income tax revenues. Fourth, the VAT system was expanded with VAT remaining a central government tax, but to allow SNGs to share in the growth of their lost tax base, the central government committed to giving back 30 percent of its increased revenue from VAT and consumption tax (CT) each year. Fifth, a national tax system was established to collect central government revenues with a separate local tax system to collect local taxes.

Finally, and perhaps for our purposes most importantly, the central government committed to rebating to each province an amount equal to the reduction in the local tax base due to TSS. This last element was a compromise made by the central government to win the support of the SNGs for the new system. The central government announced that the special transfer mechanism,

¹¹Bahl and Marinez-Vazquez (2006), Wong (2000), and Zhang (1999).

called *shui shou fan huan* (tax repayment) would differ from normal central grants in that its expenditure would not be subject to central control. As a result, even though with the TSS the collection of the majority tax revenue was to be transferred from the provinces to the center, the central government would allow the provinces to automatically claim much of these funds.

The current view of the informal literature on the effects of TSS two decades after its introduction is that it successfully increased tax revenue collection, simplified the tax structure, and resulted in a recentralization of revenues. Nonetheless, revenue sharing with the SNGs continues to be ad hoc. Perhaps more importantly for the question of the devolution of decision-making powers to local governments, the TSS does not appear to have constituted a significant *de facto* break with the FRS: as the World Bank (2002, p.26) puts it “the TSS reform did not address the issue of extra-budgetary revenues, but instead gave implicit support to redoubled efforts by local governments to seek supplementary resources to support local economic development, in the process undermining the authority of the budget.” However, to go beyond the informal literature and to determine the consequences of the TSS for economic outcomes requires a formal econometric analysis that takes causal identification seriously. This is where we turn in the next section.

3 Empirical Analysis

In this section, we study the effects of 1994 tax reform empirically. The policy change that the TSS ushered in poses a unique experience for analyzing the impact of fiscal federalism because of the scale of institutional adjustment as well as the staggered implementation across localities.

In what follows, we will show that focusing on endogenous measures of the stance of fiscal federalism can be deeply misleading. The ratio of revenues received by the local government to total revenues received by the central government and the ratio of expenditures by the local government to total expenditures by central government are typically assumed to be proxies for fiscal federalism, but in the case of China, these two measures move in extreme opposite directions during the policy period under consideration, suggesting they cannot both be proxies for fiscal

federalism. This merely highlights the problem with using endogenous variables to proxy for the object of interest, fiscal federalism.

Instead, we exploit a quasi-natural experiment in the rollout and implementation of the TSS. Specifically, we take advantage of the fact that the reform was carried out in two stages. Our treatment group (denoted Group A), implemented reform in 1992, while our control group (denoted Group B) implemented the TSS in 1994. Group A consisted of nine regions, including Zhejiang Province, Liaoning Province, Xinjiang Autonomous region, Tianjin municipality, Shenyang City, Dalian City, Qingdao City, Wuhan City, and Chongqing City.¹² All other regions are included in Group B. Both Shenyang and Dalian belongs to Liaoning Province, Qingdao belongs to Shandong Province, Wuhan belongs to Hubei Province and Chongqing was a city affiliated with Sichuan Province until 1997. Since Shandong, Hubei, and Sichuan Provinces implemented TSS in 1994, but the cities of Qingdao, Wuhan, and Chongqing implemented TSS in 1992, we will consider alternative classifications of treatment in the analysis below for robustness.

This staggered introduction of the TSS allows us to use a difference-in-difference strategy to identify the impact of the reform on economic activity. While the process of selection of localities may not have been purely random, the pre-trends of observable characteristics suggest that there are no noticeable trend differences between the localities that received the treatment and those that did not. As such, we treat the application of treatment as random, and interpret the difference-in-difference estimates as the true causal impact of the TSS on local economic growth.

Our empirical approach begins by first following the traditional path in the fiscal federalism literature by considering the impact of fiscal federalism using typical proxy variables on growth rates at the locality level. Using this approach, we find conflicting results in the case of China, with revenue-based proxies of fiscal federalism implying that greater fiscal decentralization leads to lower growth rates, but expenditure-based proxies of fiscal federalism imply higher growth rates. These results are robust to the inclusion of controls for possible omitted variables, with the estimated effects becoming larger in magnitude once controls for locality heterogeneity are

¹²Shu-Ki and Yuk-Shing (1994).

included. Furthermore, point estimates vary wildly (and lose statistical significance) when the sample is split in to before and after reform periods. These results justify the concern over the use of endogenous proxies for fiscal federalism.

We then estimate the impact of the TSS by exploiting the unique policy environment and utilizing a difference-in-difference (DID) approach. Our DID estimation results suggest that policy treatment did increase economic growth. In our preferred specification, treatment increased per capita growth by 2.2 percentage points relative to the counterfactual of no TSS reform. To put this estimate into context, the average per capita growth rate across localities during the FRS period was 11.9%, meaning that the introduction of TSS increased per capita growth rates by roughly 18% relative to this baseline. For robustness, using alternative specifications, the range of estimates for the impact of TSS varies between 1.7 and 4.0 percentage points.

Finally, we consider an alternative measure of economic activity by considering income per capita. Once again, endogenous proxies for fiscal federalism tell conflicting stories, none of which are robust to sample period selection. When we exploit the natural experiment, however, our DID estimation identifies a significant and robust effect of treatment on the level of economic activity. Our estimates imply that TSS increased per capita income by 11.7% relative to the counterfactual of no TSS reform. To help put this number in context, the median locality per capita income in 1991 was 1,657 yuan. Our point estimate implies that per capita incomes were 194 yuan higher than they would have otherwise been in the absence of TSS reform.

3.1 Data

The regional data is drawn from the Department of Comprehensive Statistics (1999), and covers all 31 provinces of Mainland China between 1980 and 1999.¹³ Variables include information on revenues and expenditures of each locality, disaggregated by source and type. The dataset also includes information on locality GDP and population, as well as exposure to international forces

¹³In Mainland China, there are three different names used for effectively the same administrative level: provinces, municipalities, and autonomous regions. For comparative purposes, these are similar in scope to U.S. states. Presently, there are 23 provinces, 5 autonomous regions, and 4 municipalities in China.

through locality imports and exports. The dataset gives a complete portrait of economic conditions at the locality level.

Of the 31 localities, 4 were dropped due to incomplete or unreliable data. Tibet and Hainan were dropped because of missing data. Chongqing was a city affiliated with the Sichuan province, and was not directly supervised by the central government until 1997, when it became one of four municipalities. For these reasons, Tibet, Hainan, Chongqing, and Sichuan were dropped from the analysis. This leaves us with 27 localities, 20 years of data, and 540 total observations. Summary statistics of key variables are reported in Table 1.

One particularly noteworthy aspect of the dataset is the amount of heterogeneity observed across these localities and over time. There is significant variation in GDP, population, local expenditures and revenues, openness to international trade, and local expenditure priorities. Figure 1 shows heterogeneity in growth rates over time for each locality, while Figure 2 emphasizes the significant cross-sectional variation over time. Furthermore there is significant variation in locality measures of expenditures of agricultural support programs, expenditure share on education, expenditure share on government administration, and capital construction expenditure shares. Given so much observable heterogeneity, our approach will take care to control for individual locality effects that may influence economic growth.

In our difference-in-difference implementation, we rely on the structure of the timing of reform to identify the effect of the TSS relative to the FRS. To this end, we restrict our dataset to the years 1985-1999, which covers both the FRS and TSS phases of fiscal federalism in China. To interpret our estimates as causal, we would ideally want random assignment of treatment and control. Unfortunately, the exact selection process for treatment has not been made public, and was likely not explicitly random. However, observable pre-treatment characteristics of the localities that implemented TSS in 1992 are statistically similar to the pre-treatment characteristics of those localities that implemented TSS in 1994.

Table 2 provides comforting evidence about the observable differences between control and treatment groups. When all types of treatment are grouped together, a number of observable

variables are significantly different than those in the control group, raising the specter that unobservable variables are also significantly different. However, some localities (Shandong and Hubei Provinces) did not implement TSS fully in 1992. Rather, two important cities (Qingdao and Wuhan) within these localities did implement TSS in 1992, while the rest of the Provinces adopted TSS in 1994. We further distinguish types of treatment by separating out these two provinces as "partial treatment" since only a part of the province adopted TSS in 1992. When treatment is broken up into full and partial treatment groups, it becomes clear that the partial treatment group is driving the observed differences. We will take care in the empirical analysis below to consider the impact of different types of treatment.

With the exception of population size, there are no statistically significant differences between the twenty-one control localities and the four full treatment localities, supporting our interpretation of full treatment as random. In our analysis below, we will be careful to consider per capita economic measures to account for potential differences in size across localities. For the two partial treatment localities, there are however significant differences compared to the control group. As such, we will take care to consider separate treatment effects for partial and full treatment, as well as include control variables to account for observable differences across localities.

While there are some statistical differences in observable means, especially for partial treatment localities, Figure 3 shows that pre-treatment trends in $\ln(\text{GDP per capita})$ for control and treatment (any treatment) groups do not differ. Figure 4 shows growth rates per capita in the pre-treatment period, and although there are notable year to year fluctuations in growth rates, the treatment group is not statistically different from the control group in this regard. The group mean evidence presented in Table 2 along with the parallel trends evidence in Figures 3 and Figure 4 motivates our use of locality and year fixed effects in our DID specification, and lends support to our identification assumption of random application of treatment.

3.2 Empirical Implementation

We begin our analysis by exploring the relationship between measures of fiscal federalism and economic growth, following the traditional approaches in the fiscal federalism literature. In the process, the concerns over the interpretation of traditional approaches will be highlighted.

To get a first cut at the data, consider Table 3, which reports regression results for three different proxies of fiscal federalism and their impact on local economic growth rates per capita. Our first proxy considers the ratio of local revenue to total central revenue in China. The idea is that as the share of local revenue increases relative to central revenue, the local government gains more economic freedom over local choices through higher available local revenue. Results in Table 3 imply that a higher ratio of local to central revenue leads to slower economic growth per capita, suggesting that fiscal decentralization actually reduces economic growth.

The second standard measure of fiscal federalism is the ratio of local tax revenue to central tax revenue, an alternative measure of fiscal decentralization based on revenue flexibility. As reported in Table 3, the estimated effect of enhanced fiscal decentralization reduces economic growth per capita, consistent with the measure based on total revenue rather than just tax revenue. In both cases, revenue based proxies for fiscal federalism suggest that decentralization hurts local economic growth.

When we consider expenditure-based proxies of fiscal federalism, however, we find evidence in support of the opposite conclusion. The ratio of local expenditure to central expenditure has a positive effect on growth per capita, and is also statistically significant. This suggests that as localities gain more control over local expenditures relative to the central government, growth rates increase.

One limitation with the results in the first part of Table 3 is that omitted variables may be correlated with the measures of fiscal federalism, and this may be influencing the observed estimates leading to conflicting results. The standard approach to an omitted variables bias is to include sufficient relevant control variables, although specifying which control variables to include is problematic. To attempt to correct for omitted time varying factors at the locality

level, the second part of Tables 3 reports estimates for fiscal federalism proxies after controlling for observable locality factors such as spending on social programs (education, health, culture, and science), government administration, capital construction, as well as spending on agriculture support programs.

Consider first a revenue-based proxy of fiscal federalism, the ratio of local to central revenues. Our raw point estimate is negative and statistically significant (column (1)). A series of regression results are reported in columns (4) through (8), as we systematically include relevant locality specific factors such as spending on government administration, capital construction, and social services such as health and education, and agricultural programs. In all cases, the coefficient on the total revenue proxy of fiscal federalism is negative and statistically significant. In fact, when all relevant controls are included (column (8)), the point estimate increases nearly 50% relative to the raw coefficient reported in column (1), with even stronger statistical significance. We estimate that an increase in the ratio of local revenue to central revenue by one unit leads to a 1.1 percentage point decrease in the growth rate per capita, a non-trivial effect.

We next conduct the same basic exercise, except we look at only local tax revenue, which strips out any central government transfers that are often earmarked for specific spending projects. Once again, the inclusion of controls has no effect on the sign or significance of the key variable of interest, but interestingly the point estimate is nearly identical in the preferred specification with all relevant controls (Column (9)). The impact of a one-unit increase in the ratio of local tax revenue to central tax revenue results in a 1.7 percentage point decrease in economic growth with high statistical significance. At least for revenue-based proxies for fiscal federalism, the result is quite robust: greater fiscal decentralization leads to lower economic growth.

Lastly, we conduct the same exercise, but focus on an expenditure-based measure of fiscal federalism. And once again, the results originally found in Column (3) stand up to the inclusion of locality specific time varying controls. In all cases, the point estimate is positive and statistically significance (results not reported, but available on request). When all relevant controls are included (Column (10)), the point estimate is nearly 4 times as large as that found in the simple bivariate

case, and highly statistically significant. A one-unit increase in the ratio of local expenditure to central expenditure is associated with a 4.5 percentage point increase in economic growth per capita.

These conflicting results, robust to a battery of controls, suggest that endogenous proxies for fiscal federalism are unsuitable for estimating the effect of fiscal federalism on economic outcomes such as economic growth. Traditional approaches for measuring and estimating the effect of fiscal federalism are likely to be conflicting, confusing, and unreliable because they don't directly measure fiscal federalism, but rather rely on endogenous proxies for fiscal federalism.

To reinforce this concern, Table 4 considers these three proxies before and after the policy experiment. Not only do revenue and expenditure based measures tell different stories, each measure is not robust to sample period selection, losing significance or even switching sign before and after the policy experiment under consideration. This should not be surprising since these are endogenous proxies, and the policy experiment under consideration had differential impacts on revenues and expenditures.

While the literature has traditionally focused on the impact of fiscal federalism on growth rates, to provide a richer portrait of economic consequences of the TSS reform, we also explore the impact on GDP levels. Next, we show that changing the economic outcome of interest does not resolve the concerns about using endogenous proxies for fiscal federalism.

Following the same basic argument as with growth rates, we start with the simple relationships between GDP per capita and proxies for fiscal federalism using revenue and expenditure measures in Table 5. As with growth rates, revenue-based measures of fiscal decentralization are correlated with lower levels of GDP per capita, while our expenditure-based measure of fiscal decentralization is associated with higher GDP levels per capita. In all cases, measures are highly statistically significant.

Focusing more specifically on each proxy of fiscal federalism, Columns (4) through (8) demonstrates that local to central total revenue is robust to additional controls. The patterns are very similar to those found for growth rates above. While controls reduce the estimated effect by about

half, and explain nearly 80% of observed variation in GDP per capita, the general conclusion holds – fiscal decentralization decreased GDP per capita levels. Column (9) reports the preferred specification focusing on tax revenue instead of total revenue. The estimated coefficient on this revenue-based measure of fiscal federalism is negative and highly statistically significant, and while controls reduce the point estimate by about two-thirds (and explain about 80% of the variation), the negative relationship between revenue-based proxies of fiscal federalism and GDP per capita levels is confirmed.

Expenditure-based proxy results, however, tell the opposite story. As with the growth rate analysis, revenue and expenditure proxies differ in the estimated effects. The point estimate for local to central expenditure is always positive and highly statistically significant, robust to the inclusion of control variables, and accounts for nearly 80% of observed variation in log GDP per capita (results not reported). In the most preferred specifications (Column 10), the absolute value of the estimated effect based on expenditures is nearly identical to the absolute value of the estimated effect based on tax revenues – but with the opposite sign. Once again, endogenous proxy measures of fiscal federalism tell conflicting and unreliable stories about the economic impact of reform since the reform under consideration had differential effects on revenue and expenditures.

To reinforce these concerns, Table 6 reproduces the before and after approach utilized above for growth rates. Again, proxies for fiscal federalism are not robust to sample period selection. In some cases, estimated coefficients lose significance, while in other cases, estimated coefficients switch signs. Overall, the conflicting evidence strongly pushes against trying to use endogenous proxy variables to measure the impact of fiscal federalism. Instead, we proceed using our preferred quasi-experimental approach.

We move beyond approaches in the traditional literature by leveraging the unique policy experience of Chinese localities, exploiting a quasi-experimental environment and using a difference-in-difference estimation approach. As outlined above, localities differed in the timing of policy, and we consider the first localities to adjust as the treatment group compared to a control group of localities that adjusted later. We estimate the following:

$$Growth\ Rate_{st} = \alpha + \delta_{DID}(Treatment * Post)_{st} + \beta Controls_{st} + FE_s + FE_t + \epsilon_{st}$$

where subscript s refers to the province and t refers to the year. Based on observed locality heterogeneity across space and time, we include locality and year fixed effects as well as time-varying controls. Our DID estimate is captured by the coefficient δ_{DID} .

The difference-in-difference approach sidesteps concerns over endogenous proxies, but requires a large enough national experience to precisely estimate effects (if they exist). Figure 5 shows the magnitude of the TSS reform when it was implemented nationally in 1994. The share of local revenue to central revenue decreased by nearly 30 percentage points, emphasizing just how large a reform experience TSS truly was for fiscal federalism. Anecdotal evidence suggests that not only was the reform large, but that it did in fact matter for growth. The simple growth rate average for the treatment group was 16.7% from 1992-1999, compared to the control group, which grew at 15.7%. Such simple descriptive statistics are intriguing, but to tease out causal estimates requires a more careful identification strategy.

We start our DID analysis by considering a locality as treated if the locality received at least some treatment. This includes localities where some cities received treatment even though the entire locality did not receive treatment initially. Table 7 estimates the treatment effect to be 0.6 percentage points in Column 1, but not statistically significant. As discussed above, however, there is significant heterogeneity in localities, and over time, which should not be attributed to treatment directly. Columns (2) through (6) include time-varying controls, with no effect on statistical significance, although the point estimate more than doubles when all controls are included.

However, once year and province fixed effects are included, the point estimate of treatment on locality growth rates per capita is positive and statistically significant. The full DID model (Column (8)) estimates the impact of treatment to be 1.7 percentage points and is statistically

significant. For comparative purposes, average per capita growth rates during the FRS period were 11.9%, implying that TSS treatment increased per capita growth rates by around 14%. Localities that experienced the TSS treatment first grew faster than localities that received treatment later, and the estimated effect is economically significant.

While the results in Table 7 are notable, one concern is that treatment heterogeneity might be confounding the estimated effects. To address this concern, we restrict treatment to only those localities that fully received treatment and consider partially treated localities as part of the control group. The same basic patterns emerge for full treatment only (unreported results available upon request). The preferred specification, reported in Column (9) of Table 7, suggests that full treatment localities saw their growth rate increase by 2.1 percentage points compared to localities that adopted TSS later. Compared to FRS average per capita growth rates, full TSS treatment increased per capita growth rates by over 17%.

To separately consider the effect of partial treatment (since these localities experience some treatment early), we break treatment up in to three distinct groups: full treatment, partial treatment, and no treatment (control). Again, unreported patterns are nearly identical to the previous results. The most preferred specification (Column (10)) confirms the estimated full treatment effect of 2.2 percentage points (and statistically significant). The partial treatment effect is estimated at 1.0 percentage points, but is not statistically significant. Given the heterogeneity within localities, the result suggestively reinforces the view that TSS treatment positively affected per capita growth, even when only partially applied.

An additional concern remains regarding the definition of treatment timing. Treatment localities implemented TSS in 1992, while all other localities implemented TSS in 1994. Thus far, we have defined post-treatment as any year after 1992 even though all localities adopted TSS by 1994. This is a conservative approach in that it allows for first-mover or hysteresis effects, or if TSS adoption impacts growth with a lag, but if no such effects exists, our definition of treatment would underestimate the true effect since it would reduce across-group heterogeneity. An alternative approach would be to define treatment as occurring only in 1992 and 1993, since in

1994 all localities adopted TSS. However, this approach would bias the estimates if the impact of TSS was felt with a lag. For anecdotal evidence that TSS adoption did impact growth with a lag, Figure 6 plots year-by-year growth rates for treatment and control groups starting in 1992. While not definitive, the descriptive statistics suggest that TSS adoption did not impact growth for treatment groups until 1993, and then for control groups in 1995.

To show the robustness of the results, we estimate the treatment effect using two alternative definitions of treatment, first with the most strict definition (1992 and 1993), and then a looser definition allowing for a year lag in implementation (1993 and 1994). The estimate reported in column (11) shows that our more conservative approach may underestimate the true effect substantially. When we restrict treatment to just 1992 and 1993, the estimated impact is 3.99 percentage points and is statistically significant. In column (12), we allow for a one-year lag between TSS implementation and growth impacts, and estimate the effect of TSS to be 3.06 percentage points. Both estimates are notably larger than our preferred estimate, but as we do not want to take a stand on the lag structure of the impact of TSS, we prefer the estimate in column (10), which can be treated as lower bound for the impact of TSS on growth rates.

Summing up the results on per capita economic growth rates, we find that while traditional approaches to estimating the impact of fiscal federalism depend upon unreliable endogenous proxies, when the institutional structure of TSS reform is exploited with a difference-in-difference estimator, the causal effect of TSS on per capita growth rates is large and statistically significant, representing a 18% increase compared to per capita growth rates during the FRS period.

As an alternative measure of economic activity, Table 8 considers all localities that received at least some treatment as treated, and we find that fiscal federalism treatment increases GDP per capita levels. While the same general tendency was identified with growth rates, the levels analysis has the added benefit that in the most preferred specification (Column (9)), TSS treatment is significant at the 0.1% level. The DID estimate implies that TSS treatment lead to a 8% increase in per capita GDP, or 133 Yuan per person compared to median per capita GDP level of 1657 Yuan in 1991.

Columns (10) and (11) break treatment up into full and partial treatment, and confirm the growth findings. The full treatment effect is larger (11.5% and 11.7% respectively) and highly statistically significant. The partial treatment effect is positive (2.6%), but not statically significant, mirroring the growth results. Overall, the model explains nearly all of the observed variation in locality GDP per capita levels. The implied causal TSS full treatment effect equates to 194 Yuan per person (compared to 1991 per capita GDP levels).

4 Conclusion

This paper looked at a particular fiscal experiment in the evolution of the system of fiscal federalism in China. The importance of the question derives from the fact that fiscal federalism has been credited for the growth trajectory that took China above Japan to rank as the second largest economy in the world. Taking everything together, a coherent story concerning the particular fiscal experiment, namely the TSS reform, emerges.

While traditional approaches focused on endogenous proxies of fiscal centralization find conflicting and fragile results, our quasi-experimental approach utilizing a large-scale policy experiment in China with staggered timing to implementation provides consistent and robust estimates. Following the traditional literature's focus on growth rates, we find that treatment increased growth rates by economically meaningful amounts, and estimates were robust to alternative specifications and controls.

Having shown that the TSS reform was in fact successful in raising growth rates and levels of economic activity, this paper remains silent about the particular mechanisms responsible for this causal link. The increase in revenues for the central government brought about by TSS reform allows the central government to internalize externalities associated with public goods provision, coordinate economic activity across disparate regions, provide fiscal support for macroeconomic stability, and provide insurance in response to asymmetric shocks. Future work will explore these mechanisms in greater depth to better understand the optimal level of fiscal federalism in China.

References

Ahmad, E., L. Keping, T. Richardson, and R. Singh (2002), “Recentralization in China?” IMF Working Paper 02/168 (Washington: International Monetary Fund).

Bahl, R. and Jorge Martinez-Vazquez (2006), “Fiscal Federalism and Economic Reform in China,” in J.S. Wallack and T.N. Srinivasan (eds.) *Federalism and Economic Reform: International Perspectives*, Cambridge University Press, Cambridge.

Besley, T. and S. Coate (2003). “Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach,” *Journal of Public Economics* 87, 2611–2637.

Blanchard, O. and A. Shleifer (2000) *Federalism with and without political centralization: China versus Russia*. No. w7616. National Bureau of Economic Research.

Brennan, G. and J. M. Buchanan (1980). *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.

Brueckner, J.K. (2005) “Fiscal Federalism and Economic Growth,” *Journal of Public Economics*, 90, 2107–2120.

Dabla-Norris, E. (2005), “Issues in Intergovernmental Fiscal Relations in China,” IMF Working Paper 05/30 (Washington: International Monetary Fund).

Ding, Y. (2008), *Fiscal Decentralization and Economic Growth in China, 1994–2002*, *Journal of Chinese Economic and Business Studies*, 5:3, 243–260.

Jin, H., Qian, Y., and B.R. Weingast (2005). *Regional decentralization and fiscal incentives: Federalism, Chinese style*. *Journal of Public Economics*, 89, 1719–1742.

Jin, J. and Zou, H. (2005), “Fiscal decentralization, revenue and expenditure assignments, and growth in China,” *Journal of Asian Economics* 16 1047–1064.

Lin J.Y. and Z. Liu (2000) “Fiscal Decentralization and Economic Growth in China,” *Economic Development and Cultural Change* 49:1, 1–21.

Montinola, G., Qian, Y., and B.R. Weingast (1995). *Federalism, Chinese style: The political basis for economic success in China*. *World Politics*, 48, 50–81.

Oates, W. E. (1972). *Fiscal Federalism*. New York: Harcourt Brace Jovanovich.

Oates, W. E. (2005), “Toward a second-generation theory of fiscal federalism,” *International Tax and Public Finance*, 12, 349–373.

Oi, J. (1992), “Fiscal reform and the economic foundations of local state corporatism in China,” *World Politics*, 45, 99–126.

Shu-Ki, T. and C. Yuk-Shing (1994), “China’s tax reforms of 1994: breakthrough or compromise?” *Asian Survey*, 34(9), 769–788.

Shirk, S. (1993), *The political logic of economic reforms in China*. Berkeley: University of California Press.

Stegarescu, D. (2005) "Public Sector Decentralisation: Measurement Concepts and Recent International Trends," *Fiscal Studies*, 26, 301–333.

Weingast, B. R. (1995) "The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development," *Journal of Law and Economic Organization* 11, 1–31.

Weingast, B. R. (2009). Second generation fiscal federalism: the implications of fiscal incentives. *Journal of Urban Economics*, 65, 279–293.

Weingast, B. R. (2014), "Second generation fiscal federalism: political aspects of decentralization and economic development," *World Development*, 53, 14-25.

Wong, C. P. (2000). Central-local relations revisited: the 1994 tax-sharing reform and public expenditure management in China. *China Perspectives*, 52-63.

World Bank (2002), "China National Development and Sub-National Finance: A Review of Provincial Expenditures," World Bank Report No: 22951-CHA.

Zhang, L. (1999) Chinese Central-Provincial Fiscal Relationships, Budgetary Decline and the Impact of the 1994 Fiscal Reform: An Evaluation," *The China Quarterly*, No. 157 (Mar., 1999), pp. 115-141.

Zhang, T. and H. Zou (1998), "Fiscal decentralization, public spending and economic growth in China," *Journal of Public Economics*, 67, pp. 221–240