Discussion Paper
No. 2018-56 | July 17, 2018 | http://www.economics-ejournal.org/economics/discussionpapers/2018-56

Please cite the corresponding Journal Article at
http://www.economics-ejournal.org/economics/journalarticles/2018-65

Storage infrastructure and agricultural yield: evidence from a capital investment subsidy scheme

Somdeep Chatterjee

Abstract
In a developing economy, the availability of storage infrastructure is considered essential for two purposes; the reduction of post-harvest losses resulting in food shortage, and allowing for gains from inter-temporal trade due to potential arbitrage opportunities arising out of volatility in food grain prices. This paper provides empirical evidence on a lesser studied impact of storage infrastructure, viz, agricultural yield. The author exploits potentially exogenous variation generated by the intensity of access to a capital investment subsidy program for construction and renovation of rural godowns in India to identify causal effects of better storage on yield. He finds that the program led to an increase in rice yield by 0.3 tons per hectare, approximately a 20% increase compared to the baseline. A potential mediating channel for such an effect would be reduced storage costs facilitating better investments in productive inputs. As supportive evidence, the author finds that fertilizer consumption increased by 21% in response to the intervention.

JEL Q12 Q18 O12 O13
Keywords Storage; yield; fertilizer consumption; Grameen Bhandaran Yojana

Authors
Somdeep Chatterjee, Business Environment Group, Indian Institute of Management Lucknow, Lucknow, India, somdeep@iiml.ac.in

The author thanks Chon-Kit Ao and Indranil Biswas for useful discussion in conceptualizing the framework underlying this paper. Research assistance from Shreya Kapoor is duly acknowledged.

Citation Somdeep Chatterjee (2018). Storage infrastructure and agricultural yield: evidence from a capital investment subsidy scheme. Economics Discussion Papers, No 2018-56, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2018-56

Received July 3, 2018 **Accepted as Economics Discussion Paper** July 17, 2018 **Published** July 17, 2018

1 Introduction

The United Nations Development Programme's (UNDP) '2030 agenda' includes 17 new sustainable development goals, one of which is *responsible consumption and production* (Goal 12).[1] Among the stated targets of Goal 12 is the objective to *...reduce food losses along production and supply chains, including post-harvest losses* by 50%. Gustavsson et al. (2011) estimate that about 1.3 billion tons of food produce are wasted or lost each year globally. Much of this is due to lack of adequate storage facilities and other associated logistical concerns. The availability of storage infrastructure in terms of adequate warehouses or godowns to preserve crop produce is therefore of immense global importance, especially in light of the UNDP agenda.

Alternately, from the point of view of farmers, especially in developing countries, a major motivation for acquiring storage services would be to cope with price volatility and seasonal fluctuations. Many farmers in developing countries rely on credit and successful repayment of their debts are likely to be contingent upon the possibility of selling their produce at favorable prices. However, in the absence of storage facilities, lean season and peak season variation coupled with market price dyanmics may compel farmers to sell a lot of produce at low prices as determined by market conditions and conversely, not being able to sell enough when the prices are higher. Therefore, from a pure welfare perspective of the farmers, access to storage infrastructure seems important as well.

While the above motivations should be sufficient to encourage the development of storage infrastructure worldwide in general, and developing countries in particular, it may still be interesting to explore if access to better storage has other potential benefits. For instance, does improved storage infrastructure increase or decrease the motivation of farmers to produce? We know that better storage can help reduce loss of food grains but if farmers already accounted for the potential of losses due to lack of proper storage facilities, will they increase or decrease their production once they have access to storage? In this paper, I study a

[1] More about the 2030 agenda can be found here: `http://www.in.undp.org/content/india/en/home/post-2015/sdg-overview/`

capital investment subsidy scheme for construction and renovation of rural godowns in India to answer this question.

The answer to this question is important from two perspectives. First, from the perspective of a policy maker in a developing country, if better storage can lead to better productivity, then policies and incentives to promote storage facilities should be worth taking up as it would potentially increase the welfare of agents in the economy. Second, from the perspective of entrepreneurs in the warehousing business, farmers maybe encouraged to acquire their services if they can be convinced that access to storage leads to better yield. At the end of the day all the farmer cares about is his own productivity and if if his output per unit of land cultiavted can go up due to some intervention or innovation, it is likely that the farmer would adopt that technology or make use of that service.

The *Grameen Bhandaran Yojana* (GBY) was such a program introduced in India in 2001-02 to provide the famring community with *scientific storage* facilities and enable them to *avoid deterioration* and wastage of the food grain production.[2] I exploit the intensity of penetration of this scheme along with potentially exogenous variation provided by the access to this program in terms of time and institutional features of implementation to find the causal effects of storage facilities on agricultural yield for rice, which is a major food crop of the country.

The empirical framework relies on the natural experiment of introduction of the GBY. This facilitates the inference about the counterfactual scenario, ie, what would have happened in the absence of GBY. I find that enhanced access to GBY leads to an increase in rice yield. This helps answer the above research question because this finding implies that in the absence of GBY, rice yield would have been lower. Therefore, with better storage facilities the farmers appear to be motivated to improve their production and in the counterfactual scenario with no (or less) access to storage, rice yields would have been lower. This suggests that farmers do account for post-harvest losses when making their production decisions and input choices. I use several years of pre-GBY data to perform falsification exercises in

[2] The National Bank for Agricultural and Rural Development, India has laid out the details of this scheme on their website: `https://www.nabard.org/content.aspx?id=593`

support of the counterfactual assumptions and identification strategy. I also present a simple theoretical framework with concave production functions and storage costs to motivate and support the empirical analysis.

The rest of the paper is organized as follows. In section 2, I present some background of this research by relating the study to existing literature and providing information about the policy implementation. I also discuss the theoretical premise of this study by presenting a very simple conceptual model. In section 3, I discuss the estimation of the causal effects of the program. I first lay down the identification strategy and then discuss data sources and present the regression specification. Section 4 presents the main results and falsification tests. Section 5 describes a robustness check and section 6 concludes.

2 Background

In this section, I present some details about the GBY program and discuss the literature to which this paper adds. Broadly speaking, the GBY might facilitate gains from intertemporal trade by allowing for storage which has been considered a problem in the literature in this field. Much of the lack of such arbitrage gains are attributed to price volatility, excess consumption in current periods, giving away of food grains to friends and relatives, many of these being related to the idea of not having adequate access to storage infrastructure. I conclude this section by describing a very simple conceptual model which can help understand the theoretical channel through which better storage may affect production and yield.

2.1 The Grameen Bhandaran Yojana - A Capital Investment Subsidy Program

The GBY was launched as a capital investment subsidy program by the government of India in 2001-02. The idea was to subsidize the setup of new godowns and renovate existing ones which were becoming dysfunctional. The government realized that a major problem faced by farmers in the country was lack of retention infrastructure for their food grain produce and this may lead to hughe post-harvest losses and increased consumption by the producers themselves to prevent spoilage. Among the stated objectives of the program, capacity enhancement for storage with associated facilities in farming areas was prioritized to enable quality control of the output and better marketability. From this it is fairly obvious that the government was aware of the issues associated with price volatility and strived to protect the farmers' interests in this regard. Whether the policy eventually worked or not remains an empirical question examined in this paper, but the advent of the program seems to be grounded by some theoretical premise.

An associated objective of the government was to encourage private sector entities and cooperative societies to invest in agriculture by allowing them to setup warehouses for storing food grains via this program. The government invited participation in the program to construct rural godowns from self-help groups, non-government organizations, farmers or a group of farmers, any private proprietary or partnership firm and so on.[3] Indian districts are sub-divided into rural and urban areas based on the local governments of those areas. Municipalities and municipal corporations essentially look after issues in urban regions whereas rural areas have local self governments known as *panchayats* in a three-tier federal structure. The only geographic restriction in setting up these godowns was that it had to be outside of the municipal corporation areas, ie, must be restricted to rural regions.

The individual private entrepreneur who would construct a rural godown would decide on the storage capacity of the unit independently with the condition of having a minimum capacity of 100 tonnes and a maximum capacity of 10,000 tonnes of grain, to be eligible for subsidy under the scheme.[4] There were some specific engineering requirements which were laid down as well to be eligible for subsidies, such as adequate protection from rodents, birds etc, sufficient ventilation, robust structure conducive to weather conditions in the area and so on.

The major highlight of this program was the assistance from the government in terms of providing subisides for setting up such godowns. The subsidy was linked to institutional

[3]Details found here: http://agricoop.nic.in/sites/default/files/1_0.pdf

[4]based on geographic and topographic conditions, special discretionary relaxations were sometimes considered under the consideration of NABARD for setting up smaller godowns, particularly in hilly areas.

credit and would only be made available to projects funded by banks and other recognized formal financial institutions. Loans would be given out to the entrepreneurs and the banks would be responsible for collecting the repayments. The subisidy would be provided by the government through the banks in terms of favorable loan terms. The subsidy rates were fixed as described in Table 1. From the initial disbursement of the first installment of the loan, a term limit of roughly 15 months was considered for the completion of the project.

Table 1: Credit-Linked Subsidy Assistance under GBY

Category	Subsidy Rate (as percentage of capital cost)	Max Ceiling (in 100,000 INRs)
Women Farmers	33.33%	62.50
Self help groups and co-operatives (women run)	33.33%	62.50
Scheduled Castes and Scheduled Tribes (SC/ST) Entrepreneurs	33.33%	62.50
Co-operatives run by SCs/STs	33.33%	62.50
All Farmers (other than women)	25%	46.87
Agriculture graduates, cooperatives and State/ Central Warehousing Corporations	25%	46.87
All other individuals, firms etc	15%	28.12
Renovation of godowns of cooperatives	25%	N.A.

Notes: The category column indicates the category of entrepreneurs undertaking the godown construction project. There were special considerations for historically marginalized sections of the society, ie, SCs and STs and also for women.

2.2 Related Literature

The traditional literature on the effects of storage mainly focusses on the risk-response of households in terms of storage as a savings or investment device (Saha and Stroud 1994). More recently, researchers have started taking into account the impact of programs and policies that promote storage on various outcomes of interest. For example, Femenia (2015), in a general equilibrium setup, shows that providing subsidies for setting up private storage facilities (much like the GBY model) would eventually destabilize the market and Gouel (2013) presents a framework to comapre various storage policies in a standard setting with well defined risk preferences of consumers.

There is a large body of literature, particularly in the domain of agricultural economics, acknowledging the problem of post-harvest losses, summarized in a meta-analysis by Affognon et al. (2015). However, convincing empirical estimates of the magnitude of these losses are largely unavailable and one has to rely on reports coming out of the World Bank and other similar orgranizations for descriptive data on these numbers.[5] It is nonetheless well established that the reduction in post-harvest losses is a very important objective particularly for developing countries. This paper contributes to this strand of literature by providing empirical estimates of such an intervention targeting reducing post-harvest losses on actual yield of food grains.

The other strand of literature that this paper relates to is in the field of access to new agricultural technology. Specifically, papers like Conley and Udry (2010) focus on adopting a new technology by many farmers in a community and study the role of learning, Duflo, Kremer and Robinson (2011) study investment decisions of farmers in Kenya with respect to fertilizer purchases concluding that subsidies may help procastinating present-baised farmers to make a welfare improving investment decision, Emerick (2018) in the context of India finds that there exist trading frictions among farmers in terms of adoption of new technology and interventions like door-to-door sales may alleviate some of these concerns.

This paper is most closely related to Basu and Wong (2015) and Aggarwal, Francis and Robinson (2018) which study the effects of experimentally providing farmers with some storage facilities. Aggarwal, Francis and Robinson (2018), in the context of interventions in Kenya, conclude that access to storage may act as an investment instrument rather than a savings instrument and based on their estimations suggest that interventions that help farmers store grain could have better welfare implications compared to encouraging savings in bank. A potential channel through which these effects operate is embedded in an idea from behavioral economics, known as, *mental accounting* (Thaler 1999). Mental accounting, in this context, involves earmarking the saved grains for future use rather than falling prey

[5] For instance, the World Bank published a “Missing Food” report on post-harvest losses in the context of sub-Saharan Africa to be found here: `http://siteresources.worldbank.org/INTARD/Resources/MissingFoods10_web_final1.pdf`

to temptations of consumption at current periods. Basu and Wong (2015) find that their storage intervention (providing free weather-sealed drums and sacks for food grain storage in West-Timor) had no overall effect on consumption smoothing or health but find that it led to increases in reported income and non-food consumption. This is indicative of the fact that if a large-scale well intentioned storage intervention is directed towards farmers, at the least one would expect the budget set to expand opening up many more consumption-possibilities.

2.3 Conceptual Framework

While some of the results above look at the effects of storage from the perspective of better consumption, in this paper, I approach the issue a little differently. The basic premise of this paper is to estimate if better storage can lead to better production. Going by the findings of Basu and Wong (2015), it is understandable that the mediating channel would be a favorable budget constraint. If this is so, it is natural to also expect that farmers may be encouraged to invest more in inputs (as suggested by Aggarwal, Francis and Robinson 2018). If the investment is made in productive inputs, one would expect that output and hence yield per unit of land, would rise.

In this section, I present a simple model of a profit maximizing producer (who is also the seller) choosing inputs in the presence of storage costs, to describe this potential channel. Consider a standard concave production function $y = f(X)$ where y denotes the output and x denotes any representative input (or vector of inputs) in the production process. Concavity of production function implies $f\prime(x) > 0$ and $f\prime\prime(x) < 0$. Assume that the market price is denoted p and the producer sells a fraction α of the produce in the market and stores $(1-\alpha)$ times the produce. Without loss of generality, assume that input x has a per-unit cost of c. Additionally, the producer incurs a per-unit cost ψ towards storage of the product. The producer chooses input x to maximize profit π. We can write out the objective function of the producer as follows:

$$\underset{x}{\text{maximize}} \quad \pi = p \cdot \alpha \cdot f(x) - c \cdot x - \psi \cdot (1-\alpha) \cdot f(x) \tag{1}$$

The optimal choice of x must satisfy the first order condition of profit maximization given by:

$$\frac{\partial \pi}{\partial x} = p \cdot \alpha \cdot f\prime(x) - c - \psi \cdot (1 - \alpha) \cdot f\prime(x) = 0 \tag{2}$$

Solving for x we can find the following condition:

$$f\prime(x) = \frac{c}{p \cdot \alpha - \psi \cdot (1 - \alpha)} \tag{3}$$

Clearly, there is a positive relationship between the marginal productivity term $f\prime(x)$ and the storage cost ψ. If storage costs decline, ψ falls and as a result $f\prime(x)$ falls. Given the concavity of the production function, a lower $f\prime(x)$ implies a higher value of x. So with decrease in storage costs, optimal input choice increases and given that the input has positive productivity, the output, and hence yield, should be higher. The idea to be explored here is that an intervention like GBY leads to better storage infrastructure and hence may play an important role in reducing the ψ, which eventually implies, following this model, that output should go up through the mediating channel of productive investments in input. Such investments are now facilitated by lower storage costs leading to favorable budget constraints.

3 Estimation

In this section, I first discuss the identification strategy of this paper to estimate causal effects of the storage intervention on agricultural yield and other outcomes of interest. Then I go on to discuss the data sources and present the regression specifications in the empirical framework section.

3.1 Identification Strategy

Empirically identifying the effects of improved storage infrastructure on agricultural production and yield is difficult. Firstly, there maybe several factors correlated with better storage infrastructure which may also affect output and hence linear regressions using a simple OLS framework is likely to suffer from omitted variables bias. One such omitted variable could be the presence of better non-storage infrastructure in general which may lead to overall impacts on the agricultural production in the region. The other major concern is that better storage infrastructure maybe available in those areas which already have better yield. In other words, entrepreneurs may selectively choose to locate in areas that are flourishing and hence a simple OLS model will suffer from selection bias and reverse causality issues.

The ideal way to estimate such an effect would be to randomly assign better storage facilities to some regions. Then a means comparison of the average output in these regions could provide reliable causal estimates of storage on productivity. Even though the GBY was a targeted program towards improving and upgrading storage facilities in rural India, the policy was not randomly allocated to regions, as is usual with a national level program. So essentially the program in principle affected all regions of the country, ruling out the possibility of cross-sectional variation in terms of rolling out of the policy. However, the institutional features of implementation of this program provides an interesting quasi-experimental setup that can allow us to identify the causal effects of GBY on outcomes of interest using the intensity of the program reach.

Since assistance was linked to institutional sources of credit, the availability of banking infrastructure in a given region is likely to be heavily correlated with the intensity of the program. To avoid issues arising out of the unlikely scenario of banking infrastructure improving in a region in response to this policy, I restrict attention to the cumulative infrastructure available at the start of the prorgram. So I use the number of existing bank branches at the district levels as of 2001-02 as a potential source of exogenous variation in access to GBY. However, just comparing regions with more banks to those with less banks would not give true causal estimates of GBY on outcomes like agricultural yield because bank infrastructure potentially leads to several other financial and economic benefits that may spur on agricultural growth. Doing a difference-in-difference estimation seems attractive exploiting the time dimension of the policy introduction and comparing more and less banked regions before and after 2001-02. However, even such a setup is not sufficient to make a causal claim of GBY because it is quite possible that with other competing programs like the Kisan Credit Card scheme, for instance, yield would naturally be higher in more banked regions post-2002 and this may not entirely be attributable to the GBY (see Chatterjee 2018).

To address these issues, while continuing with the cross-sectional bank infrastructure variation and intertemporal variation, I incorproate a third dimension to the identification strategy, ie, an intensity measure of the GBY program using actual reach of the program. This idea is similar to a lot of applied econometrics studies done in the field of development economics (like Duflo 2001). I look at the intensity of the program in the first 3 years, ie, 2002-2005 at the state levels to get an idea as to which states were ahead in terms of GBY implementation. State level variation can occur because of several reasons ranging from the alignment of the state government with the central government, available state level infrastructure, market conditions in the state, taxation policies etc. I use a fuzzy triple difference estimation by comparing districts with more banks as of 2001-02 to others in states with more GBY access as of 2005 to others, before and after the introduction of the program. I define GBY access in terms of total number of new godowns constructed in the first 3 years of the program at state levels.

Therefore, based on the intensity measure of program reach, define *GBY districts* as the ones with above-median number of rural bank branches in states with above-median number of new godowns constructed as of 2005. All other districts would be like the control group of this quasi-experiment. The identifying assumption of this fuzzy triple difference estimation paradigm is that the difference in average yield in *GBY districts* before and after 2001-02 would be no different from the difference in average yields for the non-GBY districts between the same time periods, in the absence of the program, after controlling for other differences between the cross-sections contained within the definition of GBY and non-GBY districts. Since this is an assumption about the counterfactual, there is no clear way to test for the validity of this. However, since I have multiple years of data, I perform a falsification exercise and provide support for this claim at a later section. So, any changes in average yield over time between GBY and non-GBY districts would be the causal effect of GBY on yield, based on this assumption and empirical framework.[6]

3.2 Data

Data used in this paper comes from three sources. First, the Village Dynamics of South Asia (VDSA) database by ICRISAT which provides a district level panel dataset for various agriculture related variables.[7] I use data from 15 major states (comprising above 80% of the population of the country) for the years 1990 to 2010 to account for a period of roughly 20 years, such that we have approximately the same number of years before and after the program. I focus on production of rice as a major food grain. The data also includes variables representing total production, total area under production, gross and net cropped and irrigated areas,, number of markets in district, rainfall etc. To check for robustness of estimates, I also find the effects of GBY on a second food grain of great importance in the country, viz, wheat. Finally, I look at the potential effects of the storage program on usage of other associated inputs like fertilizers. The second data source is the website of India's central bank, Reserve Bank of India (RBI) which regularly publishes data on branches of banks. The data on number of rural bank branches for each year at the district level have been manually compiled from the statistics published on the RBI website. The third and final data source comes from a complied dataset published on the portal, indiastat.com on the total new construction of godowns under the GBY scheme at state level. The original data source as per the listed publication is the answer given during the question hour in the upper house of the Indian parliament on August 12, 2005. The main outcome of interest is rice yield which has been calculated as the total production of rice divided by the total rice cropped area under cultivation. Similarly, wheat yield has been calculated and used as another outcome alongwith consumption of fertilizers as the third outcome of interest. The mean rice yield for the apportioned data used in the analysis is around 1.77 tonnes per hectare. Mean wheat yield, which is also considered as an additional outcome, is roughly

[6] It is worth noting that non-GBY districts do not mean that these districts have zero access to GBY, it is just because of the way we have defined a *GBY district* for expositional simplicity.

[7] The appendix presents some of the summary statistics of variables used as outcomes or controls in the regression that follow, from the VDSA dataset.

around 2 tonnes per hectare. Mean consumption of Nitrogen, Phosphate and Potash (NPK) fertilizers is around 52 thousand tonnes per year per district.

3.3 Empirical Framework

I implement the above strategy exploiting potentially exogenous variation in reach of the GBY as described above and run the following specification running regressions for district d in state s and time t:

$$Y_{dst} = \alpha + \beta_1 \cdot GBY_{dst} + \beta_2 \cdot (moreGD_s * moreBNK_d) + \beta_3 \cdot (after_t * moreBNK_d)$$
$$+ \beta_4 \cdot (after_t * moreGD_s) + \beta_5 \cdot after_t + \beta_6 \cdot moreGD_s + \beta_7 \cdot moreBNK_d + \gamma \cdot X_{dst} + \epsilon_{dst} \quad (4)$$

This is the baseline specification where α is a constant, $moreGD_s$ is a dummy variable which takes the value 1 if the are above-median number of new godowns constructed under GBY in state s and zero otherwise, $moreBNK_d$ is the dummy variable which takes the value 1 if district d has above-median number of rural bank branches and zero otherwise, $after_t$ is a dummy for time which takes the value 1 is year t is post-2002 and zero otherwise. The interaction of the intertemporal and cross-sectional dummy variables captured by GBY generates potentially exogenous variation in intensity of the program reach under the assumption that the difference in the difference-in-differences of means of outcome Y along the 3 dimensions is statistically indistinguishable from zero in the absence of the intervention.

As a result, GBY takes the value 1 if all the three above dummy variables take the value 1, and it takes the value zero of either of the above dummy variables takes the value zero, for the given district at the given state and for the particular point in time. The coefficient β_1 would therefore give us the causal effect of storage access due to GBY on outcomes of interest Y. Some of the outcomes studied in this paper are rice yield, wheat yield and fertilizer consumption. The controls for gross cropped and irrigated areas, annual rainfall, number of markets, area under cultivation, crop-specific cultivation area that is irrigated etc are represented by X_{dst}. In some departures from this standard specification, I also use year and district-fixed effects.

4 Results

The regression results are reported in this section. I follow the above empirical framework to run variants of the regression specification described above to get a robust causal estimate of the program on outcomes of interest. I first report the findings from the main regressions using rice yield as the outcome of interest. As part of a falsification exercise, I then describe why the identifying assumption may be reasonable. Finally, I present results from regressions using other outcomes like an alternate food grain, wheat, and input usage like fertilizers.

4.1 Effects of GBY on Rice Yield

The main outcome variable considered in this paper is the yield of the major food grain, rice. Rice yield has been calculated as the total production in a give district divided by the total area cultivated specific to rice production and therefore the units of measurement of yield would be tonnes per hectare. On average the rice yield for the sample seems to be 1.77 tonnes per hectare. Table 2 presents the results from regression of Equation 4 with variations in the model specification.

In column 1, I report results from a baseline regression without using any of the controls and even though a positive impact of GBY on rice yield is estimated, it suffers from lack of precision due to the large enough standard errors. Introducing controls clearly alleviates some potential omitted variables bias as the coefficient becomes a lot larger in column 2. It appears that with enhanced intensity of GBY penetration and more availability of bank branches and higher number of godowns, rice yield can increase by 0.4 tonnes per hectare on average which compared to the mean seems like a very large effect on productivity. The finding is in line with the conceptual framework section above and it confirms the predictions of the theoretical model. Introducing year fixed effects in column 3 does not change the

Table 2: Effects of GBY on Rice Yield

	Dependent Variable: Rice Yield			
	(1)	(2)	(3)	(4)
GBY	0.127	0.423***	0.427***	0.343***
	(0.102)	(0.144)	(0.145)	(0.114)
District FE				Yes
Year FE			Yes	Yes
Controls		Yes	Yes	Yes
R^2	0.17	0.48	0.48	0.91
Observations	4914	3375	3375	3375
Mean of Dep Var	1.77 tonnes per hectare			

Notes: The sample contains the population of districts for 15 major states of India for a period between 1990-2010 coming from the VDSA database. All columns report results from different regressions. The coefficient *GBY* is the causal effect of the GBY program on outcomes, as described in the estimation strategy section. All regressions control for relevant baseline dummy variables and double interactions. Controls include gross cropped area, gross irrigated area, total number of markets, annual actual rainfall, area under cultivation and irrigation specific to rice production. The outcome for this table is rice yield (total production divided by total area under rice cultivation). Robust standard errors clustered at the district level are reported in parentheses. *** $p<0.01$ **$p<0.05$ *$p<0.1$

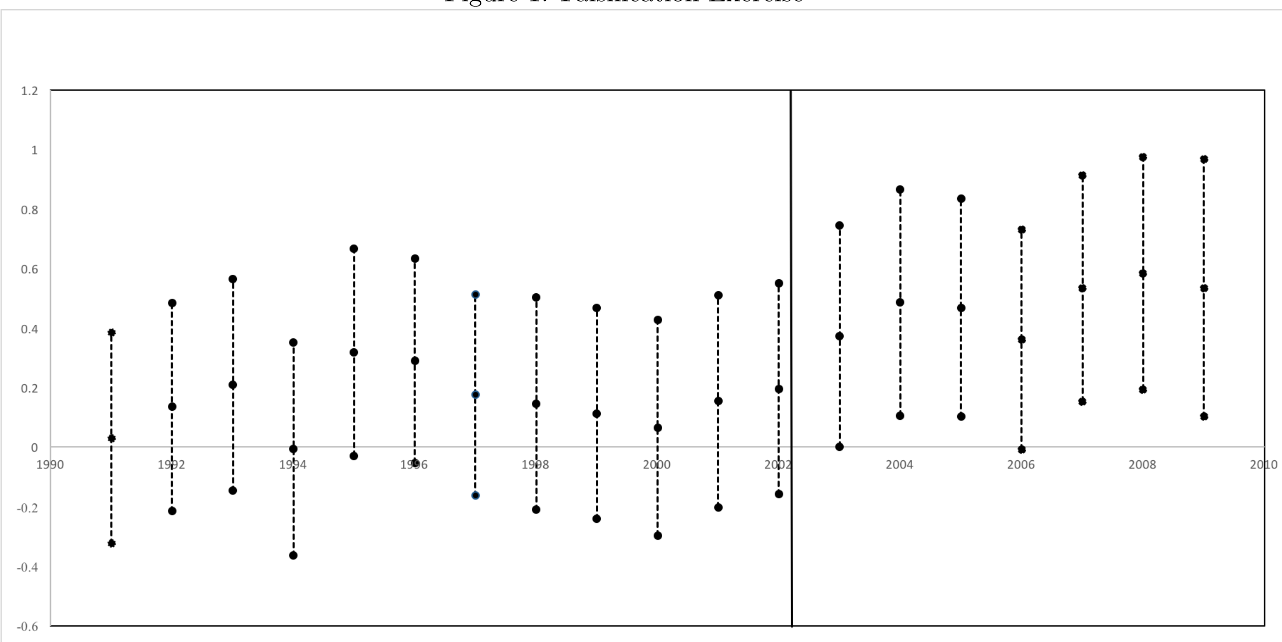
coefficient much. The results are also robust to introduction of district fixed effects, as in column 4, but the coefficient is slightly smaller. It appears that after controlling for time-invariant district specific characteristics, rice yield can increase by 0.343 tonnes per hectare on average with enhanced exposure to GBY program. Compared to the mean, this is roughly a 20% increase which is definitely a very large impact of storage.

The results suggest that with access to better storage infrastructure, rice yield is likely to go up. Potential channels have been explored in the conceptual framework section, where a program like GBY can be thought of as an intervention to reduce the storage costs ψ of the farmer. With a relaxed budget constraint, now the farmer can make better input investments and hence output per hectare might go up. Essentially, such a model allows us to get a near-estimate of the elasticity of storage costs on agricultural yield. However, the 20% figure estimated above should not be considered an actual elasticity estimate because it is implausible to quantify how much ψ would have declined with access to GBY infrastructure. Future research in this domain may be directed towards estimating the impacts of GBY on costs, which could further enable the prediction of yield-to-storage cost elasticity, which maybe a very relevant statistic for policy formulation as well as an important metric for marketing storage and warehouses by private entities.

4.2 Falsification Exercise

One concern about the above regressions would be that the effect picked up is not necessarily that of GBY but could be something else. In essence the claim we are making is that in the absence of GBY, the estimated coefficients would be statistically indistinguishable from zero. Since this is a claim about the counterfactual, it is not testable. However, it is possible to do a falsification exercise by exploiting the fact that we have multiple years of data at our disposal. Essentially, the triple difference estimate was giving us the difference in outcomes for GBY and non-GBY districts (as defined above with the caveats mentioned) before and after the policy was introduced. This implies, that the references period is the before-policy span of the data. In this section, I restrict the reference point to just 1 year instead of 10, ie, I consider 1990 as the omitted reference year and calculate pair wise triple differences for the GBY districts minus non-GBY districts between year t and year 1990. Figure (1) plots all such coefficients of rice yield estimated using the same framework as above, but only for pair of years instead of all years. The dotted lines signify the 95% confidence intervals for these estimated. Clearly, none of these estimates seem to be significant prior to the policy year, ruling out the possibility that there were pre-existing differences among these regions. The estimates become significant only post-2002 suggesting that something must have happened during that time period and therefore it is not implausible to think that GBY is the policy that has driven these estimates upwards. From the figure, it is also likely that in the absence of the program, the trend would have continued and hence the counterfactual predictions that have been made seem rather innocuous. Overall, this falsification exercise provides support to the identification assumption and reaffirms confidence in the empirical specification.

Figure 1: Falsification Exercise



Notes: The above figure plots the coefficients for rice yield using 1990 as the omitted year and reports coefficients for all triple interactions by year. The dotted lines represent 95% confidence intervals. The interpretation is as if the regression is run for pairs of years (t and 1990) instead of all years and therefore the dummy $after$ is now replaced by $t90$. The estimated coefficients are for $(moreGD_s * moreBNK_d) * t90$ where t90 takes the value 1 for year=t and takes the value 0 for year = 1990. All controls are used as in the main specification. Robust standard errors clustered by district are used to calculate the confidence intervals.

4.3 Effects on Other Food Grains and Inputs

In this section, I find impacts of the program on other food grains and input choices. The primary reason is to confirm if the results are mostly generalizable in the Indian agricultural setting or a phenomenon just specific to rice. The natural choice of outcome is therefore yield of the next most popular food crop of the country, wheat. I calculate wheat yield in the same way as rice yield by dividing the total production by the total area under production. As mentioned above, the average district produces about 2 tonnes of wheat per hectare over the span of the data. Apart from wheat yield, I also consider outcomes like fertilizer consumption. I use the total consumption of nitrogen, potash and phosphate (NPK) fertilizers with the idea that with the relaxation of budget constraints with a declining ψ (storage costs), it maybe possible for the farmers to now make a productive investment in terms of buying more fertilizers. This could be a potential channel through which the effects on yield operate.

Table 3: Effects of GBY on Wheat Yield and NPK-Fertilizers Use

	Wheat Yield		**Fertilizer Consumption**	
	(1)	(2)	(3)	(4)
GBY	0.266**	0.117	11.638*	11.075**
	(0.129)	(0.106)	(6.459)	(5.284)
District FE		Yes		Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
R^2	0.59	0.91	0.68	0.93
Observations	3132	3132	4079	4079
Mean of Dep Var	2.07 tonnes per hectare		52.37 thousand tonnes	

Notes: The sample contains the population of districts for 15 major states of India for a period between 1990-2010 coming from the VDSA database. All columns report results from different regressions. The coefficient *GBY* is the causal effect of the GBY program on outcomes, as described in the estimation strategy section. All regressions control for relevant baseline dummy variables and double interactions. Controls include gross cropped area, gross irrigated area, total number of markets, annual actual rainfall, area under cultivation and irrigation specific to wheat production for columns 1 and 2. The wheat specific area cropped is not used in columns 3 and 4. The outcomes for this table are wheat yield (total production divided by total area under wheat cultivation) and total consumption of NPK fertilizers in 1000 tonnes. Robust standard errors clustered at the district level are reported in parentheses. *** p<0.01 **p<0.05 *p<0.1

Table 3 reports the results. Like rice yield, wheat yield also seems to go up in response to the GBY program by 0.27 tonnes per hectare as evident from column 1. However, introduction of district fixed effects in column 2 brings this magnitude down to an imprecisely estimated level of 0.12 tonnes per hectare. Therefore, although there seems to be suggestive evidence of an impact on wheat (at least in terms of the direction of the effect), precision is lost when conditional on district level time invariant characteristics are controlled for. Columns 3 and 4 suggest strong positive impacts on fertilizer consumption. Enhanced exposure to GBY seems to increase fertilizer consumption by around 11 thousand tonnes. This estimate is statistically significant and seems to be robust to district fixed effects. Compared to the mean, the effect size is roughly around 21% which is considerably large.and very similar to the approximately 20% increase found in rice yield. This result helps suggest a mediating channel for the effects of storage on rice yield, ie, through access to better inputs like use of fertilizers perhaps facilitated by the reduction in storage costs, modelled as ψ in the conceptual section above.

5 Robustness Checks

Even though the falsification exercise presented above provides supportive evidence in favor of the identifying assumption, there may still remain concerns that the changes observed in yield and consumption of fertilizers are just inter-temporal even after accounting for the cross sectional variations and the empirical strategy just gives us some spurious inference.

The other concern could be about the interpretation of the findings. While we do observe increasing yields, could it be partially owed to the fact that farmers are now putting more land to use rather than through the mediating channel of cost as claimed in the above analysis? Even though I control for land put to cultivation, the regression does not rule out the possibility of new land being put to cultivation or changing land use patterns leading to substitution between different types of land used for agriculture vis-a-vis other activities.

Both the above concerns can be alleviated by looking at two specific outcomes, viz, amount of barren land and amount of land used for non-agricultural purposes. If the same empirical framework is used and amount of land put to non-agricultural use changes, that may mean that the storage program has affected land use patterns and a deeper investigation of the issue is required before making predictive comments about the effects of the intervention on agricultural yield. Similarly, if amount of barren land changes we may be concerned about spurious predictions from the model because this outcome is potentially an unrelated variable and should in principle be unaffected by the program.

Table 4: Robustness Checks

	Barren Land		**Non Agricultural Land**	
	(1)	(2)	(3)	(4)
GBY	-1.043	7.875	5.567	2.591
	(7.733)	(4.913)	(8.904)	(5.364)
District FE		Yes		Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
R^2	0.04	0.95	0.25	0.95
Observations	4028	4028	4028	4028
Mean of Dep Var	47.88 thousand hectares		62.95 thousand hectares	

Notes: The sample contains the population of districts for 15 major states of India for a period between 1990-2010 coming from the VDSA database. All columns report results from different regressions. The coefficient *GBY* is the causal effect of the GBY program on outcomes, as described in the estimation strategy section. All regressions control for relevant baseline dummy variables and double interactions. Controls include gross cropped area, gross irrigated area, total number of markets, annual actual rainfall. Robust standard errors clustered at the district level are reported in parentheses. *** p<0.01 **p<0.05 *p<0.1

Results from Table 4 confirm that concerns about spurious inference and interpretation of results may be ignored. It is clear that the program, using the identification framework used above, has no impact either on barren land or the amount of land put to non-agricultural use.

6 Conclusions

In this paper, I study the Grameen Bhandaran Yojana which was a capital investment subsidy scheme introduced by the Indian government in 2002 for construction and renovation of rural godowns. I find that the program led to increase in rice yield and fertlizer consumption.

I do not find any changes in land use patterns suggesting that the mediating channel for increased yield could be reduced costs of storage facilitating investment in productive inputs. While previous studies in the field have recognized the importance of reducing post harvest losses from the point of view of reducing food wastage and improving storage infrastructure to allow intertemporal arbitrage benefits due to price volatility of food grains, the idea of better storage infrastructure increasing productivity remains little explored. This paper fills that void by presenting, to the best of my knowledge, the first empirical estimates of the impact of the GBY intervention on agricultural yield. To identify the causal effects of the program, I used a fuzzy triple difference estimation design exploiting the inter-temporal variation in the reach of the program coupled with institutional features of implementation and actual program intensity. Future research may be directed towards estimating the impact of GBY on production costs which may enable calculating the elasticities of the GBY effect on yield through the cost reduction channel.

References

1. Affognon, Hippolyte, Christopher Mutungi, Pascal Sanginga and Christian Borgemeister (2015) “Unpacking Postharvest Losses in Sub-Saharan Africa: A Meta-Analysis”, World Development, Vol. 66, pp. 49-68
2. Aggarwal, Shilpa, Eilin Francis and Jonathan Robinson (2018) “Grain today, gain tomorrow: Evidence from a storage experiment with savings clubs in Kenya”, Journal of Development Economics, Vol. 134, pp. 1-15
3. Basu, Karna and Maisy Wong (2015) “Evaluating seasonal food storage and credit programs in east Indonesia”, Journal of Development Economics, Vol. 115, pp. 200-216
4. Chatterjee, Somdeep (2018) “The Curious Case of Farmer Credit Cards: Higher Output with No Increase in Borrowing”, mimeo
5. Conley, Timothy and Christopher Udry (2010) “Learning about a New Technology: Pineapple in Ghana”, American Economic Review, Vol. 100, No. 1, pp. 35-69

6. Duflo, Esther (2001) “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment”, American Economic Review, Vol. 91, No. 4, pp. 795-813
7. Duflo, Esther, Michael Kremer and Jonathan Robinson (2011) “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya”, American Economic Review, Vol. 101, No. 6, pp. 2350-2390
8. Emerick, Kyle (2018) “Trading frictions in Indian village economies”, Journal of Development Economics, Vol. 132, pp. 32-56
9. Femenia, Fabienne (2015) “The effects of direct storage subsidies under limited rationality: a general equilibrium analysis ”, Agricultural Economics, Vol. 46, pp. 715-728
10. Gouel, Christophe (2013) “RULES VERSUS DISCRETION IN FOOD STORAGE POLICIES ”, American Journal of Agricultural Economics, Vol. 95, No. 4, pp. 10291044
11. Gustavvson, Jenny, Christel Cederberg, Ulf Sonnesson, Robert van Otterdijk and Alexandre Meybeck (2011) “Global Food Losses and Food Waste”, FAO (UN) Study Report
12. Thaler, Richard (1999) “ Mental accounting matters”, Journal of Behavioral Decision Making, Vol. 12, pp. 183206
13. Saha, Atanu and Janice Stroud (1994) “A Household Model of On-Farm Storage under Price Risk”, American Journal of Agricultural Economics, Vol. 76, No. 3, pp. 522-534

A Appendix: Summary Statistics

Table A.1: Summary Statistics

	Observations	Mean	Standard Deviation
Rice Production (in ’000 tonnes)	11708	175.646	284.257
Wheat Production (in ’000 tonnes)	11708	170.956	320.411
Gross Cropped Area (in ’000 hectares)	9295	578.746	360.023
Rice Cropped Area (in ’000 hectares)	11708	106.762	149.949
Wheat Cropped Area (in ’000 hectares)	11708	78.213	97.219
Gross Irrigated Area (in ’000 hectares)	9295	155.784	215.979
Irrigated Area for Rice (in ’000 hectares)	10787	43.954	78.221
Irrigated Area for Wheat (in ’000 hectares)	10787	65.814	96.140
Total Number of Markets Available	10878	13.625	16.387
Annual Rainfall (in millimeters)	9344	1125.234	749.792
Consumption of Fertilizers (in ’000 tonnes) (Nitrogen, Potash and Phosphate)	11610	32.232	40.910

Notes: The results are from the VDSA database using the full available span of the data for all states. I use only a subsample of the data from here restricting the analysis to 15 major states (comprising above 80% of the population of the country) for the years 1990 to 2010.