LAHORE SCHOOL OF ECONOMICS

THESIS FOR THE DEGREE OF MPHIL

RESOURCE MISALLOCATION AND AGGREGATE PRODUCTIVITY IN PUNJAB

Submitted by:

Muhammad Haseeb

Thesis Supervisor:

Dr. Theresa Thompson Chaudhry

Abstract

We follow Hsieh and Klenow (2009) to study the role of misallocation in aggregate productivity for manufacturing plants in Punjab. Data for manufacturing plants in Punjab is taken from the Census of Manufacturing 2000-01 and 2005-06. In this paper, we essentially look at the extent to which marginal products differ across firms within each industry. We then imitate the liberalization settings by allowing marginal products to equalize across the plants in each industry. We find relatively more productivity dispersion in Punjab as compared to Hsieh and Klenow (2009) in India and China. Furthermore, we find that moving to the US efficiency level would boost manufacturing TFP in Punjab by 22.33% and 55.83% in years 2000-01 and 2005-06 respectively. We also explore the potential sources of productivity dispersion for manufacturing plants in Punjab.

Contents

Introduction

Total Factor Productivity (TFP) is that fraction of a country's economic growth which is unexplained by the conventional input factors: capital and labor. It indicates the efficiency of a production system in translating inputs into outputs.

Solow (1957), in his pioneering contribution to the productivity literature, lays down the basis for growth accounting. He uses the aggregate, economy-wide production function to separate out the role of factors of production; capital and labor, and residual (TFP) in the economic growth of an economy. In order to find the residual, he subtracts the weighted average growth of capital and labor from overall output growth. He carries out this exercise on U.S. data from 1909-1949 finding that TFP growth was the key factor responsible for the remarkable U.S. economic growth.

His work was followed by a number of papers in which TFP become the centre of discussion. The evolution of productivity measures along with availability of broad sets of data resulted in cross country comparisons of TFP. Many of these studies find that the key distinction between rich and poor countries lies in the productivity differences. Klenow and Rodriguez-Clare (1997) find that TFP growth explains 90 percent of the cross-country differences in output growth. Hall and Jones (1999) study the role of social infrastructure in explaining the large cross-country differences in productivity.

In recent years, the increasing availability of plant-level data has provided a valuable micro foundation for understanding aggregate productivity. This stream of research exploits a plant-level production function to compute an individual firm productivity measure. Plant-level productivity is then aggregated to arrive at an expression for economy wide productivity. Productivity dynamics at the micro-level have unveiled key sources of changes in aggregate productivity. These papers mainly look at the importance of entry and exit dynamics of firms, movement of individual plants in productivity cohorts, and allocation of resources across plants, in explaining the changes in aggregate productivity.

Productivity is often measured as a ratio of output to inputs. Literature classifies productivity measures into two broad categories: single factor productivity measures and multi factor productivity measures. Single factor productivity measures signify the efficiency of a single input factor, such as capital or labor, in producing output. Any change in single factor productivity measures can represent both embodied and disembodied technical change. Any change in productivity which is not captured by factor inputs comes under disembodied technical change. For example a change in labor productivity can be attributed to a change in capital or any other input factor (embodied technical change) or it can be attributed to a shift in technical efficiency (disembodied technical change). On the other hand, multi factor productivity measures usually take into account all input factors and thus represent only disembodied technical change.

There is a broad literature that provides evidence for the sources of disembodied technical change. Factors such as managerial practice, organizational technique, research and development, learning by doing, and productivity spillovers, are few important sources of disembodied technical change[1].

Recently a number of papers have taken a separate approach and study the role of policy distortions in aggregate productivity (e.g., Melitz, 2003; Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). These papers argue that any policy distortion, which can potentially create misallocation of resources across firms in an industry, can have significant consequences on aggregate productivity.

These policies impose taxes or subsidies on output or factor inputs. For example, polices that put restrictions on the size of a firm or that provide subsidized loans to firms for non-economic reasons can create plant-level distortions in the allocation of resources. The latter operates through the misallocation of capital across plants. Profit maximization implies that any firm benefitting from subsidized loans will equate its marginal product of capital to a lower interest rate compared to the firm without a subsidized loan. This plant-level misallocation has important implications on aggregate TFP.

In this paper, we follow Hsieh and Klenow (2009) to study the role of misallocation in aggregate TFP for manufacturing plants in Punjab. The objective of this paper is twofold. First, we study the productivity distribution of manufacturing plants in Punjab. Second, we estimate the gains in aggregate TFP as a result of removing misallocation across plants. We draw the data for manufacturing plants in Punjab from the Census of Manufacturing 2000-01 and 2005-06. We

[1] For further discussion see Syverson (2011).

find relatively more productivity dispersion in Punjab as compared to Hsieh and Klenow (2009) in India and China. Furthermore, we find that moving to the US efficiency level would boost manufacturing TFP in Punjab by 22.33% and 55.83% in years 2000-01 and 2005-06 respectively.

The rest of the paper is organized as follows. In the next section, we will review the literature on productivity dynamics and resource misallocation. The subsequent section will lay down the theoretical model developed in Hsieh and Klenow (2009). Next, we will provide the details for estimation strategy and data sources. Finally, in the last two sections, we present results and conclusions.

Literature Review

In this section, we will review the literature on productivity dynamics and resource misallocation. First, I will discuss the literature on productivity dynamics and size distribution of firms. This will be followed by the discussion on resource misallocation and its impact on Total Factor Productivity (TFP).

In recent years, availability of micro-level data has allowed researchers to study the dynamics of productivity in detail. This advancement has shifted the focus towards the important questions such as the evolution and survival of the firm, sources of productivity variation, the role of productivity in the size distribution of the firms, and the role of resource misallocation in total factor productivity.

Jovanovic (1982) provides the theoretical basis for firm selection in an industry where each firm follows a particular productivity shock. In his model each entrant receives a random draw from the productivity distribution of an industry. A valuable draw will help the firm to survive and grow whereas the firm with an unfavorable draw is more likely to decline and exit. An equilibrium will be achieved where the net value of entry becomes zero. Therefore, the selection of the firm, in equilibrium, rests in the firm specific productivity shock. In his model, small firms have a variable and higher growth rate and they are also more likely to leave the industry. Hopenhayn (1992) uses the same framework and develops the conditions for a steady state equilibrium. In his model firms enter and exit in the equilibrium. In the steady state, entry and exit rates are equal and the firm size distribution is stationary.

Researchers exploit these theoretical models to study the dynamics of productivity with micro-level datasets. An important paper by Olley and Pakes (1996) studies the evolution of establishment-level productivity in the telecommunication equipment industry of the U.S. The primary thesis of the paper is to measure the impact of technological change and deregulation on productivity. During the 1970s and 1980s the telecommunication industry in the U.S. went through major restructuring attributable to rapid technological development and liberalization of the regulatory environment. Authors find two sources of bias in estimating the production function parameters required for productivity estimations. The first one arises due to the simultaneity between productivity and input choices. Secondly, authors observed a higher rate of entry and exit during the restructuring period. This high iteration can cause selection bias in the estimation.

Olley and Pakes (1996) use structural techniques to establish a proxy variable for the unobserved productivity variable. They use the assumption that investment is a strictly increasing function of firm productivity. Therefore the inverse of the investment function can be used for the identification. They find that the aggregate productivity, measured as the output share weighted average of individual plant productivity, increased significantly following the restructuring of the telecommunication equipment industry. The authors further decompose the results and find that the primary source of productivity gain is the reallocation of capital towards more productive plants rather than the increase in average productivity.

Bartelsman and Dhrymes (1998) exploit micro-level data on U.S. manufacturing firms in high-tech industries to study the dynamics of productivity in detail. The authors use several measures of TFP and find that aggregate productivity followed a sustained decline from 1972 to 1984 and then experienced sharp increase after 1984. They attributed this TFP gain to reallocation of resources from less productive firms to more productive firms. The authors also use transition probabilities, based on estimated productivity, to study the movements of plants within productivity cohorts. They find a high level of uncertainty in the survival of entrant firms. Another interesting observation is that new firms enter at the upper levels of productivity cohorts. The authors also find that larger firms sustain their productivity ranking and they are also less likely to fail as compared to smaller firms. However, these findings are highly sensitive to the measure of productivity used.

An important finding of the micro-level productivity literature is the significant heterogeneity among the productivity levels of firms. Syverson (2004) employs several measures of productivity to compute the productivity distribution for four-digit U.S. manufacturing industries. He finds that average difference in total factor productivity between firms in 90th percentile (efficient firms) and 10th percentile (inefficient firms) is between 1.91 and 2.68. These enormous differences are robust to the different measures of productivity employed.

Syverson (2004) also studies the role of product substitutability in limiting the dispersion of productivity within an industry. In the perfect product substitutability setting, efficient firms are capable of capturing all the product demand in the market; thus, driving out the less efficient firms from the market. He tests this conjecture on U.S. manufacturing industries and finds a negative relationship between product substitutability and productivity dispersion.

Evidence on these enormous productivity differences has motivated researchers to study the sources of heterogeneity among plants. Recent literature documents the role of factors such as technology, research and development, competition, and market structure in explaining the productivity dispersion[2].

Another branch of the productivity literature deals with the role of resource misallocation caused by policy distortions inhibiting growth in Total Factor Productivity. These papers incorporate both specific and generic policy distortions at the plant-level problem to study its impact on TFP.

Hopenhayn and Rogerson (1993) document the long-run impact of policies related to severance pay on employment and average productivity. Policies that restrict firms from firing employees create distortions that promote less efficient use of resources. The paper extends the model developed in Hopenhayn (1992) and introduces a policy distortion (fixed payment for each job destroyed) using an adjustment cost function. The authors calibrate the model with the plant-level data on manufacturing firms in the U.S. to carry out the policy experiments. They find that moving from zero tax on dismissal (benchmark model) to a 20 percent tax on job destruction decreases employment by 2.5 percent. Apart from this, equivalent severance pay policy reduces average labor productivity by 2.1 percent. These numbers demonstrate that such policies have a significant impact on the aggregate economy.

Melitz (2003) studies the impact of exposure to trade on the measures of Total Factor Productivity. He first develops the closed economy model based on the Hopenhayn (1992) framework. In moving from the state of being a closed economy to an open economy, he introduces a trade friction in the form of a variable and fixed trading cost. These frictions separate exporting from non-exporting firms. After observing their productivity, firms decide whether to incur the trading cost and participate in trade. Trade offers relatively profitable prospects for firms and therefore encourages productive firms to enter the market. This process

[2] See Syverson (2011) for detailed discussion on factors affecting productivity growth

will continue until the zero profit condition is achieved once again. Therefore, exposure to trade leads to reallocation of resources among firms and drives out the less efficient firms from the market. The reallocation of resources towards more efficient firms will have a positive effect on aggregate total factor productivity. Another important finding of the paper is that increased exposure to trade leads to a welfare gain.

Parente and Prescott (1999) develop a framework to study the impact of monopoly right protection policies on total factor productivity. Government can protect a group of factor suppliers by a set of laws that prohibit other firms from employing effective changes in work practices. Regulations such as severance pay, restrictions on the entry and expansion of the firm, limits on the adoption of new technology are a few examples of such policies. They develop a game-theoretic framework in which monopoly rights restrict firms from entering into the industry. Entrants have to make a large amount of investment to overcome this restriction. In each stage of the game, an entrant decides whether to overcome the coalition or enter the coalition. The authors illustrate that for a sufficiently large size of coalition, it is not feasible for entrants to overcome the restriction. Further, they carry out a thought experiment to estimate the impact of moving from the monopoly setting to free enterprise arrangements on productivity, and eradicating monopoly rights would increase total factor productivity by a factor of 2.72, a significant figure.

Schmitz (2001) studies the impact of policies that restrict private firms from entering or expanding in an industry on aggregate labor productivity. He models an economy where government, rather than private firms, produces the investment goods. The intuition is that government production, being less efficient, will have a negative effect on aggregate productivity. The government sector receives a subsidy that is financed by a tax on the private sector. He calibrates this setting with data from the United States and Egypt. The purpose is to compare an economy with excessive government involvement in the production of investment goods (Egypt) with an economy with almost negligible government involvement in production (United States). He finds such polices explain 30 percent of the aggregate labor productivity gap between the U.S. and Egypt in the 1960s.

Bergoeing et al. (2002) study the role of policy distortions in Chile and Mexico following the severe economic crisis of the 1980s.They find that the relatively fast recovery of Chile is attributable to the differences in Total Factor Productivity rather than capital accumulation. Further examination reveals that divergence in TFP growth between the two countries was an outcome of differences in policy reforms. They find that the banking and bankruptcy laws in Mexico created distortions in the market that led to lower aggregate total factor productivity. The banking system in Mexico, which remained nationalized until 1990s, provided subsidized loans to certain sectors. These subsidies created distortions in the efficient allocation of capital among firms. Likewise bankruptcy law protected poorly performing (inefficient) firms from exiting the industry and at the same time prevented efficient firms from entering the market. Authors incorporate these policy distortions in their model to explain the relatively lower aggregate TFP in Mexico.

In another paper, Restuccia and Rogerson (2008) build up on the model developed in Hopenhayn (1992) to examine the impact of policies that create misallocation in resources on Total Factor Productivity. They argue that settings where government or private institutions provide favors to individual firms create distortions in the efficient allocation of resources. They develop a single good industry model in an entry and exit framework. In the firm’s optimization problem, policy distortions were introduced as a tax/subsidy on output. The consumer’s optimization problem determines the equilibrium rental rate of capital, which, along with the zero profit condition for the entry of plants, establishes the steady state equilibrium of the model. The authors calibrate this model with plant-level data for U.S. manufacturing firms. First, based on different assumptions and parameters, the benchmark case for no-distortion was estimated. Then using different distortion parameters, they mimic an economy where individual firms face heterogeneous prices. They find that, in settings where inefficient firms are provided with subsidies, an output subsidy of 40 percent can reduce TFP by considerable amount of 31 percent. This drop in TFP is sensitive to the number of inefficient firms in the market. If just 10 percent of the firms are subsidized and 90 percent are taxed, as opposed to 50 percent (benchmark case), then the drop in TFP is found out to be 49 percent.

In a recent paper, Hsieh and Klenow (2009) follow the Restuccia and Rogerson (2007) approach to compare the policy distortions in the United States with India and China. They develop a monopolistic competition model where individual firms face idiosyncratic policy distortions. They consider two separate distortion parameters: output distortion and capital distortion. An output distortion affects the marginal products of capital and labor by the same proportion. Capital distortions, on the other hand, increase/decrease the marginal product of capital relative to the marginal product of labor. The authors introduce these distortions in the firm optimization framework to capture the potential loss in aggregate productivity.

Hsieh and Klenow (2009) calibrate their model with plant-level data in India, China and United States. The first set of estimations include a comparison of the dispersion in total factor productivity in each country. They find that productivity distributions in India and China are more widely distributed, consistent with the conjecture of relatively more policy distortions in India and China. The thicker tail of the productivity distribution in India, relative to United States, also provided the evidence for survival of inefficient firms. Authors also estimate the potential sources of TFP variation by regressing it on a set of age, ownership, size and region dummies. They find that ownership is relatively an important factor in explaining TFP variation for China. In another set of estimations, effects of liberalization on TFP have been estimated. The authors mimic the liberalization settings by allowing marginal products to equalize across the plants in each industry. They find that liberalization can provide relative TFP gains of 86 percent in China, 127 percent in India, and 42.9 percent in United States for the latest year in each country's sample. They also find that the size distribution of firms for the full liberalization case is relatively much more dispersed for each country. For size measured as the value added, this illustrates that small and large size plants should produce more than what they are currently producing.

In the last set of estimations, the authors perform a thought experiment in which relative TFP gains in India and China were computed for the United States' efficiency level in 1997. Given that India and China move to United States efficiency level, they find 30-50% and 40-59% TFP gains for China and India respectively. Interestingly, they find no improvement in TFP levels took place for India over the time period of 1987 to 1994.

Following Foster et al. (2008), this paper also exploits an important distinction between revenue based and physical quantity based productivity measures. Foster et al. (2008) employ a rare set of plant-level data where producer-level prices were observed separately. They use information on 11 homogenous product manufacturers in the U.S. to study the role of producer-level prices on productivity measures. The intuition is that if prices reflect the idiosyncratic demand shifts then revenue based productivity measures will yield biased estimates. They find that physical output based productivity distribution is much more dispersed than revenue based productivity distribution. This reflects that physical quantity based productivity is negatively correlated with producer-level prices while revenue based productivity is positively related to prices. They observe that the primary reason for this discrepancy lies in the price setting behavior between young and incumbent producers. They find that, even though entrants are more productive than incumbent firms, young producers charge relatively lower prices as compared to incumbents. When productivity is measured with revenue, this price setting behavior eradicates the differences in productivity between young and mature firm.

In this paper, we will follow Hsieh and Klenow (2009) to study the role of policy distortions in Pakistan. Khawaja and Mian (2005) find that such distortions are widespread in the banking sector of Pakistan. They use loan-level data to estimate the extent of political rents in banking sector. They identify the political connections of firms by matching the data with national and state level election results. They observe highly preferential behavior of public banks in lending to politically connected firms. They find that, even though politically connected firms show 50 percent higher default rate, loans extended to politically connected firms are 45 percent larger than the lending volume to other firms.

Theoretical Framework

This section will provide a brief version of the monopolistic competition with heterogeneous firms model developed in Hsieh and Klenow (2009). They make use of an optimization framework to model the effect of policy distortions on firm level marginal products of capital and labor. Later, they derive an expression for industry level TFP as a function of resource misallocation.

In this framework, a representative firm produces a single final good Y in a perfectly competitive market. This firm makes use of S different intermediate goods in a Cobb-Douglas production technology. Intermediate goods are produced by Sdifferent manufacturing industries with each having an output of Y_S:

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s} \tag{1}$$

Industry output Y_S combines M_S differentiated products with constant elasticity of substitution:

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{2}$$

Each differentiated product is produced with the following firm level production function:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \tag{3}$$

Here, A_{si} represents the Total Factor Productivity, K_{si} and L_{si} represent capital and labor respectively. It is important to note that capital and labor shares are the same across all the firms in an industry.

Hsieh and Klenow (2009) exploit two separate distortion factors; output distortion and capital distortion. Any distortion that has a same impact on both marginal product of capital and labor comes under output distortions. For example, policies that impose an output tax/subsidy on an establishment affects the marginal products of capital and labor by the same amount. The other factor covers all those distortions that affect the marginal product of capital relative to the marginal product of labor. Examples are policies that provide subsidized loans for non-economic reasons, that decrease the marginal product of capital relative to marginal product of labor. The extent of the policy distortion will be reflected in the marginal product of labor and capital heterogeneity across establishments. Suppose τ_y and τ_k represent the output and capital distortions respectively. Then profits are given by the following function:

$$\pi_{si} = \left(1 - \tau_{Y_{si}}\right) P_{si} Y_{si} - w L_{si} - \left(1 + \tau_{K_{si}}\right) R K_{si} \tag{4}$$

Profit maximization implies:

$$\frac{K_{si}}{L_{si}} = \frac{\alpha_s}{1 - \alpha_s} \cdot \frac{w}{R} \cdot \frac{1}{1 + \tau_{K_{si}}} \tag{5}$$

$$L_{si} \propto \frac{A_{si}^{\sigma-1}\left(1 - \tau_{Y_{si}}\right)^{\sigma}}{\left(1 + \tau_{K_{si}}\right)^{\alpha_s(\sigma-1)}} \tag{6}$$

$$Y_{si} \propto \frac{A_{si}^{\sigma}\left(1 - \tau_{Y_{si}}\right)^{\sigma}}{\left(1 + \tau_{K_{si}}\right)^{\alpha_s \sigma}} \tag{7}$$

Marginal revenue products of capital and labor are given by:

$$MRPL_{si} = \frac{w}{(1 - \tau_{Y_{si}})} \tag{8}$$

$$MRPK_{si} = R\,\frac{(1 + \tau_{K_{si}})}{(1 - \tau_{Y_{si}})} \tag{9}$$

Then the weighted average marginal revenue products of capital and labor in a sector can be expressed as:

$$\overline{MRPL_s} \triangleq \frac{w}{\sum_{i=1}^{M_s}\left(1 - \tau_{Y_{si}}\right)\frac{P_{si}Y_{si}}{P_s Y_s}} \tag{10}$$

$$\overline{MRPK_s} \triangleq \frac{R}{\sum_{i=1}^{M_s} \frac{(1-\tau_{Y_{si}})}{(1+\tau_{K_{si}})} \frac{P_{si}Y_{si}}{P_s Y_s}} \tag{11}$$

Following Foster et al. (2008) revenue based productivity ($TFPR$) and physical output based productivity ($TFPQ$) are defined as:

$$TFPQ_{si} \triangleq A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \tag{12}$$

$$TFPR_{si} \triangleq P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \tag{13}$$

Furthermore, establishment-level $TFPR$ is proportional to the geometric mean of the plant's marginal products of capital and labor[3]:

$$TFPR_{si} \propto (MRPK_{si})^{\alpha_s}(MRPL_{si})^{1-\alpha_s} \tag{14}$$

Intuitively, in the no-distortion case, revenue based productivity should equalize across establishments. Large and efficient establishments, with higher $TFPQ$, will have a higher level of output and a relatively smaller price. Therefore, more resources will be allocated towards the efficient producer until $TFPR$ equalizes across firms. This distinction implies following expression for industry TFP[4]:

$$A_s = \left[\sum_{i=1}^{M_s} \left(A_{si} \ \frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{15}$$

Where A_{si} and $TFPR_{si}$ are defined in (12) and (14) respectively.

[3] A similar exercise with aggregate marginal revenue products of capital and labour will yield an expression for $TFPR_s$

[4] Please refer to Hsieh and Klenow (2009) for detailed derivations

In the no-distortion case, marginal products of capital and labor should equalize across establishments. In this scenario, each establishment will have the same $TFPR$ and expression (15) will become:

$$\overline{A}_s = \left[\sum_{i=1}^{M_s} (A_{si})^{\sigma - 1} \right]^{\frac{1}{\sigma - 1}} \tag{16}$$

These two expressions will be used to carry out the liberalization experiments. Expression (15) implies that greater the difference between sector average and individual plant-level $TFPR$, the lower will be the industry TFP.

Data Description

The plant-level data for Punjab is taken from the Census for Manufacturing Industries (CMI) 2000-01 and 2005-06.The CMI is conducted after every five years; it is intended to cover all the registered manufacturing firms in Pakistan with 10 or more employees. The province of Punjab is covered by Punjab Directorate of Industries. The CMI 2005-06 has information on 6,417 manufacturing establishment from all over Pakistan. The Punjab level CMI 2005-06 contains information on only 3,528 manufacturing plants, whereas CMI 2000-01 covers 4,809 establishments at the national level, out of which 2,357 are based in province Punjab.

The coverage of firms was improved in CMI 2005-06 by enhancing survey frame. CMI 2000-01 has only covered the firms listed in industrial directory; whereas in 2005-06 more firms were added to the frame by consulting the results of the the Economic Census 2001. There were nearly 50% more firms in the 2005-06 sample compared to the 2000-01 sample. This does not only reflect growth of firms; it could also be a sign of better coverage and less non-response rate. Looking at the registration date of the firms, we find that growth has been significant during this period. In CMI 2005-06 only 39% firms report their registration date; out of these firms 5.6% were born after 2001. In order to put things in perspective, we took the ratio of this number to total number of establishment in CMI 2000-01. We find that addition of at least 7.7% firms, in CMI 2005-06, was purely due to growth of new firms[5].

This paper will make use of following variables from each census: labor compensation, nominal output (revenue), expenditure on input materials, energy cost, book value of capital, date of registration (for computing the age of the firm), and form of ownership.

In Punjab, in both years, one of the major industries is cotton ginning. However, according to International Standard Industrial Classification (Rev. 3.1), cotton ginning is no longer considered to be a manufacturing activity; therefore, in our analysis, we exclude these establishments. Following this, 221 (9.3%) and 455 (12.9%) firms are dropped in CMI 2000-01 and 2005-06 respectively.

---

[5] This analysis only includes those firms that were born after 2000-01. We wanted to look at the growth of firms due to the addition of new firms between 2000-01 and 2005-06. We used registration date of a firm to impute its birth year. Since registration date was not available for all the firms, we could only perform this exercise for a fraction of 2005-06 firms.

We also get rid of all those firms that provide either missing or negative values on capital stock, labor compensation and value added. This is mainly because most of our analysis makes use of expressions with logarithmic transformations. We drop 163 (6.9%) and 286 (8.1%) firms in CMI 2000-01 and 2005-06 respectively.

Furthermore, we trim outliers from each industry to make our estimations robust. We pool both years and trim tails of capital and output distortions, $\log\left(TFPR_{si}/\overline{TFPR_S}\right)$, and $\log(A_{si}/\overline{A_s})$. Following this, we drop total of 346 firms (7.2% of cleaned dataset) in both years[6].

After cleaning we are left with 1,840 establishments from the CMI 2000-01 and 2,546 establishments from the CMI 2005-06. Table B in Appendix provides information on the distribution of firms in each year.

We are also borrowing an important piece of information from Camacho & Conover (2010). In order to calculate the distortion parameters, we require the undistorted capital and labor shares. Hsieh and Klenow (2009) assume that labor and capital shares are comparatively undistorted in the U.S. Therefore, they exploit U.S. labor and capital shares for India and China. In their paper, Camacho & Conover (2010) have provided the US labor shares for three digit industries.

[6] In order to keep our estimations consistent, we have followed Hsieh and Klenow (2009) in identifying outliers. In this exercise, we flagged all the firms that fall with-in 1-2% of top and bottom extremes based on four different variables discussed above. We then dropped each firm that was identified as an outlier in the first step. Therefore, drop of 7.2% is a combined trimming of four different measures.

Estimations

For the current study, estimations will mainly be based on the calculations of the following four productivity measures:

$$TFPR_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}} \tag{17}$$

$$A_{si} = \kappa_s \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}} \tag{18}$$

Where $\kappa_s = \frac{(P_sY_s)^{-\frac{1}{\sigma-1}}}{P_s}$

$$\overline{TFPR_s} = \frac{\sigma}{\sigma-1}\left(\frac{\overline{MRPK_s}}{\alpha_s}\right)^{\alpha_s}\left(\frac{\overline{MRPL_s}}{1-\alpha_s}\right)^{1-\alpha_s} \tag{19}$$

Where $\overline{MRPK_s} = \frac{R}{\sum_{i=1}^{M_s}\frac{(1-\tau_{Y_{si}})}{(1+\tau_{K_{si}})}\frac{P_{si}Y_{si}}{P_sY_s}}$ and $\overline{MRPL_s} = \frac{w}{\sum_{i=1}^{M_s}\left(1-\tau_{Y_{si}}\right)\frac{P_{si}Y_{si}}{P_sY_s}}$

$$A_s = \left[\sum_{i=1}^{M_s}\left(A_{si}\,\frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{20}$$

The first expression (17) measures the plant-level Total Factor Revenue Productivity. Expression (18) measures the plant-level $TFPQ$ with nominal output $P_{si}Y_{si}$. In the data sets, plant-level real output is unobserved. Therefore, observed nominal output is raised to the power $\frac{\sigma}{\sigma-1}$ to impute the real output Y_{si}. This exercise makes use of a scalar κ_s, which is unobserved and therefore assumed as $\kappa_s = 1$. This assumption will not affect our calculations for relative productivities. Expression (19) is a measure of industry-level Total Factor Revenue Productivity. This expression is derived by taking a geometric mean of industry-level $\overline{MRPK}$ and $\overline{MRPL}$. Finally the last productivity measure (20) is an industry-level $TFPQ$.

In order to compute the $\overline{MRPK}$ and $\overline{MRPL}$ we need information on plant level distortions. Hsieh and Klenow (2009) impute the distortion parameters in the following manner:

$$\tau_{K_{si}} = \frac{\alpha_s}{(1-\alpha_s)}\frac{wL_{si}}{RK_{si}} - 1 \tag{21}$$

$$\tau_{Y_{si}} = \frac{\sigma}{(\sigma - 1)} \frac{wL_{si}}{(1 - \alpha_s)P_{si}Y_{si}} - 1 \tag{22}$$

Expression (21) implies that a capital distortion is observed where the ratio of a plant's wage bill to its capital stock is different than the ratio of respective output elasticities. The next expression (22) implies that an output distortion is observed where the labor share is different than the elasticity of output with respect to labor. In both cases, we are comparing undistorted US labor and capital shares with the respective observed information on Punjab to infer the distortions.

This exercise requires following key parameters: labor and capital shares (α_s), elasticity of substitution between plants (σ), rental price of capital (R), and industry output shares (θ_s). We will follow the same conventions in order to maintain the comparability of our results to Hsieh and Klenow's analysis. Elasticity of substitution between plants is positively correlated with liberalization gains; therefore, to avoid the exaggeration of results, it is taken as the modest estimate of $\sigma = 3$. Undistorted rental price of capital is taken as $R = 0.10$. However, effective cost of capital will differ for each firm based on idiosyncratic capital distortion. Furthermore, since we are using relative productivity measures, choice of this parameter will not affect our liberalization experiments. Finally, industry output shares are taken as ratio of aggregate industry value added to aggregate economy-wide value added $\theta_s = \frac{P_s Y_s}{Y}$

Productivity distribution

In Figure 1, we use $\log(A_{si} M_s^{\frac{1}{\sigma-1}} / \overline{A_s})$ to plot the TFPQ distribution for each year. These distributions show roughly same amount of dispersion across time. Furthermore, we also observe a stretched left tail in year 2005-06, representing the survival of less productive firms. In Table 1, we present several dispersion measures for $\log(TFPQ)$. The first measure is a within industry standard deviation weighted by the value-added share of each industry. It shows that there was more dispersion in 2005-06 as compared to 2000-01. Next, we find the within industry difference between 75th and 25th percentiles weighted by the value-added share of each industry. Numbers presented in Table 1 are calculated on a logarithmic scale. We can use exponential functions to convert these numbers into more meaningful values. For example, in year 2000-01, firms in the 75th percentile were 9.5 times more productive then firms in 25th percentile[7]. This difference is even higher in year 2005-06. Across the measures, we have observed slightly greater dispersion in year 2005-06.

We also present Hsieh & Klenow's (2009) calculations for India, China and the United States. These calculations are given for different points in time for each country**.** We observe a relatively high level of dispersion in Punjab as compared to China and the United States.

In Figure 2, we use $\log\left(TFPR_{si}/\overline{TFPR_S}\right)$ to plot the TFPR distribution for each year. First, in comparing physical output based productivity (TFPQ) and revenue based productivity (TFPR), we observe similar results to Foster et al. (2008); the TFPQ distribution is relatively more dispersed then the TFPR distribution. This validates the negative correlation between TFPQ and producer-level prices.

Comparing TFPR distributions across time indicates greater dispersion in the year 2005-06 as compared to 2000-01. Table 2 presents TFPR dispersion statistics for each year. All three measures show relatively more within industry productivity spread in the year 2005-06. In 2000-01, firms in the 75th percentile were almost two and half times more productive than firms in the 25th percentile. This difference increases to slightly above three in 2005-06. Similarly, in 2000-01, firms in 90th percentile were almost five times more productive then firms in 25th percentile. This difference increases to a massive factor of nearly ten in the year 2005-06.

Comparing these statistics with Hsieh & Klenow's (2009) calculations indicate similar revenue based productivity dispersion in Punjab and China in 2000-01. However, in 2005-06 Punjab show relatively more productivity dispersion then China. In drawing comparison of Punjab with India and United States, it is important to note that Hsieh and Klenow's (2009) estimations for these countries are from different time periods. We find that, in all three years, India is almost as dispersed as Punjab in 2000-01. Finally, U.S. is far less dispersed than any other Country at any point in time.

---

[7] This value correspond to exponential of 1.88

Figure 1: TFPQ distribution

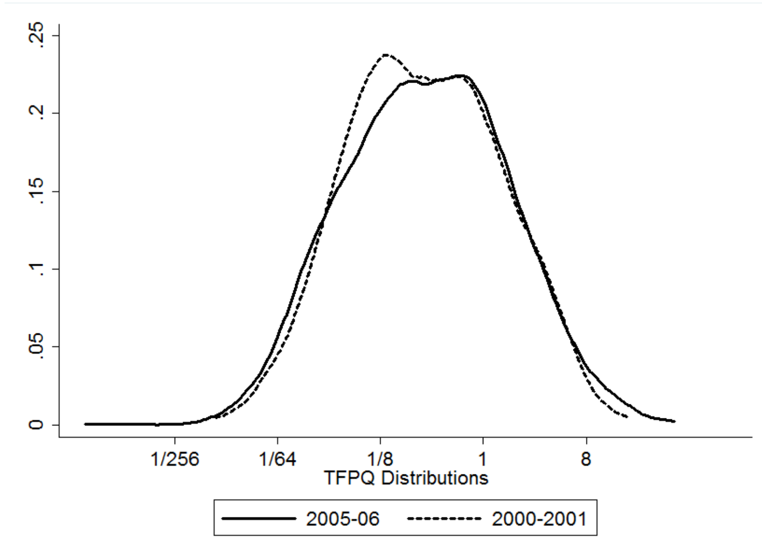


Figure 2: TFPR Dispersion

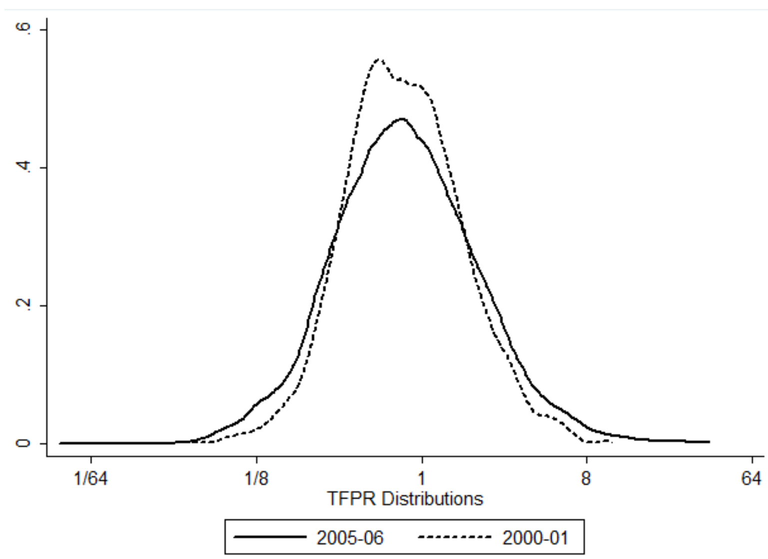


Table 1: TFPQ dispersion[8]

Punjab	2001	2005	
S.D	1.52	1.7	
75 – 25	2.25	2.56	
90 – 10	4.08	4.48	
N	*1,840*	*2,546*	
China	**1998**	**2001**	**2005**
S.D	1.06	0.99	0.95
75 – 25	1.41	1.34	1.28
90 – 10	2.72	2.54	2.44
N	*95,980*	*108,702*	*211,304*
India	**1987**	**1991**	**1994**
S.D	1.16	1.17	1.23
75 – 25	1.55	1.53	1.60
90 – 10	2.97	3.01	3.11
N	*31,602*	*37,520*	*41,006*
United States	**1977**	**1987**	**1997**
S.D	0.85	0.79	0.84
75 – 25	1.22	1.09	1.17
90 – 10	2.22	2.05	2.18
N	*164,971*	*173,651*	*194,669*

Table 2: TFPR Dispersion[8]

Punjab	2001	2005	
S.D	0.68	0.88	
75 – 25	0.90	1.17	
90 – 10	1.68	2.27	
N	*1,840*	*2,546*	
China	**1998**	**2001**	**2005**
S.D	0.74	0.68	0.63
75 – 25	0.97	0.88	0.82
90 – 10	1.87	1.71	1.59
N	*95,980*	*108,702*	*211,304*
India	**1987**	**1991**	**1994**
S.D	0.68	0.67	0.67
75 – 25	0.79	0.81	0.81
90 – 10	1.73	1.64	1.60
N	*31,602*	*37,520*	*41,006*
United States	**1977**	**1987**	**1997**
S.D	0.45	0.41	0.49
75 – 25	0.46	0.41	0.53
90 – 10	1.04	1.01	1.19
N	*164,971*	*173,651*	*194,669*

[8] *Note: India, China, and United States calculations are taken from Hsieh and Klenow (2009). Punjab figures are author's calculations.*

This analysis points up an important shift in productivity dispersion during these five years. We can have two possible explanations for this shift. First, as we have discussed above, the coverage of firms was much better in CMI 2005-06. Therefore, these differences could simply be due to more representative frame in year 2005-06. Apart from this, policy distortions could be an important source of explanation of this shift. We will study more about policy distortions in our liberalization experiment.

Productivity Variation Explained

In this section, we will use regression analysis to study sources of within industry productivity variation. We will analyze four possible explanations of variation: region, size, ownership type, and age. In each regression, we run the following specification:

$$\log TFPR_{si} - \log \overline{TFPR_s} = \beta_0 + \beta_1 X_{si} + \beta_2 YEAR + \varepsilon_{si}$$

The dependent variable is the deviation of plant-level TFPR from its industry's average in each year. X_{si} is a vector of dummies, representing region, size, ownership type, or age of the firm in the respective regressions below. For each regression, we are pooling data for both years. We are also weighting the regression by industry value-added share to control for the size effect of the industry. While interpreting the coefficients on these regressions, we are well aware of potential endogeneity bias on some of our independent variables. However, we are primarily interested in the share of total TFPR variation explained by each category.

In Table 3, we present set of regressions for the pooled dataset. In each regression, we add another set of dummies to study the cumulative explanation of TFPR dispersion. In the first regression, we have dummy variables representing ownership type on the right hand side. We are keeping "Domestic Private Firms" as an omitted category. We observe that ownership type is not an important determinant of with-in industry productivity variation in Punjab; it only explains 0.1% of the variation in TFPR.

In second regression, we further add firm-size quartiles on the right hand side. Size is measured as firm value-added. We are taking "Bottom Size Quartile" as our omitted category. The estimated regression now explains massive 19.3% of the variation in TFPR.

Table 3: TFPR variation explained by ownership, size, and region

Variables	(1)	(2)	(3)
Public	-0.0538	-0.264***	-0.254***
	(-0.86)	(-4.64)	(-4.43)
Foreign	0.594**	0.279	0.294
	(2.09)	(1.09)	(1.15)
Collaboration	0.0941	-0.0672	-0.0977
	(0.95)	(-0.75)	(-1.09)
First Size Quartile		-1.017***	-1.029***
		(-30.21)	(-30.60)
Second Size Quartile		-0.506***	-0.510***
		(-17.32)	(-17.52)
Third Size Quartile		-0.141***	-0.140***
		(-5.00)	(-4.94)
North Punjab			0.0485
			(0.91)
South Punjab			0.270***
			(6.95)
Wes Punjab			0.0191
			(0.33)
Constant	-0.221***	0.0942***	0.0707***
	(-17.77)	(5.04)	(3.65)
N	4386	4386	4386
Adjusted R-sq	0.001	0.193	0.201

*Note: t-statistics in parentheses * p<0.10. ** p<0.05. *** p<0.01*

Table 4: TFPR variation explained by ownership, age, size, and region

Variables	(1)	(2)	(3)	(4)
Public	-0.109	-0.106	-0.246***	-0.251***
	(-1.48)	(-1.23)	(-3.02)	(-2.98)
Foreign	0.858*	0.808	0.615	0.617
	(2.24)	(1.57)	(1.29)	(1.29)
Collaboration	-0.0135	-0.0118	-0.145	-0.155
	(-0.26)	(-0.06)	(-0.83)	(-0.88)
First Age Quartile		-0.0784	0.0588	0.0623
		(-1.01)	(0.80)	(0.85)
Second Age Quartile		0.0316	0.0561	0.0597
		(0.42)	(0.81)	(0.85)
Third Age Quartile		-0.0616	-0.0933	-0.0924
		(-0.80)	(-1.31)	(-1.29)
First Size Quartile			-1.138***	-1.130***
			(-11.03)	(-10.93)
Second Size Quartile			-0.584***	-0.584***
			(-7.69)	(-7.66)
Third Size Quartile			-0.0699	-0.0671
			(-1.18)	(-1.13)
North Punjab				0.0198
				(0.17)
South Punjab				-0.181
				(-1.52)
West Punjab				0.0214
				(0.20)
N	993	993	993	993
Adjusted R-sq	0.002	0.002	0.141	0.140

*Note: t-statistics in parentheses * p<0.10. ** p<0.05. *** p<0.01*

Lastly, we add dummies representing the four different regions in Punjab[9]. We are taking "Central Punjab" as omitted category. We find that region explains very little of TFPR's variation.

We do not observe registration date for firms in CMI 2000-01. Therefore, we report another set of regressions in Table 4 with only firms that report registration date. We are taking "Bottom Age Quartile" as an omitted category. We find that age does not explain any significant variation in TFPR.

Collectively, these three categories explain 20.1% of the within industry TFPR variation in Punjab. In contrast, in Hsieh and Klenow (2009) analysis, all four categories explain 4.71 % and 10.01% of the within-industry TFPR variance in India and China respectively. In our analysis, we find that size the most important driver of TFPR variation in Punjab. They find similar results for India and China; however, the magnitude is still relatively small. On the other hand, ownership is the key driver in China's with-in industry productivity dispersion. In contrast, we find it relatively less important for Punjab's case.

Furthermore, in the last column of Table 3, we find that public firms have values of TFPR on average 25.4% less than private firms. Hsieh and Klenow (2009) found a somewhat smaller difference for India (28.5%) and relatively larger difference for China (41.5%). Moreover, we find no difference between foreign and domestic private firms after controlling for size. One important explanation for this result could be very low statistical power for foreign private firms.

We also find that firms in the bottom size quartile have much higher TFPR then rest of the firms. It shows a clear evidence of economies of scale. These results are statistically highly significant. It shows that a firm's growth and TFPR have an important relationship. Moreover, we find that firms in southern Punjab have greater TFPR on average than firms in central Punjab. However, we find no differences in the other two regions. One possible reason for this result could be relatively less statistical power for imputing regional variation since, in our final

[9] Note: Districts in each region are as follows: North Punjab: Rawalpindi, Attock, Jhelum, and Chakwal
South Punjab: Bahawalpur, Bahawalnagar, R.Y. Khan, Multan, Khanewal, Lodhran, and Vehari
West Punjab: D.G.Khan, Layyah, Muzaffargarh, Bhakkar, Khushab, Rajanpur, and Mianwali
Central Punjab: Faisalabad, Jhang, T.T. Singh, Nankana Sahib, Gujranwala, Gujrat, Mandi Baha-ud-Din, Hafizabad, Sialkot, Narowal, Sheikhupura, Kasur, Okara, Sahiwal, Pakpattan, Sargodha, Lahore

sample, more than 85% of the firms fall in Central Punjab. Finally, we find that, on average, TFPR is higher in 2005-06 than 2000-01. This result is consistent across all three specifications in Table 3.

This analysis has pointed out a very important relationship between the size of a firm and TFPR for manufacturing industries in Punjab. In a later section, we will study more about this relationship by comparing the efficient (hypothetical) and actual size distributions of firms. First, we will carry out the liberalization experiment in Punjab.

Liberalization Experiment

We will now recall our derivation of firm TFP function:

$$A_s = \left[\sum_{i=1}^{M_s}\left(A_{si}\ \frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{23}$$

In their liberalization experiment, Hsieh and Klenow (2009) argue that if marginal products equalize across plants in an industry then we will observe the same (revenue based) Total Factor Productivity (TFPR) in each plant within an industry. This is because firms with more (output based) Total Factor Productivity (TFPQ) are more likely to charge lower prices in order to gain market share. Following this intuition, under perfect efficiency, the TFP function would be:

$$\overline{A_s} = \left[\sum_{i=1}^{M_s}(A_{si}\ )^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{24}$$

Combining equation (23) and (24) and making use of our CES and Cobb-Douglas aggregator, we can look at the economy-wide change in output due to equalization of marginal products across plants:

$$\frac{Y}{Y_{efficient}} = \prod_{s=1}^{S}\left[\sum_{i=1}^{M_s}\left(\frac{A_{si}}{\overline{A_s}}\ \frac{\overline{TFPR_s}}{TFPR_{si}}\right)^{\sigma-1}\right]^{\frac{\theta_s}{\sigma-1}} \tag{25}$$

In Table 8, we report the percentage gain in total output for Punjab and the corresponding estimates of Hsieh and Klenow (2009) for China, India and the US. We computed these statistics by estimating equation (25), then taking its reciprocal to arrive at the ratio of efficient to actual output, then subtracting 1 from it, and then multiply the resultant by 100.

Table 8: Full Liberalization Case

Punjab	
2005-06	**2000-01**
112.1%	66.5%
India	
100.4-127%	
China	
86.6-115.1%	
United States	
36.1-42.9%	

Note: India, China, and United States calculations are taken from Hsieh and Klenow (2009). Punjab figures are author's calculations.

Table 9: TFP Gains Relative to US

Punjab	
2005-06	**2000-01**
55.83%	22.33%
India	
40.2-59.2%	
China	
30.5-50.5%	

Note: India, China, and United States calculations are taken from Hsieh and Klenow (2009). Punjab figures are author's calculations.

We find that by fully equalizing TFPR within industries, Punjab can gain 66.5% and 112.1% in aggregate manufacturing TFP in year 2000-01 and 2005-06 respectively. Interestingly, it indicates a higher level policy distortion in Punjab in the latter year. This is consistent with our findings of higher TFPR variation in Punjab in 2005-06. In comparing these estimates with Hsieh and Klenow's (2009) calculations, we find that, on average, Punjab is relatively less distorted than India and China. In a similar way to Punjab, India also shows higher aggregate TFP gains in the latter years. Finally, we can also see that the United States is far less distorted then the remaining three countries.

We now compare our results to the US efficient output level to discover aggregate productivity gains in Punjab if it moves to the US efficiency level. To avoid the exaggeration of our results, we choose the year 1997 where the US observed its highest aggregate TFP gains, following Hsieh and Klenow. In Table 9, we report these estimates along with Hsieh and Klenow's (2009) estimates for India and China. For each year, we calculate efficient to actual output for Punjab, we then compute the same ratio for the US in 1997, and then we divide these two ratios to find aggregate TFP gains in Punjab relative to the US. We find that moving to US efficiency levels would raise aggregate TFP in Punjab by 22.33% and 55.83% in the years 2000-01 and 2005-06 respectively. These statistics again verify relatively greater distortions in the latter year for Punjab. We also observe the same pattern for India and China.

In order to check the consistency of our results with Hsieh and Klenow's (2009) analysis, we perform the same robustness check by varying the elasticity of substitution. We find that our results vary substantially when we set $\sigma = 5$ instead of $\sigma = 3$. Hsieh and Klenow's (2009) report similar results for China and India.

While Pakistan has introduced a number of liberalization polices for the manufacturing sector, this analysis indicates that allocative efficiency in the manufacturing sector of Punjab is less in 2005-06 than in 2000-01. However, as we have earlier noted, the coverage of firms is significantly higher in 2005-06, so that it would be difficult to interpret this as a true decline.

Size distribution of Firms

In our last set of estimations, we compare the actual to efficient distributions of firms by size in Punjab. We are measuring size by value-addition of a firm. We compute both of these expressions as the deviation from industry mean on a logarithmic scale. We also account for the number of firms in an industry.

We calculate actual size using the following expression:

$$Actualva_{si} = PS_{si} * M_s^{\frac{1}{\sigma-1}}$$

Here, PS_{si} is the value added of a firm and $M_s^{\frac{1}{\sigma-1}}$ is an adjustment factor for the number of industries in a sector.

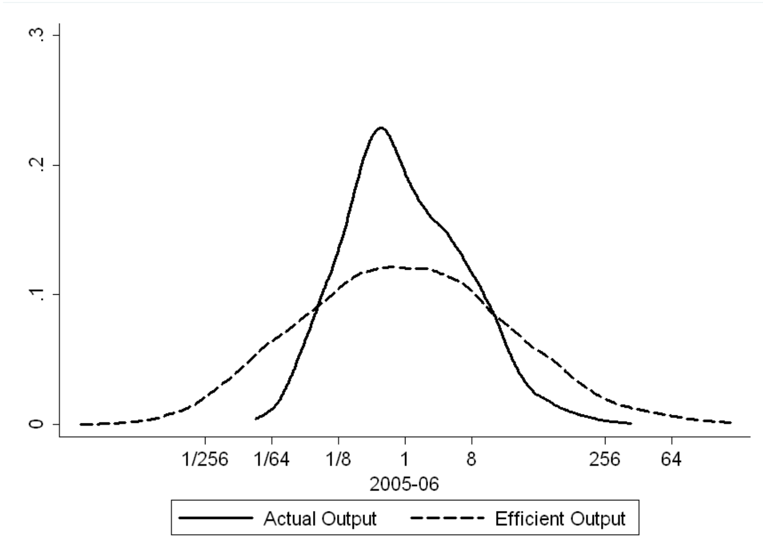
We calculate efficient size by disintegrating expression (25) for a plant level efficient output:

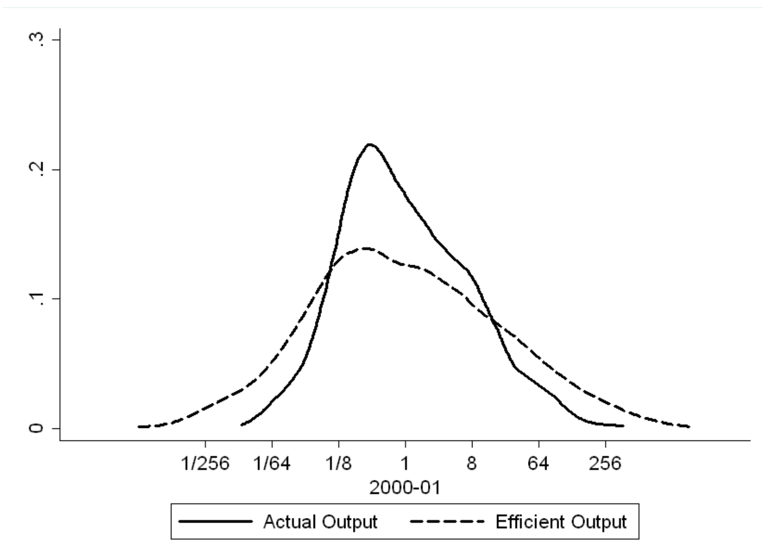
$$Efficientva_{si} = \left[\frac{A_{si}}{A_s}\frac{TFPR_s}{TFRP_{si}}\right]^{\sigma-1} * M_s^{\frac{1}{\sigma-1}} = \left[\frac{A_{si}}{A_s}\right]^{\sigma-1} * M_s^{\frac{1}{\sigma-1}}$$

In Figure 3, we draw the efficient and actual distributions using $log\,\frac{Actualv\ a_{si}}{Actualv\ a_s}$ and $log\,\frac{Efficien\ tva_{si}}{Efficientv\ a_s}$ respectively. In both years, the hypothetical distribution is much more dispersed than the actual distribution. It indicates that there should fewer medium sized firms and more small and large sized firms. Hsieh and Klenow (2009) find similar patterns for India and China.

In Figure 3, we draw efficient and actual distributions using $log\,\frac{Actualv\ a_{si}}{Actualv\ a_s}$ and $log\,\frac{Efficien\ tva_{si}}{Efficientv\ a_s}$ respectively. In both years, the hypothetical distribution is much more dispersed than the actual distribution. It indicates that there should more medium sized firms and less small and large sized firms. Hsieh and Klenow (2009) find similar patterns for India and China.

Figure 3: Efficient vs. Actual Output





Conclusion

In this paper we have exploited Hsieh and Klenow's (2009) methodology to analyze firm level data from Pakistan. We have used information from CMI 2000-01 and 2005-06 to study productivity dispersion and policy distortions in Punjab.

We find that productivity dispersion, measured by TFPQ, is higher in Punjab as compared to India and China. However, these differences become much less if productivity is measured by TFPR. Moreover, in comparing Punjab across time, we find relatively more dispersion in the year 2005-06, although this may be an artifact of the greater coverage of firms in the more recent dataset.

In the next set of estimations, we have used regression analysis to study potential sources of variation in TFPR. We find that size, age, and region explain nearly 19.3% of the TFPR variation. This figure is large compared to Hsieh and Klenow's (2009) calculations for India and China. Results indicate that size of a firm, measured by value added, is the major driver of TFPR variation. We also find that age of a firm is not important in explaining TFPR variation in Punjab.

Furthermore, we find that firms with public ownership have much lower TFPR than private domestic firms. However, this difference is still greater in China. We also find clear evidence of economies of scale. Firm in the bottom size quartile are found to have much higher TFPR as compared to larger firms. Finally, we could not find any significant variation in age quartiles and regional dummies.

In the next section, we performed a liberalization experiment to compare "efficient" output with actual output. We find that moving to absolute efficiency can boost manufacturing TFP in Punjab by 66.5% and 112.1% in years 2000-01 and 2005-06 respectively. Likewise, moving to US efficiency level would increase manufacturing TFP by 22.33% and 55.83% in years 2000-01 and 2005-06 respectively. On average, these gains are smaller than Hsieh and Klenow's (2009) estimates for India and China. This indicates relatively less distortion in Punjab.

We are also well aware of potential limitations of our results. First, in both years a significant proportion of firms did not respond to the survey. That raises the question of representativeness of our datasets. However, we believe that coverage of firms was improved in year 2005-06. Therefore, it portrays a relatively true picture of manufacturing sector in Punjab.

Secondly, we are also not certain on the exact magnitude of the measurement errors in CMI in both years.

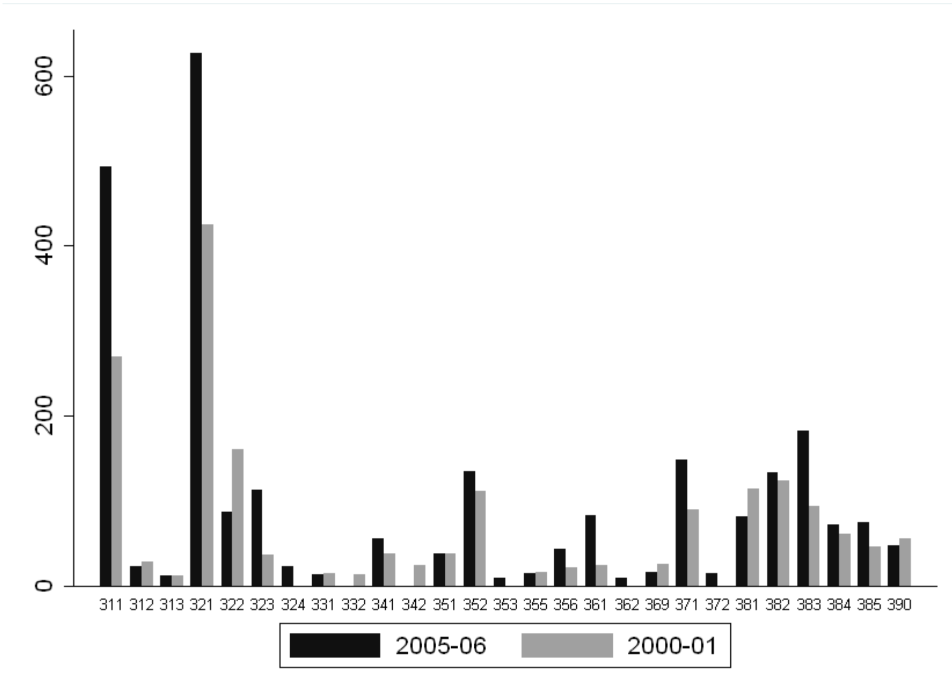
Appendix

Table A: Distribution of firms

Sr.	Industries	ISIC (Rev. 2) Codes	CMI 2005-06 (Frequency of firms)	CMI 2000-01 (Frequency of firms)	US Labor Share (%)
1	Food manufacturing	311	494	270	52%
2	Food manufacturing	312	23	28	36%
3	Beverage industries	313	12	12	42%
4	Tobacco industries	314	627	425	22%
5	Manuf. of textiles	321	87	161	76%
6	Manuf. of wearing apparel	322	113	36	75%
7	Manuf. of leather and products of leather	323	22	0	74%
8	Manuf. of footwear	324	13	15	74%
9	Manuf. of wood and wood products	331	0	13	77%
10	Manuf. of furniture and fixtures	332	55	37	76%
11	Manuf. of paper and paper products	341	0	24	66%
12	Printing, publishing and allied industries	342	38	38	67%
13	Manuf. of industrial chemicals	351	135	111	42%
14	Manuf. of other chemical products	352	9	0	34%
15	Petroleum Refineries	353	14	16	33%
16	Manuf. of Petroleum products	354	43	21	49%
17	Manuf. of rubber products	355	83	24	73%
18	Manuf. of plastic products	356	9	0	65%
19	Manuf. of pottery, china and earthenware	361	16	26	79%
20	Manuf. of glass and glass products	362	148	90	62%
21	Manuf. of other non-metallic mineral products	369	15	0	62%
22	Iron and steel basic industries	371	82	114	76%
23	Non-ferrous metal basic industries	372	133	123	53%
24	Manuf. of fabricated metal products	381	182	94	74%
25	Manuf. of machinery except electrical	382	72	61	73%
26	Manuf. of electrical machinery apparatus	383	74	46	70%
27	Manuf. of transport equipment	384	47	55	59%
28	Manuf. of scientific equipment	385	494	270	64%
29	Other Manufacturing industries	390	23	28	67%

Note: US labore shares are taken from Camacho & Conover (2010)
Frequency of firms is generated on post-cleaning data

Table B: Frequency Distribution of Firms in Punjab



References

Bartelsman, E. J., & Dhrymes, P. J. (1998). Productivity Dynamics: U.S Manufacturing Plants, 1997-1986. *Journal of Productivity Analysis,* 9, 5-34.

Bergoeing, R., Kehoe, P. J., Kehoe, T. J., & Soto, R. (2002). Policy-Driven Productivity in Chile and Mexico in the 1980's and 1990's. *The American Economic Review*, 92(2), 16-21.

Camacho, A., & Conover, E. (2010). Misallocation and Productivity in Colombia's Manufacturing Industries. IDB Working Paper Series IDB-WP-123. Inter-American Development Bank

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1), 394–425.

Hall, R. E., & Jones, C. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others?. *The Quarterly Journal of Economics*, 114(1), 83-116.

Hopenhayn, H. A. (1992). Entry, Exit and Firm Dynamics in Long Run Equilibrium. *Econometrica,* 60(5), 1127-1150.

Hopenhayn, H., & Rogerson, R.(1993). Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy*, University of Chicago Press, 101(5), 915-38.

Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4), 1403-1448.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica*, 50(3), 649-670

Khwaja, A. I., & Mian, A. (2005). Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market. *The Quarterly Journal of Economics*, 120(4), 1371-1411.

Klenow, P. J., & Rodríguez-Clare, A. (1997). The Neoclassical Revival in Growth Economics: Has It Gone Too Far? *NBER Macroeconomics Annual*, 12, 73-103.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725.

Olley, G. S., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica,* 64(6), 1263-1297.

Parente, S. L., & Prescott, E. C. (1999). Monopoly Rights: A Barrier to Riches, *American Economic Review*, 89(5), 1216-1233.

Restuccia, D., & Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics,* 11 (4), 707–720.

Schmitz, J. A. (2001). Government Production of Investment Goods and Aggregate Labor Productivity, *Journal of Monetary Economics*, 47(1), 163-187.

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics*, 39(3), 312-320.

Syverson, C. (2004). Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*, 86(2), 534-550.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature,* 49(2), 326-365.