Ann Reg Sci
DOI 10.1007/s00168-017-0824-7

ORIGINAL PAPER

Economic geography and misallocation in Pakistan's manufacturing hub

Theresa Chaudhry[1] · **Muhammad Haseeb**[2] · **Maryiam Haroon**[3]

Received: 16 May 2016 / Accepted: 10 March 2017

Abstract In this article, we explore whether localization of industries can reduce economic distortions and dispersion in total factor productivity (TFP) among firms in Punjab, Pakistan's largest province economically. We consider two types of misallocation: (i) dispersion in the distribution of output-based TFP (TFPQ), in particular the survival of low productivity firms in the left tail; and (ii) dispersion in revenue-based TFP (TFPR), indicative of allocative inefficiency. The results are mixed: On the one hand, we find that the distribution of TFPQ is less dispersed in more agglomerated areas, measured by the localization quotient, local productive concentration, and average firm size. At the same time, we find that average TFPQ is also positively related to localization, especially the presence of small firms in the same sector, even though own-firm TFP is lowest for small firms. On the other hand, we do not find evidence that agglomeration improves allocative efficiency measured as deviations in TFPR from the sector average, concluding rather that greater localization of small firms is associated with firms being more output and capital constrained.

✉ Theresa Chaudhry
theresa@lahoreschool.edu.pk

Muhammad Haseeb
M.Haseeb@warwick.ac.uk

Maryiam Haroon
maryiamharoon@gmail.com

1 Department of Economics, Lahore School of Economics, Intersection Main Boulevard, Phase VI, DHA and Burki Road, Burki, Lahore 53200, Pakistan

2 Department of Economics, University of Warwick, Coventry CV4 7AL, UK

3 Centre for Research in Economics and Business, Lahore School of Economics, Intersection Main Boulevard, Phase VI, DHA and Burki Road, Burki, Lahore 53200, Pakistan

Published online: 20 March 2017

JEL Classification D24 · R12 · L11 · L25 · L6

1 Introduction

Differences in total factor productivity (TFP) are among the most important factors in understanding differences in output per worker across countries (Banerjee and Duflo 2005). One type of misallocation lowering aggregate output-based TFP (TFPQ) is greater dispersion in its distribution, in particular the survival of low productivity firms in the left tail. Persistent dispersion has important consequences for competitiveness: When low productivity firms remain in business, high productivity firms' access to scarce resources is limited, which lowers aggregate TFP. Substantial dispersion in the distribution of TFP has been documented for both developing and developed economies (Hsieh and Klenow 2009; Syverson 2011; Vollrath 2009) but it tends to be higher in developing countries.

Using an alternative measure of productivity, Hsieh and Klenow (2009) suggest that there is another type of misallocation represented by dispersion in *revenue*-based (as opposed to output-based) total factor productivity (TFPR). Deviations in TFPR represent a failure of the economy to achieve allocative efficiency, in other words the correct allocation of inputs to their appropriate use based on the marginal revenue product of inputs. Policies, regulations, or other factors differentially affecting a subset of firms can distort input choices by altering their cost of capital or output levels, thus moving them away from their efficient levels. Examples of these distorting policies are taxes, differential credit access, and restrictions on location and size.[1] For example, if a firm has sub-par access to credit and faces a higher interest rate than average, this leads to a downward distortion in the amount of capital employed by the firm. In the other direction, some firms may get preferential access to credit, increasing the use of capital in production beyond its efficient level. Earlier research indicates that firms in Punjab, Pakistan, indeed face differing prices for labor inputs based on location (Naveed 2015; Tirmazee 2012). Hsieh and Klenow (2009) calculated that moving to the levels of efficiency obtained by relatively undistorted US firms would boost aggregate TFP significantly in both China (by 30–50%) and India (by 40–60%). Using the same methodology, Haseeb and Chaudhry (2014) determine that the distributions of TFP (both TFPQ and TFPR) were even more dispersed in Pakistan than in China and India.

In this article, we build on the results of Haseeb and Chaudhry (2014) to relate the dispersion of the TFP distribution in Pakistan to its economic geography. We aim to understand whether agglomeration, in the form of localization, can help to reduce unfavorable dispersion in total factor productivity, either through the distributions in output-based and revenue-based TFP or through the measures of output and capital distortions derived by Hsieh and Klenow (2009).[2] We find evidence that localization reduces dispersion in output-based TFP but has no impact on allocative efficiency

[1] Migration restrictions, such as those facing both labor and firms in China, can lead to firms losing out from scale and agglomeration economies (Au and Henderson 2006).

[2] TFPR is proportional to a ratio of output and capital distortions (see Hsieh and Klenow 2009, p. 1410).

Table 1 Agglomeration index for selected countries. *Source*: World Development Report 2009, The World Bank

Countries	Agglomeration index
USA	71.9
India	52.4
Pakistan	53.6
China	37.2

Agglomeration index ranges from 0 (low) to 100 (high)

implied by the dispersion in revenue-based TFP distribution. There is some evidence that localization, particularly of small firms, is significantly related higher levels of output and capital restrictions on firms. Further, though firm size (by value added) is positively related to output-based TFP, the localization of small firms in the same sector tends to boost average firm-level TFPQ.

We focus on two types of misallocation: (i) dispersion in output-based TFP (TFPQ), especially the survival of low productivity firms, which tends to lower average TFP; and (ii) dispersion in revenue-based productivity measures (TFPR) indicating a lack of allocative efficiency as identified by Hsieh and Klenow (2009). Given that economic geography is known to impact firm-level performance, usually in positive ways,[3] an interesting question is whether agglomeration has a role in the dispersion in TFP and the misallocation of resources. Along these lines, our research explores whether agglomeration is associated with a reduction in the dispersion of output-based and/or revenue-based TFP measures. Further, we measure the extent to which agglomeration is associated with output and capital distortions individually.

Pakistan is a relevant case to study because, like India, it has an intermediate level of agglomeration in contrast to the USA, which is highly localized, and China, which is, at the aggregate level, relatively less agglomerated (see Table 1). Countries like Pakistan and India provide an opportunity to study the role of localization in the phase of development where agglomeration economies are high but before congestion has set in Desmet et al. (2015).

Agglomeration can impact the dispersion in firm-level output-based productivity (TFPQ) through the type of firms that *enter* a market, the type of firms that *survive*, and how the benefits of agglomeration are allocated across firms in concentrated areas. If agglomeration reduces the number of low productivity firms in the left tail of the distribution, then misallocation will also go down. This is because resources that had been employed by these exiting, low productivity firms will now be put to more productive uses.

[3] The findings of the existing empirical literature on agglomeration have largely documented that agglomeration, in the form of localization and/or urbanization, enhances the productivity of firms (Andersson and Lööf 2011; Ehrl 2013; Henderson 1986, 2003; Hu et al. 2015; Henderson et al. 1995; Hansen 1990; Hanson 1996, 1997). On the other hand, other aspects of agglomeration might lower productivity, for example, if congestion sets in, or if firms in agglomerated sectors remain small and, due to their large number, are unable to achieve economies of scale. According to Shaver and Flyer (2000), localization economies disproportionately enhance the competitiveness of weak firms possibly at the expense of stronger firms, leading the latter to avoid concentrated areas. Folta et al. (2006) find diminishing returns to agglomeration as a cluster grows. Spatial proximity, by facilitating communication among firms, may even enable non-competitive behavior such as collusion. Evidence from Chinese manufacturing firms reveals non-competitive pricing in industrial clusters (Brooks, Kaboski and Li).

Combes et al. (2012) build a model combining agglomeration economies (that right-shift the productivity distribution but also possibly dilate it if more productive firms benefit disproportionally) in addition to a selection effect of agglomeration (where greater competition in larger cities forces weaker firms out of the market). The selection effect, due to greater competitive pressure, should cause the distribution of firm-level TFP to be more truncated on the left in more agglomerated areas, that is, less dispersed, by raising the minimum level of TFP required for survival, thus reducing misallocation. Combes et al. (2012), however, did not find the selection effect to be strong in France. Evidence from Pakistan suggests exit and entry rates are both higher in more agglomerated areas (Haroon and Chaudhry 2014; Nasir 2013). But even if exit rates are higher in clusters, it is guaranteed neither that only low productivity firms will exit nor that new firms will be more productive.[4]

On the other hand, as alluded to above, the distribution of TFPQ could be more dispersed if the benefits of agglomeration are not evenly distributed but instead are increasing in the firm's own productivity. Combes et al. (2012) allow for this possibility in their model and find evidence in French data to support the idea that agglomeration economies "dilate" the productivity distribution in addition to right-shifting it.

Moving on to the second type of misallocation, any dispersion in the revenue-based productivity measure, TFPR, represents a lack of allocative efficiency, that is, the failure of the economy to allocate inputs to their most productive use. This can happen when firms are treated differently, for instance, if some firms have greater access to finance (leaving other firms capital constrained), or when labor regulations are unevenly applied (causing the affected firms to be labor constrained, that is, capital subsidized in relative terms), or when well-connected firms are less subject to red tape (leaving other firms output constrained). Barriers to reallocation of inputs from less productive to more productive firms can also result in dispersion in TFPR. One may expect agglomeration to reduce dispersion in TFPR if geographic proximity and dense networks that allow information diffusion in agglomerated areas facilitate the movement of labor toward more productive firms and loosen credit constraints to give firms more access to credit. On the other hand, if some agglomerated areas are also special economic zones, the firms there may have access to incentives, finance, or superior infrastructure as compared to firms in the same sector located outside the zone, driving a wedge between their TFPR and the sector average and increasing dispersion.

Competition, which reduces output prices charged by firms with higher output-based productivity, is also necessary for TFPR to converge across firms in the same sector (Hsieh and Klenow 2009). This competition effect may be stronger or weaker in agglomerated areas, as the empirical evidence is mixed. Siba et al. (2012) found evidence that increased competition in agglomerated areas both lowered prices and raised total factor productivity in Ethiopia, while Brooks et al. (2016) discovered evidence of collusion in Chinese industrial clusters.

Some recent work has explored the impact of agglomeration on firm-level TFP, where TFP is calculated as a residual from the estimation of a firm-level production

[4] Duranton (2015) survey of the current literature does not identify any strong evidence to date that productivity growth in urban areas is responsive to higher rates of entry and exit.

function (Cingano and Schivardi 2004; Graham 2009; Martin et al. 2011; Combes et al. 2012; Ehrl 2013; Hu et al. 2015). We have instead followed Hsieh and Klenow's methodology to calculate TFP directly, rather than as a production function residual.[5] We also use Hsieh and Klenow's approach to calculate misallocation of resources, including firm-level output and capital distortions, and link them to agglomeration. Their techniques have been applied in studies on productivity and capital misallocation (Chen and Irarrazabal 2015; Inklaar et al. 2015; Kalemli-Ozcan and Sørensen 2014). However, few if any studies have looked at the impact of the localization of different size firms on TFP dispersion and misallocation, as we do here in our research.

The remainder of the article is organized as follows. Section 2 describes the data sources and the methodology developed by Hsieh and Klenow (2009) for measuring firm-level TFP and capital and output distortions. Section 3 presents the results of the analysis in two subsections: (i) district-level distributions of firm-level TFP and (ii) correlation of TFP dispersion, output, and capital distortions with agglomeration. Section 4 concludes.

2 Description of the data and measurement of firm-level TFP and distortions

The main source of data is the Census of Manufacturing Industries (CMI) 2005–2006 for the Punjab province of Pakistan, which contains data on 3528 manufacturing units, a response rate representing approximately 35–40% of firms in Punjab.[6] We focus on Punjab since it is the largest province in Pakistan population-wise and economically, and it is responsible for nearly half of the manufacturing value added in the country (Pakistan Bureau of Statistics akistan Bureau of Statistics).

Our empirical analysis additionally uses data from the Government of Punjab's Directory of Industries (DOI) for 2002, 2004, and 2006 to calculate the measures of agglomeration/localization. The DOI is a firm-level data set, which includes more than 16,000 firms in each year and includes information on each firm's year of establishment, employment level, and location.

In what follows, we use the methodology of Hsieh and Klenow (2009) for the calculation of total factor productivity and distortions. Their framework considers two kinds of distortions: those that affect the output level ($\tau_{Y_{si}}$), and those affecting amount of capital used relative to labor ($\tau_{K_{si}}$).

Hsieh and Klenow (2009) write the firm's profit function for firm i in sector s as:

$$\pi_{si} = \left(1 - \tau_{Y_{si}}\right) P_{si} Y_{si} - wL_{si} - \left(1 + \tau_{K_{si}}\right) RK_{si} \qquad (1)$$

[5] We lack the panel data for firms in Pakistan required to apply semiparametric methods such as Olley and Pakes (1996) and Levinsohn and Petrin (2003), or generalized methods of moments (GMM) dynamic panel models.

[6] Discussions with the agency responsible for collecting the data indicate that the sample is skewed toward smaller units in the sampling frame. After dropping cotton-ginning activities (as it is no longer considered a manufacturing activity according to most international studies), the raw data set is left with just over 3000 firms.

where P is the output price, Y is output, w is the wage, L is labor, R is the rental rate of capital, and K is capital. After a number of manipulations (which the interested reader can find in Hsieh and Klenow's paper), the distortion parameters for capital and output, respectively, are imputed as:

$$1 + \tau_{K_{si}} = \frac{\alpha_s}{(1-\alpha_s)} \frac{wL_{si}}{RK_{si}} \tag{2}$$

$$1 - \tau_{Y_{si}} = \frac{\sigma}{(\sigma-1)} \frac{wL_{si}}{(1-\alpha_s)\, P_{si} Y_{si}} \tag{3}$$

where σ is the elasticity of substitution between plants and α_s and $(1-\alpha_s)$ are the sector-specific shares of capital and labor in output, respectively. Following Hsieh and Klenow (2009), the elasticity of substitution takes a value of $\sigma = 3$ and the US shares of labor are used for $(1-\alpha_s)$, as they are considered relatively undistorted.

A positive (negative) output distortion $\tau_{Y_{si}}$ signifies that the firm is output restricted (subsidized). Similarly, a positive (negative) capital distortion $\tau_{K_{si}}$ means that the firm is capital restricted (subsidized) relative to labor. In both cases, we are comparing undistorted US labor and capital shares with the respective observed information for Punjab to infer the distortions present.[7]

The derivation of the productivity measure begins with a Cobb–Douglas firm-level production function:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \tag{4}$$

where K_{si} and L_{si} represent capital and labor respectively, and A_{si} represents the firm's individual *output-based* total factor productivity measure. Naturally, each firm will have its own value of A_{si} (called TFPQ here), because firms, depending on the characteristics of the owners and managers, location, and even luck, will differ in levels of entrepreneurial ability, organizational capital, market power, and access to customers, among other factors. The distribution of A_{si} for the firms who enter and survive may also depend on the level of agglomeration as discussed earlier. Using the methodology of Hsieh and Klenow (2009), TFPQ, or A_{si}, is computed from the data using the following formula:

$$\text{TFPQ} = A_{si} = \kappa_s \frac{(P_{si} Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}} \tag{5}$$

where κ_s, which is unobserved, is a scalar and following Hsieh and Klenow (2009) is assumed without loss of generality to take the value $\kappa_s = 1$.

An alternative productivity measure, which we will see can also be used to measure misallocation or distortions in the economy, is the *revenue-based* total factor productivity measure, or $P_{si} A_{si}$, referred to as TFPR for firm i in sector s, and is defined from the production function as:

[7] According to Hsieh and Klenow (2009), an output distortion is observed where the labor share is different from the elasticity of output with respect to labor and a capital distortion is observed where the ratio of a plant's wage bill to its capital stock differs from the ratio of the respective output elasticities.

$$\text{TFPR} = P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}L_{si}^{1-\alpha_s}} \tag{6}$$

The idea for the revenue-based TFPR measure of productivity is that firms that are more efficient (with higher A_{si} or TFPQ) will have higher production, but will have to sell their output at a lower price because of downward sloping demand for their particular product variety (Hsieh and Klenow 2009). If all firms receive equal treatment (with respect to credit allocation, regulations, etc.), then more resources will be allocated toward the efficient producers with higher A_{si} until the marginal products of capital and labor are equated across firms. Through this process, TFPR should become equalized across firms in the economy, regardless of the individual values of A_{si} (or TFPQ). However, if we detect dispersion in TFPR, then we can conclude that there is misallocation in the economy; in other words, some firms are relatively favored, while others are relatively disadvantaged.

We use the data from the CMI 2005–2006 to calculate TFPQ, TFPR, and capital and output distortions following the methods of Hsieh and Klenow (2009) for the analysis that follows.[8]

3 Results and analysis

We first look graphically at how the distributions of both output-based and revenue-based total factor productivities, TFPQ and TFPR, vary with agglomeration in Punjab. The second set of results involves estimation of the relationship between agglomeration and the total factor productivity distributions as well as output and capital distortions.

3.1 Dispersion of total factor productivity and agglomeration

We start by looking to see whether a graphical relationship exists between agglomeration and the distributions of TFPQ and TFPR, that is, the output- and revenue-based total factor productivities respectively. As defined in Hsieh and Klenow (2009), we plot the log difference of firm TFP from the sector average TFP according to the level of agglomeration (in quartiles) by district.[9]

As we can see from Fig. 1, the dispersion in TFPQ does not appear fall monotonically as the level of agglomeration rises. However, when comparing the distribution of the least agglomerated to that of the most highly agglomerated, the former seems to have a lower mean productivity and greater mass of underperforming firms in the left tail. Dispersion in TFPR in Fig. 2, indicative of allocative inefficiency according to

[8] The data requirements include labor compensation, nominal output (revenue), expenditures on input materials and energy, book value of capital, and the industry level cost shares for labor and capital. Details of calculations of the firm-level TFPQ and TFPR in Pakistan can be found in Haseeb and Chaudhry (2014).

[9] Punjab's districts were ranked by their location quotient and divided into quartiles. The firms in those districts in each quartile of agglomeration were included in the distribution. For example, the firms in districts considered most agglomerated were included in the distribution for "very highly agglomerated" areas.

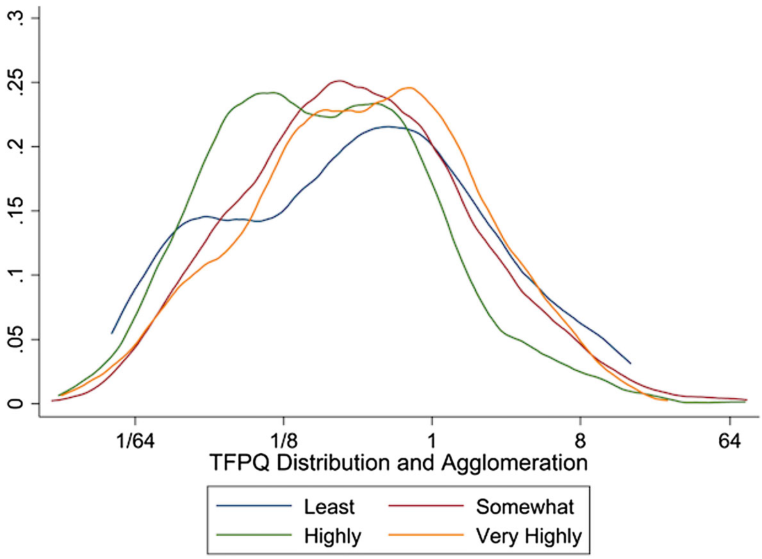

Fig. 1 Distribution of TFPQ by level of agglomeration (in quartiles). *Source*: Authors' calculations based on CMI Punjab, 2005–2006, trimmed data

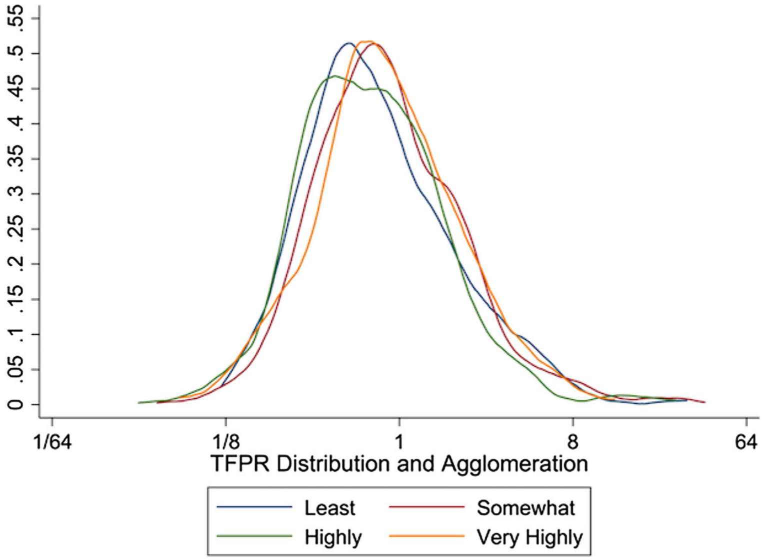

Fig. 2 Distribution of TFPR by level of agglomeration (in quartiles). *Source*: Authors' calculations based on CMI Punjab, 2005–2006, trimmed data

Hsieh and Klenow (2009), seems to bear little relation to the level of agglomeration.[10] We do not see a clear relationship between the intensity of agglomeration and either of

[10] That dispersion in TFPR is in general less than that for TFPQ is an expected result since firms that are more productive (with higher A_{si} or TFPQ), should be producing more output at lower prices (Hsieh and Klenow 2009).

Table 2 Agglomeration and dispersion in total factor productivity. *Source*: Authors' calculations

Agglomeration intensity	TFPQ			TFPR		
	SD	90–10	95–5	SD	90–10	95–5
Low	1.59	4.18	5.56	0.81	2.12	2.84
Somewhat	1.56	4.08	5.19	0.88	1.97	2.66
High	1.67	4.54	5.41	0.89	1.95	2.83
Very high	1.56	4.05	5.32	0.89	2.36	3.06

the TFP distributions when we consider various attributes of the distributions, including the standard deviation or percentile ratios such as 90/10 and 95/5 ratios (Table 2). However, if we again limit ourselves to comparing the TFP distributions where the agglomeration intensity is "low" with those areas where it is "very high", it appears that the dispersion of TFPQ falls with increased agglomeration intensity while the opposite is true for TFPR. In other words, one type of misallocation may be falling while the other is rising. Kolmogorov–Smirnov tests for the equality of distributions indicate that the differences between the TFP distributions for "low" versus "very high" agglomeration are statistically significant (p values are 0.064 and 0.026 for TFPQ and TFPR, respectively; see Table 8 in "Appendix").

3.2 Correlation of agglomeration with misallocation and TFP

In this section, we look at how misallocation, measured as dispersion the distribution of output- and revenue-based TFP, and firm-level output and capital distortions, are related to agglomeration. Recall that dispersion in TFPQ represents the distribution of firm-level total factor productivities, A_{si}. The distribution of TFPQ depends on natural factors (such as the distribution of entrepreneurial talent) but can also depend on other factors including access to markets, market power, and frictions that impact firm entry and exit, including agglomeration. Any dispersion in TFPR represents a failure of allocative efficiency such that marginal products have failed to equalize. The output and capital distortion calculations were based on derivations from Hsieh and Klenow (2009), though they did not use them in a similar regression exercise.

The regression equations for the TFP distributions (estimated separately) are:

$$\log \mathrm{TFPQ}_{si} - \log \overline{\mathrm{TFPQ}_s} = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{7a}$$

$$\log[abs(\mathrm{TFPQ}_{si} - \overline{\mathrm{TFPQ}_s})] = \alpha + \beta L_{si} + \sum_j \gamma_j X_{jsi} + \varepsilon_{si} \tag{7b}$$

$$\log\left[abs\left(\mathrm{TFPR}_s i - \overline{\mathrm{TFPR}_s}\right)\right] = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{8}$$

where the dependent variable in Eq. (7a) is the log deviation from sector average TFPQ, L is the measure of localization/agglomeration and X_j is a vector representing firm-level characteristics considered (including firm size by value added and ownership type) and district fixed effects. The dependent variables in Eqs. (7b) and (8) are the log of total TFP dispersion, where total TFP dispersion is calculated as the absolute value of the difference between firm-level TFP (for firm i in sector s) and geometric mean TFP for sector s, for TFPQ and TFPR in 2006, respectively. The difference between Eqs. (7a) and (7b) is that the first equation will tells us whether agglomeration has a relationship with average TFPQ (in other words, if agglomeration right-shifts the TFPQ distribution), whereas (7b) will indicate how agglomeration relates to total dispersion (since all deviations from mean TFPQ, both positive and negative, are treated as dispersion). Hsieh and Klenow (2009) carried out a regression similar to Eq. (7a) for China and India, but they only looked at the relationship of TFP with the type of firm ownership or firm exit; they did not consider any measures of agglomeration in their estimations.

Next we examine individually the output and capital distortions derived by Hsieh and Klenow (2009). Similar to the TFPQ analysis, we will consider the relationship of localization with total distortions (the absolute values) as well as the average distortion (which tells us whether firms are on average constrained or subsidized). The equations we estimate are:

$$ln \left|\tau_{Y_{si}}\right| = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{9}$$

$$\ln \left(1 - \tau_{Y_{si}}\right) = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{10}$$

$$ln \left|\tau_{K_{si}}\right| = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{11}$$

$$\ln \left(1 + \tau_{K_{si}}\right) = \alpha + \beta L_{si} + \sum_j \gamma X_{jsi} + \varepsilon_{si} \tag{12}$$

where the dependent variables in Eqs. (9) and (11) are the log of the absolute value of output and capital distortions and in Eqs. (10) and (12) are the log values of the firm-level output distortion, $\left(1 - \tau_{Y_{si}}\right)$, and capital distortion, $(1 + \tau_{K_{si}})$, respectively; L is the measure of localization/agglomeration, X_j is a vector representing firm-level characteristics considered (including firm size by value added and ownership type) and district fixed effects. Recall that $\tau > 0 (\tau < 0)$ represents an output or capital constraint (subsidy). When we take the absolute value of τ for Eqs. (9) and (11), we are treating both subsidies and constraints equally as evidence of distortions (misallocation), and the estimations will tell us if agglomeration is associated with greater or lower levels of total misallocation. In Eq. (10), the dependent variable $\ln\left(1 - \tau_{Y_{si}}\right)$ will take a negative value when a firm is output constrained, while in Eq. (12) $\ln\left(1 + \tau_{K_{si}}\right)$ will take a positive value when a firm is capital constrained; the signs will be reversed in the case that a firm is output or capital subsidized.[11] The results for Eqs. (10) and (12) will tell us whether firms in more agglomerated areas are on average constrained or subsidized in their output and capital usage.

We use three different measures of agglomeration considered in the literature, and we construct them for three different periods: 2006, 2004, and 2002. The first measure is the *location quotient* (Beaudry and Schiffauernova 2009) and is constructed as (employees in region x and industry y)/(total employees in industry y), where total employment is based on the total manufacturing employment in Punjab. We also disaggregate the location quotient into localization of small, medium, and large firms.[12] Second, we calculate the *local productive concentration* as (employment in firm z/employment in region x and industry $y)^2$, summed over all firms z in region x and sector y (Combes et al. 2004). Finally, we use the region's *average firm size* in the sector, which is *inversely* related to localization, calculated as (employees in region x and industry y)/(number of firms in region x and industry y), as described in Glaeser et al. (1992).

The relationship between localization and productivity is likely to run in both directions. Localization can increase productivity, but at the same time more productive firms are more likely to grow, and grow larger, leading to greater agglomeration (Graham et al. 2010). Given these concerns, our preferred specifications for the main results, which we report in Tables 3 and 4, are instrumental variable regressions, where localization in 2006 is instrumented with its lagged values from 2004 and 2002.[13]

3.2.1 Results using aggregate localization measures

First, we look at the results of estimating how the distributions of output-based and revenue-based TFP, that is, TFPQ and TFPR, are related to localization. Excess dispersion on the left side of the TFPQ distribution and any dispersion in the TFPR distribution is indicative of misallocation. Beginning with the TFPQ distribution, recall that according to Combes et al. (2012), agglomeration economies can right-shift and possibly dilute the distribution if the highly productive firms benefit most, but the selection effect of greater competition should force out less productive firms, reducing the density in the left tail of the distribution. We find strong evidence in Table 3 that localization is associated with both higher average TFPQ (cols. 1, 4, and 7) and lower dispersion in the TFPQ distribution (cols. 2, 5, and 8) for all three measures of

[11] Like our Eq. (12), Kalemli-Ozcan and Sørensen (2014) have used the log value of the Hsieh and Klenow capital distortion in a regression analysis, but they were not examining the role of agglomeration but rather institutional failures in Africa.

[12] For example, the small firm location quotient is defined as (employees of small firms in region x and industry y)/(total employees in industry y). Small firms are defined as firms with less than 10 workers, medium size as firms with 10–49 workers, and large firms having employment level greater than 49 workers.

[13] Reed (2015) has suggested that instrumenting current values with lagged values may be a better solution to simultaneity than using lagged values alone.

Table 3 Instrumental variable regressions of deviations of TFPQ and TFPR from sector averages on localization in 2006. Source: Authors' calculations based on CMI Punjab 2005–2006, trimmed data, and Directory of Industries 2006, 2004, 2002

Localization measured as:	Location quotient			Local productive concentration			Average firm size		
	(1) Average TFPQ deviation	(2) Absolute value of TFPQ deviation	(3) Absolute value of TFPR deviation	(4) Average TFPQ deviation	(5) Absolute value of TFPQ deviation	(6) Absolute value of TFPR deviation	(7) Average TFPQ deviation	(8) Absolute value of TFPQ deviation	(9) Absolute value of TFPR deviation
Log localization 2006 (IV)	0.470***	−0.590***	−0.101	0.370***	−0.742***	−0.278***	−0.281***	0.254***	−0.0752
	(4.479)	(−9.303)	(−0.849)	(4.145)	(−11.62)	(−2.615)	(−7.560)	(15.72)	(−1.612)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2284	2284	2284	2284	2284	2284	2285	2285	2285
R-squared	0.613	N/A	0.044	0.634	N/A	0.004	0.312	0.312	0.048

Dependent variables are $\log \text{TFPQ}_{si} - \log \overline{\text{TFPQ}_s}$, $\log[abs(\text{TFPQ}_{si} - \overline{\text{TFPQ}_s})]$, $\log\left[abs(\text{TFPR}_{si} - \overline{\text{TFPR}_s})\right]$

IV: Instruments for Localization in 2006 are Localization in 2004 and 2002

t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; standard errors clustered at the district level

Regressions have also controlled for status as public enterprise, foreign collaboration, firm size (value added) quartile

Table 4 Instrumental variable regression of distortions in output (Y) and capital (K) on localization in 2006. *Source*: Authors' calculations based on CMI Punjab 2005–2006, trimmed data, and directory of industries 2006, 2004, 2002

Localization measured as:	Location quotient			Local productive concentration				Average firm size				
	(1) Absolute value of Y distortion	(2) Average Y distortion	(3) Absolute value of K distortion	(4) Average capital distortion	(5) Absolute value of Y distortion	(6) Average Y distortion	(7) Absolute value of K distortion	(8) Average K distortion	(9) Absolute value of Y distortion	(10) Average Y distortion	(11) Absolute value of K distortion	(12) Average K distortion
Log localization 2006 (IV)	0.132	−0.117*	0.249*	1.002*	0.129	0.319	0.947**	0.829	−0.037	0.0294	−0.083	−0.127
	(1.216)	(−1.755)	(1.683)	(1.845)	(0.634)	(1.475)	(2.461)	(1.618)	(−1.231)	(0.636)	(−1.121)	(−0.951)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2284	2284	2284	2284	2285	2285	2285	2285	2285	2285	2285	2285
R-squared	0.086	0.31	0.03	0.27	0.08	0.318	0.05	0.213	0.081	0.318	0.043	0.217

Dependent variables are $ln\left|\tau_{Y_{si}}\right|$, $\ln\left(1n-\tau_{Y_{si}}\right)$ for the absolute value and average output distortions respectively, and $ln\left|\tau_{K_{si}}\right|$, $\ln\left(1+\tau_{K_{si}}\right)$ for the absolute value and average capital distortions, respectively

IV: Instruments for Localization in 2006 are Localization in 2004 and 2002

t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; standard errors clustered at the district level

Regressions have also controlled for status as public enterprise, foreign collaboration, firm size (value added) quartile localization.[14] Moving to the TFPR distribution, dispersion may be lower in agglomerated areas if competition brings down producer prices and efficient input markets allocate labor and capital efficiently, or higher if agglomerated producers collude more or receive special treatment. In our specifications, localization is related to lower TFPR dispersion in two out of three specifications, but it is only signification when measured as local productive concentration (Table 3, cols. 3, 6, and 9).

When output distortions are present, some firms are restricted from growing to their efficient size ($\tau_{Y_{si}} > 0$), so that the dependent variable $\ln\left(1 - \tau_{Y_{si}}\right)$ will take a negative value in Eq. (10). Other firms are favorably treated and grow beyond the size that is most efficient ($\tau_{Y_{si}} < 0$), so that the dependent variable will take a positive value. The results of estimating of Eq. (10) in Table 4 will therefore inform us as to whether firms are on average output restricted or output subsidized. The estimation of Eq. (9), on the other hand, treats both output constraints and subsidies as distortions, for a measure of total misallocation. These results are reported in Table 4. When we estimate Eq. (9), we do not find that absolute output distortions are related in a statistically significant way to any of our localization measures (Table 4, cols. 1, 5, and 9). We find that localization is associated with firms being on average output constrained (Eq. 10) but significant only when the location quotient is used (Table 4, col 2).

We perform a similar exercise for capital distortions. In estimating Eq. (12), a positive value of the dependent variable $\ln\left(1 + \tau_{K_{si}}\right)$ indicates that a firm is capital constrained relative to labor ($\tau_{K_{si}} > 0$) and a negative value indicates that a firm is capital subsidized, that is, labor constrained ($\tau_{K_{si}} < 0$). The results of estimating of Eq. (12) in Table 4 will therefore inform us as to whether firms are on average capital restricted or subsidized. The estimation of Eq. (11), on the other hand, treats both capital constraints and subsidies as distortions, for a measure of total misallocation. We find that localization is significantly related to total capital distortions in two out of three specifications (Table 4, cols. 3, 7, and 11). We find that localization is related to firms being on average capital constrained (Eq. 12) but significant only when the location quotient is used (Table 4, col. 4).

While not reported in Table 4, a very robust result across specifications using the location quotient or local productive concentration is that, in comparison with the largest firms (measured by value added), firms in first three firm size quartiles are relatively capital restricted but output subsidized, and these differences are large and statistically significant. The smallest quartile of firms (in value added terms) is the most output subsidized relative to the largest quartile.

3.2.2 Results using disaggregated localization measures

Next, we decompose one of the agglomeration measures, the location quotient, into localization of small, medium, and large firms separately and look at their individual correlations with TFPQ dispersion, output distortions, and capital distortions (Tables 5,

[14] We include district dummies in most specifications to control for levels of infrastructure and human capital and proximity to markets.

Table 5 Regression of log TFPQ (relative to sector mean) on the disaggregated location quotient. *Source*: Authors' calculations based on CMI Punjab 2005–2006, trimmed data, and directory of industries 2006, 2004, 2002

	Log avg TFPQ deviation in 2006			Log absolute value of TFPQ deviation in 2006		
	(1)	(2)	(3)	(4)	(5)	(6)
Log small firm loc Q (2006)	0.110***			−0.183***		
	(6.746)			(−35.71)		
Log medium-size firm loc Q (2006)	−0.0433**			0.0494***		
	(−2.383)			(6.488)		
Log large firm loc Q (2006)	−0.125***			0.0403***		
	(−11.62)			(8.249)		
Log small firm loc Q (2004)		0.178***			−0.192***	
		(8.751)			(−33.04)	
Log medium-size firm loc Q (2004)		−0.00764			−0.00734	
		(−0.382)			(−0.991)	
Log large firm loc Q (2004)		−0.133***			0.0674***	
		(−10.99)			(13.86)	
Log small firm loc Q (2002)			0.145***			−0.119***
			(7.339)			(−46.83)
Log medium-size firm loc Q (2002)			0.0394*			−0.0804***
			(1.947)			(−20.80)
Log large firm loc Q (2002)			0.0873***			−0.0533***
			(4.804)			(−16.24)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2692	2681	2285	2672	2681	2284
Adjusted *R*-squared	0.696	0.687	0.685	0.531	0.564	0.626

Dependent variable is $\log \text{TFPQ}_{si} - \log \overline{\text{TFPQ}_s}$ or $\log[abs(\text{TFPQ}_{si} - \overline{\text{TFPQ}_s})]$

t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; standard errors clustered at the district level

Regressions have also controlled for status as public enterprise, foreign collaboration, firm size (value added) quartile

6, 7). The regressors of interest are current and lagged localization, and estimation is by OLS.

The localization of small firms is positively related to average total factor productivity (TFPQ) in Table 5 (cols. 1–3). Hu et al. (2015) found that the localization of upstream firms supported firm-level TFP, and although we are not able to prove it, this may also be the case here, since large vendor segments comprising small firms

Table 6 Regression of distortions in output on the disaggregated location quotient. *Source*: Authors' calculations based on CMI Punjab 2005–2006, trimmed data, and directory of industries 2006, 2004, 2002

	Average output distortions in 2006			Absolute value of output distortions in 2006		
	(1)	(2)	(3)	(4)	(5)	(6)
Log small firm loc Q (2006)	−0.00546			0.0220*		
	(−0.317)			(1.847)		
Log medium-size firm loc Q (2006)	−0.0468*			0.0415**		
	(−1.758)			(2.086)		
Log large firm loc Q (2006)	0.0448**			−0.0284		
	(2.442)			(−1.452)		
Log small firm loc Q (2004)		−0.0425*			0.0612***	
		(−1.758)			(3.412)	
Log medium-size firm loc Q (2004)		0.00502			−0.0134	
		(0.142)			(−0.491)	
Log large firm loc Q (2004)		0.0194			−0.0202	
		(0.707)			(−0.833)	
Log small firm loc Q (2002)			−0.0221***			0.0166*
			(−2.986)			(1.857)
Log medium−Size firm loc Q (2002)			0.0126			−0.0133
			(1.160)			(−1.023)
Log large firm loc Q (2002)			0.0114			0.00292
			(1.686)			(−0.236)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2692	2681	2285	2672	2681	2284
Adjusted R-squared	0.330	0.324	0.322	0.085	0.08+	0.082

Dependent variable is $\ln\left(1-\tau_{Y_{si}}\right)$ or $ln\left|\tau_{Y_{si}}\right|$
t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; standard errors clustered at the district level
Regressions have also controlled for status as public enterprise, foreign collaboration, firm size (value added) quartile

are known to support many of the clusters in Punjab.[15] In two specifications, TFPQ is negatively related to employment in large-scale firms in the same industry (col 1 and 2) even though previous research including Greenstone et al. (2010) and Hu et al. (2015) had shown benefits from large-scale firms. In unreported results, we note that TFP is increasing monotonically in firm size, so that the smallest firms (measured through value added) also have the lowest TFPQ. In spite of this, we also find that the localization of small firms is associated with decreased dispersion in TFPQ (Table 5,

[15] Examples of where large vendor segments support clusters include electric fans in Gujrat/Gujranwala and surgical goods in Sialkot.

Table 7 Regression of distortions in capital on the disaggregated location quotient.
Source: Authors' calculations based on CMI Punjab 2005–2006, trimmed data, and directory of industries 2006, 2004, 2002

	Average capital distortions in 2006			Absolute value of capital distortions in 2006		
	(1)	(2)	(3)	(4)	(5)	(6)
Log small firm loc Q (2006)	0.124***			0.0520**		
	(3.207)			(2.529)		
Log medium-size firm loc Q (2006)	−0.0185			0.00646		
	(−0.389)			(0.159)		
Log large firm loc Q (2006)	−0.0369			−0.0477**		
	(−1.291)			(−2.121)		
Log small firm loc Q (2004)		0.109***			0.0722**	
		(2.936)			(2.520)	
Log medium-size firm loc Q (2004)		0.0670			0.0196	
		(1.128)			(0.443)	
Log large firm loc Q (2004)		−0.0527			−0.0417	
		(−1.256)			(−1.306)	
Log small firm loc Q (2002)			0.0937***			0.0389***
			(8.959)			(3.330)
Log medium-size firm loc Q (2002)			−0.0109			0.0139
			(−0.725)			(0.493)
Log large firm loc Q (2002)			0.00787			0.0236
			(0.753)			(1.597)
District FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2692	2681	2285	2672	2681	2284
Adjusted R-squared	0.242	0.250	0.256	0.049	0.052	0.047

Dependent variable is $\ln\left(1+\tau_{K_{si}}\right)$ or $ln\left|\tau_{K_{si}}\right|$
t-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; standard errors clustered at the district level
Regressions have also controlled for status as public enterprise, foreign collaboration, firm size (value added) quartile

cols. 4–6). The signs on the coefficients for localization of medium and large size firms are also mostly significant, but the sign is not consistent across specifications.

The localization of small firms, whether measured in current or lagged values, is associated with firms being on average both output and capital restricted, a finding that is statistically significant in five out of six specifications (Tables 6, cols. 1–3; 7, col 1–3). The localization of small firms is also associated with higher total output and capital distortions when we take the absolute values of the distortions as our dependent variable, and is statistically significant in all specifications (Tables 6, cols. 4–6; 7, cols. 4–6).

4 Conclusions

In this article, we have documented the dispersion in firm-level TFP and the limits of its association with agglomeration across Punjab, Pakistan's largest province economically and population-wise.

The forces of agglomeration have not had a measurable impact on reducing the dispersion in revenue-based total factor productivity nor on eliminating output and capital distortions. It may be that other factors, such as superior quality and large market shares for some firms lead to persistently higher markups within sectors so that no amount of additional competition offered by agglomeration will eliminate it (Atkin et al. 2015). On the other hand, we find a robust correlation between agglomeration and average output-based TFP in addition to reducing the dispersion in TFPQ.

We also find some evidence pointing to agglomerated firms being output and capital constrained. Both of these effects appear to be connected to the localization of small firms. Small firms contribute to the productivity of the cluster, but they themselves are lower in productivity and use relatively little capital. When we mapped the geographic locations of the firms encountering high levels of capital and output distortions, we find that firms in both the well-known Sialkot district (home of important export-oriented clusters and an entrepreneurial hub) in the northeast and surprisingly also the less developed southern districts have the highest shares of firms constrained in both output and capital.

These findings should be considered in the formation of industrial policy to promote special economic zones in Punjab, especially since the development of industrial estates is one of the main items on the agenda of the government's Punjab Industries Sector Plan 2018.

As new data sets become available, we plan to test some of the specific channels through which agglomeration relates to the distribution of firm-level productivities, including the roles of price competition, exit, and unequal distribution of agglomeration spillovers.

Appendix

See Table 8.

Table 8 Kolmogorov–Smirnov tests for the equality of distributions by level of agglomeration

TFPQ				TFPR			
Agglomeration intensity	Low	Somewhat	High		Low	Somewhat	High
Low	x	x	x	Low	x	x	x
Somewhat	0.085	x	x	Somewhat	0.027	x	x
High	0.001	0.000	x	High	0.480	0.000	x
Very high	0.064	0.093	0.000	Very high	0.026	0.776	0.000

References

Andersson M, Lööf H (2011) Agglomeration and productivity: evidence from firm-level data. Ann Reg Sci 46(3):601–620. doi:10.1007/s00168-009-0352-1

Atkin A, Chaudhry A, Chaudry S, Khandelwal A, Verhoogen E (2015) Mark-up and cost dispersion across firms: direct evidence from producer surveys in Pakistan. Am Econ Rev Pap Proc 105(5):537–544. doi:10.1257/aer.p20151050

Au CC, Henderson JV (2006) How migration restrictions limit agglomeration and productivity in China. J Dev Econ 80(2):350–388. doi:10.1016/j.jdeveco.2005.04.002

Banerjee A, Duflo E (2005) Growth Theory through the Lens of Development Economics. In: Durlauf S, Aghion P (eds) Handbook of economic growth, vol 1A. Elsevier Science Ltd, North Holland, pp 473–552

Beaudry C, Schiffauernova A (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate. Res Policy 38:318–337. doi:10.1016/j.respol.2008.11.010

Brooks WJ, Kaboski JP, Li, YA (2016) Growth policy, agglomeration, and (the lack of) competition (No. w22947). National Bureau of Economic Research. doi:10.3386/w22947

Cingano F, Schivardi F (2004) Identifying the sources of local productivity growth. J Eur Econ Assoc 2(4):720–744

Chen K, Irarrazabal A (2015) The role of allocative efficiency in a decade of recovery. Rev Econ Dyn 18(3):523–550. doi:10.1016/j.red.2014.09.008

Combes PP, Magnac T, Robin JM (2004) The dynamics of local employment in France. J Urban Econ 56(2):217–243. doi:10.1016/j.jue.2004.03.009

Combes PP, Duranton G, Gobillon L, Puga D, Roux S (2012) The productivity advantages of large cities: distinguishing agglomeration from firm selection. Econometrica 80(6):2543–2594. doi:10.3982/ECTA8442

Desmet K, Ghani E, O'Connell S, Rossi-Hansberg E (2015) The spatial development of India. J Reg Sci 55(1):10–30

Duranton G (2015) Growing through cities in developing countries. World Bank Res Obs 30(1):39–73

Ehrl P (2013) Agglomeration economies with consistent productivity estimates. Reg Sci Urban Econ 43(5):751–763. doi:10.1016/j.regsciurbeco.2013.06.002

Folta TB, Cooper AC, Baik YS (2006) Geographic cluster size and firm performance. J Bus Ventur 21(2):217–242. doi:10.1016/j.jbusvent.2005.04.005

Glaeser EL, Kallal HD, Scheinkman JA, Schleifer (1992) Growth in cities. J Polit Econ 100(6),1126–1152. http://links.jstor.org/sici?sici=0022-3808%28199212%29100%3A6%3C1126%3AGIC%3E2.0.CO%3B2-D

Graham DJ (2009) Identifying urbanisation and localisation externalities in manufacturing and service industries. Pap Regional Sci 88(1):63–84

Graham D, Melo P, Jiwattanakulpaisarn P, Noland R (2010) Testing for causality between productivity and agglomeration economies. Reg Sci 50(5):935–951. doi:10.1111/j.1467-9787.2010.00676.x

Greenstone M, Hornbeck R, Moretti E (2010) Identifying agglomeration spillovers: evidence from winners and losers of large plant openings. J Polit Econ 118(3):536–598. doi:10.1086/653714

Government of the Punjab (2015) Punjab Industries Sector Plan 2018: Promoting Industrial Development and Investment, Planning & Development Department, supported by International Growth Centre, UK

Hansen ER (1990) Agglomeration economies and industrial decentralization: the wage–productivity trade-offs. J Urban Econ 28(2):140–159. doi:10.1016/0094-1190(90)90047-Q

Hanson GH (1996) Localization economies, vertical organization, and trade. Am Econ Rev 86(5):1266

Hanson GH (1997) Increasing returns, trade and the regional structure of wages. Econ J 113–133. http://www.jstor.org/stable/2235274

Haroon M, Chaudhry AA (2014) Where do new firms locate? The effects of agglomeration on the formation and scale of operations of new firms in Punjab. Economics Discussion Papers, No 2014–2021, Kiel Institute for the World Economy

Haseeb M, Chaudhry TT (2014) Resource misallocation and aggregate productivity in Punjab (no. 1-2014). Centre for Research in Economics and Business, The Lahore School of Economics

Henderson JV (1986) Efficiency of resource usage and city size. J Urban Econ 19(1):47–70. doi:10.1016/0094-1190(86)90030-6

Henderson V, Kuncoro A, Turner M (1995) Industrial development in cities. J Polit Econ 103(5):1067–1090. doi:10.1086/262013

Henderson V (2003) The urbanization process and economic growth: the so-what question. J Econ Growth 8(1):47–71. doi:10.1023/A:1022860800744
Hsieh CT, Klenow PJ (2009) Misallocation and manufacturing TFP in China and India. Q J Econ 124(4):1403–1448. doi:10.1162/qjec.2009.124.4.1403
Hu C, Xu Z, Yashiro N (2015) Agglomeration and productivity in China: firm level evidence. China Econ Rev 33(2015):50–66. doi:10.1016/j.chieco.2015.01.001
Inklaar R, Lashitew AA, Timmer MP (2015) The role of resource misallocation in cross-country differences in manufacturing productivity. Forthcoming in Macroeconomic Dynamics
Levinsohn J, Petrin A (2003) Estimating production functions using inputs to control for unobservables. Rev of Econ Stud 70(2):317–341
Kalemli-Ozcan S, Sørensen BE (2014) Misallocation, property rights, and access to finance: evidence from within and across Africa. In: African Successes: Modernization and Development, vol 3. University of Chicago Press, Chicago
Martin P, Mayer T, Mayneris F (2011) Spatial concentration and plant-level productivity in France. J Urban Econ 69(2):182–195
Nasir M (2013) Agglomeration and firm turnover (No. 2-2013). Centre for Research in Economics and Business, The Lahore School of Economics
Naveed R (2015) Relative factor abundance and relative factor price equality in Punjab. Lahore J Econ 20(1):105–133
Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunication equipment industry. Econometrica 64(6):1263–1297
Pakistan Bureau of Statistics (2009) Census of manufacturing industries 2005–2006: executive summary. Islamabad (industry_mining_and_energy/publications/cmi200506/Executive_Summary.pdf). Accessed 31 August 2015
Shaver JM, Flyer F (2000) Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. Strateg Manag J 21(12):1175–1194. http://www.jstor.org/stable/3094452
Siba E, Söderbom M, Bigsten A, Gebreeyesus M (2012) Enterprise agglomeration, output prices, and physical productivity: firm-level evidence from Ethiopia. WIDER working paper no. 2012/85
Syverson C (2011) What determines productivity? J Econ Lit 49(2):326–365. doi:10.1257/jel.49.2.326
Tirmazee Z (2012) Relative wage variation and industry location within districts of Punjab. Dissertation, Lahore School of Economics, mimeo
Vollrath D (2009) How important are dual economy effects for aggregate productivity? J Dev Econ 88(2):325–334. doi:10.1016/j.jdeveco.2008.03.004
World Bank (2009) World development report 2009: reshaping economic geography. The World Bank, Washington, DC