Increasing productivity dispersion: Evidence from light manufacturing in Brazil

Gonzales-Rocha, Erick and Mendez-Guerra, Carlos

Kwansei Gakuin University, Kyushu University

15 August 2018

Increasing Productivity Dispersion: Evidence from Light Manufacturing in Brazil

Erick Gonzales-Rocha
Kwansei Gakuin University

Carlos Mendez-Guerra
Kyushu University

August 15, 2018

Abstract

Large productivity dispersion within narrowly defined sectors is widely documented. However, across studies, several statistics are used to assess dispersion and there is not enough discussion about differences among them. Using firm-level data for the textile and furniture sectors in Brazil over the 2003-2009 period, we estimated different TFP measures according to four methods: Ordinary Least Squares (OLS for short), the stochastic frontier model of Battese and Coelli (1988, 1992)(STCH for short), the control function approach of Levinsohn and Petrin (2003) (LP for short), and the corrected control function approach of Ackerberg et al. (2015) (ACF for short). Next, we calculated three dispersion statistics: Standard Deviation (SD); Coefficient of Variation (CV); and Interquartile Range (IQR). After confirming the existence of large productivity dispersion within the studied sectors, we analyzed if the dispersion is increasing or decreasing over time. For both sectors, SD and CV convey an increasing productivity dispersion message, but they do so at different rates (CV is seven times higher than SD). On the contrary, IQR suggests less productivity dispersion over time for textiles and mixed results for furnitures. Overall, in terms of characterizing the increasing productivity dispersion, the CV statistic combined with the ACF method define an upper bound while the IQR with LP method define a lower bound. Considering these results, the article underlines that there are non-trivial differences in the use of dispersion statistics. Thus, their use could not be interchangeable and should consider methodological issues, behavior in the tails of the firm productivity distribution, sample sizes and scenarios of divergence/convergence, among others.

JEL Codes: D24, O47, O54,
Keywords: total factor productivity, dispersion, manufacturing firms, Brazil

1 Introduction

There is a growing consensus that Total Factor Productivity (TFP) explains most of the differences in income per capita across countries (Caselli 2005, Hsieh and Klenow 2010, and Pages 2010), and its importance can be traced back at least to the seminal work of Solow (1957).[1] More recently, the increasing availability of firm-level datasets has brought additional light to issues of selection, dispersion and allocative efficiency that occur at the micro level and are shown to have aggregate effects (Bartelsman et al. 2010, Hsieh and Klenow 2009, and Restuccia, and Rogerson 2017). Among those, the existence of dispersion within narrowly defined sectors is well documented (Syverson 2011). However, less attention has been paid to the selection of statistics used to measure productivity dispersion. Across studies, the most commonly used measures are Standard Deviation (SD), Coefficient of Variation (CV) and Interquartile Range (IQR), with the latter having more prominence.[2] These statistics have different methods of calculation, units of measurement and mathematical properties (Ram 2018). But, despite these differences, there is not enough discussion on whether such dispersion statistics could be used interchangeably, convey the same message when applied to the same sample, or have numerical differences in their values and rates of change.

In this context, our paper aims to study the behavior of productivity dispersion statistics within two narrowly defined light manufacturing sectors in Brazil: textiles and furnitures. We used panel data from the World Bank Enterprise Surveys to estimate TFP at the firm level under different estimation frameworks, including: Ordinary Least Squares (OLS for short), the stochastic frontier model of Battese and Coelli (1988, 1992) (STCH for short), the control function approach of Levinsohn and Petrin (2003) (LP for short), and the corrected control function approach of Ackerberg et al. (2015) (ACF for short). Estimated results—across all methods—suggest large productivity differences within each of the two studied sectors. Nevertheless, determining if dispersion is increasing or decreasing over time may vary depending on the used statistic. On the one hand, SD and CV unequivocally suggest that dispersion has increased over time, even though CV does so at six to seven times higher rates of change as compared to SD (CV with the ACF method set and upper bound for higher rates of change). On the other hand, IQR registers a reduction in dispersion for the textile sector and a mixed (increase/decrease) message for the furniture sector (IQR with the LP method set a lower bound for rates of change). We argue that these are non-trivial differences. The differences are based not only on how the statistics are constructed but also on how dispersion occurs in the extremes of the productivity distribution, samples sizes as well as scenarios of convergence or divergence. The rest of the paper is organized as follows. Section 2 briefly describes the methods for estimating firm-level productivity, dispersion statistics and the data source. Section 3 presents the results for each method and statistic. Finally, Section 4 offers some concluding remarks.

[1]It is argued that Tinberger (1942) introduced the concept in a study published in German which might not have been translated to English until 1959 according to Chen (1997). Also, other researchers could have measured TFP before Solow (1957) but it is the latter who integrated the concept into economic theory using calculus as noted by Griliches (1996).

[2]See, for example, Cunningham, Foster, Grim, Haltiwanger, Pabilonia, Stewart and Wolf 2018, Bartelsman and Wolf 2017, Cette, Corde and Lecat 2018, Foster, Grim, Haltiwanger and Wolf 2018, 2017, 2016, and Ito and Lechevalier 2009.

2 Methods and Data

2.1 Productivity Measurement

One common procedure for measuring Total Factor Productivity (TFP) relies on the econometric estimation of an aggregate production function.[3] This function relates outputs to inputs and a residual measure of efficiency. Within this framework, we measure productivity at the two digit level ISIC code in order to have less heterogeneity across outputs and inputs of different firms. In the database for Brazil, the sectors with the highest number of observations were textiles and furnitures.[4] In addition, the textiles sector was selected given its pivotal role in both providing employment for low skilled labor and creating dynamic effects that foster industrialization (Adhikari and Yamamoto 2007, and Keane and te Velde 2008).

In this article, we calculated TFP applying different methods: Ordinary Least Squares (OLS); the stochastic frontier of Battese and Coelli (1988, 1992, STCH); the control function approach of Levinsohn and Petrin (2003, LP); and the corrected control function approach of Ackerberg et al. (2015, ACF). Using a Cobb-Douglas production function, the equations below describe the estimation of TFP in a setting where technology is Hicks neutral, capital (K) and labor (L) are paid the value of their marginal products, and production value added (VA) is the amount of sales (Y) minus the cost of intermediate inputs (M). Additionally, for presentation purposes, variables in logarithms are depicted by lower case letters.

$$Y_{it} = A_{it} K_{it}^{\beta_k} L_{it}^{\beta_l} {M_{it}}^{\beta_m}$$

$$VA = Y - M$$

$$va_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + v_{it} \tag{1}$$

$$tfp_{it} = va_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} \tag{2}$$

The way in which the parameters for capital and labor are estimated in equation 1 varies according to the estimation method.

The Stochastic Frontier Model of Battese and Coelli (1988, 1992, STCH)

Historically, several production functions assume producers to be successful optimizers. However, Kumbhakar and Lovell (2000) highlighted that not all producers use the minimum inputs under a given technology (technical efficiency), allocate inputs in a cost-effective manner given input prices (cost efficiency), or allocate their outputs in a revenue-maximizing manner given output prices (profit efficiency). Thus, there are several instances where cost is not

[3]Another common method is the growth accounting approach. See, for instance, Stigler (1947), Abramovitz (1956), Kendrick (1956, 1961) and Denison (1962, 1967, 1972, 1974). In this method, the elasticities of the inputs are typically calibrated to the national income shares.

[4]Textile and clothing corresponds to ISIC 18 (manufacture of wearing apparel; dressing and dyeing of fur) and furniture to ISIC 36 (manufacture of furniture; manufacturing n.e.c.).

minimized or profit is not maximized. In stochastic frontier models, the analysis is conducted relaxing the successful optimizers assumption and accounting for the differences of firms that operate at a production frontier (using inputs and outputs in the most efficient manner) and those firms that do not. In the frontier approach, error terms are composed of a traditional random noise component (as in various least squares techniques where errors terms are assumed to be symmetrically distributed with zero means) and a new one-sided inefficiency component that aims to capture the effects of inefficiency.[5] The random variation in the operating environment generating inefficiencies that, in turn, create one-sided variations make the production frontiers stochastic.

The specific model is is an extension of Battese and Coelli (1988) that can be applied to panel data. Battese and Coelli (1992) defined the model as:

$$Y_{it} = f(x_{it};\beta)exp(V_{it} - U_{it})$$

and

$$U_{it} = \eta_{it}U_i = \{exp[-\eta(t-T)]\}\,U_i,\ \ t\in\zeta(i); i = 1, 2, ..., N;$$

where Y_{it} is the production function for firm i in period t. Then, $f(x_{it};\beta)$ is an appropriate function with vectors $x_{it};\beta$ of factor inputs and unknown parameters, respectively. The measurement and specification of random errors V_{it} are assumed to be independently and identically distributed $N(0,\sigma_V^2)$. While, one-sided inefficiencies U_{it} are independent and identically distributed non-negative truncations of the $N(\mu,\sigma^2)$ distribution. An unknown scalar parameter is represented byη_{it} and $\zeta(i)$ represents a set of T_i time periods (out of a T total period) for which there are observations for a firm i. The STCH model is suitable for unbalanced panel datasets (Daude and Fernández-Arias 2010 and Ibarrarán, Maffioli and Stucchi 2009).

The Control Function Approach of Levinsohn and Petrin (2003, LP)

The LP method is an extension of the method developed by Olley and Pakes (1996) (OP for short) originally aimed at avoiding potential serial correlations between input levels and unobserved (but, potentially observable or predictable) firm-specific processes that affect the production technology.[6] OP used investment as a variable that could reveal and control for these unobserved shocks. Investment, however, is not easily found in firm level datasets (several times is equal to zero), moreover it is prone to data quality problems and it is likely to be affected by unobservable shocks. Investment also requires the analysis of a dynamic programming problem to verify the strict monotonicity assumption in terms of marginal products of capital and productivity shocks. Hence, LP proposed a method that uses intermediate inputs instead of investment to account for the unobserved shocks. Intermediate inputs are proposed because they are non-dynamic inputs (which makes it easier to satisfy the strict monotonicity assumption), they affect current profits and do not rule

[5] The composed error terms are not symmetric and do not have zero means

[6] In other words, factors or shocks hidden in the error term that might be affecting decisions about capital and labor and, thus, make the estimated parameters inconsistent.

out shocks to the investment demand function (prices or other unobservables), and they are more commonly available in firm level datasets.

The Corrected Control Function Approach of Ackerberg et al. (2015, ACF)

ACF suggested that the OP and LP methods suffer from functional dependence problems, particularly in their first stages, where the estimated coefficient of labor is correctly identified only under a few and specific instances (for instance, when shocks to the price of labor or output occur after levels of investment or intermediate inputs are defined, but before decisions about labor are made).[7] Therefore, ACF proposed a method that uses inverted demand functions conditional on decisions about labor inputs avoiding the functional dependence problem and also relaxing other assumptions. Specifically, the ACF method allows exogenous, serially correlated, unobserved shocks to the price and/or amount of labor while also accommodating labor to have dynamic effects (such as hiring and firing costs). The basic procedure, described by Ackerberg et al. (2015), uses the following value added production function (ω_{it} represent unobserved productivity shocks that are potentially observable or predictable and ε_{it} represent shocks that are not observable):

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it}$$

ACF requires a number of assumptions when intermediates inputs are used as a proxy for unobserved shocks. The first assumption is concerned with the timing of decisions about inputs and establishes the accumulation of capital under the function below (capital and investment are decided in period $t-1$ while labor (l_{it}) can be decided in periods t, or $t-b$, with $0 < b \leq 1$):

$$k_{it} = \kappa(k_{it-1}, i_{it-1})$$

The next assumptions are scalar unobservable and strict monotonicity (in ω_{it}) in the following intermediate input demand function:

$$m_{it} = \widetilde{f}_t(k_{it}, l_{it}, \omega_{it})$$

Subsequently, the above function is inverted $\omega_{it} = \widetilde{f}_t^{-1}(k_{it}, l_{it}, m_{it})$ and introduced into the production function:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \widetilde{f}_t^{-1}(k_{it}, l_{it}, m_{it}) + \varepsilon_{it} = \widetilde{\Phi}_t(k_{it}, l_{it}, m_{it}) + \varepsilon_{it}$$

Then, using semiparametric methods, the first stage moment condition is defined as follows ($\hat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it})$ is estimated, but β_l is not):

$$E\left[\varepsilon_{it} \mid I_{it}\right] = E\left[y_{it} - \hat{\widetilde{\Phi}}_t(k_{it}, l_{it}, m_{it})\right] = 0$$

The production function parameters (including β_l) are estimated in the second stage moment condition:

[7]ACF mentioned that in practice, the functional dependence problem would not be literally observed.

$$E\left[\xi_{it} + \varepsilon_{it} \mid I_{it-1}\right] = 0$$

$$E\left[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g\left(\hat{\tilde{\Phi}}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\right) \mid I_{it-1}\right] = 0$$

Because the second moment requires the estimation of one additional parameter (β_l, besides β_0 and β_k), one more unconditional moment needs to be defined:

$$E\left[\left(y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - \rho\left(\hat{\tilde{\Phi}}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}\right)\right) \begin{pmatrix} 1 \\ k_{it} \\ l_{it-1} \\ \hat{\tilde{\Phi}}_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \end{pmatrix}\right] = 0$$

In this way, the ACF method accounts for shocks related to k_{it} and l_{it} whereas LP allows only shocks to k_{it} and OP does not account for any of those. In any case, we used four methods (OLS, STCH, LP and ACF) to calculate TFP and study the distribution dynamics.

2.2 Productivity Dispersion Measurement

Commonly used statistics to describe productivity dispersion are Standard Deviation (SD), Coefficient of Variation (CV) and Interquartile Range (IQR) which are estimated using the following well-known expressions.[8]

$$SD_t = [(1/n-1)\textstyle\sum_i (tfp_{it} - \overline{tfp_t})^2]^{0.5} \tag{3}$$

$$CV_t = SDTFP_t / \overline{TFP_t} \tag{4}$$

$$IQR_t = Q_3 - Q_1 \tag{5}$$

where tfp_{it} is logarithm of real TFP for firm i in time t, $\overline{tfp_t}$ is the mean value of logarithm of TFP, $SDTFP_t$ is the standard deviation of TFP, $\overline{TFP_t}$ is the mean of real TFP in year t and $Q_3 - Q_1$ is the difference between the third and the first quartile.

Subsequently, we estimated growth rates (GR) and rates of change (ROC)[9] as

$$GR_{SD} = (SD_{t+1} - SD_t)/SD_t \tag{6}$$

[8] The CV coefficient is based on real values as log-transformed data require a variation in the formula for CV to be correctly estimated (Canchola et al., 2017).

[9] By using rates of change we aim to analyze the behavior of dispersion statistics not only in terms of their values but also in terms of their rates of change over time to see if there are different interpretations.

$$ROC_{sd} = (sd_{t+1} - sd_t)/((t+1) - t) \tag{7}$$

where sd stands for logarithm of SD. GR and ROC can be calculated for CV and IQR in a similar way (just replace sd for cv or iqr). Ram (2018) argued that methods of calculation, units of measurement and mathematical properties among these measures are different.[10] To study the differences, a direct comparison of linear rates of change in the previous measures present problems because linear trends are unit dependent and the measurement units of SD, CV and IQR are largely different. For that reason, logarithmic rates of change are estimated so that comparisons can be made. Negative rates will indicate a reduction in productivity dispersion (that is, convergence) while positive rates would suggest an increase in dispersion (that is, divergence).

Among the previous measures, IQR seems to be more used because it is easier to interpret and it would be more robust to outliers (see Cunningham et al. 2018, Bartelsman and Wolf 2017, or Foster et al. 2018). However, we argue that in narrowly defined sectors within already small sample sizes, IQR may leave aside valuable observations points. Moreover, it is important to look at the entire distribution as there is considerable dispersion within the upper and lower tails.[11]

2.3 Data

Data comes from the World Bank's Enterprise Surveys Project. This project is well known for collecting firm-level data across countries using a common methodology, standardized questionnaires and a stratified random sample that covers the non-agricultural economy of main cities and nearby business areas. In addition to adequately covering different cities, sectors and firm sizes, the dataset also provides a wide arrange of variables that allow the calculation of firm performance indicators such as productivity. One limitation, however, is that the project only focuses on formal firms with more than 5 employees. Thus, considering that about 40 percent of the Brazilian economy is informal (Schneider et al. 2011), economy-wide generalizations should be taken with caution.

Using data for the years 2003 and 2009, a balanced panel dataset was constructed for each sector. As expected, this construction implied a large reduction in the sample size. After structuring the panel dataset, the data cleaning process included a post-estimation of outliers procedure for all main variables.[12] Only 48 observations in textiles sector and 62 observations in the furniture sector had the required variables to compute TFP in both years. Table 1 summarizes the datasets that resulted from this construction and cleaning process.

[10] Bartelsman and Wolf (2017) noted that dispersion calculated from value-added measures of production tends to be higher when compared to gross output measures.

[11] Cunningham et al. (2018) argued that the behavior between the tails could be different and noted the important implications of the dynamics of the most productive firms for the whole distribution.

[12] We computed several residual statistics to identify outliers and conflicting observations.

Table 1: Summary Statistics - Textiles and Furnitures

		Textiles					Furnitures			
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Sales 2003	48	11,400,000	35,700,000	150,847	181,000,000	62	11,300,000	28,000,000	151,062	160,000,000
Sales 2009	48	20,200,000	71,400,000	6,000	420,000,000	62	13,900,000	33,300,000	2,400	212,000,000
Labor 2003	48	2,117,498	7,368,616	65,918	41,500,000	62	1,605,594	4,093,550	48,065	27,400,000
Labor 2009	48	3,880,581	15,100,000	1,500	93,000,000	62	1,807,292	4,449,864	1,200	29,000,000
Capital 2003	48	895,799	2,850,071	1,285	16,000,000	62	1,649,194	5,261,774	2,087	39,100,000
Capital 2009	48	1,755,485	6,459,338	1,200	43,000,000	62	2,160,630	6,698,165	220	50,000,000
Int. Inputs 2003	48	5,404,904	18,300,000	3,131	98,300,000	62	4,862,486	10,700,000	15,654	59,600,000
Int. Inputs 2009	48	7,052,957	22,900,000	1,440	120,000,000	62	6,631,881	15,100,000	900	82,000,000
Electricity 2003	48	64,582	262,431	571	1,821,798	62	104,235	234,240	14	1,598,513
Electricity 2009	48	176,889	732,142	72	5,000,000	62	145,859	380,130	60	2,400,000

Notes: Monetary values are expressed in 2009 local currency units and were deflated using the National Index of Consumer Prices (INPC in Portuguese) obtained from the Brazilian Institute of Geography and Statistics (IBGE in Portuguese). Statistics for balanced panel data only (firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys

3 Results

3.1 Productivity

Table 2 presents the results for the estimation of the production function. Overall, the estimated parameters are within similar ranges despite using different methods.[13] The differences in the results for the method of ACF (columns 4 and 8) could be explained because of the strict use of balanced panel data.

It is worth noting that the estimated parameters in Table 2 will, in some instances, produce negative values of productivity which, according to Fernández-Arias and Rodríguez-Apolinar (2016), are not uncommon to find in the LAC region. They also contested the hypothesis about the existence of these negative values because of measurement errors in the factor of labor (quality of education, specifically) and further showed that a possible bias in human capital accumulation will actually strengthen the productivity shortfall in the region. Using the results of the production function estimation, we calculated TFP. Summary statistics for these values are presented in Tables 3 and 4. It can be observed that the OLS and STCH methods consistently report similar values. ACF and LP provide TFP values in a different order of magnitude but with consistency within the methods.

Although these summary statistics are informative, they have limitations at the moment of understanding the behavior of TFP. In terms of central tendency indicators, median values suggest the textiles sector is experiencing a productivity deterioration over time, independently of the method used to calculate TFP. However, mean values for the same sector suggest slight improvements overall. Meanwhile, for the furniture sector, both median and mean values suggest improvements of TFP over time.

[13]Estimations were also conducted for the case of OLS considering the information available for each year (2003 and 2009) separately as well as including a year dummy variable. Results of the estimated coefficients proved to be robust to these modifications and are not included for the sake of brevity. The methods of ACF and LP require at least two points in time (panel data) to be executed, so no further explorations were done. Additional results exploring other considerations are presented in Appendix A.

Table 2: Production Function Estimation

	Textiles				Furnitures			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	OLS	STCH	LP	ACF	OLS	STCH	LP	ACF
Capital	0.147**	0.147***	0.207*	-0.0264	0.177***	0.158***	0.0749	-0.307
	(0.0619)	(0.0563)	(0.119)	(0.213)	(0.0484)	(0.0409)	(0.126)	(0.397)
Labor	0.885***	0.885***	0.723***	1.079***	0.876***	0.886***	0.768***	1.455***
	(0.0610)	(0.0637)	(0.103)	(0.302)	(0.0704)	(0.0498)	(0.103)	(0.506)
Constant	0.727	0.730			0.392	266.9***		
	(0.534)	(0.543)			(0.556)	(0.466)		
Observations	96	96	96	48	124	124	124	62
R-squared	0.866				0.895			
Number of panelid		48				62		

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The coefficients were estimated using the balanced panel data (only firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009). For the textiles sector, 48 firms comply with this condition and 62 firms for the furniture sector. Regarding the acronyms of the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Table 3: Detailed TFP Summary Statistics - Textiles

	ln TFP								TFP							
stats	ols		stch		lp		acf		OLS		STCH		LP		ACF	
	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009
Mean	0.7238	0.7302	0.7242	0.7307	2.1148	2.1067	0.2661	0.3248	2.6454	3.0442	2.6464	3.0458	10.7674	11.9972	1.7075	2.1978
Min	-0.4223	-0.4825	-0.4216	-0.4818	0.8579	0.3816	-1.0217	-0.7731	0.6555	0.6173	0.6560	0.6177	2.3581	1.4646	0.3600	0.4616
Median	0.7211	0.6063	0.7214	0.6071	2.1362	1.9652	0.2257	0.1420	2.0571	1.8337	2.0577	1.8353	8.4687	7.1362	1.2533	1.1526
Max	2.3769	3.0903	2.3771	3.0909	3.7925	4.5046	1.9961	2.9229	10.7711	21.9829	10.7740	21.9959	44.3692	90.4353	7.3606	18.5958
SD	0.6743	0.7249	0.6743	0.7249	0.7008	0.7584	0.6985	0.7748	2.2450	4.2587	2.2459	4.2622	9.2640	15.9449	1.5281	3.5850
CV	0.9317	0.9927	0.9311	0.9921	0.3314	0.3600	2.6248	2.3852	0.8486	1.3990	0.8487	1.3993	0.8604	1.3290	0.8949	1.6312
IQR	0.7748	0.5804	0.7747	0.5804	1.0232	0.6220	0.9477	0.6892	1.4590	1.0473	1.4593	1.0478	8.4587	4.5867	1.1192	0.8651
Skewness	0.6038	1.6597	0.6039	1.6604	0.4584	1.0941	0.6042	1.7746	2.1052	3.5271	2.1056	3.5277	2.1716	3.6006	2.2062	3.8240
Kurtosis	2.8994	5.9985	2.8999	6.0009	2.7735	5.0508	3.0162	6.3930	7.0711	14.9573	7.0722	14.9608	7.6344	16.2838	7.4955	17.0390
N	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48

Notes: TFP values were calculated using the coefficients estimated for the balanced panel data (only firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009). Variables in logarithms are represented by lowercase letters in the left side of the table. Regarding the acronyms of the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015). Among the TFP summary statistics, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range (the rest are self-explanatory).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Table 4: Detailed TFP Summary Statistics - Furnitures

	ln TFP								TFP							
stats	ols		stch		lp		acf		OLS		STCH		LP		ACF	
	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009
Mean	0.3200	0.4634	0.4250	0.5774	3.0057	3.1170	-1.3122	-0.7700	1.7022	2.2177	1.9019	2.4714	27.7540	31.4847	0.5086	1.4371
Min	-1.2502	-0.8941	-1.1492	-0.7938	1.3996	1.4350	-3.0222	-2.2549	0.2865	0.4090	0.3169	0.4521	4.0536	4.1995	0.0487	0.1049
Median	0.2980	0.3492	0.4206	0.4775	2.9523	3.1330	-1.4033	-0.9338	1.3472	1.4180	1.5229	1.6120	19.1596	22.9440	0.2459	0.3931
Max	2.0427	2.8996	2.1796	2.9628	4.8805	5.0867	1.7122	3.8890	7.7112	18.1667	8.8430	19.3527	131.6947	161.8498	5.5411	48.8604
SD	0.6419	0.7236	0.6482	0.7185	0.7942	0.8111	1.0100	1.0236	1.3017	2.6886	1.4897	2.9343	24.5432	29.5626	0.8743	6.1891
CV	2.0057	1.5617	1.5252	1.2445	0.2642	0.2602	-0.7697	-1.3294	0.7647	1.2124	0.7833	1.1873	0.8843	0.9390	1.7189	4.3065
IQR	0.9491	0.8654	0.9669	0.8439	1.2456	1.0967	1.2391	1.0401	1.2858	1.2868	1.4729	1.3708	26.3602	26.0717	0.3190	0.4482
Skewness	0.1879	1.0482	0.2255	1.0508	0.2415	0.1895	0.8107	1.8397	2.4574	3.9953	2.5198	3.8057	1.8810	2.2311	4.0584	7.4383
Kurtosis	3.2282	4.3489	3.2481	4.3323	2.3608	2.6708	3.5430	8.7438	10.4709	21.9839	10.7788	19.9914	7.0615	8.6918	21.2710	57.4445
N	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62

Notes: TFP values were calculated using the coefficients estimated for the balanced panel data (only firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009). Variables in logarithms are represented by lowercase letters in the left side of the table. Regarding the acronyms of the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015). Among the TFP summary statistics, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range (the rest are self-explanatory).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

3.2 Productivity Dispersion

Dispersion indicators such as SD, CV or IQR suggest that there is indeed significant dispersion in productivity within relatively narrowly defined sectors (textiles and furnitures). Taking advantage of our panel data, we also analyzed whether this dispersion is increasing (divergence) or decreasing (convergence) over time by calculating growth rates and rates of change for SD, CV and IQR.

The results in tables 5 and 6 suggest that while SD and CV unequivocally register increases in productivity dispersion for both sectors and across all methods, IQR indicates decreases in dispersion for textiles and mixed evidence for furnitures.[14] For the increasing productivity dispersion message of SD and CV, however, there are numerical differences, CV registers higher (six to seven times) rates of change in dispersion as compared to SD. Linking these two indicators to methods of productivity estimation, ACF yields the highest dispersion. For the case of IQR, a decrease of dispersion in textiles could be explained by the fact that IQR does not consider the dispersion in the tails of the distribution. We argue that IQR is an indicator that should be used with caution, in particular with small sample sizes because it might leave out valuable observations. For instance, we confirmed previous findings of large dispersion within the tails of the productivity distribution. Within this large dispersion in the tails, it is the most productive firms (fourth quartile) that have even much higher dispersion levels and rates as compared to the least productive firms (first quartile).[15] As a matter of fact, in the tails, all dispersion statistics, including IQR, display positive rates of change in dispersion overall (in particular for the fourth quartile and the method of ACF). It could be said that in terms of measuring rates of change in productivity dispersion, we find an upper bound set by the CV statistic and the ACF method and a lower bound set by

[14]We focus on ROC because they have the same units and allow comparisons across dispersion statistics.
[15]See Appendix B where we calculated dispersion statistics and rates of change for both tails.

IQR and the LP method.

Table 5: Growth Rates and Rates of Change - Textiles

Method	Growth Rates (GR)			Rates of Change (ROC)			Exponential ROC		
	GR_{SD}	GR_{CV}	GR_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$
OLS	0.0749	0.6485	-0.2822	0.0120	0.0833	-0.0553	0.0121	0.0869	-0.0538
STCH	0.0750	0.6489	-0.2820	0.0120	0.0833	-0.0552	0.0121	0.0869	-0.0537
LP	0.0821	0.5447	-0.4577	0.0132	0.0725	-0.1020	0.0132	0.0752	-0.0970
ACF	0.1092	0.8227	-0.2270	0.0173	0.1001	-0.0429	0.0174	0.1052	-0.0420

Notes: The exponential Rate of Change responds to the following equation $eROC = e^{ROC} - 1$. For ROC and eROC, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range. Regarding other acronyms in the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Table 6: Growth Rates and Rates of Change - Furnitures

Method	Growth Rates (GR)			Rates of Change (ROC)			Exponential ROC		
	GR_{SD}	GR_{CV}	GR_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$
OLS	0.1274	0.5854	0.0008	0.0200	0.0768	0.0001	0.0202	0.0798	0.0001
STCH	0.1164	0.5158	-0.0693	0.0184	0.0693	-0.0120	0.0185	0.0718	-0.0119
LP	0.0213	0.0618	-0.0109	0.0035	0.0100	-0.0018	0.0035	0.0100	-0.0018
ACF	0.0135	1.5053	0.4048	0.0022	0.1531	0.0566	0.0022	0.1654	0.0583

Notes: The exponential Rate of Change responds to the following equation $eROC = e^{ROC} - 1$. For ROC and eROC, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range. Regarding other acronyms in the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

4 Concluding Remarks

While there is ample evidence about large productivity dispersion within narrowly defined sectors, the discussion about the use of different statistics to document this dispersion is less robust. Across studies, three dispersion statistics are commonly applied: Standard Deviation (SD), Coefficient of Variation (CV) and Interquartile Range (IQR). Sometimes, these statistics are used individually, interchangeably or in parallel even though they have different methods of calculation, units of measurement and mathematical properties. Therefore, we used Brazilian data for the textile and furniture sectors over the period 2003-2009 to estimate TFP using several estimation methods: Ordinary Least Squares (OLS for short), the stochastic frontier model of Battese and Coelli (1988, 1992)(STCH for short), the control function approach of Levinsohn and Petrin (2003)(LP for short), and the corrected control function approach of Ackerberg et al. (2015)(ACF for short). Then, we calculated their dispersion statistics and rates of change over time to allow comparison among them. Our goal was to evaluate differences among dispersion statistics in terms of their use, numerical values and conclusions when they are applied to the same samples under different productivity estimation methods.

Results suggest that there are non-trivial differences among statistics when it comes to evaluate productivity dispersion at the firm level. While SD and CV suggest an increase of productivity dispersion over time for both sectors (with CV showing six to seven times higher rates of change than SD), IQR suggests a reduction of dispersion in textiles and mixed findings for furnitures. IQR does not consider extremes values, but when studying dispersion only within the tails of the productivity distribution, we found that all dispersion statistics overwhelmingly suggest the existence of even higher levels and rates of dispersion. This may suggest that the study of productivity dispersion should go beyond the analysis of specific statistics at some points in time and into the analysis of the dynamics for the entire distribution given the heterogeneous behavior of rims, complex patterns of convergence and the possible formation of multiple convergence clubs.[16] Overall, the CV statistics with the ACF method set an upper bound and IQR with LP a lower bound in terms of registering rates of change in productivity dispersion. If dispersion statistics are to be analyzed, they are not interchangeable and its use ought to consider the behavior in the tails of the firm productivity distribution (particularly with small samples sizes) and the methods of TFP estimation. Future research could add to the analysis further differences under scenarios of converge or divergence, non-manufacturing sectors, countries' income levels and employment weighted productivity dispersion statistics.

References

Abramovitz, M. (1956). Resource and output trends in the United States since 1870. *The American Economic Review*, 46(2):5–23.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Adhikari, R. and Weeratunge, C. (2007). Textiles and clothing in South Asia: Current status and future potential. *South Asia Economic Journal*, 8(2):171–203.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2010). Cross-country and within-country differences in the business climate. *International Journal of Industrial Organization*, 28(4):368–371.

Bartelsman, E. J. and Wolf, Z. (2017). Measuring Productivity Dispersion. Tinbergen Institute Discussion Papers 17-033/VI, Tinbergen Institute.

[16]See, for example, Mendez-Guerra and Gonzales-Rocha (2018).

Battese, G. E. and Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 38(3):387 – 399.

Battese, G. E. and Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis*, 3(1):153–169.

Canchola, J., Tang, S., Hemyari, P., Paxinos, E., and Marins, E. (2017). Correct use of percent coefficient of variation (cv) formula for log-transformed data. *MOJ Proteomics Bioinform*, 6(4).

Caselli, F. (2005). Chapter 9 accounting for cross-country income differences. volume 1 of *Handbook of Economic Growth*, pages 679 – 741. Elsevier.

Cette, G., Corde, S., and Lecat, R. (2018). Firm-level productivity dispersion and convergence. *Economics Letters*, 166(C):76–78.

Chen, E. K. (1997). The total factor productivity debate: Determinants of economic growth in East Asia. *Asian-Pacific Economic Literature*, 11(1):18–38.

Cunningham, C., Foster, L., Grim, C., Haltiwanger, J., Pabilonia, S., Stewart, J., and Wolf, Z. (2018). Dispersion in dispersion: Measuring establishment-level differences in productivity. Working papers, U.S. Census Bureau, Center for Economic Studies.

Daude, C. and Fernández-Arias, E. (2010). On the role of productivity and factor accumulation in economic development in Latin America. IDB Working Paper Series 155, Inter-American Development Bank.

Denison, E. F. (1962). *The sources of economic growth in the United States and the alternatives before us*. New York: Committee for Economic Development.

Denison, E. F. (1967). Sources of postwar growth in nine western countries. *The American Economic Review*, 57(2):325–332.

Denison, E. F. (1972). Classification of sources of growth. *Review of Income and Wealth*, 18(1):1–25.

Denison, E. F. (1974). *Accounting for United States economic growth, 1929-1969*. Washington D.C.: Brookings Institution.

Fernández-Arias, E. and Rodríguez-Apolinar, S. (2016). The productivity gap in Latin America: Lessons from 50 years of development. IDB Working Paper Series 692, Inter-American Development Bank.

Foster, L., Grim, C., Haltiwanger, J., and Wolf, Z. (2016). Firm-level dispersion in productivity: Is the devil in the details? *American Economic Review*, 106(5):95–98.

Foster, L., Grim, C., Haltiwanger, J. C., and Wolf, Z. (2018). Innovation, productivity dispersion, and productivity growth. Working Paper 24420, National Bureau of Economic Research.

Foster, L. S., Grim, C. A., Haltiwanger, J., and Wolf, Z. (2017). Macro and micro dynamics of productivity: From devilish details to insights. Working Paper 23666, National Bureau of Economic Research.

Griliches, Z. (1996). The discovery of the residual: A historical note. *Journal of Economic Literature*, 34(3):1324–1330.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2010). Development accounting. *American Economic Journal: Macroeconomics*, 2(1):207–23.

Ibarrarán, P., Maffioli, A., and Stucchi, R. (2009). Sme policy and firms productivity in Latin America. IZA Discussion Papers 4486, Institute for the Study of Labor (IZA).

Ito, K. and Lechevalier, S. (2009). The evolution of the productivity dispersion of firms: a reevaluation of its determinants in the case of Japan. *Review of World Economics*, 145(3):405–429.

Keane, J. and te Velde, D. W. (2008). The role of textile and clothing industries in growth and development strategies. Research reports and studies, Overseas Development Institute.

Kendrick, J. W. (1956). Productivity trends: Capital and labor. *The Review of Economics and Statistics*, 38(3):248–257.

Kendrick, J. W. (1961). *Productivity Trends in the United States*. Princeton: Princeton University Press.

Kumbhakar, S. C. and Lovell, C. K. (2000). *Stochastic Frontier Analysis*. Cambridge UK: Cambridge University Press.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Mendez-Guerra, C. and Gonzales-Rocha, E. (2018). A comparison of TFP estimates via distribution dynamics: Evidence from light manufacturing firms in Brazil. MPRA Paper 87723, University Library of Munich, Germany.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Pages, C. (2010). *The Age of Productivity*. Palgrave Macmillan US.

Ram, R. (2018). Comparison of cross-country measures of sigma-convergence in per-capita income 1960-2010. *Applied Economics Letters*, 25(14):1010–1014.

Restuccia, D. and Rogerson, R. (2017). The causes and costs of misallocation. *Journal of Economic Perspectives*, 31(3):151–74.

Schneider, F., Buehn, A., and Montenegro, C. (2010). Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. Policy Research Working Paper 5356, The World Bank.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3):312–320.

Stigler, G. J. (1947). *Trends in Output and Employment*. New York: National Bureau of Economic Research.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Tinbergen, J. (1942). Zur theorie der langfirstigen wirtschaftsentwicklung. *Weltwirts. Archiv*, (1):511–49.

Appendix

A Additional Results

Tables here present results when the coefficients for capital and labor were estimated using all the information available for each sector (tables presented in the main text were estimated using only the balanced panel data: firms with information for all main variables in the years 2003 and 2009). The balanced panel dataset is used in the main text because the study of distribution dynamics focuses solely on firms contained in it. Below, outputs for summary statistics, the production function estimation and detailed TFP summary statistics are reported.

Summary statistics suggest a substantial reduction in the number of observations for each sector in the second round of the survey. In addition, a large number of observations is lost because of missing data and firms that were not surveyed in both years. In the main text, table 1 is equivalent.

Table 7: Summary Statistics (all data) - Textiles and Furnitures

	Textiles					Furnitures				
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Sales 2003	425	5,953,067	30,100,000	46,623	494,000,000	305	9,494,395	27,700,000	88,159	261,000,000
Sales 2009	153	11,200,000	50,600,000	1,200	420,000,000	161	9,234,948	29,400,000	1,200	212,000,000
Labor 2003	424	1,007,923	4,808,694	1,154	80,200,000	305	1,259,605	3,523,766	12,104	41,700,000
Labor 2009	140	2,302,671	12,000,000	500	95,500,000	152	1,363,283	4,539,689	370	40,000,000
Capital 2003	408	981,848	10,100,000	164	200,000,000	291	1,101,871	3,334,586	2,087	39,100,000
Capital 2009	122	1,713,296	9,966,734	80	101,000,000	126	1,165,737	4,797,700	170	50,000,000
Int. Inputs 2003	391	2,456,418	11,500,000	3,131	172,000,000	303	4,294,554	11,600,000	15,654	95,100,000
Int. Inputs 2009	138	3,915,659	15,200,000	200	120,000,000	142	3,915,393	11,300,000	100	82,000,000
Electricity 2003	415	68,227	357,922	146	4,772,132	302	136,468	822,907	9	13,700,000
Electricity 2009	149	118,190	666,853	50	6,374,000	152	85,878	272,096	0	2,400,000

Notes: Monetary values are expressed in 2009 local currency units and were deflated using the National Index of Consumer Prices (INPC in Portuguese) obtained from the Brazilian Institute of Geography and Statistics (IBGE in Portuguese).
Source: Authors' elaboration.

For the production function estimation, the coefficients of capital and labor do not significantly change when using all the information available as shown in the table below (particularly for the furniture sector). Table 2 in the main text is equivalent.

Table 8: Production Function Estimation (all data)

	Textiles				Furnitures			
	(1)	(2)	(7)	(8)	(9)	(10)	(15)	(16)
VARIABLES	OLS	STCH	LP	ACF	OLS	STCH	LP	ACF
Capital	0.158***	0.160***	0.279***	-0.117	0.216***	0.209***	0.0839	-0.316
	(0.0293)	(0.0253)	(0.0952)	(0.277)	(0.0322)	(0.0272)	(0.109)	(0.441)
Labor	0.872***	0.869***	0.758***	1.252***	0.802***	0.804***	0.727***	1.467**
	(0.0346)	(0.0314)	(0.0469)	(0.405)	(0.0452)	(0.0315)	(0.0533)	(0.572)
Constant	0.653**	0.803			0.898**	313.4***		
	(0.320)	(2.376)			(0.384)	(0.281)		
Observations	498	498	490	56	411	411	404	67
R-squared	0.793				0.849			
Number of panelid		445		56		349		67

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The coefficients were estimated using all the information available for each sector. OLS stands for Ordinary Least Squares, STCH for the non-parametric estimation, LP for Levinshon and Petrin and ACF for Ackerberg, Caves and Frazer.
Source: Authors' elaboration.

Tables below show detailed TFP summary statistics only for firms with information in years 2003 and 2009, calculated using coefficients of capital and labor estimated with all the information available. For the textiles sector, table 9 reports no major changes in the results for the OLS and STCH methods as compared to results obtained using only the balanced panel dataset. However, the LP and ACF methods register changes as reflected in central tendency indicators. This may suggest they are more sensitive to extreme values or the inclusion/elimination of firms, particularly in small samples. On the contrary, table 10for the furniture sector indicates changes in central tendency indicators for the OLS and STCH methods while the LP and ACF do not drastically change. These are also additinal reasons for using only the balanced panel dataset in the main text. Tables 3 and 4 are equivalent.

Table 9: Detailed TFP Summary Statistics (all data) - Textiles

	log TFP								TFP							
stats	ols		stch		lp		acf		OLS		STCH		LP		ACF	
	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009
Mean	0.7475	0.7746	0.7648	0.7915	0.8092	0.7906	-0.936	-0.8139	2.9549	3.1423	3.007	3.1948	3.1965	3.151	0.5704	0.7415
Min	-1.065	-0.4739	-1.0435	-0.4643	-0.814	-0.6283	-3.4071	-2.4534	0.3447	0.6226	0.3522	0.6286	0.4431	0.5335	0.0331	0.086
Median	0.7634	0.6104	0.7844	0.6241	0.7812	0.6656	-0.9206	-1.0411	2.1456	1.8412	2.1911	1.8666	2.184	1.9458	0.3983	0.3531
Max	2.7347	3.1118	2.7527	3.1251	2.8247	3.036	0.9538	1.9618	15.4046	22.461	15.6856	22.7614	16.8552	20.8222	2.5956	7.1123
SD	0.7994	0.728	0.7995	0.7281	0.8113	0.7478	0.8741	0.8515	2.9239	4.193	2.9763	4.2563	3.2394	3.7783	0.5704	1.2493
IQR	0.8496	0.6218	0.8542	0.6099	1.0176	0.614	0.9847	0.8264	1.6381	1.1686	1.6782	1.1629	2.1582	1.1312	0.3796	0.3505
CV	1.0695	0.9398	1.0453	0.9199	1.0026	0.9458	-0.9338	-1.0462	0.9895	1.3344	0.9898	1.3323	1.0134	1.1991	1	1.6849
Skewness	0.3002	1.4853	0.3013	1.4799	0.4399	1.1353	-0.0307	1.2925	2.3586	3.5291	2.3601	3.5235	2.2953	3.0961	2.078	4.0331
Kurtosis	3.1481	5.3226	3.145	5.3078	2.8744	4.2028	3.443	5.1431	8.8565	15.4127	8.8671	15.3727	8.4984	12.91	6.7708	19.2928
N	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53

Notes: TFP values were calculated with the coefficients estimated using all the information available for each sector. Variables in logarithms are represented by lowercase letters in the left side of the table. OLS stands for Ordinary Least Squares, STCH for Stochastic, LP for Levinsohn and Petrin, and ACF for Ackerberg, Caves and Frazer. Among the TFP summary statistics, Std. Dev. stands for standard deviation, IQR for interquartile range and CV for coefficient of variation (the rest are self-explanatory).
Source: Authors' elaboration.

Table 10: Detailed TFP Summary Statistics (all data) - Furnitures

	log TFP								TFP							
stats	ols		stch		lp		acf		OLS		STCH		LP		ACF	
	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009	2003	2009
Mean	0.8196	0.9164	0.8797	0.9792	3.4255	3.5145	-1.3597	-0.8096	2.7973	3.4514	2.9762	3.6638	42.8212	47.8985	0.4918	1.4324
Min	-0.7473	-0.4544	-0.6888	-0.3980	1.8181	1.6748	-3.0907	-2.3395	0.4736	0.6348	0.5022	0.6717	6.1600	5.3376	0.0455	0.0964
Median	0.8170	0.8217	0.8665	0.8894	3.3495	3.5562	-1.4497	-0.9637	2.2640	2.2747	2.3787	2.4341	28.5120	35.0334	0.2347	0.3815
Max	2.4680	3.2205	2.5399	3.2573	5.2802	5.5392	1.6968	3.9096	11.7988	25.0418	12.6788	25.9806	196.4058	254.4846	5.4565	49.8778
SD	0.6506	0.7326	0.6527	0.7307	0.8152	0.8489	1.0203	1.0353	2.0205	3.8404	2.1658	4.0259	38.1145	45.7397	0.8579	6.3160
IQR	0.9463	0.8947	0.9650	0.8990	1.2939	1.1799	1.2414	1.0349	2.2415	2.1101	2.4135	2.2802	42.5022	39.9518	0.3051	0.4283
CV	0.7938	0.7994	0.7420	0.7462	0.2380	0.2415	-0.7504	-1.2789	0.7223	1.1127	0.7277	1.0988	0.8901	0.9549	1.7443	4.4094
Skewness	0.0279	0.8486	0.0389	0.8395	0.2196	0.0871	0.8095	1.8361	2.1143	3.5162	2.1373	3.4279	1.7511	2.2461	4.0893	7.4599
Kurtosis	2.9525	3.7169	2.9547	3.6777	2.2856	2.6391	3.5435	8.7633	8.8733	18.0800	8.9944	17.2712	6.2442	9.0308	21.5625	57.6768
N	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62

Notes: TFP values were calculated with the coefficients estimated using all the information available for each sector. Variables in logarithms are represented by lowercase letters in the left side of the table. OLS stands for Ordinary Least Squares, STCH for Stochastic, LP for Levinsohn and Petrin, and ACF for Ackerberg, Caves and Frazer. Among the TFP summary statistics, Std. Dev. stands for standard deviation, IQR for interquartile range and CV for coefficient of variation (the rest are self-explanatory).
Source: Authors' elaboration.

B Productivity Dispersion in the Distribution's Tails

We estimated productivity dispersion statistics for the 25th (first quartile) and 75th (fourth quartile) percentile of each distribution for all TFP estimation methods.

Table 11: Detailed TFP Summary Statistics for the Tails of the Distribution - Textiles

	OLS				STCH				LP				ACF			
Stats	2003		2009		2003		2009		2003		2009		2003		2009	
	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th
Mean	0.9594	5.6055	1.0885	7.4627	0.9597	5.6076	1.0890	7.4671	3.6868	22.7329	3.9198	29.1645	0.5915	3.7111	0.7091	5.7085
Min	0.6555	2.7206	0.6173	2.4815	0.6560	2.7228	0.6177	2.4823	2.3581	13.3840	1.4646	9.9923	0.3600	1.8626	0.4616	1.8167
Median	0.9340	4.7176	1.1175	3.7100	0.9342	4.7201	1.1183	3.7106	3.8838	15.8784	3.9569	15.4335	0.6399	3.1652	0.7291	2.3630
Max	1.1765	10.7711	1.2911	21.9829	1.1768	10.7740	1.2921	21.9959	4.7388	44.3692	5.3108	90.4353	0.6973	7.3606	0.8588	18.5958
SD	0.1610	2.7277	0.2022	6.9659	0.1610	2.7292	0.2022	6.9723	0.7154	11.3918	1.0282	25.3941	0.1066	1.8988	0.1399	6.0583
CV	0.1678	0.4866	0.1857	0.9334	0.1678	0.4867	0.1857	0.9337	0.1940	0.5011	0.2623	0.8707	0.1802	0.5117	0.1972	1.0613
IQR	0.1971	4.3101	0.2702	6.9437	0.1974	4.3137	0.2704	6.9472	1.1599	15.4098	1.2577	22.5510	0.1581	3.1906	0.2617	4.5869
skewness	-0.4284	0.6841	-1.0138	1.3212	-0.4276	0.6842	-1.0147	1.3215	-0.3438	0.9103	-0.9386	1.5195	-0.9645	0.7380	-0.3392	1.4909
kurtosis	2.4166	2.1275	3.3688	3.1989	2.4141	2.1270	3.3703	3.1991	2.0208	2.3458	3.8136	4.0038	2.7200	2.1524	1.6778	3.5873
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12

Notes: TFP values were calculated using the coefficients estimated for the balanced panel data (only firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009). The 25th indicates that data is provided only for the first quartile and 75th for the fourth quartile only. Regarding the acronyms of the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015). Among the TFP summary statistics, SD stands for standard deviation, IQR for interquartile range and CV for coefficient of variation (the rest are self-explanatory).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Table 12: Detailed TFP Summary Statistics for the Tails of the Distribution - Furnitures

	OLS				STCH				LP				ACF			
Stats	2003		2009		2003		2009		2003		2009		2003		2009	
	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th	25th	75th
Mean	0.6441	3.3353	0.7484	5.0597	0.7114	3.7601	0.8443	5.6088	8.1429	62.2579	8.9753	70.3704	0.0938	1.3997	0.1818	4.6208
Min	0.2865	2.0979	0.4090	2.2220	0.3169	2.3766	0.4521	2.4049	4.0536	37.0110	4.1995	39.1444	0.0487	0.4491	0.1049	0.6931
Median	0.6880	2.7051	0.7733	3.5288	0.7619	3.0679	0.8555	3.9041	8.3163	51.5461	9.4787	56.2534	0.0940	0.8369	0.1993	1.1868
Max	0.8121	7.7112	0.9352	18.1667	0.9037	8.8430	1.0342	19.3527	10.6508	131.6947	13.0727	161.8498	0.1301	5.5411	0.2450	48.8604
SD	0.1513	1.5761	0.1347	4.1454	0.1653	1.8310	0.1457	4.4853	2.0516	23.7691	2.7494	33.9243	0.0254	1.3896	0.0530	11.8801
CV	0.2350	0.4725	0.1799	0.8193	0.2324	0.4869	0.1726	0.7997	0.2519	0.3818	0.3063	0.4821	0.2712	0.9928	0.2917	2.5710
IQR	0.1892	0.9286	0.1625	4.1296	0.1908	1.1156	0.1823	4.8444	3.0170	28.3175	3.3351	41.6356	0.0380	1.1145	0.0986	1.3454
skewness	-1.1056	1.7331	-0.9411	2.1297	-1.1205	1.7227	-1.1762	1.9672	-0.5865	1.5951	-0.3535	1.3709	-0.3993	1.9813	-0.3389	3.5326
kurtosis	3.2778	4.9089	3.6607	7.2871	3.3279	4.8899	4.3328	6.5130	2.3497	5.4738	2.2091	4.2414	2.1238	6.1415	1.5056	13.6816
N	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16

Notes: TFP values were calculated using the coefficients estimated for the balanced panel data (only firms with complete data for sales, capital, labor, intermediate inputs and electricity for the years 2003 as well as 2009). The 25th indicates that data is provided only for the first quartile and 75th for the fourth quartile only. Regarding the acronyms of the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015). Among the TFP summary statistics, SD stands for standard deviation, IQR for interquartile range and CV for coefficient of variation (the rest are self-explanatory).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Based on these results, the tables below present the growth rates and rates of change in productivity dispersion statistics for both sectors.

Table 13: Growth Rates and Rates of Change for the Tails of the Distribution - Textiles

Method	Growth Rates (GR)						Rates of Change (ROC)						Exponential ROC					
		25th			75th			25th			75th			25th			75th	
	GR_{SD}	GR_{CV}	GR_{IQR}	GR_{SD}	GR_{CV}	GR_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$
OLS	0.2560	0.1071	0.3709	1.5537	0.9182	0.6110	0.0380	0.0170	0.0526	0.1563	0.1086	0.0795	0.0387	0.0171	0.0540	0.1691	0.1147	0.0827
STCH	0.2557	0.1067	0.3698	1.5547	0.9186	0.6105	0.0380	0.0169	0.0524	0.1563	0.1086	0.0794	0.0387	0.0170	0.0538	0.1692	0.1147	0.0827
LP	0.4373	0.3518	0.0843	1.2292	0.7376	0.4634	0.0605	0.0502	0.0135	0.1336	0.0921	0.0635	0.0623	0.0515	0.0136	0.1429	0.0965	0.0655
ACF	0.3123	0.0945	0.6547	2.1905	1.0741	0.4376	0.0453	0.0151	0.0839	0.0453	0.1216	0.0605	0.0463	0.0152	0.0876	0.0463	0.1293	0.0624

Notes: The exponential Rate of Change responds to the following equation $eROC = e^{ROC} - 1$. For ROC and eROC, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range. Regarding other acronyms in the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.

Table 14: Growth Rates and Rates of Change for the Tails of the Distribution - Furnitures

Method	Growth Rates (GR)						Rates of Change (ROC)						Exponential ROC					
		25th			75th			25th			75th			25th			75th	
	GR_{SD}	GR_{CV}	GR_{IQR}	GR_{SD}	GR_{CV}	GR_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	ROC_{SD}	ROC_{CV}	ROC_{IQR}	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$	$eROC_{SD}$	$eROC_{CV}$	$eROC_{IQR}$
OLS	-0.1102	-0.2343	-0.1412	1.6302	0.7338	3.4472	-0.0195	-0.0445	-0.0254	0.1612	0.0917	0.2487	-0.0193	-0.0435	-0.0251	0.1749	0.0961	0.2824
STCH	-0.1185	-0.2573	-0.0444	1.4497	0.6423	3.3424	-0.0210	-0.0496	-0.0076	0.1493	0.0827	0.2447	-0.0208	-0.0484	-0.0075	0.1611	0.0862	0.2773
LP	0.3401	0.2158	0.1054	0.4272	0.2627	0.4703	0.0488	0.0326	0.0167	0.0593	0.0389	0.0642	0.0500	0.0331	0.0168	0.0611	0.0396	0.0664
ACF	1.0853	0.0753	1.5923	7.5493	1.5897	0.2072	0.1225	0.0121	0.1588	0.1225	0.1586	0.0314	0.1303	0.0122	0.1721	0.1303	0.1719	0.0319

Notes: The exponential Rate of Change responds to the following equation $eROC = e^{ROC} - 1$. For ROC and eROC, SD stands for standard deviation, CV for coefficient of variation and IQR for interquartile range. Regarding other acronyms in the table, OLS stands for Ordinary Least Squares estimation; STCH stands for the stochastic frontier model of Battese and Coelli (1988, 1992); LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Ackerberg et al. (2015).
Source: Authors' calculations using data from the World Bank's Enterprise Surveys.