



Munich Personal RePEc Archive

A Comparison of TFP Estimates via Distribution Dynamics: Evidence from Light Manufacturing Firms in Brazil

Carlos Mendez-Guerra and Erick Gonzales-Rocha

4 July 2018

Online at <https://mpra.ub.uni-muenchen.de/87723/>
MPRA Paper No. 87723, posted 11 July 2018 16:04 UTC

A Comparison of TFP Estimates via Distribution Dynamics: Evidence from Light Manufacturing Firms in Brazil

Carlos Mendez-Guerra
Kyushu University

Erick Gonzales-Rocha
Kwansei Gakuin University

July 4, 2018

Abstract

This article compares the distribution dynamics of two commonly used TFP estimation frameworks: the control function approach of Levinsohn and Petrin (2003) (LP for short) and the corrected control function approach of Akerberg et al. (2015) (ACF for short). Using Brazilian firm-level data for the textile and furniture sectors, we estimate the transitional dynamics and long-run equilibrium of the TFP distribution for each framework over the 2003-2009 period. Results of this comparison are as follows. In the textile sector, the distribution dynamics for both frameworks are to some extent qualitatively similar. In the furniture sector, however, the distribution dynamics are largely different. While the LP framework shows relatively less mobility, two convergence clusters in the transition stage, and a bumpy distribution in the long run; the ACF framework shows relatively more backward mobility, a unique convergence cluster in the transition, and a highly symmetric distribution in the long run. In light of these results, the article concludes urging researchers not to rely too heavily on one or the other framework. It seems more appropriate to consider both frameworks for drawing inferences on productivity convergence and dispersion dynamics.

JEL Codes: D24, O47, O54,

Keywords: total factor productivity, control function approach, distribution dynamics, manufacturing firms, Brazil

1 Introduction

The estimation of total factor productivity (TFP) is a fundamental issue in applied economics. In the growth and development literature, for instance, TFP accounts for a large share of the per-capita income differences across countries (Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010). Motivated by the premise that aggregate TFP depends on the TFP of individual firms, an increasing amount of research has focused on the aggregate implications of firm-level productivity and its dispersion (Syverson, 2004; Hsieh and Klenow, 2009; Bartelsman et al. 2010). Most articles in this literature typically start by estimating a production function, computing a residual TFP measure, and then documenting large (and increasing) productivity dispersion within narrowly defined sectors.

Computing TFP at the firm level, however, is a challenging task. Different methods have been suggested to deal with the endogeneity problems that arise due to the positive correlation between observable input levels and unobservable productivity shocks. Building on the original control function method of Olley and Pakes (1996)(OP for short), two related TFP frameworks are most commonly used in the literature. On the one hand, the framework of Levinsohn and Petrin (2003)(LP for short) used intermediate inputs as a proxy variable for unobservable productivity shocks. On the other, Akerberg, Caves and Frazer (2015) (ACF for short) proposed a methodological correction that deals with the functional dependence problems that may arise in the LP framework.

Motivated by the methodological similarities and differences of the two previously described frameworks, this article compares their TFP estimates through the lens of the nonparametric distributional analysis of Quah (1993, 1997) and Johnson (2000, 2005). In particular, using firm-level data from the World Bank's Enterprise Surveys project, we compare both the transitional dynamics and the long-run equilibrium of the TFP distribution within two narrowly defined manufacturing sectors in Brazil: textile and furniture. The main purpose of this analysis is to have a nonparametric assessment of the dynamics of productivity dispersion, the mobility of firms within the productivity distribution, and the formation of multiple convergence clusters that are associated to each TFP framework.

Results of this comparative exercise indicate that in the textile sector, the distributional dynamics of both TFP frameworks are to some extent qualitatively similar. Both indicate a similar pattern of distributional convergence that is largely driven by the backward mobility of most productive firms. In the furniture sector, however,

the distributional dynamics of both frameworks largely differ. The LP framework shows relatively less productivity mobility, the formation of two convergence clusters in the transition phase, and a bumpy distribution in the long run. In contrast, the ACF framework shows larger backward mobility, a unique convergence cluster in the transition, and a more symmetric distribution in the long run.

The rest of the article is organized as follows. Section 2 briefly explains the two frameworks for estimating TFP, the distribution dynamics analysis, and the data source. Section 3 presents the transitional dynamics and long-run equilibrium results for each framework and sector. Finally, Section 4 offers some concluding remarks.

2 Methods and Data

2.1 Two Estimation Frameworks of TFP

The measurement of TFP requires the estimation of a production function. Among different alternatives, variations of the original control function approach of Olley and Pakes (1996) are ubiquitous in the literature. This approach is well known for using proxy variables and a control function to address the endogeneity problems that arise due to the positive correlation between input levels and the unobservable productivity shocks. Building on this line of research, two modern incarnations have been suggested by Levinsohn and Petrin (2003) and Akerberg, Caves and Frazer (2015).

The Control Function Approach of Levinsohn and Petrin (2003)

The original OP framework uses investment as proxy for productivity. This identification strategy, however, has one major limitation: investment decisions are typically accumulated for some years before being implemented and, as a result, such delay violates the monotonicity assumption of the framework. Motivated by this limitation, Levinsohn and Petrin (2003) suggested intermediate inputs as an alternative proxy for productivity. This proxy not only handles the concerns associated with the monotonicity assumption, but it is also easier to implement in practice, since it takes advantage of more readily available data.

The Corrected Function Approach of Akerberg, Caves and Frazer (2015)

The ACF framework proposes a functional dependence correction to the LP framework. In particular, the ACF framework incorporates the more general notion that

firms are not able to instantly adjust their labor input when subject to productivity shocks. In this new setting, the inputs demand functions are now conditional on the choice of both labor and capital inputs.

2.2 Distribution Dynamics Analysis

The distribution dynamics analysis was first introduced in the economic growth econometrics literature by Quah (1993, 1997). It is a nonparametric density¹ representation of the dynamics of a system. In particular, it studies the evolution of the entire cross-sectional distribution of productivity and its intra-distribution dynamics.² The evaluation of changes in both its external shape and its intra-distribution dynamics provides valuable information about the persistence of productivity differences, stratification patterns, and convergence clubs.

From a dynamic system model conceptualization, the transitional dynamics of the system are characterized by a continuous state-space stochastic process that is graphically presented as a (bivariate) conditional distribution function, known in the literature as the stochastic kernel.³ The long-run equilibrium of the system, on the other hand, is typically characterized by the shape of an estimated ergodic distribution. In what follows, we briefly sketch some essential methodological aspects of these two components of the framework.⁴

Transitional Dynamics via the Stochastic Kernel

Denote $d_t(x)$ as the initial productivity⁵ distribution across firms at time t . Next, denote $d_{t+s}(y)$ as the same distribution at some future time $t + s$. To model the change from time t to time $t + s$, we assume a time-homogeneous Markov chain process. Also, similar to a first-order autoregressive process, the transition between an initial state, x , and a future/final state, y , is assumed to be mapped by a transitional probability operator, $P(y | x)$. Under this setting, the distribution at time $t + s$ is given by a

¹The terms density and distribution are used as equivalent in this article.

²While the work of Quah (1993, 1996, 1997) focuses on the economic performance of countries, this article focuses on the economic performance of firms.

³This stochastic kernel distribution is also known as the continuous-time representation of the transition matrix of a Markov process.

⁴For a more comprehensive presentation, please refer to Dal Bianco (2016), Epstein, Howlett, and Schulze (2003) and Magrini (2009).

⁵Productivity is measured in relative terms. That is, the total factor productivity of each firm is normalized by the cross-sectional average of the sample.

transition probability operator, $P(y | x)$, and the initial distribution. That is,

$$\underbrace{d_{t+s}(y)}_{\text{Final Distribution}} = \int_{-\infty}^{+\infty} \underbrace{P(y | x)}_{\text{Transitional Operator}} \underbrace{d_t(x)}_{\text{Initial Distribution}} dx. \quad (1)$$

The transition probability operator $P(y | x)$, also known as the stochastic kernel function in the growth econometrics literature, is a conditional distribution that can be calculated as:

$$P(y | x) = \frac{d_{t,t+s}(y, x)}{d_t(x)}, \quad (2)$$

where $d_{t,t+s}(y, x)$ is a bivariate distribution that is typically estimated via nonparametric methods.

The nonparametric estimator for the joint density is

$$d_{t,t+s}(y, x) = \frac{1}{nh_y h_x} \sum_{i=1}^n K_y \left(\frac{y - y_i}{h_y} \right) K_x \left(\frac{x - x_i}{h_x} \right),$$

where y and x denote the (relative) productivity of each firm at time $t + s$ and t respectively, K_y and K_x denote kernel functions, and h_y and h_x denote the smoothing parameters for y and x respectively. Under the premise that this estimator is not particularly sensitive to the most common distributional functions,⁶ most articles in the literature⁷ have adopted a Gaussian functional form:

$$d_{t,t+s}(y, x) = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{\sqrt{2\pi}h_y} e^{-\frac{1}{2}\left(\frac{y-y_i}{h_y}\right)^2} \frac{1}{\sqrt{2\pi}h_x} e^{-\frac{1}{2}\left(\frac{x-x_i}{h_x}\right)^2} \right]. \quad (3)$$

In contrast to the relative agreement on the functional form, there is less consensus on the estimation of smoothing parameters (h_y and h_x). This is largely due to the challenges of finding an optimal balance between variance and bias. Following Magrini (1999, 2009) and Kar, Jha, and Kateja (2011), the estimation of optimal bandwidths in this article is based on the minimization of the asymptotic mean integrated square error (AMISE).

Long-run Equilibrium via the Ergodic Distribution

The long-run (ergodic) distribution represents an equilibrium that results from a large number of iterations on the stochastic kernel applied to the initial distribu-

⁶See Silverman (1986) and Wand and Jones (1995) for further details.

⁷See for example the work of Andrade et al. 2004; Epstein et al. 2003; Quah, 1997.

tion. As $s \rightarrow \infty$, Equation 1 is used to estimate such equilibrium and the ergodic distribution, $d_\infty(y)$, becomes the solution to the following problem:

$$d_\infty(y) = \int_{-\infty}^{+\infty} P(y | x) d_\infty(x) dx = d_\infty(x). \quad (4)$$

To solve this problem, we follow the analytical approach suggested by Johnson (2000, 2005) and use the MATLAB library developed by Magrini (2009). If the solution to this problem exists, the shape of this ergodic distribution provides valuable information regarding the convergence patterns of firms. For instance, if $d_\infty(y)$ shows a tendency towards a unique point of mass or mode, then it is indicative of distributional convergence. On the other hand, if $d_\infty(y)$ shows a divergent tendency towards multiple modes, then it is indicative of multiple convergence clubs (Galor, 1996).

2.3 Data

The original data for this article comes from the World Bank's Enterprise Surveys project. This project is well known for collecting firm-level data across countries using a stratified random sample that covers the non-agricultural economy of the main cities of different countries. In particular, we use the panel dataset that is documented in Gonzales-Rocha and Mendez-Guerra (2018). In this work, the authors use data for the years 2003 and 2009 to build a balanced panel dataset for the textile and furniture sectors. As expected, the balanced-panel construction and the focus on quite narrowly defined sectors required a large reduction in sample size. Moreover, only 48 observations in textile sector and 62 observations in the furniture sector had the required variables to compute TFP in both years.

3 Results

3.1 Distribution Dynamics of TFP in Textiles

Panels A and B of Figure 1 compare the transitional dynamics of TFP for each of the two estimation frameworks described in Subsection 2.1. To facilitate the interpretation of relative distances, the cross-sectional mean is normalized to zero⁸ in each period and the natural logarithm of TFP is applied.⁹ The 45-degree line indicates

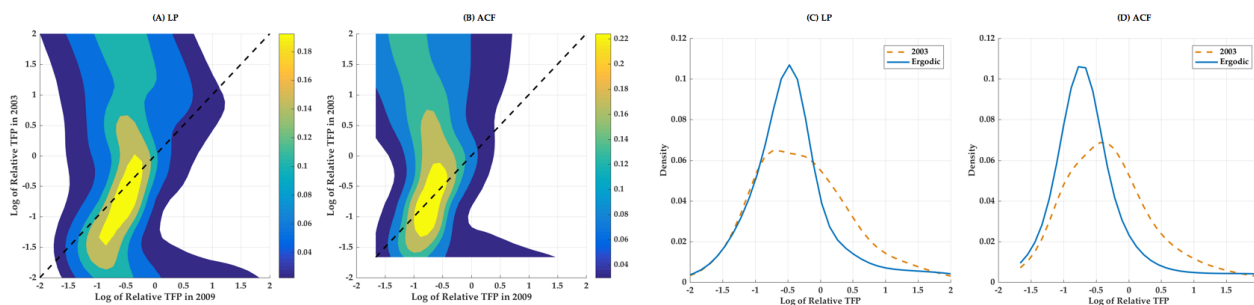
⁸Using TFP in relative terms is also helpful to control for aggregate shocks that might affect all firms within the industry.

⁹Note that this rescaling of variables is also applied in Panels C and D of Figure 1 as well as in all the panels of Figure 2.

productivity stagnation (or lack of intra-distribution mobility). Transitional dynamics are characterized by the position and shape of the estimated stochastic kernel relative to the 45-degree line. As shown by Panels A and B of Figure 1, there is a counterclockwise rotation in the stochastic kernel that indicates both the forward mobility of the less productive firms and the backward mobility of the more productive firms. Although this convergence pattern is qualitatively similar between the two TFP estimation frameworks, relatively larger backward mobility is observed through the lens of the ACF framework. Regarding the shape of the stochastic kernel, both frameworks clearly show only one mode of high density. Taken together, these results imply that the transitional dynamics of the textiles sector appear to be characterized by the formation of one convergence cluster that is largely below the mean TFP level of the year 2009. Moreover, this characterization is consistent across the two TFP frameworks of the study.

Panels C and D of Figure 1 compare the long-run equilibrium distribution of TFP for each estimation framework in the textile sector. Note that this distribution should not be considered as a forecast, instead its main purpose is to clarify and magnify the observed transitional dynamics (Quah, 1997). To some extent, the long-run (ergodic) distribution is qualitatively similar for the two TFP estimation frameworks. Although the ACF framework shows a reduction of relative productivity arising from the lower tail, the dynamics of the industry are largely driven by the reduction of relative productivity arising from the upper tail of the ergodic distribution. Moreover, in both frameworks, the ergodic distribution is clearly unimodal and much narrower when compared to the shape of the initial (2003) distribution.

Figure 1: Transitional Dynamics and Long-Run Equilibrium in Textiles



Notes: LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Akerberg et al. (2015).

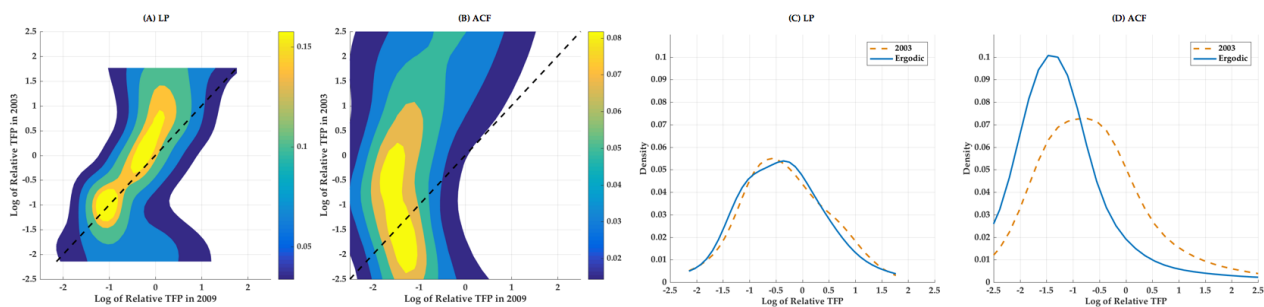
Source: Authors' calculations using data from the World Bank's Enterprise Surveys

3.2 Distribution Dynamics of TFP in Furnitures

In contrast to the results of the textile sector, the transitional dynamics of the two frameworks are largely different in the furniture sector (Panels C and D of Figure 2). While the LP framework shows two density clusters, the ACF framework shows only one cluster. Moreover, the position and shape of those clusters highlight two different convergence dynamics. While the LP framework suggests that the high-productivity cluster is slowly transitioning towards the low-productivity cluster, the ACF framework suggests a more completed convergence process.

The long-run equilibrium of the two TFP measures are also largely different in the furniture sector (Panels C and D of Figure 2). While the LP framework suggests a relatively asymmetric and bumpy distribution, the ACF framework suggests a more symmetric distribution in the long run. In addition, both tails of ACF ergodic distribution are notoriously behind those of the initial (2003) distribution. Taken together, these results reemphasize what was previously observed in the study of transitional dynamics, the ACF framework systematically reports a larger reduction of relative productivity in the furniture sector.

Figure 2: Transitional Dynamics and Long-Run Equilibrium in Furnitures



Notes: LP stands for the control function approach of Levinsohn and Petrin (2003); and ACF stands for the corrected control function approach of Akerberg et al. (2015).

Source: Authors' calculations using data from the World Bank's Enterprise Surveys

4 Concluding Remarks

Increasing availability of firm-level data and advances in methods for estimating production functions have stimulated a fast-growing literature that studies the dynamics of productivity dispersion within narrowly defined sectors. Two of the most common frameworks for estimating productivity at the firm level are the control function approach of Levinsohn and Petrin (2003) and the corrected control function approach

of Akerberg et al. (2015). Since these frameworks share common methodological foundations (that is, the control function approach), very often researchers have used them interchangeably, or at least without a detailed analysis of robustness.

Using Brazilian data for the textile and furniture sectors over the 2003-2009 period, this article compares the robustness of the TFP estimates of these two frameworks through the lens of the distribution dynamics analysis of Quah (1993, 1997) and Johnson (2000, 2005). Results indicate that only in the textile sector, the productivity dynamics of both frameworks are—to some extent—qualitatively similar. In contrast, in the furniture sector, the productivity dynamics implied by each framework largely differ. In this sector, the LP framework indicates relatively less productivity mobility and the formation of two convergence clusters whereas the ACF framework indicates relatively more productivity mobility and the formation of a unique convergence cluster. Overall, the potential existence of such differences across frameworks should encourage further research using other sectors and larger datasets. More importantly, they should also caution researchers in treating these frameworks as interchangeable for drawing inferences on productivity convergence and dispersion dynamics.

References

- Akerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.
- 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.
- Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741.
- Dal Bianco, S. (2016). Going clubbing in the eighties: convergence in manufacturing sectors at a glance. *Empirical Economics*, 50(2):623–659.
- Epstein, P., Howlett, P., and Schulze, M.-S. (2003). Distribution dynamics: stratification, polarization, and convergence among OECD economies, 1870–1992. *Explorations in Economic History*, 40(1):78–97.
- Galor, O. (1996). Convergence? inferences from theoretical models. *The Economic Journal*, pages 1056–1069.

- Gonzales-Rocha, E. and Mendez-Guerra, C. (2018). Increasing productivity dispersion in Brazil: Evidence from light manufacturing firms. *mimeo*.
- Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1):83–116.
- 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.
- Johnson, P. A. (2000). A nonparametric analysis of income convergence across the US states. *Economics Letters*, 69(2):219–223.
- Johnson, P. A. (2005). A continuous state space approach to convergence by parts. *Economics Letters*, 86(3):317–321.
- Kar, S., Jha, D., and Kateja, A. (2011). Club-convergence and polarization of states: A nonparametric analysis of post-reform India. *Indian Growth and Development Review*, 4(1):53–72.
- Magrini, S. (1999). The evolution of income disparities among the regions of the European Union. *Regional Science and Urban Economics*, 29(2):257–281.
- Magrini, S. (2004). Regional (di) convergence. In *Handbook of regional and urban economics*, volume 4, pages 2741–2796. Elsevier.
- Magrini, S. (2009). Why should we analyse convergence using the distribution dynamics approach? *Scienze Regionali*.
- Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.
- Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, pages 427–443.
- Quah, D. T. (1997). Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2(1):27–59.
- Syverson, C. (2004a). Market Structure and Productivity: A Concrete Example. *Journal of Political Economy*, 112(6):1181–1222.

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