UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

GUILHERME ANTONIO CALABRIA BAYMA COSTA

Estimating Firm-Level Productivity under Size-dependent Distortions

Recife

2020

GUILHERME ANTONIO CALABRIA BAYMA COSTA

Estimating Firm-Level Productivity under Size-dependent Distortions

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Centro de Ciências Sociais Aplicadas da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre em Economia.

Área de Concentração: Teoria Econômica

Orientador: Paulo Henrique Vaz

Recife

2020

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

C837e Costa, Guilherme Antonio Calabria Bayma
Estimating firm-level productivity under size-dependent distortions / Guilherme Antonio Calabria Bayma Costa. - 2020.
53 folhas: il. 30 cm.

Orientador: Prof. Dr. Paulo Henrique Vaz.
Dissertação (Mestrado em Economia) – Universidade Federal de Pernambuco, CCSA, 2020.
Inclui referências e apêndices.

1. Produtividade. 2. Produtividade administrativa. 3. Produção. I. Vaz, Paulo Henrique (Orientador). II. Título.

336 CDD (22. ed.) UFPE (CSA 2023 – 043)

GUILHERME ANTONIO CALABRIA BAYMA COSTA

ESTIMATING FIRM-LEVEL PRODUCTIVITY UNDER SIZE-DEPENDENT DISTORTIONS

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Teoria Econômica.

Aprovado em: 20/02/2020.

BANCA EXAMINADORA

____
Profº. Dr. Paulo Henrique Vaz (Orientador)
Universidade Federal de Pernambuco

____
Profº. Dr. Rafael da Silva Vasconcelos (Examinador Interno)
Universidade Federal de Pernambuco

____
Profº. Dr. Fabio Miessi Sanches (Examinador Externo)
INSPER

Dedico a minha mãe, Márcia Maria.

AGRADECIMENTOS

Agradeço aos professores do PIMES pelos ensinamentos e conhecimentos compartilhados. Aos colegas de turma pela companhia e parceria. Aos colegas do GAPPE por compartilharem o dia a dia da produção do trabalho. E ao meu orientador, Paulo Henrique Vaz, pelo apoio, ensinamento e boas conversas.

Agradeço à minha família e amigos por me acompanharem e ajudarem durante este período. E à minha mãe por estar ao meu lado nos momentos mais difíceis.

RESUMO

Políticas públicas que dependem dos tamanhos das empresas estão presentes em todo o mundo e geralmente estão associadas à má alocação de recursos e à perda de produtividade agregada. Apesar da estimação e identificação de funções de produção terem recebido muitas contribuições recentemente, não existe uma discussão da performance de métodos recentes sob distorções endógenas. Neste trabalho é mostrado que nessas circunstâncias uma hipótese dos métodos recentes para a estimação de funções de produção é violada e a estimação se torna problemática. Empiricamente, é mostrado que esse problema na identificação gera estimadores de produtividade de firmas e dispersão de produtividade sistematicamente viesados. Ignorar esses efeitos durante estimações pode levar a conclusões precipitadas.

Palavras-chaves: Distorções endógenas; produtividade de firmas; funções de produção.

ABSTRACT

Size-dependent policies are present worldwide and are usually associated with *misallocation* of resources and to loss of aggregate productivity. Although the estimation and identification of production functions have received many contributions recently, there is not a discussion on the performance of recent methods under severe size-dependent distortions. We show that under these circumstances a necessary assumption of recent production function estimators is violated and estimation became problematic. Empirically, we show that this identification problem creates systematically biased estimates of firm-level productivity and TFP dispersion. Ignoring these effects may drive misleading policy conclusions.

Keywords: Size-dependent policies; size distortions, productivity; production functions.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUCTION

Over the last 25 years, firm-level data have become more available in developing and developed countries. With this empirical gain, firm-level productivity (TFP) has been largely used in many fields of economics. In particular, the literature on the estimation and identification of production functions has evolved vastly since the pioneering work by Olley e Pakes (1996). Simultaneously, parallel literature has been given much attention to the *misallocation* of resources and its different causes: financial frictions, trade restrictions, size-dependent policies, and a big variety of regulations (Hsieh e Klenow (2009); Restuccia e Rogerson (2013); Restuccia e Rogerson (2017)).

This paper connects these two types of literature by discussing the estimation of production functions in an environment with size-dependent distortions. In particular, we analyze the performance of recently proposed production function estimators, such as[1]: OLLEY; PAKES (1996, henceforth OP), LEVINSOHN; PETRIN (2003, henceforth LP), WOOLDRIDGE (2009, henceforth WD), and ACKERBERG; CAVES; FRAZER (2015, henceforth ACF). In possession of the production function coefficient estimates, it is possible to calculate the firm level TFP and investigate the efficiency of estimation under size-dependent policies. In the last 20 years, many contributions have been made to the estimation and identification of production functions, in a way that this paper does not cover all relevant methodologies[2].

The main challenge of these new methods is to solve the simultaneity and selection problem that arises when estimating production functions. Firm-level TFP is not observed by the econometrician but is partially observed by the firm. Simultaneity bias occurs due to the endogenous choice of inputs by firm managers. If capital and labor are chosen with some knowledge of their productivity, then the choice of inputs will be partially determined by TFP, leading to endogeneity when estimating production functions with OLS. To solve this issue, Olley e Pakes (1996) proposed a semi-parametric estimation using investment as a proxy for unobserved productivity, while Levinsohn e Petrin (2003) follow the same approach suggesting to use of intermediate inputs as a proxy. Moreover, firms may exit the market due to low

[1] For instance, those methods were used by Pavcnik (2002); Fernandes (2007); Blalock e Gertler (2004); Ozler e Yilmaz (2009); Topalova e Khandelwal (2011); Kasahara e Rodrigue (2008); Dollar, Hallward-Driemeier e Mengistae (2005); Javorcik (2004); Amiti e Konings (2007); Dodlova et al. (2015); Halpern, Koren e Szeidl (2015); Yu (2015); Bernard, Moxnes e Saito (2019); Atkin, Khandelwal e Osman (2017); Levine e Warusawitharana (2019); and many others.

[2] See Blundell e Bond (2000), Loecker (2007), Katayama, Lu e Tybout (2009), and Gandhi et al. (2017) for alternative methodologies not covered in this paper.

productivity draws or high cost to operate, adding also a selection bias.

Size-dependent policies are policy distortions that depend on firms' size, which that could be in terms of sales, labor, or capital (GUNER; VENTURA; XU, 2008). This type of distortion is pointed out by Restuccia e Rogerson (2017) as one of the main sources of *misallocation*. Such policies could occur via a discouragement on production of large firms, or via support to small ones. Size-dependent distortions can appear in many different ways, where examples include financial constraints, subsidies, taxes, and a multitude of options depending on the creativity of the policymaker.

Size-dependent distortions are very common worldwide. The World Bank (2016) highlights that small and medium firms are disproportionately affected by inefficiencies in the business environment. In extreme cases, reports in countries such as Sierra Leone show that more than 65% of firms have financial constraints, where small and medium establishments represent 96% of this group. As a result, policymakers usually choose this group of firms as the target for public policies, thus, creating size-dependent distortions.

Many recent papers discuss size-dependent policies in varied circumstances. Guner, Ventura e Xu (2008) study the impacts of such policies and find positive effects on the number of firms and negative effects on output. Gourio e Roys (2014) and Garicano, Lelarge e Reenen (2016) investigate labor regulations in France that depend on the number of employees in each firm. Braguinsky, Branstetter e Regateiro (2011) show that labor market restrictions in Portugal lead firms to reduce their employment. Almunia e Lopez-Rodriguez (2018) analyze the effects of size-dependent tax enforcement on firms' tax compliance in Spain. Martin, Nataraj e Harrison (2017) study small-scale industry promotion in India and its effects on job creation. Cavalcanti e Vaz (2017) show that in Brazil low-revenue firms are the target of a policy that subsidizes interest rates for long-term investment. In summary, these policies are very frequent among countries and will be the object of many future empirical studies.

Although estimation and identification of production functions have received many contributions recently, as far as we know, there are no published papers raising up a discussion on the performance of recent methods under size-dependent distortions. This topic has many policy implications since the evidence shows that these distortions are spread among countries and these methods have been hugely used empirically worldwide. We show that in such an environment, a necessary assumption is violated. Thus, estimation should be less efficient, and inputs' coefficients not identified. This follows from the fact that the distortion is heterogeneous between firms and that this component is not taken into account by the methods during estimation. Estimated firm-level TFP should also be biased. As productivity is calculated as the part of the output that is not explained by the use of inputs, any unobserved heterogeneity between firms will be captured by the estimated TFP. So, as size-dependent distortions create heterogeneity between firms, this should bias estimates of firm-level productivity.

To our knowledge, this is the first paper that analyzes the performance of estimation production function methods under size-dependent distortions. The empirical strategy relies on simulated firm-level data following Syverson (2001), Biesebroeck (2007), and Ackerberg, Caves e Frazer (2015). We generate two main samples of profit-maximizing firms with the same firm-TFP and initial capital in both, but in only one of them, we impose a size-dependent distortion, affecting firms' decisions as a consequence. With these two samples, we compare estimations of the four methods with and without size-dependent policies[3]. As theory suggested, our empirical results confirm that under size-dependent distortions estimation is problematic and the estimates are not consistent. Additionally, our main finding is that this identification problem creates systematically biased estimates.

The contributions of this paper cover both the literature on misallocation and on production function estimation. Our first contribution is to show that, under size-dependent distortions, traditional methods for estimating production functions underperform. This occurs through a positive bias in the labor input coefficient and a negative bias in the capital input coefficient. This finding is new and needs to be investigated in future research.

One second contribution is to find heterogeneous patterns on estimated firm-level TFP bias. When the inputs' coefficient estimates are biased, the direction of the estimated productivity bias is not straightforward. Since the directions of bias on estimates of the production function are opposite, bias on the estimated TFP will be heterogeneous between firms that have different shares of inputs. We also show that firm-level TFP is overestimated for small firms while it is underestimated for large firms in the majority of methodologies analyzed.

Policy implications from this finding are straightforward and empirical studies of firm-level productivity may drive misleading conclusions. For example, suppose that in a country there are no size-dependent policies and that the policymaker desires to insert a distortion focusing on small firms and aiming to boost its productivity. If firm-level productivity is estimated before and after the policy, it should be observed an increase in the estimated productivity of small firms even if the real effect of the policy is null. An increase in estimated TFP could be only the positive bias observed in our study and not an increase in real productivity.

[3] We use Monte Carlo simulations as robustness for the results.

A third contribution is related to *misallocation* measures. Hsieh e Klenow (2009) suggest that dispersion of TFP reflects *misallocation* of resources. Bartelsman, Haltiwanger e Scarpetta (2013) argue that the within-industry covariance between size and productivity is a more robust measure to assess the impact of *misallocation* distortions. We show that both measures are biased in the presence of endogenous distortions. In particular, the dispersion of TFP and the covariance term are underestimated.

This finding brings new evidence to the discussion of cross-country differences in firm-level productivity. Countries face different degrees of endogenous distortions and this could lead to heterogeneous bias in cross-country measured dispersion. Estimation of firms' TFP in countries with more severe endogenous distortions leads to higher underestimation of dispersion of productivity. This result alerts us to caution for this cross-country comparison. Countries with more distortions are probably the ones with more dispersion in actual TFP. Thus, its estimated dispersion of TFP should be underestimated and could be similar to an estimated dispersion of a country with fewer distortions and less dispersion in actual productivity.

The paper proceeds as follows. In Chapter 2 we include a quick review of the literature on firm-level TFP estimation and the identification problem in the presence of size-dependent distortions. In Chapter 3 we describe the data-generating process, starting from a representative firm optimization problem. In Chapter 4 we present the results analyzing the estimation of production functions and TFP. Finally, the conclusion presents our concluding remarks.

2 REVIEW OF LITERATURE ON TFP ESTIMATION

In this section, we briefly review the literature on productivity (TFP) estimation and include the identification of complications in the presence of size-dependent distortions. We start with a discussion of the reason why OLS estimates of the production function are biased and introduce recent approaches that aim to solve this issue. In particular, we include the approaches of Olley e Pakes (1996), Levinsohn e Petrin (2003), Wooldridge (2009), and Ackerberg, Caves e Frazer (2015). We detail the OP method, while for LP, WD, and ACF we focus on their improvements. Most of the content in this section we borrow liberally from the survey on TFP estimation written by Beveren (2012) and from the review on production function estimation by Ackerberg, Caves e Frazer (2015).

First, consider a simple Cobb-Douglas production function as

$$Y_{it} = A_{it} L_{it}^{\beta_l} K_{it}^{\beta_k}, \tag{2.1}$$

where Y_{it} is the output, K_{it} is the capital input, L_{it} is the labor input, β_l is the labor coefficient, β_k is the capital coefficient, and A_{it} is the Hicksian neutral productivity. In this equation, A_{it} is the only variable unobserved by the econometrician. Taking natural logs of (2.1) results in

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \varepsilon_{it}, \tag{2.2}$$

where the observed variables are the log of output y_{it}, the log of labor input l_{it}, and the log of capital input k_{it}. Here note that the log of the unobserved variable A_{it} is divided into three terms, that is,

$$\log A_{it} = \beta_0 + \omega_{it} + \varepsilon_{it},$$

where ε_{it} measures productivity shocks that are not predictable by the firm, β_0 is the mean productivity level across firms and over time, and ω_{it} is productivity shocks observed by firms when making input decisions.

In possession of the coefficient estimates, firm-level TFP can be estimated as

$$\hat{\omega}_{it} = y_{it} - \hat{\beta}_l l_{it} - \hat{\beta}_k k_{it}. \tag{2.3}$$

2.1 ORDINARY LEAST SQUARES (OLS)

The parameters β_l and β_k in (2.2) could be easily estimated using OLS. However, it is known that OLS estimation leads to biased estimates of β_l and β_k Marschak e Andrews (1944). There are two main problems when trying to estimate (2.2): simultaneity bias and selection bias. Simultaneity bias occurs due to the endogenous choice of inputs by the firms. If the firm chooses k_{it} and l_{it} with some knowledge of ω_{it}, then the choice of inputs will be partly determined by the productivity, leading to the correlation between ω_{it} and (k_{it}, l_{it}). Selection bias arises from the fact that the exit decision from the firm is not taken into account, introducing an endogeneity of attrition.

The literature on the estimation of production functions has given the most attention to the simultaneity problem. In the last 25 years many new methodologies have tried to solve the endogeneity problem of inputs choice. We review the approaches by Olley e Pakes (1996), Levinsohn e Petrin (2003), Wooldridge (2009), and Ackerberg, Caves e Frazer (2015) [1] .

The selection problem received focus in the literature with the work of Olley e Pakes (1996), taking the exit decision explicitly into account. However, since then this problem has been less discussed. Levinsohn e Petrin (2003) argue that there is no necessity to focus on selection issues since the selection correction makes little difference when the simultaneity correction is been used.

2.2 OLLEY-PAKES (OP)

Olley e Pakes (1996) (OP) introduces a three-step estimator that aims to identify the parameters of the production function taking into account both the selection and simultaneity problems. The key to solving the simultaneity problem is to use investment as a proxy for productivity shocks. Behind the identification problem, they develop a discrete-time dynamic model of firm behavior. In this model, the optimal decision of investment by the firms is policy functions resulting from a dynamic optimization problem. In order to obtain consistent estimates, assumptions need to be satisfied. Ackerberg, Caves e Frazer (2015) summarizes these in five assumptions [2]:

[1] This literature has been received many recent contributions and there are relevant approaches that were not covered in this paper. For more information see Blundell e Bond (2000), Loecker (2007), Katayama, Lu e Tybout (2009), and Gandhi et al. (2017).

[2] For more details about the assumptions in OP and LP see Ackerberg, Caves e Frazer (2015).

Assumption 1 *Information Set: The firm's information set at t, that is, I_{it} , includes current and past productivity shocks $\{\omega_{i\tau}\}_{\tau=0}^{t}$ but does not include future productivity shocks $\{\omega_{i\tau}\}_{\tau=t+1}^{\infty}$. The transitory shocks ε_{it} satisfy $E[\varepsilon_{it}|I_{it}]=0$.*

Assumption 2 *First Order Markov: Productivity shocks evolve according to the distribution*

$$p\left(\omega_{it+1}|I_{it}\right)=p\left(\omega_{it+1}|\omega_{it}\right).$$

This distribution is known to firms and stochastically increasing in ω_{it}.

Assumption 3 *Timing of input choices: Firms accumulate capital according to*

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

where investment i_{it} is chosen in period $t1$. Labor input l_{it} is non-dynamic and chosen at t.

Assumption 4 *Scalar Unobservable: Firms' investment decisions are given by*

$$i_{it}=f_t\left(k_{it},\omega_{it}\right).$$

Assumption 5 *Strict Monotonicity: $f_t(k_{it},\omega_{it})$ is strictly increasing in ω_{it}.*

Assumptions 1 and 2 imply that productivity is not observed until t but the firm has knowledge about the distribution of next productivity socks. Assumption 3 guarantees that labor is a non-dynamic input, while capital is dynamic under an investment process. Assumption 4 implies that investment decisions are a function of variables k_{it} and ω_{it}. Assumption 5 states the monotonicity of the investment function regarding the productivity shock.

Given assumptions 4 and 5, one can invert the investment function and obtain $\omega_{it}=f_t^{-1}\left(i_{it},k_{it}\right)$. Substituting this into (2.2) we obtain

$$y_{it}=\beta_0+\beta_k k_{it}+\beta_l l_{it}+f_t^{-1}\left(k_{it},i_{it}\right)+\varepsilon_{it}=\beta_l l_{it}+\Phi_t\left(k_{it},i_{it}\right)+\varepsilon_{it}, \tag{2.4}$$

where $\Phi_t\left(k_{it},i_{it}\right)$ is approximated by a high-order polynomial in i_{it} and k_{it}. The first stage estimation gives a consistent estimate of the labor input coefficient and its moment condition is given by

$$\mathbb{E}\left[\varepsilon_{it}|I_{it}\right]=\mathbb{E}\left[y_{it}-\beta_l l_{it}-\Phi_t\left(k_{it},i_{it}\right)|I_{it}\right]=0 \tag{2.5}$$

The second step aims to solve the selection problem controlling for the exit decision. To include this step is necessary to add an exit rule where the firm does not exit the market if the observed productivity is above some threshold that depends on the level of capital input:

$$\chi_{it} = \begin{cases} 1 & \text{if } \omega_{it} \geq \underline{\omega}_{it}(k_{it}) \\ 0 & \text{otherwise.} \end{cases}$$

The intuition is that the exit rule is a function of productivity and the minimum condition of survival, and both are functions of investment and capital. Follows that a nonparametric probit estimation of a survival dummy on i_{it} and k_{it} produces the estimates of the probability of survival $\hat{P}_{it}$, that is,

$$\Pr_{it}(\chi_{t+1} = 1) = \sum_{x=1}^{3} \gamma_{1x}(k_{it})^x + \sum_{x=1}^{3} \gamma_{2x}(i_{it})^x + \sum_{x=1}^{3} \gamma_{3x}(k_{it} \times i_{it})^x .$$

In the third step, only the coefficient of the capital input is estimated. To achieve this, it is necessary to use Assumptions 1 and 2 to decompose ω_{it} into its conditional expectation at time $t-1$, and an innovation term ξ_{it}, that is,

$$\omega_{it} = \mathbb{E}[\omega_{it}|\omega_{it-1}, \chi_{it}] + \xi_{it} = g(P_{it-1}, \omega_{it-1}) + \xi_{it},$$

where substituting this into (2.2) we obtain

$$\begin{aligned} & y_{it} - \beta_t l_{it} \\ & = \beta_0 + \beta_k k_{it} + \mathbb{E}(\omega_{it}|\omega_{it}, \chi_{it}) + \xi_{it} + \varepsilon_{it} \\ & = \beta_0 + \beta_k k_{it} + g(P_{it-1}, \Phi_{t-1}(k_{it-1}, i_{it-1}) - \beta_k k_{it-1}) + \xi_{it} + \varepsilon_{it}. \end{aligned}$$

The third stage moment condition is given by

$$\begin{aligned} & \mathbb{E}[\xi_{it} + \varepsilon_{it}|I_{it-1}] \\ & \quad = \mathbb{E}[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} \\ & \quad -g(P_{it-1}, \Phi_{t-1}(k_{it-1}, i_{it-1}) - \beta_0 - \beta_k k_{it-1})|I_{it-1}] = 0, \end{aligned} \tag{2.6}$$

where estimation proceeds by substituting the estimated coefficient of labor input and the estimated survival probability from the first and second stages.

2.3 LEVINSOHN-PETRIN (LP)

The approach in LP is almost identical to OP, but the key difference is that they propose to use intermediate inputs as a proxy for unobserved productivity. Thus, they use the demand function for an intermediate input m_{it} to invert for ω_{it}. In order to, assumptions 4 and 5 are substituted by new assumptions:

Assumption 6 *Scalar Unobservable: Firms' intermediate input demand is given by*

$$m_{it} = f_t(k_{it}, \omega_{it}).$$

Assumption 7 *Strict Monotonicity:* $f_t(k_{it}, \omega_{it})$ *is strictly increasing in* ω_{it}.

Estimation proceeds as in OP, only substituting the investment function for the demand function for an intermediate input. There are two main advantages to using intermediate inputs instead of investment. First, it is common to have many firms reporting investment as null. This turns impossible to invert investment for unobserved productivity. Second, it is easy to verify the assumption of strict monotonicity for intermediate inputs, since investment it requires an analysis of a dynamic problem.

2.4 WOOLDRIDGE (WD)

The contribution by Wooldridge (2009) is to estimate the two-moment conditions in (2.5) and (2.6) jointly in a GMM framework. The key to this strategy is to set the moment conditions in terms of two equations with the same dependent variable, but different instruments across equations.

Wooldridge (2009) argues that joint estimation of the parameters leads to simple inference and more efficient estimators. Estimation in two steps should be inefficient because they do not take into account correlation in the errors across the two equations, and they ignore serial correlation or heteroskedasticity in the errors. In contrast, GMM estimation enhances efficiency through cross-equation correlation, as well as serial correlation and heteroskedasticity are taken into account by the optimal weighting matrix.

2.5 ACKERBERG-CAVES-FRAZER (ACF)

Ackerberg, Caves e Frazer (2015) argues that there may be a functional dependence problem with the identification of the labor coefficient in the first stage of OP and LP. Thus, ACF propose a new estimator that inverts the demand functions of investment or intermediate input conditional on the labor input. In order to do so, it is necessary to substitute Assumptions 3, 4, and 5 for the next assumptions:

Assumption 8 *Timing of Input Choices: Firms accumulate capital according to*

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

where investment $i+it1$ is chosen in period $t1$. Labor input l_{it} has potential dynamic implications and is chosen at period t, period $t1$, or period tb (with $0 < b < 1$).

Assumption 9 *Scalar Unobservable: Firms' intermediate input demand is given by*

$$m_{it} = \tilde{f}_t\left(k_{it}, l_{it}, \omega_{it}\right)$$

Assumption 10 *Strict Monotonicity: $\tilde{f}_t\left(k_{it}, l_{it}, \omega_{it}\right)$ is strictly increasing in ω_{it}.*

2.6 IDENTIFICATION UNDER SIZE-DEPENDENT POLICIES

In the presence of size-dependent distortions, production function estimation methods fail to identify input coefficients. Hence, this problem also affects the estimation of firm-level productivity. This occurs because of the violation of the assumption of Scalar Unobservable.

Suppose a size-dependent policy creates a distortion ϕ_{it} that is heterogeneous between firms and affects investment and intermediate input demand functions. This distortion is not observed by the econometrician and it is not taken into account by OP, LP, WD, and ACF when inverting the demand function. Note that under these circumstances, the investment function and intermediate input demand function are given by, respectively,

$$i_{it} = f_t\left(k_{it}, \omega_{it}, \phi_{it}\right), \tag{2.7}$$

and

$$m_{it} = \tilde{f}_t\left(k_{it}, l_{it}, \omega_{it}, \phi_{it}\right). \tag{2.8}$$

In (2.7) and (2.8), both functions depend on more than one unobserved variable. This fact violates the Scalar Unobservable Assumption (assumptions 4, 6, and 9) for OP, LP, WD, and ACF. Thus, in the presence of size-dependent distortions, these methods cannot invert correctly for unobserved productivity ω_{it}, since ϕ_{it} is also unobserved.

3 DATA GENERATION

We use simulated firm-level data, constructed from a representative firm model, to estimate firm-level TFP in situations with and without a size-dependent policy. The content in this section follows the work by Syverson (2001), Biesebroeck (2007), and Ackerberg, Caves e Frazer (2015). In the data generation process (DGP), the model allows dynamics only in the capital input, while the labor and materials inputs are assumed to be not dynamic. This assumption is crucial to obtain an analytical solution for the optimal level of investment, which demands much less computational resources than other options and benefits the efficiency of the simulation process [1].

The simulated firm-level data aims to represent a sample of firms from a developing country.[2] In particular, the parameters of the model and the initial distribution of capital are chosen to approximate a sample of Colombian firms. Since the work by Roberts e Tybout (1996), this dataset of firms has been extensively used empirically in development (e.g. Midrigan e Xu (2014), Eslava et al. (2004), Brooks (2006)) , and in international economics (e.g. Garcia-Marin e Voigtländer (2019), Clerides, Lach e Tybout (1998), Eslava et al. (2013) , Isgut (2001)).

The DGP applied intends to allow for all methodologies to be consistent. In general, the assumptions of each approach are rarely contradictory and allow the existence of one benchmark DGP. Following Ackerberg, Caves e Frazer (2015), to avoid functional dependence problems, it's necessary to assume that the labor input is chosen at period t and that there is an optimization error in L_{it} used by the firm in the production function (e.g. randomly sick workers). In this case, it's needed that the materials input that is observed is chosen before the realization of the optimization error in labor, in a way that the researcher observes planned materials.

When applying the size-dependent policy, the adopted strategy was to insert a distortion in the variable used to proxy the unobserved productivity. In order to do so, when using OP we apply the distortion in the investment variable while when using LP, WD, and ACF we apply the distortion in the intermediate input variable. These types of distortions are very common both in developing and developed countries.

[1] For a numerical solution, consider the model presented in Bond e Söderbom (2005), which allows dynamics in both capital and labor inputs.
[2] For a reality check see appendix section A.2 in Biesebroeck (2007).

The main advantage of using simulated data is that we can compare the true productivity of the firm (generated by the DGP) with the estimated productivity. This approach is used to contrast the performance of the different estimators under situations with and without size-dependent policies.

3.1 BENCHMARK MODEL

3.1.1 Representative Firm and Investment Choice

Following Biesebroeck (2007), the benchmark model is from a representative firm that chooses investment and labor over time while optimizing the net present value of profits, subject to a Cobb-Douglas production function[3] and a capital accumulation equation:

$$\begin{aligned} \max_{L_{it},I_{it}} = \mathbb{E}_t &\left\{\textstyle\sum_{t=0}^{\infty} \beta^t \left[Y_{it} - W_t L_{it} - c\left(I_{it}\right)\right]\right\} \\ \text{s.t.} \quad Y_{it} &= A_{it} L_{it}^{\beta_l} K_{it}^{\beta_k} \\ K_{it+1} &= (1-\delta) K_{it} + I_{it}, \end{aligned} \tag{3.1}$$

where Y_{it} is output, W_t is wage rate, L_{it} is labor input, I_{it} is investment, A_{it} is productivity level, and K_{it} is capital input. The parameters in (3.1) are: β as an intertemporal discount factor; δ as capital depreciation; and, β_j as the coefficient of the input j. In the model is also included a convex cost function given by $c(I_{it}) = \frac{b}{2} I_{it}^2$, where b is a homogeneous adjustment cost.

The solution of the firm maximization problem for the labor input is

$$L_{it} = \left(\frac{\beta_l A_{it}}{W_t}\right)^{\frac{1}{1-\beta_l}} K_{it}^{\frac{\beta_k}{1-\beta_l}}, \tag{3.2}$$

and the Euler equation that describes the evolution of investment along the optimal path is

$$c'(I_{it}) = \beta \beta_k \beta_l^{\frac{1}{1-\beta_l}} \mathbb{E}_t \left[A_{it+1}^{\frac{1}{1-\beta_l}} W_{t+1}^{\frac{\beta_l}{1-\beta_l}} K_{it+1}^{\frac{\beta_l+\beta_k-1}{1-\beta_l}} \right] + \beta(1-\delta) \mathbb{E}_t [c'_{t+1}(I_{it+1})] \tag{3.3}$$

[3] As in Ackerberg, Caves e Frazer (2015), a structural value-added production function is used. That is, intermediate inputs do not enter the production function. The DGP is not subject to the critique of Gandhi et al. (2017), since the production function could be described as a Leontief in materials, with this input proportional to output. For more details see Ackerberg, Caves e Frazer (2015) and Gandhi et al. (2017).

A series of assumptions are made with the objective to express the investment as a function of only one state variable, the current productivity level A_{it}. These assumptions are:

Assumption 11 *Constant returns of scale:*

$$\beta_k + \beta_l = 1.$$

Assumption 12 *Wages not propagated over time and identical for each firm:*

$$W_t \overset{i.i.d.}{\sim} N\left(1, \sigma_w^2\right).$$

Assumption 13 *Log-productivity follows an AR(1) process:*

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it},$$

where $\omega_{it} = \log A_{it}$*,* $|\rho| < 1$*, and* $\xi_{it} \overset{i.i.d.}{\sim} N\left(0, \sigma_\xi^2\right)$.

Thus, joining these assumptions with the first order conditions (3.2) and (3.3) the optimal investment rule is given by

$$I_{it} = \frac{\beta\beta_k\beta_l^{\frac{1}{1-\beta_l}}}{b}\sum_{\tau=0}^{\infty}[\beta(1-\delta)]^{\tau}\left[A_{it}^{\frac{1}{1-\beta_l}}\right]^{\rho^{\tau+1}}\left[\prod_{s=0}^{\tau}e^{\frac{1}{2}\left(\frac{\sigma_\xi\rho^s}{1-\beta_l}\right)^2}\right]$$

3.1.2 Optimization Error in the Labor Input and the Choice of Material Input

Following Ackerberg, Caves e Frazer (2015), we add an optimization error to L_{it} to avoid functional dependence problems. That is,

$$L_{it} = L_{it}^{plan}e^{\xi_{it}^l},$$

where $\xi_{it}^l \overset{\text{i.i.d.}}{\sim} N\left(0, \sigma_{\xi^l}^2\right)$, and L_{it}^{plan} is planned labor input.

The material input M_{it} observed by the econometrician is assumed to be reported by the firm prior the realization of the optimization error, that is,

$$M_{it} = A_{it}K_{it}^{\beta_k}L_{it}^{plan\beta l}.$$

3.1.3 Exit rule

We also include an exogenous and an endogenous exit rule. The exogenous exit rule is determined in a way that every year about one percent of the firms are randomly shut down. As in Biesebroeck (2007), we model the endogenous exit as depending on capital. Firms below the fourth percentile of capital exit the industry. Note that we do not consider potential future exit in the representative firm optimization, since this would turn impossible to find an analytical solution for the investment function.

3.2 MODEL WITH SIZE-DEPENDENT POLICY

We use a model with distortion to compare the performance of different production function estimators. This distortion is inputted into the model via a size-dependent policy. Size-dependent policies are very common in developing and developed countries.

Following Guner, Ventura e Xu (2008), we use capital to determine firm size. A firm is treated by the policy if its amount of capital in period t is below some exogenously predetermined level $\overline{K}$. The idea here is to generate an observed heterogeneity between firms in the variable used to proxy for productivity. Thus, since OP uses investment as a proxy while LP, WD, and ACF use materials, it is needed two different types of distortions to evaluate the different estimators.

There are many types of size-dependent policies. This could occur via financial frictions, subsidies, taxes, and a multitude of options depending on the creativity of the policymaker. We intend to insert the distortion in a simple way, so it could represent any type of distortion that affects the desired variable. Hence, we simply introduce a new parameter on the variable of interest.

3.2.1 Distortion in the investment variable

When assuming distortion in the model, we consider different adjustment costs for different firms. So, the new investment function is given by

$$I_{it} = [\phi_{it}] \times \frac{\beta \beta_K \beta_L^{\frac{1}{1-\beta_L}}}{b} \sum_{\tau=0}^{\infty} [\beta(1-\delta)]^{\tau} \left[A_t^{\frac{1}{1-\beta_L}} \right]^{\rho^{\tau+1}} \left[\prod_{s=0}^{\tau} e^{\frac{1}{2}\left(\frac{\sigma_a \rho^s}{1-\beta_L}\right)^2} \right],$$

where

$$\phi_{it} = \begin{cases} \phi^0 & \text{if} \quad K_{it} \geq \overline{K} \\ \phi^1 & \text{if} \quad K_{it} < \overline{K} \end{cases}$$

3.2.2 Distortion in the intermediate input

When assuming distortion in the model, the firm's choice of material input is affected by the new parameter. That is,

$$M_{it} = [\phi_{it}] \times A_{it} K_{it}^{\beta_k} L_{it}^{plan \beta l},$$

where

$$\phi_{it} = \begin{cases} \phi^0 & \text{if} \quad K_{it} \geq \overline{K} \\ \phi^1 & \text{if} \quad K_{it} < \overline{K} \end{cases}$$

3.3 DATA

Following Biesebroeck (2007), we set a panel of 1000 firms over 10 time periods. With our exit rule, we end up with 635 firms in the last period, totalizing 8074 observations. The parameters of the model and initial distribution of capital are set in a way that our data resembles a sample of firms from a developing country. The hypothesis of constant returns to scale is adopted with $\beta_l = 0.6$ and $\beta_k = 0.4$. The other parameter values are in Table 1.

Tabela 1 – Parameter Values

Parameter	**Value**
β	0.90
δ	0.10
b	0.01
σ_ω	0.50
σ_ξ	0.37
ρ	0.10

When generating the data with the size-dependent distortion our main results consider $\phi^0 = 1$, $\phi^1 = 5$, and $\overline{K} = \frac{1}{1000}\sum_{i=1}^{1000} K_{i0}$. We test the robustness of the results using different values for ϕ^1 and $\overline{K}$, including Monte Carlo simulations with measurement errors[4].

[4] See APPENDIX TABLES and APPENDIX FIGURES.

4 RESULTS

This paper aims to analyze the performance of firm-level TFP estimation under size-dependent distortions. In order to do so, we test four recent production function estimation methods that have been very used empirically. These methods estimate the input's coefficient, and with that information, we calculate firm-level TFP as in (2.3). The focus is to compare estimation performance in situations with and without distortions. Results presented here don't pretend to certify that one method is better or worse than another, since results are from the DGP considered and the conclusion may be different for other firms' data.

Tabela 2 – Coefficient estimates

	Labor Coefficient (β_l)	Capital Coefficient (β_k)
Actual	0.6000	0.4000
OP Estimation		
Without Distortion	0.6592	0.3322
With Distortion	0.8163	0.1870
LP Estimation		
Without Distortion	0.5988	0.3970
With Distortion	0.8348	0.1840
WD Estimation		
Without Distortion	0.6245	0.3775
With Distortion	0.8654	0.1853
ACF Estimation		
Without Distortion	0.5997	0.4027
With Distortion	0.6954	0.2952

We begin by comparing results from the input's coefficient estimates and then results for firm-level TFP. Table 2 contains results for coefficient estimates from the four methodologies considered. For each methodology, it is included results for the situations with and without a size-dependent distortion. When considering the situation without distortion, results for the coefficient estimates are very similar to the actual values and seem to be consistent[1]. However,

[1] As in Biesebroeck (2007), we input measurement errors of 0.1 in the variables for the OP estimation. That may be the reason why the OP estimation presents results more distant from the actual values. Besides that, OP estimation cannot be compared to the other estimators in terms of performance, since the distortion used is different.

when looking at the results in the situation with distortion the estimators underperform and differ considerably from the true values. For OP, LP, and WD the estimated coefficients for both inputs differs at least 0.2, that is, 33.33% of the actual β_l and 50% of the actual β_k. The ACF seems to be less sensitive but still presents a bias of about 16.66% for β_l and of about 25% for β_k. The estimates in Table 2, therefore, confirm that under size-dependent distortions the estimators of production functions have worse performance. [2]

Results in Table 2 also bring information about the direction of the bias for input coefficients. For all methods analyzed, there is a positive bias in the labor coefficient and a negative bias in the capital coefficient. Given these biases, it is ambiguous the bias in estimated firm-level TFP. With this pattern, we have in (2.3) that the second term pushes the bias of $\hat{\omega}_{it}$ downward while the last term pushes the bias upward. This result implies that the bias on the estimated TFP will be heterogeneous between firms, or sectors, that have different shares of inputs. In particular, for firms that rely more on labor input, the bias should be negative while for those relying on capital input, the bias should be positive.

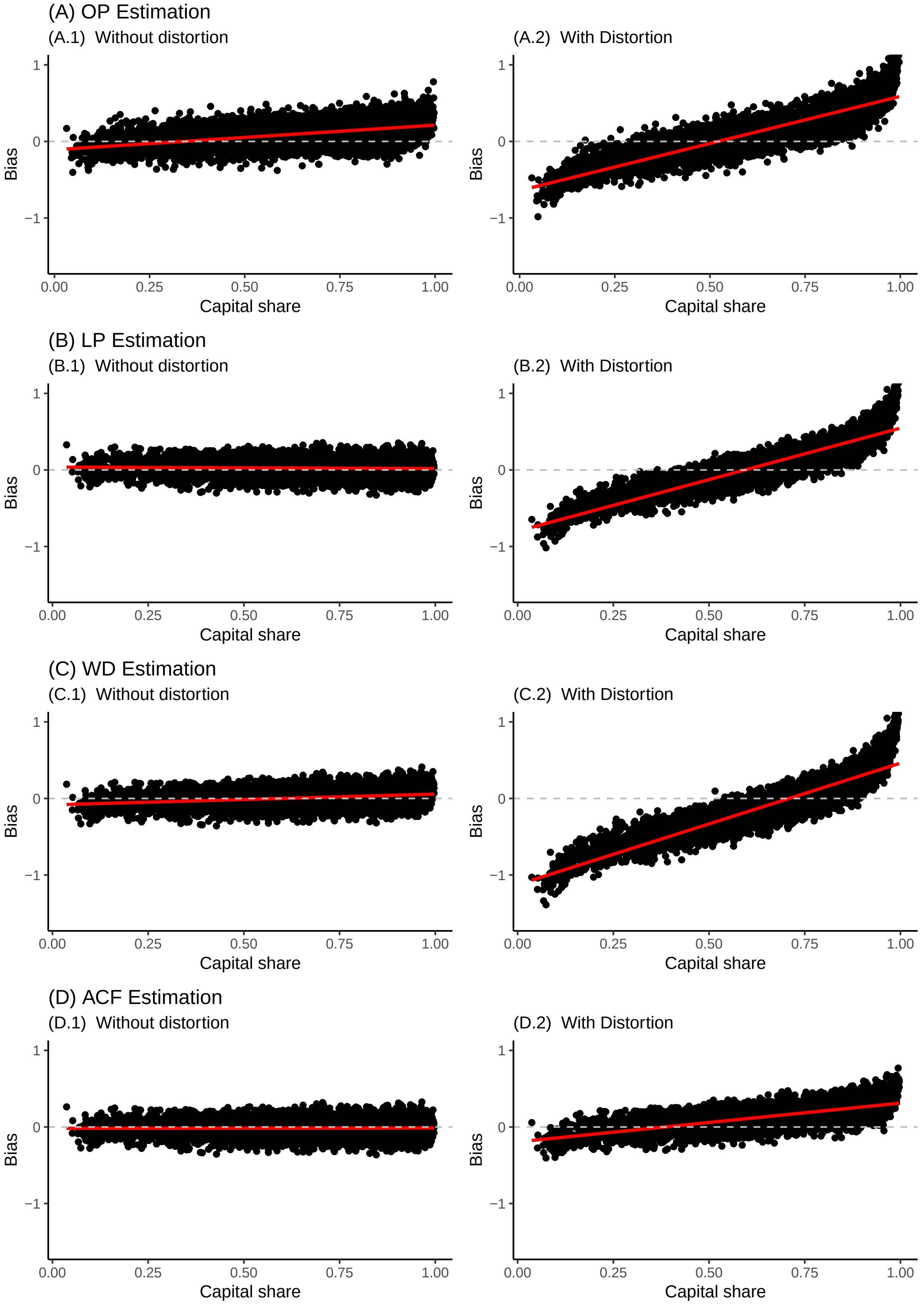
This pattern is confirmed in Figure 1. This figure plots firm-level data, where the Y-axis is the bias in estimated TFP, and the X-axis is the capital share. Results here confirm that there is a positive relationship between the bias in TFP and the share of capital input. More than heterogeneous between firms, the bias is also heterogeneous between sectors of the economy with a distinct share of inputs. This finding brings very much caution when estimating TFP under size-dependent distortions since bias is heterogeneous and estimations are biased in different directions.

Looking at firm-level data, Figure 2 plots the distribution of TFP (real and estimated) for different approaches in situations with and without distortion. Figures on the left are for the situation without distortion, while figures on the right are for situations with distortion. Each figure plots the distributions of both real and estimated TFP. When estimating firm-level TFP without distortion the two distributions are an almost perfect match for all methods. In contrast, when there is a size-dependent distortion the distributions differ noticeably. For OP, LP, and WD the range of the estimated distribution hugely decreases, while for ACF the distribution slightly shifts to the right. These visualizations highlight that when estimating firm-level TFP under size-dependent distortions the estimated distribution of productivity is imperfect. [3]

Recent literature uses productivity dispersion to explain cross-country differences in output

[2] In APPENDIX TABLES and APPENDIX FIGURES see robustness results with different distortion values,

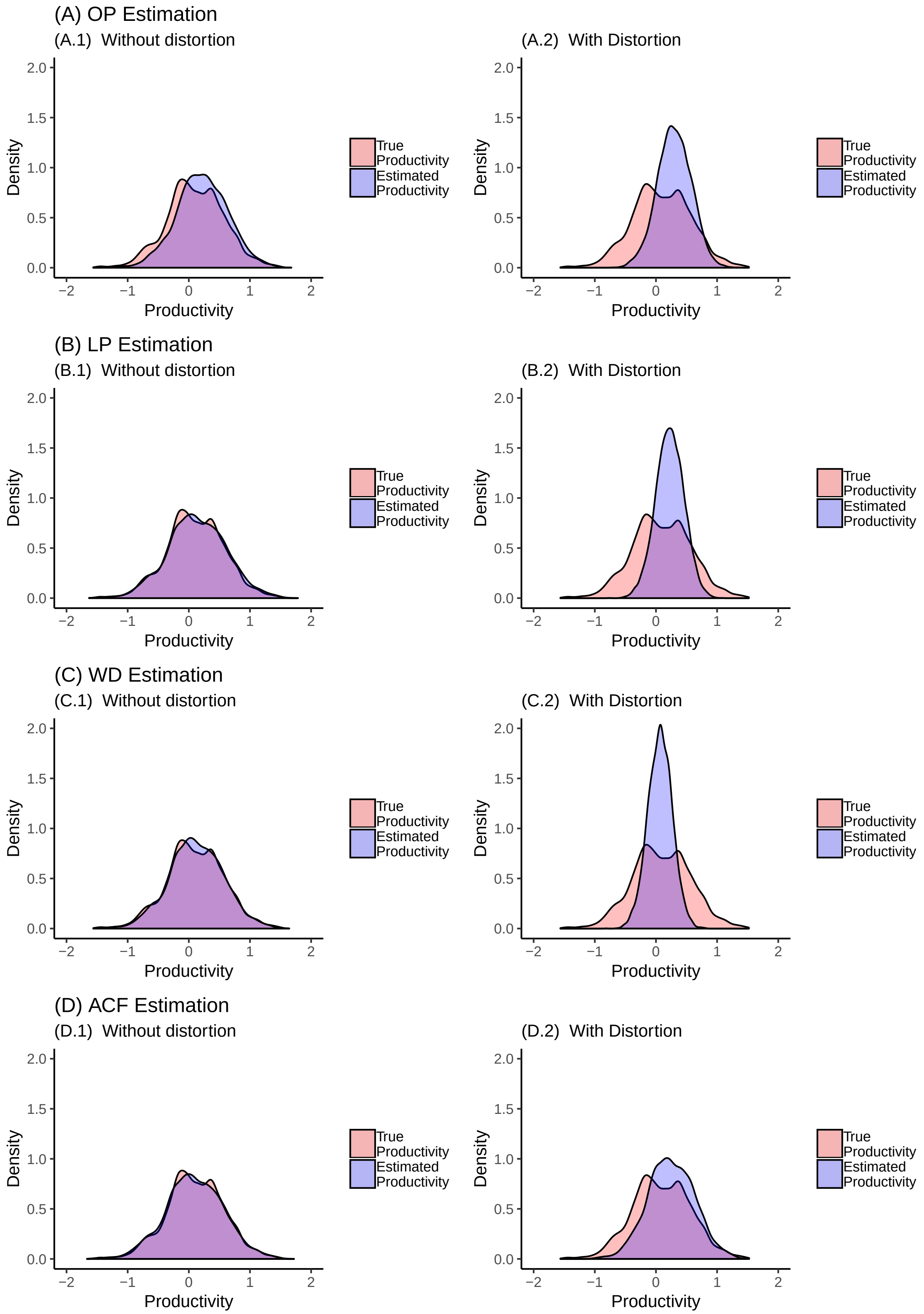
Figura 1 – Bias in estimated TFP and capital share.



different percentages of firms being affected, and Monte Carlo simulations.

[3] In Figures B.3, B.4, and B.5 we show that estimation of firm TFP in the presence of more intense size-dependent distortions leads to a higher underestimation of dispersion of productivity.

Figura 2 – Distributions of productivities



per worker through resource *misallocation* Gandhi et al. (2017). Table 3 shows different measures of dispersion of the estimated TFP for the estimators considered. In particular, these measures are the standard deviation, the 75th minus the 25th percentiles, and the 90th minus the 10th percentiles. Column 1 shows results for estimations without distortion and column 2 shows results in the presence of size-dependent policy. Results in column 1 are very similar and indicate that the dispersion of TFP estimated by different methods seems to be consistent. However, when in a situation with distortion, the dispersion of TFP differs between methods and is underestimated for all methodologies. The main implication of this table is that when estimating TFP under size-dependent distortions all measures of dispersion are smaller compared with the situation without distortions.

Tabela 3 – Dispersion of TFP

	Without Distortion	With Distortion
OP Estimation		
SD	0.4216	0.2772
75 - 25	0.5597	0.3737
90 - 10	1.0669	0.7049
LP Estimation		
SD	0.4796	0.2313
75 - 25	0.6458	0.3109
90 - 10	1.2201	0.5907
WD Estimation		
SD	0.4471	0.2011
75 - 25	0.6005	0.2720
90 - 10	1.1369	0.5094
ACF Estimation		
SD	0.4745	0.3818
75 - 25	0.6389	0.5251
90 - 10	1.2068	0.9731

In Table 4 we include estimates for aggregate productivity and the covariance between productivity and output as in Olley e Pakes (1996). While Hsieh e Klenow (2009) suggest that dispersion of TFP reflects *misallocation* of resources, Bartelsman, Haltiwanger e Scarpetta (2013) argue that the within-industry covariance between size and productivity is a more

Tabela 4 – Aggregate Productivity and Covariance between Productivity and Output

	Aggregate Productivity	Covariance between Productivity and Output
OP Estimation		
Without Distortion	0.35	0.12
With Distortion	0.36	0.06
LP Estimation		
Without Distortion	0.28	0.14
With Distortion	0.28	0.05
WP Estimation		
Without Distortion	0.26	0.13
With Distortion	0.10	0.03
ACF Estimation		
Without Distortion	0.24	0.14
With Distortion	0.37	0.11

robust measure to assess the impact of *misallocation* distortions. Estimates in column 2 show results for this covariance. For all methodologies, this measure is underestimated under size-dependent distortions. This finding is consistent with results in Table 3 and alerts for caution when measuring *misallocation* in the presence of size-dependent policies.

Table 4 brings also information about aggregate TFP. To calculate it, we follow Olley e Pakes (1996) and use a weighted average of firm-level TFP, with shares of industry output as weights. Results in column 1 show estimates of aggregate productivity and the conclusion about its performance under size-dependent distortions is not clear. WD Estimation underestimates this measure, while ACF overestimates it. At the same time, OP and LP accurately estimate the aggregate TFP. These findings need to be more investigated in future research.

Figure 3 plots the bias in the estimated firm-level TFP for groups of firms of different sizes. The idea here is to show the difference between estimated and real productivity for small, medium, and big firms based on their level of capital. Column 1 shows results without distortion and indicates that the bias is very close to zero in all estimations and seems to not depend on firms' size. Results in column 2, when the distortion is applied, show that for OP, LP, and WD it occurs to be a heterogeneous bias between firms of different sizes, in a way that the bias appears to be positive for small firms and negative for big firms. However, in the ACF estimation, this pattern is not clear and shows a relatively homogeneous bias. Hence, the results here enhance that in the presence of size-dependent distortions firm-level

TFP is overestimated for small firms and underestimated for big firms in the majority of the methodologies analyzed.

Similarly, Figure 4 plots the bias in the estimated firm-level TFPG for groups of firms of different sizes. The idea here is the same as in Figure 3 but shows results for firm-level productivity growth. Column 1 presents results for the benchmark estimation while in column 2 results are from estimations in the presence of size-dependent distortions. For both situations, it seems that there is no correlation between the bias and the firms' size. When comparing estimations with and without distortion the main difference is that the bias has a bigger variance in the second situation. Results in Figure 4 suggest that different from TFP, the bias in the estimated firm-level TFPG does not depend on firms' size.

Tabela 5 – Correlation coefficients for situations with and without distortion

	Estimated Firm-level TFP	Estimated Firm-level TFPG
OP	0.91	0.89
LP	0.88	0.75
WD	0.78	0.77
ACF	0.98	0.95

Although firm-level TFPG bias does not depend on firms' size, it's not clear that researchers should choose this measure instead of TFP. Table 5 shows correlation coefficients between estimations with and without distortion for each method. Column 1 presents results for estimated firm-level TFP, while column 2 shows results for firm-level TFPG. For all methodologies considered, the correlation is smaller for the TFPG. This result indicates that firm-level TFPG estimation may be more sensible than firm-level TFP in the presence of size-dependent distortions.

Tabela 6 – Correlation between estimated firm-level TFP from different methods without distortion

	LP	WD	ACF
OP	0.9807	0.9824	0.9807
LP	-	0.9996	0.9999
WD	-	-	0.9997

Figura 3 – Difference between estimated and real productivity per capital decile.

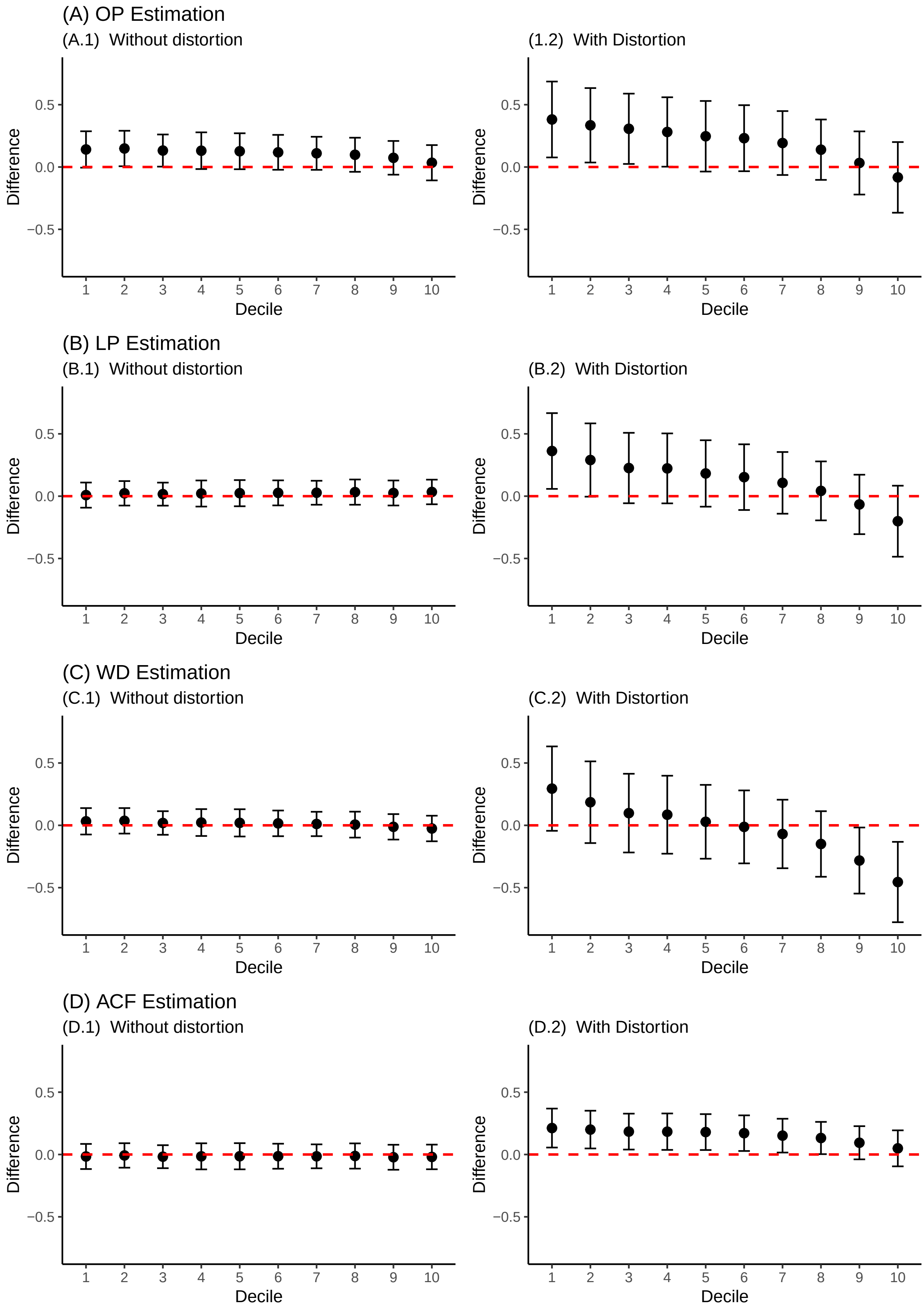


Figura 4 – Difference between estimated and real TFPG per capital decile.

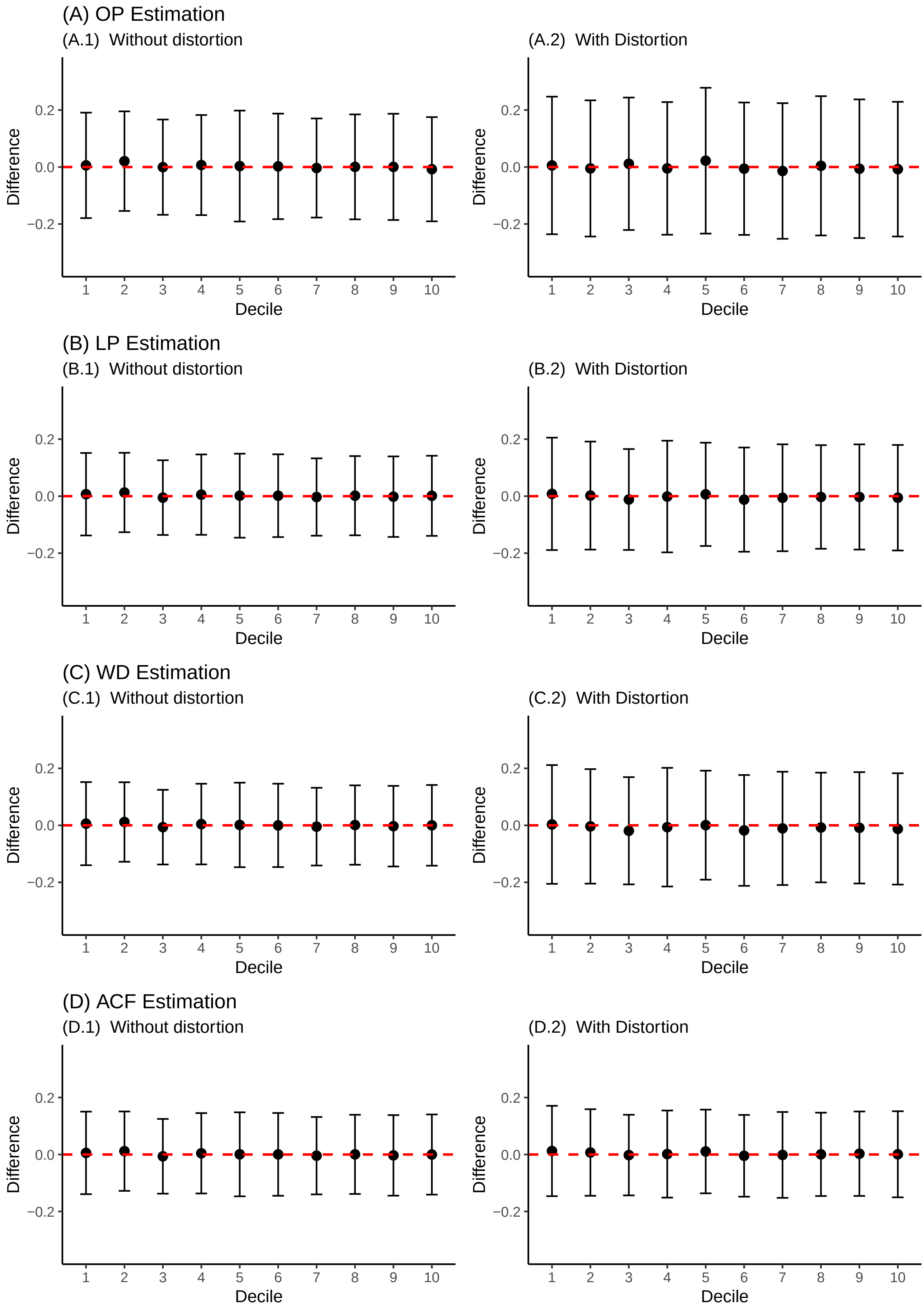


Tabela 7 – Correlation between estimated firm-level TFP from different methods with distortion

	LP	WD	ACF
OP	0.9486	0.9219	0.8847
LP	-	0.9769	0.9319
WD	-	-	0.8394

Table 6 and 7 show correlation results between estimated firm-level TFP from different methods. In Table 6, correlation coefficients are from the situation without distortion, while in Table 7 the distortion is inserted. Results indicate that in the absence of size-dependent distortion the correlations between estimations are almost perfect. However, when estimating TFP in the presence of such distortion, the correlations between methods are lower. These results highlight that, in the presence of distortions, estimations from different methods can differ notably.

CONCLUSION

This paper discusses the performance of production function estimators in the presence of endogenous distortions, such as size-dependent distortions in access to credit, and finds systematically biased estimates of firm-level productivity and TFP dispersion. Although the estimation and identification of production functions received many contributions most recently, as far as we know, there are no published papers raising a discussion on the subject of this paper and our findings are new in the literature. We show that under endogenous distortion necessary assumptions of recent production function estimators are violated. The key to our empirical approach is to simulate profit-maximizing firms in two situations: with and without endogenous distortions. In possession of these two samples, we analyze the performance of each method in estimating firm-level productivity in the presence of size-dependent distortions.

As theoretically predicts, production function coefficients and firm-level TFP are biased when estimated in the presence of severe endogenous distortions. In particular, when the distortion favors small firms, we show that productivity is overestimated for those firms and for firms that rely more on capital input. We also show that measures to assess *misallocation*, as the dispersion of TFP and the covariance between productivity and output, are underestimated. Conclusions about the effect on aggregate productivity are unclear and need more investigation.

Our results bring a new discussion to the literature and have many policy implications. Heterogeneous bias in firm-level TFP alerts for caution when doing analysis in the presence of distortions that depend on firm size. For instance, given the TFP overestimation for small firms, if firm-level productivity is estimated before and after the introduction of some size-dependent policy, it could be observed an increase in the estimated productivity of small firms even if the real effect of the policy is null.

Another result with policy implications is the underestimation of productivity dispersion, which could be used to assess *misallocation* in distorted environments. We find that estimation of firm-level TFP in the presence of more severe endogenous distortions leads to more underestimation of productivity dispersion. This result is a precautionary tale for cross-country and cross-sector comparisons. Different environments could present similarly estimated dispersion of TFP even when the actual dispersion differs. Ignoring these effects may drive misleading policy conclusions.

Future research could search for methods that solve the issue raised by this paper. Endogenous distortions are widely spread worldwide and empirical researchers need a reliable methodology to use. Meanwhile, the message is to keep caution in policy conclusions from empirical studies using these recent approaches.

REFERÊNCIAS

ACKERBERG, D. A.; CAVES, K.; FRAZER, G. Identification properties of recent production function estimators. *Econometrica*, Wiley Online Library, v. 83, n. 6, p. 2411–2451, 2015.

ALMUNIA, M.; LOPEZ-RODRIGUEZ, D. Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy*, v. 10, n. 1, p. 1–38, 2018.

AMITI, M.; KONINGS, J. Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review*, v. 97, n. 5, p. 1611–1638, 2007.

ATKIN, D.; KHANDELWAL, A. K.; OSMAN, A. Exporting and firm performance: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, Oxford University Press, v. 132, n. 2, p. 551–615, 2017.

BANK, W. *What's Holding Back the Private Sector in MENA? Lessons from the Enterprise Survey*. [S.l.]: Washington, DC: World Bank, 2016.

BARTELSMAN, E.; HALTIWANGER, J.; SCARPETTA, S. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, v. 103, n. 1, p. 305–34, 2013.

BERNARD, A. B.; MOXNES, A.; SAITO, Y. U. Production networks, geography, and firm performance. *Journal of Political Economy*, The University of Chicago Press Chicago, IL, v. 127, n. 2, p. 639–688, 2019.

BEVEREN, I. V. Total factor productivity estimation: A practical review. *Journal of economic surveys*, Wiley Online Library, v. 26, n. 1, p. 98–128, 2012.

BIESEBROECK, J. V. Robustness of productivity estimates. *The Journal of Industrial Economics*, Wiley Online Library, v. 55, n. 3, p. 529–569, 2007.

BLALOCK, G.; GERTLER, P. J. Learning from exporting revisited in a less developed setting. *Journal of development economics*, Elsevier, v. 75, n. 2, p. 397–416, 2004.

BLUNDELL, R.; BOND, S. Gmm estimation with persistent panel data: an application to production functions. *Econometric reviews*, Taylor & Francis, v. 19, n. 3, p. 321–340, 2000.

BOND, S.; SÖDERBOM, M. *Adjustment costs and the identification of Cobb Douglas production functions*. [S.l.], 2005.

BRAGUINSKY, S.; BRANSTETTER, L. G.; REGATEIRO, A. *The incredible shrinking Portuguese firm*. [S.l.], 2011.

BROOKS, E. L. Why don't firms export more? product quality and colombian plants. *Journal of development Economics*, Elsevier, v. 80, n. 1, p. 160–178, 2006.

CAVALCANTI, T.; VAZ, P. H. Access to long-term credit and productivity of small and medium firms: A causal evidence. *Economics Letters*, Elsevier, v. 150, p. 21–25, 2017.

CLERIDES, S. K.; LACH, S.; TYBOUT, J. R. Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco. *The quarterly journal of economics*, MIT Press, v. 113, n. 3, p. 903–947, 1998.

DODLOVA, M.; GÖBEL, K.; GRIMM, M.; LAY, J. Constrained firms, not subsistence activities: Evidence on capital returns and accumulation in peruvian microenterprises. *Labour Economics*, Elsevier, v. 33, p. 94–110, 2015.

DOLLAR, D.; HALLWARD-DRIEMEIER, M.; MENGISTAE, T. Investment climate and firm performance in developing economies. *Economic Development and Cultural Change*, The University of Chicago Press, v. 54, n. 1, p. 1–31, 2005.

ESLAVA, M.; HALTIWANGER, J.; KUGLER, A.; KUGLER, M. The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from colombia. *Journal of development Economics*, Elsevier, v. 75, n. 2, p. 333–371, 2004.

ESLAVA, M.; HALTIWANGER, J.; KUGLER, A.; KUGLER, M. Trade and market selection: Evidence from manufacturing plants in colombia. *Review of Economic Dynamics*, Elsevier, v. 16, n. 1, p. 135–158, 2013.

FERNANDES, A. M. Trade policy, trade volumes and plant-level productivity in colombian manufacturing industries. *Journal of International Economics*, v. 1, n. 71, p. 52–71, 2007.

GANDHI, A.; NAVARRO, S.; RIVERS, D. et al. How heterogeneous is productivity? a comparison of gross output and value added. *Journal of Political Economy, forthcoming*, 2017.

GARCIA-MARIN, A.; VOIGTLÄNDER, N. Exporting and plant-level efficiency gains: It's in the measure. *Journal of Political Economy*, The University of Chicago Press Chicago, IL, v. 127, n. 4, p. 000–000, 2019.

GARICANO, L.; LELARGE, C.; REENEN, J. V. Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, v. 106, n. 11, p. 3439–79, 2016.

GOURIO, F.; ROYS, N. Size-dependent regulations, firm size distribution, and reallocation. *Quantitative Economics*, Wiley Online Library, v. 5, n. 2, p. 377–416, 2014.

GUNER, N.; VENTURA, G.; XU, Y. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, Elsevier, v. 11, n. 4, p. 721–744, 2008.

HALPERN, L.; KOREN, M.; SZEIDL, A. Imported inputs and productivity. *American Economic Review*, v. 105, n. 12, p. 3660–3703, 2015.

HSIEH, C.-T.; KLENOW, P. J. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, MIT Press, v. 124, n. 4, p. 1403–1448, 2009.

ISGUT, A. What's different about exporters? evidence from colombian manufacturing. *Journal of Development Studies*, Taylor & Francis, v. 37, n. 5, p. 57–82, 2001.

JAVORCIK, B. S. Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages. *American economic review*, v. 94, n. 3, p. 605–627, 2004.

KASAHARA, H.; RODRIGUE, J. Does the use of imported intermediates increase productivity? plant-level evidence. *Journal of development economics*, Elsevier, v. 87, n. 1, p. 106–118, 2008.

KATAYAMA, H.; LU, S.; TYBOUT, J. R. Firm-level productivity studies: illusions and a solution. *International Journal of Industrial Organization*, Elsevier, v. 27, n. 3, p. 403–413, 2009.

LEVINE, O.; WARUSAWITHARANA, M. Finance and productivity growth: Firm-level evidence. *Journal of Monetary Economics*, Elsevier, 2019.

LEVINSOHN, J.; PETRIN, A. Estimating production functions using inputs to control for unobservables. *The review of economic studies*, Wiley-Blackwell, v. 70, n. 2, p. 317–341, 2003.

LOECKER, J. D. *Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 2007.

MARSCHAK, J.; ANDREWS, W. H. Random simultaneous equations and the theory of production. *Econometrica, Journal of the Econometric Society*, JSTOR, p. 143–205, 1944.

MARTIN, L. A.; NATARAJ, S.; HARRISON, A. E. In with the big, out with the small: Removing small-scale reservations in india. *American Economic Review*, v. 107, n. 2, p. 354–86, 2017.

MIDRIGAN, V.; XU, D. Y. Finance and misallocation: Evidence from plant-level data. *American economic review*, v. 104, n. 2, p. 422–58, 2014.

OLLEY, G. S.; PAKES, A. "the dynamics of productivity in the telecommunications equipment industry". *Econometrica*, v. 64, n. 6, p. 263–97, 1996.

OZLER, S.; YILMAZ, K. Productivity response to reduction in trade barriers: evidence from turkish manufacturing plants. *Review of World Economics*, Springer, v. 145, n. 2, p. 339–360, 2009.

PAVCNIK, N. Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of Economic Studies*, Wiley-Blackwell, v. 69, n. 1, p. 245–276, 2002.

RESTUCCIA, D.; ROGERSON, R. Misallocation and productivity. *Review of Economic Dynamics*, v. 1, n. 16, p. 1–10, 2013.

RESTUCCIA, D.; ROGERSON, R. The causes and costs of misallocation. *Journal of Economic Perspectives*, v. 31, n. 3, p. 151–74, 2017.

ROBERTS, M. J.; TYBOUT, J. R. Industrial evolution in developing countries: a preview. *Industrial evolution in developing countries: Micro patterns of turnover, productivity, and market structure*, p. 1–15, 1996.

SYVERSON, C. *Market Structure and Productivity*. Tese (Doutorado) — Ph. d. dissertation, University of Maryland, 2001.

TOPALOVA, P.; KHANDELWAL, A. Trade liberalization and firm productivity: The case of india. *Review of economics and statistics*, MIT Press, v. 93, n. 3, p. 995–1009, 2011.

WOOLDRIDGE, J. M. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, Elsevier, v. 104, n. 3, p. 112–114, 2009.

YU, M. Processing trade, tariff reductions and firm productivity: evidence from chinese firms. *The Economic Journal*, Wiley Online Library, v. 125, n. 585, p. 943–988, 2015.

APÊNDICE A – APPENDIX TABLES

Tabela A.1 – Coefficient Estimates (OP) when changing distortion value and percentage of firms affected by distortion.

	Percentage of Firms being affected				
	10%	25%	50%	75%	90%
Capital Coefficient Estimates					
$\phi^1 = 0.2$	0.283	0.227	0.155	0.126	0.142
$\phi^1 = 2$	0.325	0.312	0.299	0.292	0.310
$\phi^1 = 5$	0.282	0.242	0.202	0.188	0.225
$\phi^1 = 10$	0.241	0.197	0.149	0.137	0.167
Labor Coefficient Estimates					
$\phi^1 = 0.2$	0.699	0.745	0.804	0.836	0.817
$\phi^1 = 2$	0.670	0.685	0.704	0.717	0.710
$\phi^1 = 5$	0.708	0.752	0.787	0.818	0.814
$\phi^1 = 10$	0.744	0.795	0.829	0.862	0.862

Tabela A.2 – Coefficient Estimates (LP) when changing distortion value and percentage of firms affected by distortion.

	Percentage of Firms being affected				
	10%	25%	50%	75%	90%
Capital Coefficient Estimates					
$\phi^1 = 0.2$	0.375	0.319	0.181	0.133	0.168
$\phi^1 = 2$	0.412	0.384	0.336	0.317	0.296
$\phi^1 = 5$	0.386	0.335	0.242	0.181	0.166
$\phi^1 = 10$	0.374	0.301	0.190	0.140	0.128
Labor Coefficient Estimates					
$\phi^1 = 0.2$	0.614	0.666	0.805	0.860	0.834
$\phi^1 = 2$	0.601	0.611	0.660	0.704	0.710
$\phi^1 = 5$	0.609	0.649	0.772	0.838	0.843
$\phi^1 = 10$	0.617	0.680	0.820	0.881	0.885

Tabela A.3 – Coefficient Estimates (WD) when changing distortion value and percentage of firms affected by distortion.

	Percentage of Firms being affected				
	10%	25%	50%	75%	90%
Capital Coefficient Estimates					
$\phi^1 = 0.2$	0.363	0.343	0.214	0.182	0.203
$\phi^1 = 2$	0.378	0.391	0.382	0.310	0.319
$\phi^1 = 5$	0.378	0.400	0.277	0.182	0.159
$\phi^1 = 10$	0.382	0.376	0.207	0.119	0.094
Labor Coefficient Estimates					
$\phi^1 = 0.2$	0.631	0.654	0.817	0.874	0.832
$\phi^1 = 2$	0.625	0.626	0.660	0.715	0.729
$\phi^1 = 5$	0.628	0.635	0.788	0.867	0.888
$\phi^1 = 10$	0.629	0.660	0.854	0.921	0.936

Tabela A.4 – Coefficient Estimates from Monte Carlo Simulations

	Without Distortion				With Distortion			
Meas. Error	Mean(β_k)	SD(β_k)	Mean(β_l)	SD(β_l)	Mean(β_k)	SD(β_k)	Mean(β_l)	SD(β_l)
LP								
0.0	0.404	0.009	0.598	0.003	0.174	0.012	0.839	0.006
0.1	0.407	0.013	0.624	0.003	0.175	0.013	0.842	0.006
0.2	0.686	0.094	0.684	0.003	0.179	0.014	0.851	0.005
0.3	0.605	0.010	0.745	0.003	0.196	0.045	0.862	0.005
WD								
0.0	0.389	0.026	0.627	0.009	0.184	0.014	0.869	0.007
0.1	0.419	0.025	0.650	0.007	0.184	0.015	0.872	0.007
0.2	0.421	0.021	0.711	0.005	0.182	0.015	0.880	0.007
0.3	0.390	0.024	0.776	0.005	0.177	0.016	0.891	0.006
ACF								
0.0	0.407	0.020	0.596	0.005	0.239	0.068	0.711	0.023
0.1	0.430	0.073	0.587	0.020	0.201	0.060	0.722	0.021
0.2	0.491	0.081	0.562	0.021	0.120	0.030	0.738	0.009
0.3	0.305	0.182	0.594	0.038	0.065	0.034	0.742	0.012

Tabela A.5 – Coefficient Estimates from Monte Carlo Simulations

	Without Distortion				With Distortion			
Meas. Error	Mean(β_k)	SD(β_k)	Mean(β_l)	SD(β_l)	Mean(β_k)	SD(β_k)	Mean(β_l)	SD(β_l)
LP								
0.0	0.504	0.009	0.498	0.003	0.284	0.033	0.778	0.008
0.1	0.521	0.045	0.531	0.003	0.292	0.046	0.782	0.008
0.2	0.800	0.086	0.604	0.003	0.334	0.082	0.793	0.007
0.3	0.711	0.027	0.676	0.004	0.395	0.074	0.807	0.006
WD								
0.0	0.499	0.023	0.516	0.008	0.260	0.016	0.805	0.009
0.1	0.544	0.026	0.549	0.005	0.261	0.016	0.810	0.009
0.2	0.556	0.029	0.628	0.005	0.259	0.016	0.822	0.008
0.3	0.529	0.036	0.704	0.006	0.256	0.018	0.836	0.007
ACF								
0.0	0.494	0.006	0.499	0.003	0.386	0.109	0.607	0.049
0.1	0.428	0.010	0.510	0.003	0.340	0.107	0.625	0.048
0.2	0.315	0.017	0.529	0.004	0.219	0.072	0.667	0.022
0.3	0.144	0.023	0.532	0.009	0.132	0.039	0.691	0.010

APÊNDICE B – APPENDIX FIGURES

Figura B.1 – Bias in estimated TFP and capital share (OP)

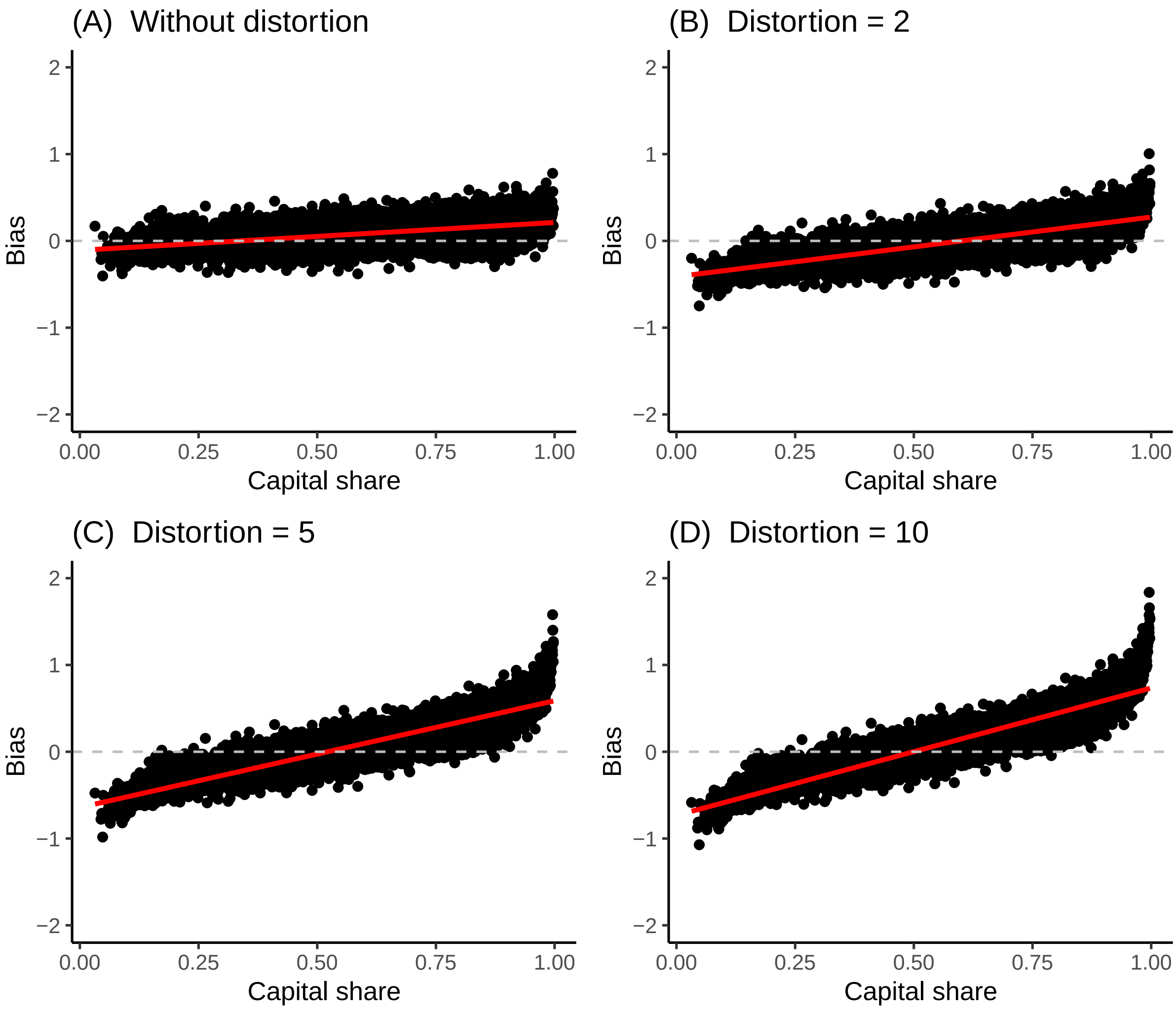


Figura B.2 – Bias in estimated TFP and capital share (LP)

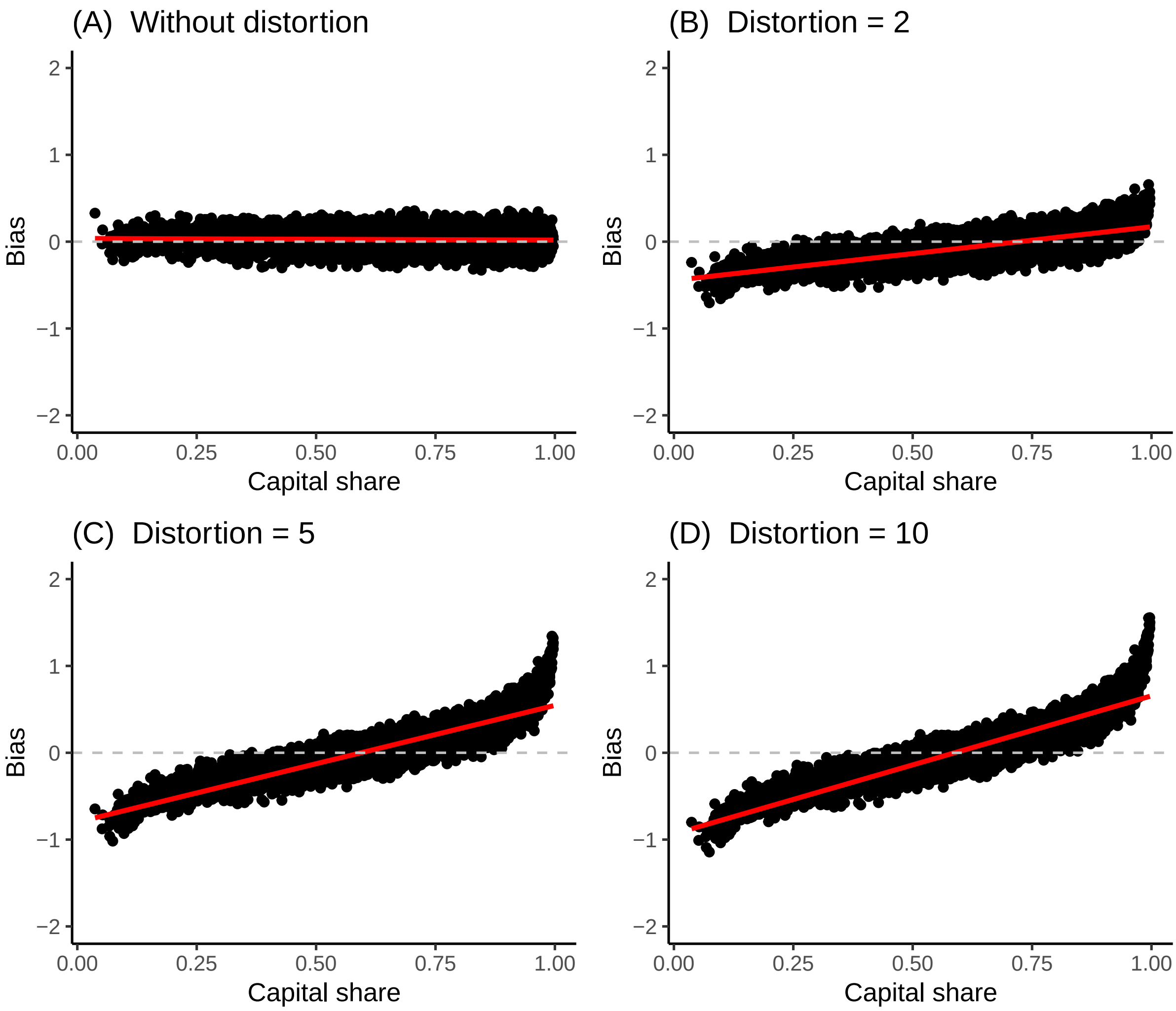


Figura B.3 – Distributions of productivities (OP)

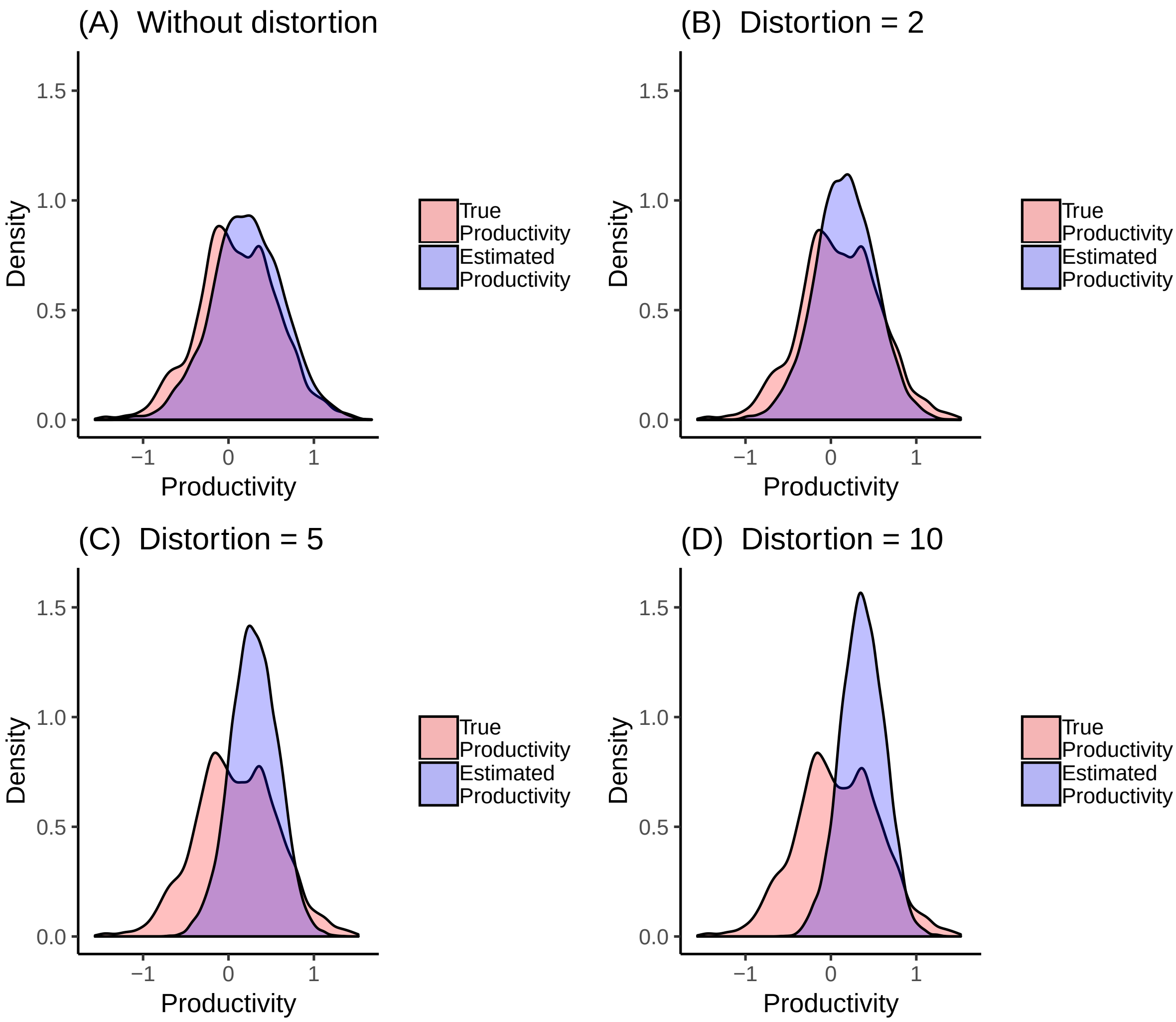


Figura B.4 – Distributions of productivities (LP)

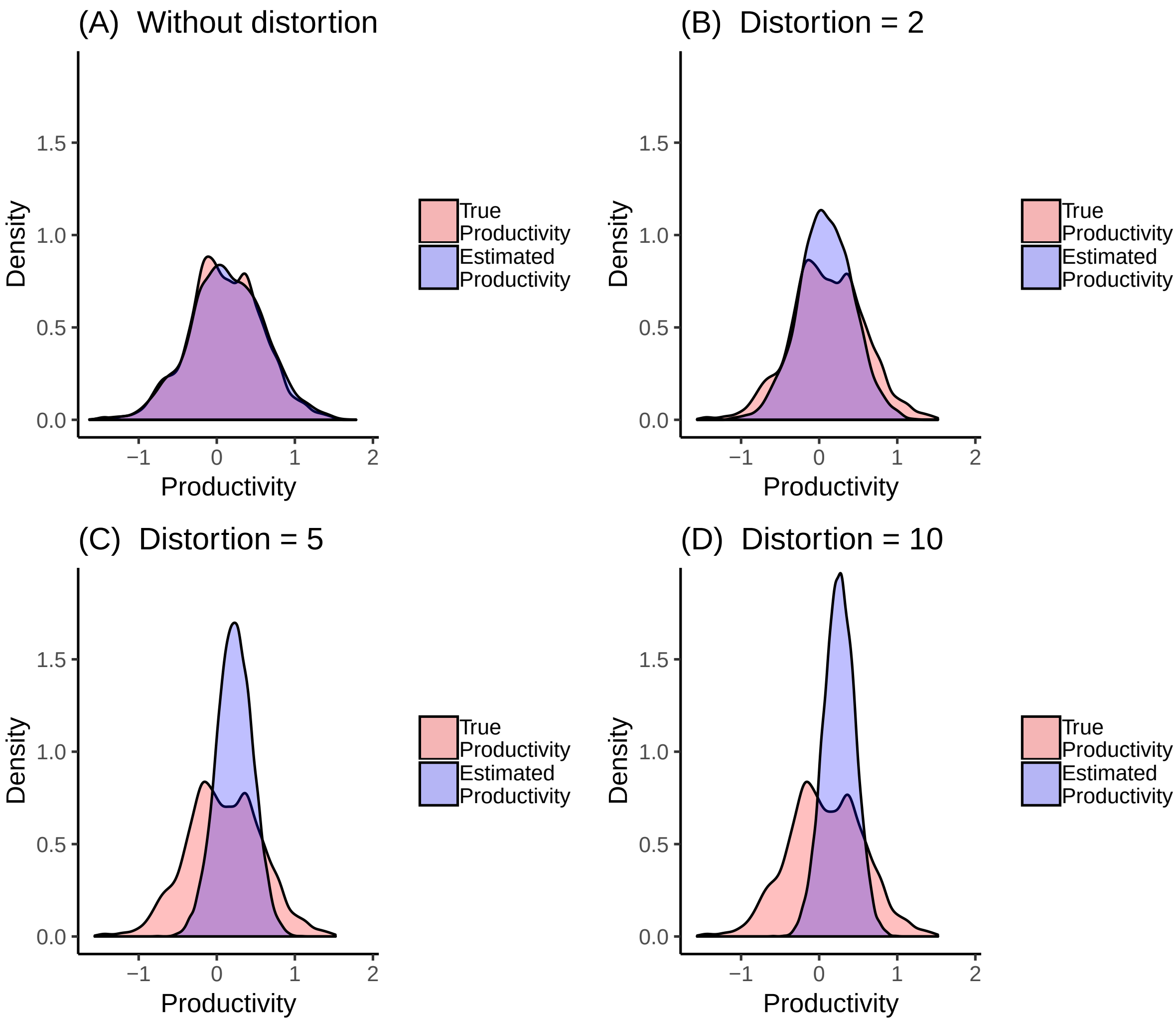


Figura B.5 – Changing Distortion Value and Relative Bias (OP)

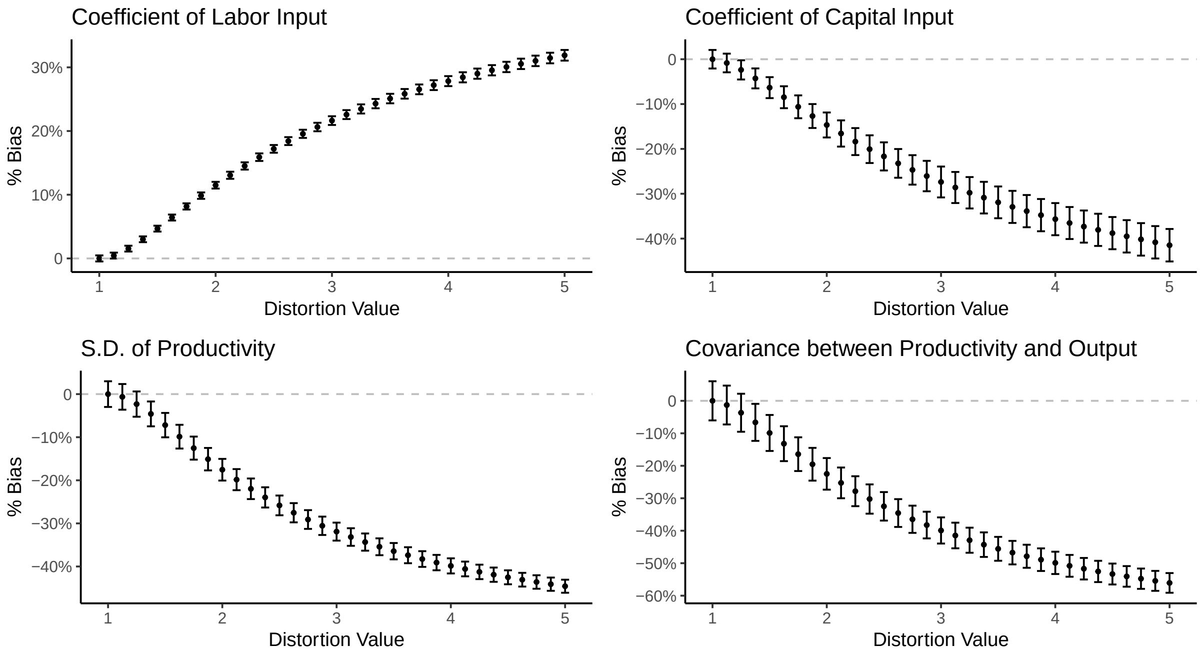


Figura B.6 – Changing Percentile of Firms affected and Relative Bias (OP)