Agglomeration Economies and Labour Misallocation in Cote d'Ivoire

BAH, Mamadou Mouminy

University of Felix Houphouët Boigny (UFHB), African Economic Research Consortium (AERC)

15 August 2021

Agglomeration Economies and Labour misallocation in Cote d'Ivoire

BAH Mamadou Mouminy
University of Felix Houphouët Boigny (UFHB), Abidjan, Cote d'Ivoire
Email : bah.mamadoumouminy@gmail.com

Abstract :

This paper analyses the effects of agglomeration economies on firm labour misallocation, using the Ivorian firm data from 2013-2016. After measuring the degree of firm labour misallocation in the first step, we assess the level of labour misallocation in denser regions in the second step. The results show on the one hand that the average labour misallocation (labour gap) at the firm level is 2,825,887 FCFA ($5,137.97[1]) over the period 2013-2016 and this gap has significantly decreased over years. On the other hand, firms located in denser regions exhibit lower labour misallocation. In terms of the magnitude, both localisation and urbanisation economies are large and statistically significant. A 10% increase in the degree of localisation in a region reduces the labour misallocation by 7.41% on average, while a 10% increase in the degree of urbanisation alters the labour misallocation by 4.26%. These findings confirm that labour misallocation has a geographical dimension, in addition to the firm characteristics. A sound policy needs to accounts for the spatial distribution of firms and the creation of active poles of development in major Ivorian regions.

Keywords: Localisation, Urbanisation, Misallocation, Total factor productivity, firm-level data
JEL code: R3, D24, L25, O4

[1] $1=550 FCFA

1- Introduction

The geographic concentration of workers and firms in the same industry or not in one area makes that area more productive, on average (Combes et al., 2012a). This can be explained by firm selection and agglomeration economies. Firm selection in larger regions may prevent the weaker firms to survive in the markets due to the level of competitiveness, while agglomeration economies allow firms located in larger regions to benefit from some productive advantages. Combes et al., (2012a) distinguish selection from agglomeration externalities and show that firms located in denser areas are, on average, 9.7% more productive than in less dense areas. They suggest that this difference is not related to firm selection but is driven by agglomeration economies.

The advantages of agglomeration economies have long been identified by Marshall (1890). In his view, agglomeration effects arise from three sources: knowledge spillovers, linkages between intermediate and final good suppliers, and labour market interactions. However, Marshall's (1890) classification has been criticised in the literature because it focuses on the channels through which agglomeration effects are observed, rather than the underlying mechanisms that determine these effects. For example, firms in denser areas may learn from other firms by observing what other firms do and copying them, by having a direct supplier/customer relationship with that firm, or by hiring workers from that firm. From this standpoint, Duranton & Puga (2004) consider three mechanisms through which agglomeration economies can occur: sharing (availability of infrastructures, service, and public good), matching (labour market, larger regions or cities can also facilitate the chances of matching) and learning (knowledge or technological spillovers).

There is a large disparity in income across countries. A large part of this disparity can be explained by differences in total factor productivity (TFP) (Hsieh & Klenow, 2010 and Caselli, 2005). From this view, several theoretical models have been developed to explain these differences in TFP. These theoretical models state that the low TFP can be attributed to the theory of resource misallocation across firms (Restuccia & Rogerson, 2008, 2013). According to this theory, frictions such as market imperfections, regulation and corruption prevent the efficient use of resources across firms, thus leading to lower aggregate TFP compared to a frictionless market (zero adjustment costs) situation (Restuccia & Rogerson, 2017). Therefore, to achieve allocative efficiency, firms within-industry can access resources such as capital and labour until their marginal products are equalised. Thus, any difference between marginal products of inputs and its costs across firms is termed resource misallocation. This paper extends this approach and argues that firms' resource misallocation should also be distributed within-industry across regions.

Hsieh & Klenow (2009) measure resource misallocation by the dispersion of TFP revenue (TFPR)[2]. They state that this TFPR does not vary between firms in the same industry unless firms face some types of distortions (for example output distortions or capital distortions[3]). In the absence of distortions, more inputs would be allocated to firms with higher physical productivity (TFPQ) to the point where their higher output results in a lower price and the same TFPR as smaller firms. In contrast, in the presence of distortions, a higher TFPR means that the firm faces barriers that raise its marginal products of inputs making the firm smaller than optimal. Alternatively, Petrin & Sivadasan (2013) propose a methodology to measure misallocation. This is the difference between the value of the marginal product of each input and its cost for the firm. Such difference or gap measures the degree of firm resource misallocation.

From this view, it is easier to look at the mechanism which makes agglomeration economies more productive, rather than focusing on the differences in allocative efficiency. Combes et al., (2012b) point out that, despite the existing frictions in the whole economy such as market imperfections, regulation, corruption, etc., firms located in larger regions can match with more productive and better-paid employees. However, concerning the difference between the value of the wage and the marginal product, a better matching is expected to reduce such difference at the firm level. Thus, the paper focuses on the matching channel and test whether, in larger regions, the thicker labour market also affects firm misallocation.

Cote d'Ivoire is an interesting case to study the problems stemming from resource misallocation. In recent years, after the successive political crises that began with the coup d'état in 1999, leading to a rebellion in 2002 and an electoral post-crisis in 2010, the country has implemented a set of reforms such as trade, fiscal and monetary, etc. These reforms succeeded in achieving macroeconomic stability, political stability, opened the economy up to trade and foreign investment

---

[2] The weighted average of the marginal revenue product of capital and labour. It means that how much revenue can be obtained from the same quantity of inputs (capital and labour).

[3] The output distortions raise the marginal product of capital and labor by the same proportions, while the capital distortions increase the marginal product of capital relative to labor. For example, the greater output distortions, the more firms face restrictions on size, and the lower output distortions, the more firms benefit from government subsidies or other preferential treatment. Similarly, the greater capital distortions, the greater firm has problems with access to credit.

and boosted educational attainment. With an urbanisation rate of 50.3% and an estimated average annual population growth rate of 3.8%[4], most formal firms in Cote d'Ivoire are concentrated in the south (mainly in the district of Abidjan). This concentration is due to Abidjan's status as the main economic hub, hosting one of the largest ports in sub-Saharan Africa, the Port of Abidjan, as well as the port of San Pedro, while the rest of the country is mainly oriented towards agriculture (Fall et al., 2016). This encourages migration to this region and its peripheries. Similarly, with 20% of the population, this district also absorbs 80% of formal employment (Fall et al., 2016). In Cote d'Ivoire, 90 per cent of the manufacturing labour force is employed by small and medium-size manufacturing firms (Africa's Pulse, 2018).

The rate of population growth in other cities has also increased considerably over time. One of the consequences is an increase in market power coupled with a rise in the cost of living in these cities, originally oriented towards agriculture. In addition, the location of firms and workers is based on the benefits and advantages that provide agglomeration externalities. In most economic industries or sectors, the establishment of first movers in a specific location has encouraged other firms to locate there as well. It is, therefore, necessary to analyse how agglomeration economics (localisation economies) affects firm-level labour misallocation.

Combining a novel census firm-level database from 2013 to 2016, the main objective of this study is to analyse the effects of agglomeration economies on firm-level labour misallocation in Cote d'Ivoire. The paper also extends the analysis to consider all sectors of the economy (agriculture, manufacturing, construction, commerce, education, etc.). Specifically, first, it measures the degree of firm labour misallocation (labour gap) using Petrin & Sivadasan (2013)'s methodology. Then, conducts an empirical analysis of the evolution of labour misallocation controlling for firm characteristics. Finally, examines whether the labour misallocation is lower in denser regions.

In this paper, the concept of firm labour misallocation refers to the gap between the value of the marginal product of labour and its marginal cost while the agglomeration economies concept, which occurs when workers and firms benefit from being near to others, will be captured by the two variables: localisation and urbanisation economies. The localisation economies are measured by the number of other employees working in the same industry and the same region. This variable captures the intra-industry externalities. While the urbanisation economies refer to the inter-

[4] According to the population census of Cote d'Ivoire conducted in 2014

industry externalities measuring the number of employees in other industries in the region where the firm is located.

The methodology employed in this paper is, according to the Restuccia and Rogerson (2013 and 2017), an indirect approach to analyzing misallocation. The indirect approach aims at measuring the full degree of misallocation in an economy without detail as to what policies or institutions may be causing it. This indirect approach differs from the direct approach, which analyses the effects of specific and observable distortions such as regulations and taxes on resource misallocation and aggregate productivity. While the direct approach, which has failed so far in finding evidence of distortions that can explain the resource misallocation, the indirect approach has been criticized for two reasons: on the one hand, its estimates of resource misallocation can reflect misspecification of production functions within industries or adjustment costs; on the other hand, estimates from different countries may not be comparable due to measurement error (Restuccia and Rogerson, 2013 and 2017).

The main results are as follows. firstly, the average labour misallocation (labour gap) at the firm level is 2,825,887 FCFA (5137.97 dollars) over the period 2013 - 2016 and this gap has significantly decreased over years when controlling for firm characteristics (age of the firm, size, competition index, etc.). Secondly, firms located in larger regions exhibit lower labour misallocation. In terms of the magnitude of the effect, a 10% increase in the degree of localisation in a region reduces the labour misallocation by 7.41% on average, while a 10% increase in the degree of urbanisation reduces the labour misallocation by 4.26%. Finally, the findings are robust and consistent evidence suggesting that the estimation of the labour gap seems not to be influenced by sample selection or outliers and these findings are not also driven by the functional form of the production function.

The paper contributes to the existing literature in three key ways. First, this study adds to misallocation literature by providing a case study on an African country. Especially, unlike Newman et al. (2019) who provide evidence on the effect of resource misallocation on South Africa manufacturing productivity, many studies on the topic address to the effect of resource misallocation on manufacturing productivity across countries. This study improves the literature by providing a deeper analysis of resource misallocation in an African context by extending the analysis to sectors other than manufacturing. Second, several studies used the methodology of Hsieh & Klenow (2009) which is a static method to measure misallocation within- industry unlike the approach of Petrin & Sivadasan (2013) which proposes a firm-level misallocation measure. This paper applies the latter method and thereby provides a monetary value to this misallocation. Finally, the effects of agglomeration economies on labour misallocation among firms within-industry, which have not been fully investigated to our knowledge in sub-Saharan Africa in general and Cote d'Ivoire in particular, are made; showing that denser areas are associated with lower labour misallocation. Hence the importance of this paper

The rest of the paper is structured as follows: Section 2 deals with the literature review, while section 3 provides the estimation methodology. Section 4 describes the data used in the analysis. The main results are presented in Section 5 along with some robustness checks. Section 6 concludes the paper.

2- Literature review

This section provides an overview of the relevant literature on misallocation. The role of firm-level resource misallocation in explaining aggregate productivity has recently been debated in the literature following the contribution of Hsieh & Klenow (2009) (HK, henceforth), who provide a methodology to assess the degree of misallocation on productivity in a monopolistic competition model.

The basic intuition of HK is that in a context where there is an allocative efficiency, the value of the marginal product of input should equate to its marginal cost. Under the HK methodology, large dispersion in marginal value product of inputs among firms within-industry can be termed as the degree of resource misallocation. Based on the monopolistic competition in product markets with heterogeneous firms, HK construct a model that allows assessing the productivity losses arising from the fact that marginal value products are not equalised across firms. The only reason why the value of the marginal products of inputs is not equalised among them is the presence of dispersion in the factors market. As a result, the dispersion in the marginal product of inputs is a measure of such distortions. HK, in their study, assess the degree of misallocation across manufacturing firm on aggregate productivity in China, India and the United States. They show that removing the distortions could, in principle, lead to an increase in aggregate productivity due to resource reallocation and conclude that TFP gains increase by 30 - 50% in China and 40 - 60% in India if resources were re-allocated to equalised marginal products to United States levels.

Following HK's methodology, several studies confirm the extent of misallocation on TFP for several countries. Examples include Cirera et al., (2020) for four (4) African countries; Inklaar et al., (2017) for 52 low and middle-income countries; Dias et al., (2016) for Portugal; Ha et al., (2016) for Vietnam; Oberfield (2013) for Chile; etc. However, these various studies do not attempt to identify the causes of resource misallocation but instead, focus on providing frameworks to analyse the consequences of misallocation that do appear to exist.

Bartelsman et al., (2013) propose an alternative measure of within-industry misallocation across countries. This is based on a decomposition of productivity originally proposed by Olley & Pakes (1996). They measure the misallocation by the covariance between firm size and productivity. This alternative measure of misallocation relies on the assumption that the higher the covariance the more efficiently resources are allocated across firms. Bartelsman et al., (2013) show significant variation across countries in the extent of within-industry misallocation and find a higher covariance between firm size and productivity in the United States than in Western European and much lower in Eastern European countries.

Later on, Petrin & Sivadasan (2013) propose an alternative methodology to measure the degree of misallocation at the firm level. This is the difference between the value of the marginal product of inputs and its cost. Their approach was used to evaluate the effect of a change in labour market regulation in Chile. They estimate that, between 1982 and 1984, reducing one unit of currency in the firm input gap leads to an increase in the value added of Chilean firms of 0.5%, on average. Fontagné & Santoni (2019) use this approach to measure the degree of misallocation among manufacturing firms in France and find a significant gap of around 9,500 euros at the firm level, on average. This measure is a useful alternative because it does not rely on the assumptions of the more extensive theoretical framework of Hsieh & Klenow (2009) and Bartelsman et al., (2013).

The effect of resource misallocation that this paper attempts to assess in terms of firm-level labour demand extends beyond labour. Restuccia & Santaeulalia-Llopis (2017) study land misallocation and productivity across farms in Malawi and show a large and significant resource misallocation in the agricultural sector due to land market restrictions. They find that reallocation of land to their efficient use among existing farmers would increase agricultural productivity by a factor of 3.6-fold. Similarly, Chen et al., (2017) assess the effects of land markets on misallocation and productivity in Ethiopia. They conclude that land rentals substantially reduce resource misallocation and increase agricultural productivity. These papers show that distortions in farm size explain a significant fraction of cross-country differences in agricultural productivity.

In addition, several studies have examined the determinants of firm performance in sub-Saharan Africa. Using the Ethiopian manufacturing firm data, Siba et al., (2012) analyse connections between agglomeration, firm-level output prices and physical productivity. They find a negative relationship between the agglomeration of firms that produce the same product in the same area and the price of that product in that area. The increasing price competition reduces the gain of firms located in the same area. However, Siba et al., (2012) find a positive relationship between the agglomeration of firms and physical productivity. This relationship is independent of the city size. Hence, the productivity level of a particular city does not depend on that city's physical characteristics. However, to some extent, a firm's location decision depends not only on the surrounding infrastructure but also on market access for reducing costs.

Similarly, Iimi & Rao (2018) explore the relationship between firm location and transport connectivity in Liberia. Transport infrastructure is among the most important factors in increasing firm productivity and supporting local businesses. The concentration of workers and firms has a considerable impact on how firms decide to locate. Their results indicate significant agglomeration economies, meaning that there are externalities of firm location choice around neighbouring areas. As long as firms are located far apart from the primary city centre, their productivity level decreased unless their intercity transport connectivity is improved. These studies have not focused on the role that the concentration of firms or workers in one area may play in explaining labour gap.

Indeed, agglomeration economies are not recent, but empirical research, about how agglomeration economies affect labour misallocation by controlling for firm characteristics and taking into account the non-random location choices of employees and firms which is little known in sub-Saharan Africa in general and Cote d'Ivoire in particular, needs to add knowledge to country and regional case studies. This study attempts to close some gap as the paper assess the effect of agglomeration economies on firm labour misallocation in Cote d'Ivoire.

3- Methodology

3.1 Measuring resource misallocation at the firm level.

To measure the labour misallocation, this paper relies on Petrin & Sivadasan (2013)'s methodology. This is the difference between the value of the marginal product of each input and its cost for the firm. Such difference or gap measures the degree of firm resource misallocation. Thus, labour misallocation refers to the gap between the value of the marginal product of labour and its cost. This gap is computed at the firm level using the estimated coefficients from the TFP analysis.

This section provides an overview of how the value of the marginal product (VMP) and the marginal cost of labour gap can be estimated using firm-level data. The estimation of this gap uses the Cobb-Douglas production function. This function is chosen as it is relatively simple, easy to handle, and generally adopted by other authors dealing with this topic. However, other alternative production function specifications (e.g. the second-order translog production function) will be used to test the robustness checks. The estimated production function for firm *i* at time *t* is:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \upsilon_{it} + e_{it}, \quad t = 1,2,...,T, \quad i = 1,2,...,N \qquad (1)$$

Where y_{it} is the log of value added, l_{it} denotes the log of the number of employees and k_{it} is the log of capital stock. The productivity shock is given as: $\varepsilon_{it} = \upsilon_{it} + e_{it}$ with υ_{it} productivity shock and e_{it} is an error term. Given value for production function and observed input levels, the marginal product of labour is given by:

$$\frac{\partial Q_{it}}{\partial L_{it}} = \beta_l e^{\varepsilon_{it}} \left(L_{it}\right)^{\beta_l - 1} \left(K_{it}\right)^{\beta_k} = \beta_l \frac{Q_{it}}{L_{it}} \qquad (2)$$

where the capitalised variables are levels of the logged variables defined above. Multiplying this marginal product of labour by the firm's output price (P_{it}) yields the value of the marginal product of labour:

$$VMP_{it}^{L} = p_{it} \frac{\partial Q_{it}}{\partial L_{it}} = \beta_l \frac{P_{it} Q_{it}}{L_{it}} \qquad (3)$$

To control price variation over time at the sector level, a year dummies variable is introduced into equation (1) as the firm's output price is generally not available in the data. Finally, the absolute value of the labour gap, the degree of labour misallocation at the firm level, is given by:

$$\left|Gap_{it}^{L}\right| = \left|VMP_{it}^{L} - w_{it}\right| \tag{4}$$

Where w_{it} is the marginal wage of workers in firm *i*. Since the marginal wage is not usually observed at the firm level, the average wage is used as a proxy. This gap is in nominal terms, so it will be deflated using the consumer price index (CPI)[5]. The GDP deflator will be used as an alternative measure to the CPI. The absolute value Gap_{it}^{L} expresses the increase in value added, induced by labour allocative efficiency. Therefore, in a world where the factors are allocated efficiently, firms within-industry (or sector) can access resources (e.g. by demanding labour) until their marginal products are equalised, thus closing the gap. Any distortions that vary this equilibrium imply a resource misallocation. These distortions may be due to market imperfections, regulatory policies, corruption, labour market rigidities, etc.

To measure labour gap, the first step is to estimate a production function and firm-level TFP. This estimate will allow obtaining the parameters (β_k et β_l). Among the recent methods used to estimate production functions[6], this paper uses the one developed by Wooldridge (2009). One of the main advantages of Wooldridge (2009)'s method is to implement Levinsohn & Petrin (2003)'s methodology in a General Method of Moments (GMM), which takes into account potential contemporaneous error correlation among the two steps procedure, as well as heteroscedasticity and autocorrelation. A brief description of Wooldridge (2009)'s method is provided in Appendix A.

The empirical investigation is conducted through two steps. The first step is to compute the degree of labour misallocation, based on firm productivity estimates. The second step performs an empirical analysis of the effects of agglomeration economies on firm labour misallocation. This has two objectives: first, the evolution of labour misallocation, controlling for firm characteristics is analysed. Then, test whether firms located in denser regions exhibit lower labour misallocation.

[5] $absolute\ real\ gap \equiv RG_{it} = \dfrac{\left|Gap_{it}^{L}\right|}{CPI_t}$

[6] Generally, the production function is estimated with the ordinary least squares (OLS) method, Olley & Pakes (1996) method, Levinsohn & Petrin (2003) method, Wooldridge (2009) method, and more recently Ackerberg et al., (2015) method. All these methods are robust depending on the data used, although some are more rigorous than others.

3.2 Empirical methodology

We consider a one-way error components unbalanced panel data model defined as follows:

$$y_{it} = \alpha + x'_{it}\beta + u_{it}\ ,\ i = 1,...,N;\ t = 1,...,T_i \qquad (5a)$$

The error term is specified as:

$$u_{it} = \mu_i + \vartheta_{it} \qquad (5b)$$

where y_{it} denotes the absolute labour gap of firm *i* at time *t*. μ_i denotes the unobservable individual-specific effect, ϑ_{it} represents the remainder disturbance and α is a scalar. This model is unbalanced in the sense that there are *N* firms observed over varying time period length $\left(T_i \text{ for } i = 1,...,N\right)$. x_{it} is a $K \times 1$ vector of explanatory variables and, β is a K-vector of parameters.

In vector form Equation (5a) can be written as:

$$Y = \alpha\iota_n + X\beta + u = Z\delta + u \qquad (6a)$$

The error term in vector form yields:

$$u = Z_\mu \mu + \vartheta \qquad (6b)$$

Where y and X are of dimensions $n \times 1$ and $n \times K$ respectively, where $n = \sum_{i=1}^{N} T_i$, $Z = \left[\iota_n,\ X\right]$ and $\delta' = \left(\alpha',\ \beta'\right)'$, ι_n is a vector of ones of dimension n , $Z_\mu = diag\left(\iota_{T_i}\right)$, ι_{T_i} is a vector of ones of dimensions T_i., β is $K \times 1$.

In the fixed effect model, the μ_i are assumed to be fixed parameters to be estimated and the remainder disturbances with ϑ_{it} independent and identically distributed $IID\left(0, \sigma_\vartheta^2\right)$; and $E\left(u_{it}\right) = 0$ $Var(u_{it}) = \sigma_u^2$, $\forall i,t$. The x_{it} are assumed independent of the ϑ_{it} for all *i* and *t*. One can substitute the disturbances given by u into (6) to get fixed effect model:

$$Y = Z\delta + Z_\mu \mu + \vartheta \qquad (7)$$

And note that Z is $n \times \left(K + 1\right)$ and Z_μ, the matrix of individual dummies, is $n \times N$. If *N* is large, Equation (7) will include too many individual dummies and β and μ are the parameters of interest.

However, in the random effect model the μ_i are assumed to be random. In this case, $\mu_i \sim IID\left(0, \sigma_\mu^2\right)$ and independent of $\vartheta_{it} \sim IID\left(0, \sigma_\vartheta^2\right)$; The x_{it} are independent of the μ_i and ϑ_{it} for all *i* and *t*. The random effect model can be written as:

$$Y = Z\delta + u \tag{8}$$

where

$$u = Z_\mu \mu + \vartheta$$

Under the above assumptions, the disturbance covariance matrix $E(uu')$ can be written as:

$$\Omega = E\left(\vartheta\vartheta'\right) + Z_\mu E\left(\mu\mu'\right)Z'_\mu = I_n + \rho Z_\mu Z'_\mu = diag(E_{T_i}) + diag\left[\left(1+\rho T_i\right)\bar{J}_{T_i}\right]$$

Where $\rho = \sigma_\mu^2 / \sigma_\vartheta^2$, I_n is an identity matrix of dimension n, $E_{T_i} = I_{T_i} - \bar{J}_{T_i}$, and $\bar{J}_{T_i} = J_{T_i} / T_i$ with J_{T_i} being a matrix of ones of dimensions T_i [7].

The empirical analysis of the evolution of labour misallocation controlling for firm characteristics is specified as:

$$Y_{it} = \alpha_0 + \phi t + X'_{it}\beta + \mu_i + \vartheta_{it} \tag{9}$$

where y_{it} denotes the absolute labour gap of firm *i* at time *t*. μ_i are firm fixed effects to control for unobserved heterogeneity, ϑ_{it} represents the remainder disturbance $\vartheta_{it} \sim IID\left(0, \sigma_\vartheta^2\right)$, α_0 is a scalar and t is a time trend. X_{it} is a vector of time-varying firms characteristics, such as firm age (linear and squared[8]), the firm size (small, medium, and large firms, with micro being considered as a reference)[9], an index for the degree of competition a firm of the industry *s* at time *t* ($comp_{it}^s$), which is measured as follows: $\ln\left(1/\text{HH}\right)_{st}$ where HH is the Herfindahl-Hirschman index of employees concentration by industry *s* and year *t*. Increases in the HH generally indicate a decrease in competition and increasing market power, whereas decreases indicate the opposite[10].

[7] See Baltagi & Chang (1994) and Baltagi (2021) for a more detailed demonstration

[8] This term captures the curvature of age.

[9] The firm size is divided into 4 groups: micro (1-9 employees), small (10-49 employees), medium (50-299 employees) and large (300 or more employees) following the World Bank stratification.

[10] Herfindahl Hirschman Index (HH) of employment concentration inside industry s and year t:

Finally, the last step examines whether firms located in larger regions have lower labour misallocation, controlling for firm characteristics and taking into account the non-random location choices of employees and firms. To achieve this, a measure of the agglomeration economies is included in the estimated equation (8) and enriches the set of controls. The estimated equation becomes:

$$\ln\left(Y_{irst}\right) = \gamma loc_{it}^{rs} + \varphi urb_{it}^{rs} + X_{it}'\beta + \xi_{rs} + \mu_i + \vartheta_{irst} \qquad (10)$$

where: $\ln\left(Y_{irst}\right)$ represents the log of labour gap of firm *i* in a region *r* by industry *s* and year *t*. μ_i are firm fixed effects to control for unobserved heterogeneity, ξ_{rs} denotes region *r* and industry *s* fixed effects, which control for any unobservable, time-invariant characteristics of the local labour market. ϑ_{irst} represents the remainder disturbance. loc_{it}^{rs}, for each firm, the number of other workers in the same industry and the same region. Its captures the localisation economies (intra-industry externalities). urb_{it}^{rs} is the number of workers in other industries in the region where the firm is located. This variable captures the urbanization economies (inter-industry externalities). These two variables (localisation economies and urbanisation economies) measure agglomeration economies. X_{it} represents a set of control variables for firms and industries (see equation 9).

Estimation strategy

Having discussed the fixed effect and random effect models and their underlying assumptions and, before also choosing the estimation strategy, which one to choose? For this, the Hausman (1978) test which is based on the difference between fixed and random effects estimators is used (see result section 5.3 and Appendix B). When the null hypothesis of no correlation between the individual effects and the X_{it} is rejected the fixed effects model will be used instead of the random effects model. Thus, in this case, the estimation strategy follows Guimaraes & Portugal (2010) and Carneiro et al., (2012)'s iterative approach for the estimation of the model with a high dimensional

---

$HH_{st} = \sum_{i \in S_t^s} \left(\frac{employees_{it}}{employees_{st}} \right)^2$ where S_t^s is the set of firms belonging to industry s at time t. The HH is a common measure of market concentration and is used to determine market competitiveness. An HH below 1% indicates a highly competitive industry. An HH below 15% indicates an unconcentrated industry, an HH of 15% to 25% indicates moderate concertation. An HH above 25% indicates a high concentration. For value equal to 100% means that all employees are fully concentrated in one industry.

fixed effect (HDFE)[11]. Controlling for firm-specific and region-industry effects requires the introduction of HDFE in the linear regression model (equations 9 and 10). The HDFE is used for three reasons: 1) the database using in this paper are getting larger, 2) estimation of the model (equations 9 and 10) with many observations and variables poses new challenges and 3) with high dimensional models explicit introduction of dummy variables to account for fixed effect is not an option because the number of units or groups for either firms or industry is large (Guimaraes & Portugal, 2010). In contrast, when the null hypothesis of no correlation between the individual effects and the X_{it} is not rejected the random effects model will be used instead of the fixed effects model.

The next section presents the dataset of Ivorian companies to be used for the analysis of agglomeration effects by means of the Petrin & Sivadasan (2013) methodology.

4- Data and summary statistics

4.1. Data and Variables

Data for this study come from the National Institute of Statistics (INS) of Cote d'Ivoire[12]. It covers the period from 2013 to 2016. The dataset covers all registered firms operating in all sectors, including agriculture, manufacturing, and services and contains detailed balance sheet information on firms' sales, value added, employment, capital stock, intermediate inputs (telecommunication fees or corporal immobilisation), labour cost (wages, benefit and bonus), etc. and information about the firm's location. The dataset also includes information on the firm's main industry of operation based on the "*nomenclatures communes pour les activités*" (NAEMA) rev1. In this paper, industries are defined at the 2-digit NAEAMA rev1 (equivalent to International Standard Industrial Classification of all economic activities-ISIC Rev4). Overall, this classification implies 39 industries in 12 sectors. Firm location follows the 2-digit level of administrative repartition. The paper follows the country's new administrative division, which contains 32 regions. The agglomeration variables are created following Martin et al., (2011):

[11] HDFE is the regression absorbing multiple levels of fixed effects

[12] National Institute of Statistics is the structure mandated to build up the database in Cote d'Ivoire

Localisation: for each firm, the number of other workers in the same industry and the same region is computed. This variable captures the intra industry externalities. For a firm *i* located in region r and operating in industry s, then localisation economies variable is defined as:

$$loc_{it}^{rs} = \ln(employees_t^{rs} - employees_{it}^{rs} + 1)$$

Urbanisation refers to the inter-industry externalities measuring is the number of workers in other industries in the region where the firm is located:

$$urb_t^{rs} = \ln(employees_t^{r} - employees_t^{rs} + 1)$$

A constraint of the procedure used to estimate the production function and TFP requires the use of lagged variables. This means that firm must be present in at least two consecutive periods for it to be included in the analysis. Eliminating observations that do not meet these criteria leads to just less than half of the viable firms exiting the sample each year. Finally, after excluding all missing, zero or negative values for value added, employees, capital stock, intermediate inputs (materials), labour cost (wages, benefit and bonus), and cleaning the data by cutting the top and bottom 1% of firms for the value added to the capital ratio in each year to eliminate outliers, we have an unbalanced panel of 20,533 firm-year observations (or 7,483 firms) over the period 2013-2016 (see appendix C, table C1). This data has its limitations. The most important one is the short time period covered. The firm's export status is not available. Some of the input data, such as materials, are not available for all years.

4.2. Summary statistics

This section presents the summary statistics. According to Table 1, the standard deviation is higher than the mean for most variables, thus implying a strong dispersion. Secondly, the average labour misallocation of the firm in the sample is 2825.887 thousand FCFA ($5137.98). The minimum value for the localisation variable is zero, which means that: some firms are the sole representative of their industry in their region. For these firms, there are therefore no localisation economies[13]. In the same table, the minimum value of the firms' number of employees is equal to 1 and the maximum is equal to 9,506.

[13] Since $localisation_{rst} = \ln(employees_{rt} - employees_{rst} + 1)$, $localisation_{rst} = 0$ when $employees_{rt} = employees_{rst}$

Table 1: Descriptive statistics

Variables	Mean	Std. Dev.	Min	Max
labour misallocation $\left\|Gap_{it}^{L}\right\|$	2825.887	4933.072	15.07218	33143.58
Value added	420615.6	3490037	17	1.88e+08
Capital	171971.1	1395698	11	5.14e+07
Labour	43.09166	252.4867	1	9506
Localisation	7.702398	2.261285	0	10.3439
Urbanisation	10.91067	2.369676	0	12.40292
Competition index	3.307005	1.036003	0.6069695	5.251218
Age[14]	9.174159	10.09166	0	75
Micro[15]	0.5954804	0.4908108	0	1
Small	0.2884625	0.4530584	0	1
medium	0.0931671	0.2906735	0	1
large	0.02289	0.1495564	0	1
Number of observations		**20,533**		

<u>Note:</u> *labour misallocation, Value-added and capital are expressed in thousands of FCFA.*
<u>Source:</u> *Authors' calculations*

The index for the degree of competition is equal 3.30 on average which means that there is an unconcentrated industry (or competitive industry) in the sample and this competition index ranges from 0.607 for mining and quarrying (higher concentrated industry) to 5.251 for education (highly competitiveness). The average age of firms in the sample is 9 years with the oldest firm being 75 years old. Finally, 59.54% of the firms are micro, meaning that they employ between 1 and 9 full-time employees

5- Results and robustness checks

5.1- Production function and TFP estimation

The coefficient estimates ($\hat{\beta}_l$ and $\hat{\beta}_k$) of the production function are presented in Appendix C, Table C3. The estimation is conducted separately for each sector. The results show significant differences in the estimated $\hat{\beta}_l$ and $\hat{\beta}_k$ across sectors. These differences reflect the heterogeneity of the production technologies used. At the sectoral level, the coefficient estimates appear reasonable and are always positive. The labour has the highest coefficient in all sectors compared to the capital. These coefficient estimates range from 0.361 to 0.893 for labour and 0.089 to 0.342 for capital. The estimated returns to scale are in the range of 0.687 to 0.994, which implies decreasing returns to scale. In the frictionless (zero adjustment cost) world, sufficient conditions

[14] Years of firm's operation.
[15] Equal to 1 if the firm is in micro size; 0 otherwise

for the optimal choice of input would require decreasing returns to scale. Figure C1 (see Appendix C) plots the distribution of firm-level TFP for the period 2014 and 2016. Generally, firm productivity increases slightly over the years (i.e. an increase of 0.25% between 2013 and 2016). In addition, the shift of this distribution to the right indicates a significant redistribution of firms to more productive levels. However, this does not mean that the use of the resource is increasingly close to allocative efficiency.

Despite this productivity growth, resource misallocation prevents the benefits of technical progress from being fully realised. Figure C1 also shows that there is a large dispersion of the TFP among firms. To have a clear explanation of this dispersion, if we consider one measure of dispersion, the interquartile ratio (the difference between the 75th and the 25th percentile), we find that this difference increases across firms within-industry (sectors). For example, as shown in Appendix C, Table C4, The interquartile ratio is 1.057 (or 187.7%)[16] in 2014, which means that a firm ranked in the 75th percentile is 187.7% more productive than a firm ranked in the 25th percentile. In 2016, this ratio increased to 1.079 (or 194.17%), which shows that the productivity improvement is not due to a re-allocation of resources (labour and capital). The dispersion of TFP across sectors is also examined (see appendix C, Table C4). The results show that the dispersion across sectors is even significantly higher. This means that at first view there is a resource misallocation.

The following section presents the labour misallocation at the aggregate and the sectoral levels.

5.2- Labour misallocation

Once the marginal product of labour coefficient has been estimated, the calculation of the labour misallocation is straightforward from equations (4). Recall that resource misallocation implies that a higher TFP is a sign that the firm faces barriers that raise the firm's marginal products of inputs, making the firm smaller than optimal: the estimated value of the marginal product of labour is higher than its cost and the gap is positive because there are some distortions in the labour market that prevent these firms from raising their level of production. Thus, the sign of the gap is important since it helps to distinguish positive and negative gaps. The summary statistics on the labour gaps by sectors, as well as for the overall sample are presented in Appendix C, Table C5.

[16] The difference between the 75th and 25th percentile firms is 1.057 in 2014, which corresponds to a TFP ratio of $e^{1.057}-1 = 1.877$

The results reveal that the labour misallocation (labour gap) at the firm level is 2,825,887 FCFA ($5137.98[17]), on average over the period 2013-2016, with a dispersion considerably higher both between and within sectors (the standard deviation is higher than the mean, the coefficient of variation is 1.745[18]). This gap varies significantly across sectors, ranging from 1,159,849 FCFA ($2,108.816) in the education to 11,315,760 FCFA ($20,574.11) in the extraction (mining and quarrying) sector. The share of firms with a positive gap over the whole period is 73.63%. The labour gap by industry is presented in Appendix C, Table C6. The results show that the gap varies considerably even across industries.

5.3- Evolution of the firm-level labour misallocation

This section presents the results of the second step, which is to analyse the evolution of labour misallocation controlling for firm characteristics. To analyse the evolution of labour misallocation the first step is to test the model (fixed effects vs random effects model). For this, the Hausman (1978) test is used. Thus, the resulting Hausman test statistic yields an observed (Chi2) χ_7^2 500.81 (see Appendix B). This is significant at the 1% level and the null hypothesis of no correlation between the individual effects and the X_{it} is rejected. This implies that the fixed effects model is used instead of the random effects model in this study.

The main results for the evolution of labour misallocation (equation 9) are reported in Table 2. In all the columns of Table 2, standard errors are double clustered at the firm level to deal with serial correlation and at the industry-year level to control for cross-sectional dependence. The result shows that the time variable (trend) is statistically significant at 1%. During the sample period (2013-2016), labour misallocation decreases by an average of 556,880 FCFA ($1012.51) per year. In column 2, the estimates imply that labour misallocation increases with a firm's age until a peak at 20 years[19] and then decline. Firm size distribution dummies (Column 3) show that larger firms tend to exhibit lower labour misallocation, consistently with the empirical evidence showing that larger firms tend to be more productive.

[17] $1=550 FCFA

[18] Coefficient of variation CV=2825.887/4933.072=1.745

[19] $\left(130.060/\left(2*3.214\right)\right)$

Table 2: Evolution of labour misallocation, 2013-2016, baseline results

Dep. Variable	labour misallocation $\left\|Gap_{it}^{L}\right\|$			
	(1)	(2)	(3)	(4)
Trend	-559.66***	-621.957***	-543.021***	-556.880***
	(90.96)	(101.930)	(101.874)	(84.142)
Age		130.060***	121.304***	120.080***
		(41.984)	(40.548)	(41.100)
$(Age)^2$		-3.214**	-2.960**	-2.863**
		130.060***	121.304***	120.080***
Small			-2,108.818***	-2,124.552***
			(181.428)	(181.760)
Medium			-4,942.093***	-4,972.546***
			(466.812)	(462.070)
Large			-8,080.374***	-8,137.593***
			(851.000)	(836.061)
Competition				-872.916***
				(225.354)
Observations	20,533	20,533	20,533	20,533
R-squared	0.74	0.745	0.756	0.758
Fixed effect	i	i	i	i
Cluster level	i & st	i & st	i & st	i & st
Number of firms	7483	7483	7483	7483

*Note: standard errors (reported in parentheses) are clustered at firm i and industry-year st level. Significant at: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable: labour misallocation. All gaps are in thousands of FCFA (deflator used is the CPI).*
Source: Authors' estimations based on the firm-level census data.

The competition index appears to have a negative and significant effect on labour misallocation meaning that the misallocation decreases as the degree of competitiveness of the industry increases. Generally, the decline of the gap is more pronounced when controlling for firm characteristics and the standard errors are clustered at firm *i* and industry-year *st* level.

One explanation of this diminution of the labour misallocation can explain by the performance recorded in recent years after the post-electoral crisis in 2010. These performances include political stability, macroeconomic stability, trade openness, massive public investment, etc. These facts guide policymakers in designing policies that trigger private sector participation. Place-based policies need to be taken into account because the physical distribution of economic activities is important.

5.4- Robustness check for the evolution of labour misallocation

This section examines the robustness of the results on the evolution of labour misallocation reported in Table 2. The sample selection, the use of an alternative deflator for nominal gap and the alternative specification of the production functions are used to check the robustness of the results reported in Table 2. The main results are robust to all these checks.

a- Sample selection and the GDP deflator

To check the robustness of the results on the evolution of labour misallocation, the sample to firms with at least 10 employees or small firms with less than 10 employees is restricted. The results are presented in Appendix C, Table C7. The decline of labour misallocation per year is large and significant when the sample to firms with at least 10 employees (column 1) or small firms with fewer than 10 employees (column 2) is restricted. In the baseline analysis (Table 2), the CPI was used as a deflator to denote gaps in FCFA. The results from using the GDP deflator are presented in table 6 column 3 and are very similar to using the CPI deflator. Generally, these results (Appendix C, Table C7) find robust and consistent evidence suggesting that estimation of the gap seems not to be driven by the sample selection in the firm-level data or by using an alternative deflator.

b- Alternative production function specifications

The same specification is used as in equation (1), but estimate it using firm-level fixed effects that vary by period. This estimator is consistent if $\upsilon_{it} = \upsilon_{ip}$, where *p* represents one of the three periods (2013, 2014, 2015 and 2016). The results presented in column 1 of Table C8 (see Appendix C) are similar to those in the baseline Table 2. As in the base case, the labour misallocation decreases per year.

One of the weaknesses of the Cobb-Douglas production function (Eq. 1) is that the elasticities of capital and labour are restricted to be constant, and the elasticity of substitution between capital and labour is restricted to 1. To overcome this, the second-order trans-logarithmic[20] function is used as an alternative Cobb-Douglass production function. The trans-logarithmic production function is estimated using the same fixed effects. The labour gap results using the translog

[20] $y_{it} = \beta_l l_{it} + \beta_{ll} l_{it}^2 + \beta_k k_{it} + \beta_{kk} k_{it}^2 + \beta_{lk} l_{it} k_{it} + \upsilon_{it} + e_{it}$

production function are presented in column 2 of Table C8. Again, the results are similar to those found in Table 2. Table C8 confirms the empirical evidence, suggesting that the main results in the baseline analysis are not determined by the functional form of the production function.

5.5- Localisation, urbanisation economies and labour misallocation

The final step is to test whether firms located in a denser region have lower labour misallocation, (Equation 10). The results using ordinary least squares (OLS) are reported in columns 1 and 2 of Table 3. The results show that the agglomeration variables (localisation and urbanisation) are statistically significant. On average, localisation and urbanisation economies are associated with lower labour misallocation when controlling for firm characteristics and fixed effects (firms *i*, and region-industry, *sy*). In terms of the magnitude, the estimated coefficients of urbanisation are larger than localisation.

Since employee skills differ across regions, however, the OLS estimation may suffer from some endogeneity bias. The fact that employees are grouped according to their skills in the different regions may create a mitigation risk bias, because, at the firm level the value of the marginal product is unobservable and relies on the elasticities of capital and labour estimated at the sectoral levels, while the average wage at the firm level is observed. Then, employees in large regions are more productive and earn higher wages. Therefore, the gap in denser regions is overestimated.

To overcome this endogeneity problem, the following strategy is adopted: employment density[21] by region and year and employment specialisation index by industry, region and year [22] are considered as an instrument of agglomeration economies (localisation and urbanisation economies). The number of firms by region, industry and year are also included as an additional instrument to be able to perform the over-identification test (Hansen's J test statistic) for the validity of the instruments. The instrumental variables results are presented in column 3 and 4 of Table 3 and confirm the previous results for the effect of agglomeration on labour misallocation.

---

21 $density_{rt} = \frac{employees_{rt}}{\sup er_{rt}}$, where $\sup er_{rt}$ is the region's area measured in kilometre square.

22 $specia_{srt} = \frac{(employees_{srt}/employees_{rt})}{(employees_{st}/employees_{t})}$ is the share of employment of industry *s* in total employment in region *r* relative to the share of employment of industry *s* in total employment at time *t*.

The estimated coefficients of localisation and urbanisation are larger in magnitude and statistically significant. A 10% increase in the degree of localisation in a region reduces the labour misallocation by 7.41% on average, while a 10% increase in the degree of urbanisation reduces the labour misallocation by 4.26%.

Table 3: Location, urbanisation and labour misallocation

Dependent Variable :	Labour misallocation in log: $\ln\left(\left\|Gap_{it}^{L}\right\|\right)$			
	(1)	(2)	(3)	(4)
	OLS		**GMM**	
Localisation	-0.183**	-0.164***	-0.708***	-0.741***
	(0.067)	(0.054)	(0.112)	(0.076)
Urbanisation	-0.729**	-0.756***	-0.494***	-0.426***
	(0.280)	(0.185)	(0.117)	(0.071)
age	-0.004	0.015***	0.011	0.014***
	(0.022)	(0.001)	(0.009)	(0.001)
small	-0.735***	-0.222***	-0.733***	-0.249***
	(0.043)	(0.038)	(0.040)	(0.049)
medium	-1.625***	-0.190***	-1.570***	-0.243***
	(0.098)	(0.056)	(0.052)	(0.071)
large	-2.428***	-0.435***	-2.321***	-0.608***
	(0.117)	(0.067)	(0.101)	(0.128)
competition	-0.532**	-0.479**	-0.435***	-0.421***
	(0.203)	(0.189)	(0.136)	(0.146)
Observations	20,522	20,524	20,063	20,118
R-squared	0.733	0.157	0.146	0.047
fixed effects	i & rs	rs	i & rs	rs
cluster level	region	region	region	region
Wald F statistic			20.21	34.44
Hansen J (p-value)			0.144	0.169
underidentification (p-value)			0.000289	0.000222

*Note: standard errors (reported in parentheses) are clustered at firm i and region-industry rs level. Significant at: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable: log of labour misallocation. The weak identification test is based on the Kleibergen-Park Wald F statistic, the test for the overidentifying restrictions is based on Hansen's J-test, and the under-identification test is based on the Kleibergen-Paap Lagrange multiplier statistic. Dependent variable: log of labour misallocation.*
Source: Authors' estimations

In columns 1 and 3, standard errors are double clustered at the firm level to deal with serial correlation and region-industry level to control for cross-sectional dependence, while in columns 2 and 4, standard errors are clustered at the region-industry level. The reported Kleibergen-Paap

Wald statistic rejects a weak identification problem, while the p-value of the Hansen J-test supports the validity of the instruments. The p-value of the reported under-identification test rejects the under-identification hypothesis of the model pointing that the model is identified.

To test the robustness of the results reported in table 3, the second-order trans-logarithmic function is used instead of the Cob-Douglass function. The results are reported in table 4 for both OLS and GMM estimations and confirm that the main results in baseline table 3 are robust to change the functional form of the production function.

6- Conclusion

This paper analyses the effects of agglomeration economies on the firm-level labour misallocation using data collected by the “Institut National de la Statistique” (INS) of Cote d’Ivoire over the period 2013-2016. The main results are as follows. firstly, the average labour misallocation (labour gap) at the firm level is 2,825,887 FCFA ($5137.97) over the period 2013 - 2016 and this gap has significantly decreased over years when controlling for firm characteristics (age of the firm, size, competition index, etc.). Secondly, firms located in denser regions exhibit lower labour misallocation. In terms of the magnitude, both localisation and urbanisation economies are large and statistically significant. A 10% increase in the degree of localisation in a region reduces the labour misallocation by 7.41% on average, while a 10% increase in the degree of urbanisation reduces the labour misallocation by 4.26%. Finally, the findings are robust and consistent evidence suggesting that the estimation of the labour gap seems not to be influenced by sample selection or outliers and these findings are not also driven by the functional form of the production function.

These results have policy implications: The first implication suggests that similar to other developing countries, labour misallocation seems to be an important issue across sectors in Cote d’Ivoire. Secondly, labour misallocation has a geographical dimension, in addition to the firm characteristics. A sound policy needs to accounts for the spatial distribution of firms and the creation of active poles of development in major Ivorian regions.

Table 4: Localisation, Urbanisation and labour misallocation, robustness to alternative production function specifications

Dependent Variable	Log labour misallocation $\ln\left(\left\|Gap_{it}^{L}\right\|\right)$							
	Cobb-Douglas Fixed effects				Translog (order 2) fixed effects			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	GMM	GMM	GMM	GMM	GMM	GMM
localisation	-0.193**	-0.163**	-0.815***	-0.787***	-0.142**	-0.119**	-0.599***	-0.615***
	(0.080)	(0.060)	(0.131)	(0.091)	(0.066)	(0.056)	(0.104)	(0.067)
urbanisation	-0.632**	-0.679***	-0.287**	-0.273***	-0.387**	-0.396***	-0.079	-0.036
	(0.251)	(0.163)	(0.112)	(0.078)	(0.160)	(0.089)	(0.090)	(0.056)
age	-0.009	0.016***	0.007	0.016***	-0.001	0.017***	0.001	0.016***
	(0.019)	(0.001)	(0.008)	(0.001)	(0.019)	(0.001)	(0.012)	(0.001)
small	-0.746***	-0.087**	-0.752***	-0.135***	-0.706***	0.010	-0.713***	-0.038*
	(0.042)	(0.041)	(0.031)	(0.039)	(0.033)	(0.021)	(0.029)	(0.020)
medium	-1.492***	0.015	-1.452***	-0.053	-1.515***	0.089**	-1.467***	0.030
	(0.093)	(0.034)	(0.029)	(0.034)	(0.085)	(0.034)	(0.027)	(0.026)
large	-2.472***	-0.311***	-2.433***	-0.514***	-2.428***	-0.212***	-2.326***	-0.399***
	(0.181)	(0.064)	(0.123)	(0.105)	(0.078)	(0.033)	(0.072)	(0.064)
comp_rs	-0.495*	-0.392	-0.420**	-0.330*	-0.236	-0.166	-0.147	-0.058
	(0.281)	(0.234)	(0.204)	(0.177)	(0.188)	(0.150)	(0.154)	(0.124)
Observations	20,522	20,524	20,063	20,118	20,522	20,524	20,063	20,118
R-squared	0.694	0.125	0.115	0.041	0.704	0.127	0.073	0.023
fixed effects	i & rs	rs	i & rs	rs	i & rs	rs	i & rs	rs
cluster level	region	region	region	region	region	region	region	region
Wald F statistic			20.21	34.44			20.21	34.44
Hansen J (p-value)			0.381	0.391			0.0587	0.0469
underidentification (p-value)			0.000289	0.000222			0.000289	0.000222

*Note: standard errors (reported in parentheses) are clustered at firm i and region-industry rs level. Significant at: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable: log of labour misallocation. The weak identification test is based on the Kleibergen-Park Wald F statistic, the test for the overidentifying restrictions is based on Hansen's J-test and the under-identification test is based on the Kleibergen-Paap Lagrange multiplier statistic. Dependent variable: log of labour misallocation. Source: Authors' estimations*

7- References

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, *83*(6), 2411–2451. https://doi.org/10.3982/ECTA13408

Africa's Pulse. (2018). *An Analysis of Issues Shaping Africa's Economic Future* (Vol. 18). World Bank. https://doi.org/10.1596/978-1-4648-1365-8

Baltagi, B. H. (2021). Unbalanced Panel Data Models. In B. H. Baltagi (Ed.), *Econometric Analysis of Panel Data* (pp. 229–257). Springer International Publishing. https://doi.org/10.1007/978-3-030-53953-5_9

Baltagi, B. H., & Chang, Y.-J. (1994). Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model. *Journal of Econometrics*, *62*(2), 67–89. https://doi.org/10.1016/0304-4076(94)90017-5

Bartelsman, E., Haltiwanger, J., & Scarpetta, S. (2013). Cross-Country Differences in Productivity: The Role of Allocation and Selection. *American Economic Review*, *103*(1), 305–334.

Carneiro, A., Guimarães, P., & Portugal, P. (2012). Real Wages and the Business Cycle: Accounting for Worker, Firm, and Job Title Heterogeneity. *American Economic Journal: Macroeconomics*, *4*(2), 133–152. https://doi.org/10.1257/mac.4.2.133

Caselli, F. (2005). *Accounting for Cross-Country Income Differences* (pp. 679–741) [Handbook of Economic Growth]. Elsevier. https://econpapers.repec.org/bookchap/eeegrochp/1-09.htm

Chen, C., Restuccia, D., & Santaeulàlia-Llopis, R. (2017). *The Effects of Land Markets on Resource Allocation and Agricultural Productivity* (No. w24034). National Bureau of Economic Research. https://doi.org/10.3386/w24034

Cirera, X., Fattal-Jaef, R., & Maemir, H. (2020). Taxing the Good? Distortions, Misallocation, and Productivity in Sub-Saharan Africa. *The World Bank Economic Review*, *34*(1), 75–100. https://doi.org/10.1093/wber/lhy018

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012a). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, *80*(6), 2543–2594. https://doi.org/10.3982/ECTA8442

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2012b). Sorting and local wage and skill distributions in France. *Regional Science and Urban Economics*, *42*(6), 913–930.

Dias, D. A., Robalo Marques, C., & Richmond, C. (2016). Misallocation and productivity in the lead up to the Eurozone crisis. *Journal of Macroeconomics*, *49*, 46–70. https://doi.org/10.1016/j.jmacro.2016.04.009

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063–2117). Elsevier. https://ideas.repec.org/h/eee/regchp/4-48.html

Fall, M., Coulibaly, S., & Editeurs. (2016). *L'Urbanisation diversifiée: Le cas de la Côte d'Ivoire*. Washington, DC: World Bank.

Guimaraes, P., & Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *The Stata Journal*, *10*(4), 628–649.

Ha, D. T. T., Kiyota, K., & Yamanouchi, K. (2016). Misallocation and productivity: The case of Vietnamese manufacturing. *Asian Development Review*, *33*(2), 94–118.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, *46*(6), 1251–1271. https://doi.org/10.2307/1913827

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, *124*(4), 1403–1448. https://doi.org/10.1162/qjec.2009.124.4.1403

Hsieh, C.-T., & Klenow, P. J. (2010). Development Accounting. *American Economic Journal: Macroeconomics*, *2*(1), 207–223. https://doi.org/10.1257/mac.2.1.207

Iimi, A., & Rao, K. (2018). *Firm Location, Transport Connectivity, and Agglomeration Economies: Evidence from Liberia*. World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-8411

Inklaar, R., Lashitew, A. A., & Timmer, M. P. (2017). The role of resource misallocation in cross-country differences in manufacturing productiivity. *Macroeconomic Dynamics*, *21*(3), 733–756. https://doi.org/10.1017/S1365100515000668

Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, *70*(2), 317–341. https://doi.org/10.1111/1467-937X.00246

Marshall, A. (1890). *Principles of Economics*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137375261

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, *69*(2), 182–195. https://doi.org/10.1016/j.jue.2010.09.002

Newman, C., Rand, J., & Tsebe, M. (2019). *Resource misallocation and total factor productivity: Manufacturing firms in South Africa* (46th ed., Vol. 2019). UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2019/680-7

Oberfield, E. (2013). Productivity and misallocation during a crisis: Evidence from the Chilean crisis of 1982. *Review of Economic Dynamics*, *16*(1), 100–119. https://doi.org/10.1016/j.red.2012.10.005

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

Petrin, A., & Sivadasan, J. (2013). Estimating lost output form allocative inefficiency, with an application to chile and firing costs. *The Review of Economics and Statistics*, *95*(1), 286–301.

Restuccia, D., & Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, *11*(4), 707–720.

Restuccia, D., & Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, *16*(1), 1–10. https://doi.org/10.1016/j.red.2012.11.003

Restuccia, D., & Rogerson, R. (2017). The Causes and Costs of Misallocation. *Journal of Economic Perspectives*, *31*(3), 151–174. https://doi.org/10.1257/jep.31.3.151

Restuccia, D., & Santaeulalia-Llopis, R. (2017). *Land Misallocation and Productivity* (No. w23128). National Bureau of Economic Research. https://doi.org/10.3386/w23128

Siba, E., Soderbom, M., Bigsten, A., & Gebreeyesus, M. (2012). *Enterprise Agglomeration, Output Prices, and Physical Productivity: Firm-Level Evidence from Ethiopia* (WIDER Working Paper Series wp-2012-085). World Institute for Development Economic Research (UNU-WIDER). https://econpapers.repec.org/paper/unuwpaper/wp-2012-085.htm

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*(3), 112–114. https://doi.org/10.1016/j.econlet.2009.04.026

APPENDIX A: Estimation of the production function and firm level TFP

This appendix section explains the approach used to estimate the production function parameters used in the baseline analysis, which is based on Wooldridge (2009). As discussed in Section 4, the Cobb-Douglas production function with two factors of production: capital and labour without imposing the nature of the returns to scale is given by Equation A1:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \upsilon_{it} + e_{it}, \quad t = 1,2,...,T,\ i = 1,2,...,N \tag{A1}$$

Where y_{it} is the log of value added, l_{it} denotes the log of the number of employees and k_{it} is the log of capital stock. The productivity shock is given as: $\varepsilon_{it} = \upsilon_{it} + e_{it}$ with υ_{it} productivity shock (observed by the firm but not by the econometrician) and e_{it} is the unanticipated shock or an error term.

As in Olley & Pakes (1996) (OP) or Levinsohn et Petrin (LP) (2003), Wooldridge (2009) expresses unobserved productivity, υ_{it} as a function of state variables which is a capital stock and a proxy. The proxy variable is investment in OP and intermediate inputs such telecommunication fees or corporal immobilisation, m_{it} in LP. To limit attrition bias, this paper uses intermediate inputs (materials) as a proxy, which is less systematically zero compared to investment in firm-level data. Thus for some function, $g(.)$:

$$\upsilon_{it} = g(k_{it}, m_{it}), \quad t = 1,2,...,T \tag{A2}$$

Assuming that $E(e_{it} | k_{it}, m_{it}) = 0$, and substituting for υ_{it} into equation (A1), the production function can be written as :

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + g(k_{it}, m_{it}) + e_{it}, \quad t = 1,2,...,T \tag{A3}$$

The parameters β_l and β_k will not be separately identified, the former owing to collinearity between labour and productivity (Ackerberg et al., (2015)) and the latter owing to the inclusion of k_{it} in $g(.)$. Thus, to overcome this problem, LP state the assumption as:

$$E(\upsilon_{it} | \upsilon_{it-1}, \upsilon_{it-2}, ..., \upsilon_{i1}) = E(\upsilon_{it} | \upsilon_{it-1}),\ t = 2,3,...,T \tag{A4}$$

The innovation is defined as follows:

$$a_{it} = \upsilon_{it} - E(\upsilon_{it} | \upsilon_{it-1}) \tag{A5}$$

Equations (A4) and (A5) imply that the innovation will be independent of the information set at time t-1 (i.e. υ_{it}). As capital (k_{it}) is determined at the previous period, it will therefore be uncorrelated with innovation a_{it}, $E\left(a_{it}\left|k_{it}\right.\right)=0$. Note that the current, but not lagged, of labour (l_{it}) is correlated with innovation, $E\left(a_{it}\left|l_{it-1}\right.\right)=0$.

$$E\left(\upsilon_{it}\left|k_{it},m_{it-1},...,k_1,m_1\right.\right)=E\left(\upsilon_{it}\left|\upsilon_{it-1}\right.\right)\equiv f\left[g\left(k_{it-1},m_{it-1}\right)\right] \quad \text{(A6)}$$

Substituting $\upsilon_{it}=f\left[g\left(k_{it-1},m_{it-1}\right)\right]+a_{it}$ into equation (A1) provides us with equation that can be used to identify the two parameters (β_l and β_k):

$$y_{it}=\beta_l l_{it}+\beta_k k_{it}+f\left[g\left(k_{it-1},m_{it-1}\right)\right]+u_{it},\quad t=2,3,...,T \quad \text{(A7)}$$

Where $u_{it}=a_{it}+e_{it}$.The moment conditions for identifying the parameters are:

$$E\left(\upsilon_{it}\left|k_{it},l_{it-1},k_{it-1},m_{it-1}\right.\right)=0$$

The unknown function $f(.)$ and $g(.)$ are approximated using a general second-order polynomial.

In the estimation of Equation (10), a full set of time dummies is included to control for heterogeneity over time in the production function and TFP. The estimation is undertaken separately for each 1-digit sector. The productivity can be estimated:

$$\hat{\upsilon}_{it}=y_{it}-\hat{\beta}_l l_{it}-\hat{\beta}_k k_{it} \quad \text{(A8)}$$

Appendix B: Fixed effects vs Random effects

Fixed effects model

```
Fixed-effects (within) regression               Number of obs     =     20,533
Group variable: id                              Number of groups  =      7,483
R-sq:                                           Obs per group:
     within  = 0.0896                                         min =          2
     between = 0.0004                                         avg =        2.7
     overall = 0.0010                                         max =          4
                                                F(7,13043)        =     183.43
corr(u_i, Xb)  = -0.4007                        Prob > F          =     0.0000
------------------------------------------------------------------------------
grwslk_aga.. |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         tps |  -556.8803   31.08485   -17.91   0.000    -617.8111   -495.9494
         age |   120.0798   27.85592     4.31   0.000     65.47818    174.6815
        age2 |  -2.863099    .602519    -4.75   0.000    -4.044124   -1.682073
       small |  -2124.552   123.5341   -17.20   0.000    -2366.697   -1882.407
      medium |  -4972.546   228.7931   -21.73   0.000    -5421.014   -4524.078
       large |  -8137.593   449.5132   -18.10   0.000    -9018.704   -7256.481
     comp_ln |  -872.9161   81.26535   -10.74   0.000    -1032.208   -713.6241
       _cons |   7319.976   315.2338    23.22   0.000     6702.072     7937.88
-------------+----------------------------------------------------------------
     sigma_u |  4677.1631
```

```
     sigma_e |  3043.6732
         rho |  .70250463   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(7482, 13043) = 5.27                 Prob > F = 0.0000
```

Random effects model

```
Random-effects GLS regression                   Number of obs     =     20,533
Group variable: id                              Number of groups  =      7,483
R-sq:                                           Obs per group:
     within  = 0.0711                                         min =          2
     between = 0.0102                                         avg =        2.7
     overall = 0.0204                                         max =          4
                                                Wald chi2(7)      =     931.49
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000
------------------------------------------------------------------------------
grwslk_aga.. |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         tps |  -560.7451   24.29485   -23.08   0.000    -608.3622   -513.1281
         age |   53.92219   11.55586     4.67   0.000     31.27312    76.57127
        age2 |   .0857463   .2573624     0.33   0.739    -.4186747    .5901673
       small |  -1051.023    88.2149   -11.91   0.000    -1223.921   -878.1251
      medium |    -1485.3   149.8151    -9.91   0.000    -1778.932   -1191.668
       large |  -2759.691   289.7418    -9.52   0.000    -3327.575   -2191.807
     comp_ln |  -590.8648   42.61176   -13.87   0.000    -674.3823   -507.3473
       _cons |   5666.521   164.6413    34.42   0.000      5343.83    5989.212
-------------+----------------------------------------------------------------
     sigma_u |   3759.217
     sigma_e |  3043.6732
         rho |  .60403139   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Hausman test

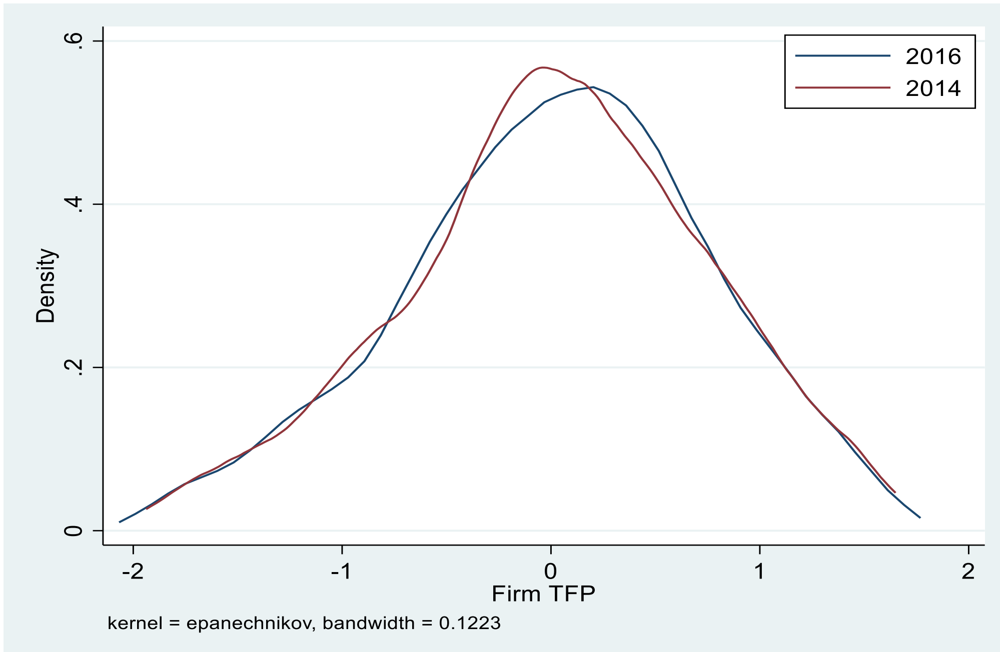
```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          re         Difference          S.E.
-------------+----------------------------------------------------------------
         tps |   -556.8803    -560.7451        3.864863         19.9959
         age |    120.0798     53.92219        66.15765        25.72028
        age2 |   -2.863099     .0857463       -2.948845        .5529363
       small |   -2124.552    -1051.023       -1073.529        88.62744
      medium |   -4972.546      -1485.3       -3487.246        176.6111
       large |   -8137.593    -2759.691       -5377.902        350.8425
     comp_ln |   -872.9161    -590.8648       -282.0513        70.36345
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test:  Ho:  difference in coefficients not systematic
                  chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      500.81
                Prob>chi2 =      0.0000
```

Source: Authors' estimations

Appendix C: TFP dispersion and labour gaps

Figure C1: Distribution of firm-level TFP



Source: *Authors' calculation based on the firm-level census data*

Table C1. Number of firms per year

	2013	2014	2015	2016
All firms (all raw)	15,438	17,467	19,239	20,714
With sales	13,215	14,921	16,150	16,743
With value added	10,689	11,921	12,883	13,285
With capital stock	11,546	16,284	17,998	19,930
With labour cost	12,747	16,201	17,913	19,392
With labour (employees)	15,007	17,465	19,239	19,451
Sample	3,407	5,646	6,384	5,096

Source: *Authors' compilations based on the firm-level census data.*

Table C2: Number of firms by region

Regions	Freq	Percent	Regions	Freq	percent
ABIDJAN	15,987	77.86	INDIENIE-DJUABLIN	190	0.93
BAFFING	7	0.03	IFFOU	38	0.19
BAGOUE	30	0.15	KABADOUGOU	30	0.15
BELIER	56	0.27	LA ME	102	0.50
BERE	4	0.02	LOH DJIBOUA	229	1.12
BOUNKANI	20	0.10	MARAHOUE	167	0.81
CAVALLY	127	0.62	MORONOU	36	0.18
GBEKE	330	1.61	NAWA	210	1.02

GBOKLE	58	0.28	N'ZI	76	0.37
GOH	304	1.48	PORO	298	1.45
GONTOUGOU	66	0.32	SAN-PEDRO	676	3.39
GUEMON	66	0.32	SUD-COMOE	234	1.14
GRAND PONTS	90	0.44	TCHOLOGO	35	0.17
HAMBOL	9	0.04	TONKPI	112	0.55
HAUT SASSANDRA	410	2.00	WORODOUGO	39	0.19
IGNEBY-TIASSA	151	0.74	YAMOUSSOKRO	266	1.30

Source: Authors' calculations based on the firm-level census data

Table C3: Cobb-Douglass Production Function Coefficient Estimates

Sectors	β_l	β_k	RTS	N
Agriculture, forestry and fishing	0.669	0.260	0.929	457
Mining and quarrying	0.677	0.118	0.795	72
Manufacturing	0.662	0.256	0.918	2344
Construction	0.786	0.165	0.951	1613
Commerce, Transportation and Accommodation and food service activities	0.817	0.158	0.975	8466
Financial and insurance activities	0.362	0.342	0.704	337
Real estate activities	0.361	0.326	0.687	313
Information and communication	0.460	0.261	0.721	637
Administrative and support service activities	0.587	0.174	0.761	1097
Education	0.800	0.089	0.889	1959
Human health and social work activities	0.842	0.101	0.943	915
Other Services	0.893	0.101	0.994	2323
Total	**0.758**	**0.171**	**0.929**	**20533**

Note : β_l and β_k represent the elasticities of labour and capital, respectively. RTS represents the return to scale.

Source: Authors' calculations

Table C4: Dispersion of TFP (2014 and 2016)

	2014			2016		
	S.D.	**75-25**	**90-10**	**S.D.**	**75-25**	**90-10**
Total	0.974	1.058	2.23	0.977	1.080	2.26
Agriculture, forestry and fishing	1.274	1.781	3.253	0.987	1.221	2.220
Mining and quarrying	0.622	0.987	1.486	1.947	1.65	4.81
Manufacturing	0.968	1.114	2.187	1.015	1.092	2.322
Construction	1.044	1.069	2.219	1.073	1.171	2.444
Commerce, Transportation and Accommodation and food service activities	0.916	0.936	2.154	0.940	1.025	2.182
Financial and insurance activities	0.891	0.889	2.273	0.934	1.315	2.267
Real estate activities	1.061	1.205	3.065	1.170	1.628	2.953

Information and communication	1.329	1.441	2.739	1.171	1.260	2.644
Administrative and support service activities	0.992	1.189	2.584	1.131	1.170	2.536
Education	0.949	1.023	2.097	0.832	.982	2.152
Human health and social work activities	0.898	1.004	1.821	0.807	.8722	1.750
Other Services	1.037	1.277	2.395	0.985	1.092	2.269
Number of Firms		**5646**			**5096**	

Note: The statistics are related to the distribution of TFP. S.D. is the standard deviation, 75-25 is the difference between the 75th and 25th percentiles, and 90-10 is the difference between the 90th and 10th percentiles.

Table C5: Absolute value of the labour gap by sector and year

Sectors	Mean	SD	Positive%	N
Agriculture, forestry and fishing	2661.667	5394.71	72.43	457
Extraction (mining and quarrying)	11315.76	12273.38	81.94	72
Manufacturing	2658.452	5177.237	63.91	2344
Construction	3368.183	5479.546	82.02	1613
Commerce, Transportation and Accommodation and food service activities	3124.09	5008.501	83.59	8466
Financial and insurance activities	3003.137	4857.825	24.33	337
Real estate activities	4303.833	7276.532	15.72	846
Information and Communication	3097.04	5611.662	35.16	637
Administrative and support service activities	2351.153	5047.602	44.21	1097
Education	1159.849	2146.264	72.18	1959
Human health and social work activities	2190.709	3208.601	86.12	915
Other Services	2880.828	4547.078	73.40	2323
Total	**2825.887**	**4933.072**	**73.63**	**20533**

***Note:** All gaps are in thousands of FCFA (deflator used is the CPI). S.D. is the standard deviation; N is the number of firms. Positive % share of firms with a positive gap. **Source:** Authors' calculations*

Table C6: Absolute value of the labour gap by industry and year, 2013-2016

Industries	Mean	SD	Posit%	N
Agriculture, forestry and fishing	2661.667	5394.71	72.43	457
Extraction (mining and quarrying)	11315.76	12273.38	81.94	72
Manufacture of food products	2641.998	5503.19	67.23	589
Manufacture of beverages and Tabacco products	6253.235	10409.3	74.51	51
Textile, wearing apparel and leather	1628.808	3430.432	63.16	114
Wood	1092.478	1245.65	45.98	137
Paper and printing	1893.327	3520.113	61.15	314
Coke and refined petroleum	9757.923	11130.56	93.33	15
Chemicals and pharmaceuticals	3147.26	5546.255	61.38	246
Rubber and plastics	2582.969	4272.094	65.22	161

other non-metallic minerals & Basic metals	4027.407	7066.987	71.875	224
Other manufacturing	2409.552	4098.053	62.07	493
Electricity, gaz and water supply	4625.139	8589.434	82.86	140
Construction of buildings	3835.217	6381.09	79.26	381
Civil engineering	3428.784	5078.454	82.95	610
Specialized construction activities	2557.23	3631.346	82.78	482
repair of motor vehicles and motorcycles	2651.98	4232.198	76.91	537
Wholesale and retail trade	3088.584	4925.466	85.55	6596
Transportation (land, water and air)	4515.967	6385.441	86.49	407
Warehousing, Postal and courier activities	3637.062	5636.812	76.21	601
Accommodation and food service activities	1933.088	4155.344	64.92	325
Publishing, Programming and broadcasting activities	2364.268	3213.279	36.26	182
Telecommunications	4195.25	7672.574	39.89	188
Computer programming, consultancy and related activities and Information service activities	2823.262	5058.453	31.09	267
Financial service activities	4216.396	7210.212	22.52	111
Insurance	2407.244	2956.615	25.22	226
Real estate activities	4303.833	7276.532	42.49	313
Legal and accounting activities	4303.833	7276.532	71.99	939
Activities of head offices; management consultancy activities	3069.227	4969.129	70.99	424
Architectural, R&D and Veterinary activities	3641.392	5336.036	76.22	471
Advertising and market research	3105.07	4007.313	77.21	272
Other professional, scientific and technical activities	2290.944	3419.667	71.05	152
Rental, leasing and Employment activities	3518.238	6287.958	58.21	213
Travel agency, tour operator, reservation service and related activities	1729.755	2016.763	45.79	214
Security and investigation activities	1802.582	4856.306	31.36	287
Services to buildings and landscape activities, Office administrative, office support activities	2460.37	5478.917	45.17	383
Education	1159.849	2146.264	72.18	1959
Human health and social work activities	2190.709	3208.601	86.12	915
Other Services	2685.652	4397.209	78.46	65
Total	**2825.887**	**4933.072**	**73.63**	**20533**

Note: All gaps are in thousands of FCFA (deflator used is the CPI). S.D. is the standard deviation; N is the number of firms. Positive % is the percentage of firms with a positive gap. Source: Authors' calculations.

Table C7: Evolution of labour misallocation sample sensitivity

Dep. Variable	labour misallocation $\left\|Gap_{it}^{L}\right\|$		
	L>10 workers	L<10 workers	GDP deflator
	(1)	(2)	(3)
trend	-789.812***	-399.961***	-1,072.136***
	(122.612)	(77.137)	(174.631)
Age	158.973***	82.298	186.462***
	(56.229)	(51.887)	(60.960)
$(\text{Age})^2$	-2.987**	-2.782*	-4.473**
	158.973***	82.298	(1.954)
Small			-2,873.778***
			(244.482)
Medium			-6,800.675***
			(626.923)
Large			-11,204.403***
			(1,180.978)
Competition	-1,037.141***	-686.961***	-1,212.288**
	(320.857)	(221.230)	(558.356)
Observations	7,733	11,634	20,533
R-squared	0.809	0.752	0.731
Fixed effect	i	i	i
Cluster level	i & st	i & st	i & st
Number of firms	2707	4552	7483

*<u>Note:</u> standard errors (reported in parentheses) are clustered at firm i and industry-year st level. Significant at: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable: labour misallocation. All gaps are in thousands of FCFA (deflator used is the GDP deflator in column 4).*

<u>Source:</u> Authors' estimations

Table C8: Evolution of labour misallocation, robustness to alternative production function specifications

	Labour misallocation $\left\|Gap_{it}^{L}\right\|$	
	Cobb-Douglass Fixed effect (1)	Translog (order 2) Fixed effects (2)
trend	-0.253***	-0.144***
	(0.034)	(0.040)
Age	0.013**	0.007
	(0.005)	(0.005)
Small	-0.704***	-0.689***
	(0.059)	(0.056)
Medium	-1.598***	-1.496***
	(0.122)	(0.134)
Large	-2.362***	-2.401***
	(0.221)	(0.214)
Competition	-0.397***	-0.293**
	(0.090)	(0.115)
Observations	20,533	20,533
R-squared	0.722	0.699
Fixed effect	i	i
Cluster level	i & st	i & st
Number of firms	7483	7483

*Note: standard errors (reported in parentheses) are clustered at firm i and industry-year st level. Significant at: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable:* Labour misallocation. *All gaps are in thousands of FCFA (deflator used is the CPI). Source: Authors' estimations*