Measurement of total factor productivity: Evidence from French construction firms

Abdoulaye Kané

2022-9 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

Measurement of total factor productivity: Evidence from French construction firms

By Abdoulaye KANE

EconomiX UMR 7235 CNRS Paris Nanterre University and CSTB

January 2022

Abstract

This paper reviews eight (8) methods of calculating total factor productivity (TFP) in the "construction of residential and non-residential buildings" sector in France. These include fixed-effects estimators; instrumental variables and the generalized method of moments (Blundell and Bond, 1999); Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Wooldridge, 2009; Ackerberg, Caves, and Frazer, 2015; the calibration method; and the data envelopment analysis (DEA) method. Then, using firm-level data from 2009 to 2018, we show that the market structure can be likened to an oligopoly situation and that capital intensity is also very low in this sector. Furthermore, the fixed-effects estimator provides the lowest capital coefficient and overestimates both the absolute value of the scale effect and the intermediate inputs coefficient. The highest capital coefficient is provided by the Wooldridge (2009) estimator. But there is little difference between the TFP measures, especially when semi-parametric methods are used. While the calibration of elasticities shows that the construction sector is labor intensive, the DEA method shows that on average only large firms are fully efficient. To our knowledge, the Ackerberg, Caves and Frazer method be a good estimator of TFP in the French construction sector. Finally, when comparing TFP levels, all estimation methods (fixed effects; Wooldridge, 2009; Olley and Pakes, 1996; Levinsohn and Petrin, 2003 and Ackerberg, Caves, and Frazer, 2015) are strongly positively correlated with each other (over 92%). However, the correlations between these methods and the non-parametric methods (DEA and calibration methods) are very low, even negative with the calibration method.

Keywords: French construction sector; Production function; Total factor productivity; Parametric estimation; Semi-parametric estimation; Non-parametric estimation; Market structure
JEL Classification: C13; C14; C23; D24; D43

1 Introduction

Total Factor Productivity (TFP) is generally defined as the portion of output not explained by the amount of inputs used in production. It is crucial in terms of economic fluctuations, economic growth and cross-country per capita income differences[1] insofar as it determines long-term economic growth and is a comprehensive industry-level productivity measure. However, if its theoretical definition seems comprehensive, its empirical implementation is far from being an easy task.

[1] Comin (2010)

Furthermore, TFP is evaluated in a variety of ways in the economics literature. Three main approaches to TFP are raised:[2]

- The first group argues that changes in TFP measure the rate of technical change. (Law, Statscan, Krugman, Young.)
- The second group holds that TFP measures only the free lunches of technical change, which are mainly associated with externalities and scale effects. (Jorgenson, and Griliches)
- The last group is sceptical that TFP measures anything useful. (Metcalf, and Griliches)

Following these different approaches, the authors have tried to measure TFP at the national, sectoral and firm level. These include the works of Stigler (1947), Solow (1957), Charnes et al. (1978, 1981), Caves, Christensen and Diewert (1982), Olley and Pakes (1996) denoted by (OP), Bartelsman and Doms (2000), Pavcnik (2002), Levinsohn and Petrin (2003) denoted by (LP), Ackerberg, Benkard, Berry and Pakes (2007), Ackerberg, Caves and Frazer (2015) denoted by (ACF).

Despite these various studies on TFP, very few concern the construction sector in general and French construction in particular. It is interesting to look at this French sector for two main reasons. First, the French construction sector has a significant weight in the global economy. Indeed, the sector's share of the French economy's gross value added between 2009 and 2018 has always been above 12.5%.[3] According to the French Building Federation (June 2020), the French construction sector corresponds to half of the industry and twice the banking and insurance activities. More specifically, the French construction in 2019 corresponded to 403,000 companies, 1,502,500 active people including 1,121,000 employees and 381,500 craftsmen with a production amounting to 148 billion euros excluding tax. This economic weight also extends to the European and global scale. Vinci, the French and European leader, ranks second worldwide behind the large Chinese builder China State Construction Engineering. Bouygues and Eiffage are also among the largest companies in the sector worldwide.

Second, according to Xerfi|MCI (2019), the French construction sector has faced an apparent labor productivity gap in recent years. Thus, before any analysis of the causes of this productivity deficit, it is imperative to question the measurement of productivity. In addition, the term "Productivity" in the construction industry is generally seen as the apparent labor productivity. Although easy to calculate and interpret from an economic point of view, apparent labor productivity can be misleading when we take into account the substitution between factors of production. Therefore, a much more comprehensive measure that takes into account several inputs is desired. The goal of this paper is twofold. On the one hand, we review several methods of calculating TFP, each with its strengths and weaknesses. On the other hand, we analyze the market structure of the construction sector in France via an empirical application of these methods using firm-level data.

Parametric and semi-parametric methods show decreasing returns to scale in French construction. The underlying economic interpretation would be the imposing weight of large firms on small firms. In other words, the market structure could be likened to an oligopoly situation. Moreover, while the elasticity of capital per worker is low in output per worker (ranging from 0.0395 for the Fixed Effect method to 0.0636 for the Wooldridge method), the elasticity of materials per worker is very high (ranging from 0.776 to 0.819 following the same methods). We also find that the ACF method can be considered a good estimator of TFP in the French construction sector.

The differences between the different parametric and semi-parametric estimators are very small when comparing the estimated TFP. Spearman's correlation coefficients between these TFP measures are generally greater than 0.92 and even 0.99 when comparing the results obtained by the GMM method and the three semi-parametric methods (OP, LP and ACF). However, the correlations between these methods and the non-parametric methods are very low, even negative with the calibration method. The rest of the paper is structured as follows. Section two presents the methods for the theoretical calculation of TFP. We move to the empirical part in section three which will present the data as well as the estimation results. Section 4 provides some concluding remarks.

[2]See Lipsey and Carlaw (2000, 2004)

[3]Statista, March 2021

2 Theoretical framework

If the authors are unanimous in attributing the term "total factor productivity" to the work of Solow (1957), they are less so as to its measurement.
This measure is all the more difficult as the sector under study is fragmented.
Usually the literature starts with a production function. Following this idea I assume a Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^{\beta_k} L_{it}^{\beta_l} M_{it}^{\beta_m} \tag{1}$$

Where Y_{it} represents gross output of firm i in period t, K_{it}, L_{it} and M_{it} are inputs of physical capital, labor (Total employment) and materials (intermediate inputs), respectively, and A_{it} is the Hicksian neutral efficiency level of firm i in period t. The labor, capital and intermediate inputs elasticities are given by β_l, β_k and β_m respectively. We take natural logarithm of *(1)* and in order to understand the nature of returns to scale[4] in the industry studied, we divide each variable by the labor input. Thus, equation *(1)* becomes:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \beta_0 + \beta_k \ln\left(\frac{K_{it}}{L_{it}}\right) + (\beta_l + \beta_k + \beta_m - 1)\ln L_{it} + \beta_m \ln\left(\frac{M_{it}}{L_{it}}\right) + \epsilon_{it} \tag{2}$$

Where $\ln A_{it} = \beta_0 + \epsilon_{it}$; β_0 measures the mean efficiency level across firms and over time; ϵ_{it} is the error term.
To simplify the writing of *(2)* , we rewrite it as follows:

$$y_{it} = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + \epsilon_{it} \tag{3}$$

where $y_{it} = \ln\left(\frac{Y_{it}}{L_{it}}\right); k_{it} = \ln\left(\frac{K_{it}}{L_{it}}\right); m_{it} = \ln\left(\frac{M_{it}}{L_{it}}\right); \gamma = (\beta_l + \beta_k + \beta_m - 1)$. It should be noted that γ provides us with information on the nature of the returns to scale.[5]
ϵ_{it} can be decomposed into an observable and unobservable component:

$$y_{it} = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + \upsilon_{it} + \mu_{it} \tag{4}$$

Where υ_{it} is the observable component, μ_{it} is the unobservable component and $\beta_0 + \upsilon_{it}$ represents firm-level productivity.
Thus using the ordinary least squares (OLS) method, we can estimate TFP as follows:

$$\widehat{tfp}_{it} = \hat{\beta}_0 + \hat{\upsilon}_{it} = y_{it} - \hat{\beta}_k k_{it} - \hat{\gamma} l_{it} - \hat{\beta}_m m_{it} \tag{5}$$

To obtain the level productivity, we take the exponential of $\widehat{tfp}_{it}$, i.e., $\widehat{TFP}_{it} = exp(\widehat{tfp}_{it})$

However, the OLS estimation method is automatically biased. Indeed, estimating a production function by OLS assumes that the factors of production are exogenous in the production function, i.e. determined independently of the firm's level of efficiency. However, authors such as Marschak and Andrews (1994) have already shown that the factors of production in the production function are not chosen independently, but rather determined by the characteristics of the firm, including its efficiency. We face a simultaneity bias.[6]

In addition to this bias, we have the selection bias. TFP is typically estimated with a cylindrical panel by omitting all firms that enter and exit during the sample period (Olley and Pakes, 1996). Although some economists believe that firm entry and exit are implicitly taken into account in the analysis (Fariñas and Ruano, 2005), explicitly omitting consideration of the exit decision of firms leads to selection bias. The reason is as follows: Firms' decisions on factor allocation in a particular period are made conditional on its survivors. In sum, the selection bias will cause the error term to be negatively correlated with capital, thus causing the capital coefficient to be biased.

For all these reasons, OLS estimation of a production function will provide us with inconsistent coefficients. To overcome these issues, different methods of estimating the TFP have been proposed.

[4] In its simple definition, returns to scale represent the increase in efficiency as a result of the increase in production factors.
[5] If $\gamma = 0$ then the returns are constant. If $\gamma > 0$ then the returns are increasing. If $\gamma < 0$ then the returns are decreasing.
[6] The choice of production factors is not under the control of the econometrician, but determined by the individual choices of firms (Griliches and Mairesse, 1995). According to De Loecker (2007), this simultaneity bias is defined as the correlation between the level of production factors and the unobserved productivity shock.

2.1 Fixed effects estimation

Assuming that productivity is firm-specific but time-invariant, it is possible to estimate TFP using the fixed effects estimator (Pavcnik, 2002; Levinsohn and Petrin, 2003):

$$y_{it} = \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + tfp_i + \mu_{it} \tag{6}$$

Where $tfp_i = \beta_0 + v_{it}$
Equation *(6)* can be estimated in level using the within-individual estimator or in first difference providing unbiased coefficients as long as the unobserved productivity tfp_i does not vary over time. In this respect, the simultaneity bias is eliminated because we have only the within-sector variation in the sample. The same is true for selection bias because exit decisions are made in an invariant time period.

However, in practice, the fixed effects estimator on a production function often leads to unreasonably low estimates of the capital coefficient because it imposes strict exogeneity of inputs conditional on firm heterogeneity (Wooldridge, 2009). In economic terms, this would mean that factors of production cannot be chosen in response to productivity shocks. This assumption probably does not hold in practice, especially not in the construction sector, which has been facing a decline in the rate of productivity growth in recent years, and thus a variation in productivity over time. Another estimator is often proposed to overcome these problems: the instrumental variables estimator and the Generalized Moment Method (GMM).

2.2 Instrumental Variables and Generalized method of moments

An alternative method for achieving consistency of coefficients across a production function is to instrument for the factors that cause the endogeneity problem. Unlike the fixed effects estimator, the instrumental variables (IV) estimator does not rely on strict exogeneity. Nevertheless, this method requires a certain number of conditions, notably on the variable or variables used as external instruments. First, the instruments must be correlated with the endogenous variable(s).
Second, the instruments cannot enter directly into the production function. Finally, the instruments must not be correlated with the error term. The last assumption rules out the existence of imperfect competition in the factor market if output or factor prices are used as instruments. Assuming a perfect market, then, factor and output prices are natural choices of instruments for the production function.

Nevertheless, factor prices become valid instruments if and only if the firm does not have market power. Indeed, if the firm has market power, it will set its prices at least partly according to the quantities of factors and its productivity. This makes prices endogenous. This endogeneity problem will always arise even if we use the average wage per worker (reflecting exogenous labor market conditions) because this wage often varies according to the qualification and quality of the employee.

Other instruments can be taken into account, such as weather conditions, exogenous shocks on the labor or capital market, independently of the firm's market power. However, as pointed out by Ackerberg et al. (2007), even in this case, the IV approach only deals with the endogeneity of factors, but not the endogeneity of firms' outputs. If instrument choices are correlated with firm-level output, endogenous output would invalidate the use of instruments.

Some authors lag inputs and then use them as instruments. This biases the capital coefficient, which is often not significant. To remedy this, Blundell and Bond (1999) propose the GMM estimator. For them, the poor performance of the IV estimator is due to the weakness of the instruments used for identification. We must therefore put the instruments in delayed first differences in the level equations and thus we will have good estimates.

Nevertheless, the main drawback of this method is to find a valid instrument to instrument the TFP. Apart from these parametric estimation methods, we have semi-parametric estimates that give us a time variability of the TFP.

2.3 The Olley-Pakes (1996) estimation algorithm

This semi-parametric estimation method solves the simultaneity bias by taking the firms' investment decision as a proxy for unobserved productivity shocks. The selection problem is solved by incorporating an exit rule in the model. This is a dynamic model of firm behavior where we have two fundamental assumptions:

1. Only one variable (state variable) is unobserved at the firm level. This is productivity, which is also assumed to evolve as a first-order Markov process.

2. The monotonicity of the "investment" variable so that the demand function for investment is invertible. This means that investment is increasing in productivity. One of the corollaries of this second hypothesis is that only positive values of investment are accepted.

Starting with the Cobb-Douglas production function given by equation *(4)*, the estimation procedure can be described as follows: capital is a state variable, affected only by current and past values of the productivity level tfp_{it}. Investment is described as follows:

$I_{it} = K_{it-1} - (1 - \delta)K_{it}$ where δ is the capital depreciation.
Also, investment decisions at the firm level can be written as a function of capital and productivity: $i_{it} = i_t(k_{it}, tfp_{it})$ where the lower case notation refers to the logarithmic transformation of the variables.
Since investment is an increasing function of productivity, conditional on capital, the investment decision can be inverted, allowing us to write unobserved productivity as a function of observable variables: $tfp_{it} = h_t(k_{it}, i_{it})$ where $h_t(.) = i_t^{-1}(.)$. Using this formula, equation *(4)* is rewritten as follows:

$$y_{it} = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + h_t(k_{it}, i_{it}) + \mu_{it} \tag{7}$$

Let's denote $\varphi(k_{it}, i_{it}) = \beta_0 + \beta_k k_{it} + h_t(k_{it}, i_{it})$
The estimation of equation *(7)* is done in two steps. The first step is to estimate, using OLS, the following equation:

$$y_{it} = \gamma l_{it} + \beta_m m_{it} + \varphi(k_{it}, i_{it}) + \mu_{it} \tag{8}$$

where $\varphi(k_{it}, i_{it})$ is approximated by a third-order polynomial in investment and capital.
In the second step, where we address the attrition bias problem, we identify the capital coefficient by estimating the following equation:

$$y_{it} - \gamma l_{it} - \beta_m m_{it} = \beta_0 + \beta_k k_{it} + g(\varphi_{t-1} - \beta_k k_{t-1}) + \mu_{it} \tag{9}$$

where *g(.)* is an unknown function which is again approximated by a third-order polynomial expression in φ_{t-1} and k_{t-1}. The probability of survival is normalized to 1. The capital coefficient can then be obtained by applying nonlinear least squares to equation *(9)*.

2.4 The Levinsohn-Petrin (2003) estimation algorithm

Like the Olley-Pakes (OP) method, the Levinsohn-Petrin (LP) method is also a semi-parametric estimate, but intermediate inputs are used as a proxy. Indeed, the monotonicity condition of OP requires that investment is strictly increasing in productivity. This implies that only observations with positive investment can be used when estimating equations *(8)* and *(9)*, and this may lead to a significant loss of efficiency depending on the data available. Moreover, if firms report zero investment in a significant number of cases, this casts doubt on the validity of the monotonicity condition. To do this, the LP estimation method uses intermediate inputs as a proxy for unobserved productivity. Since firms generally report positive material and energy use each year, it is possible to retain most observations; this also implies that the monotonicity condition is more likely to hold. The LP estimation algorithm differs from the OP algorithm in two ways. First, it use intermediate inputs as a proxy for unobserved productivity, rather than investment. This implies that intermediate inputs are expressed in terms of capital and productivity, i.e. , : $m_{it} = i_t(k_{it}, tfp_{it})$
Using the monotonicity condition, intermediate inputs are strictly increasing in productivity and are also invertible: $tfp_{it} = s_t(k_{it}, m_{it})$ where $s_t(.) = m_t^{-1}(.)$.
Equation *(4)* becomes:

$$y_{it} = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + s_t(k_{it}, m_{it}) + \mu_{it} \tag{10}$$

The second difference is related to the correction for selection bias. Although the OP approach allows for both a non-cylindrical panel and the incorporation of the probability of survival in the second stage of the estimation algorithm, the LP method does not incorporate the probability of survival in the second stage.

Recently, new and even more robust production function estimation techniques have emerged in an effort to correct the LP method. These include the Wooldridge (2009) and Ackerberg-Caves-Frazer (2015) methods.

2.5 Wooldridge method (2009)

Wooldridge (2009) proposed an alternative implementation of OP/LP moments that involves the simultaneous minimization of first and second stage moments. Using the LP model, he suggested estimating all the parameters simultaneously using the moment conditions :

$$\begin{aligned} & E[mu_{it}|I_{it} \\ & \quad \xi_{it} + \mu_{it}|I_{it-1}] \\ = E[& \quad y_{it} - \gamma l_{it} - \Phi_t(k_{it}, m_{it})|I_{it} \\ & \quad y_{it} - \beta_0 - \gamma l_{it} - \beta_k k_{it} - g(\Phi_{t-1}(k_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1})|I_{it-1}] = 0 \end{aligned} \tag{11}$$

As pointed out by Wooldridge, there are several advantages to this First, the joint approach avoids the functional dependence issue above. Even if l_{it} is functionally dependent on m_{it}, k_{it} and t, γ might be identified by the second set of moments.
Other benefits of the Wooldridge approach are potential efficiency gains, and simpler standard error calculations. There are also disadvantages of the joint approach; in particular, the joint approach nonlinear search over β_0, β_k, γ and the parameters representing the two unknown functions Φ_t and g. This method is more time-consuming and probably more error-prone than the two-step approach, which can often be obtained by a nonlinear search on only β_k, γ.

2.6 The Ackerberg-Caves-Frazer (2015) method

The main argument of the Ackerberg, Caves, and Frazer (ACF) method is that the labor coefficient may not be identified in the estimation procedures proposed by OP and LP. This is what the authors call the "functional dependence problem". Indeed, the authors believe that labor can be an argument of the demand function of the proxy variable and, consequently, of the unobserved productivity function.

Thus, the authors start from the same point of view as the LP method by taking the same assumptions of monotonicity of the proxy function but including labor: $m_{it} = f_t(k_{it}, l_{it}, tfp_{it})$. One interpretation of this assumption is that the gross output production function is Leontief in intermediate inputs.[7]
Given the assumption of strict monotonicity, we can invert the intermediate input demand :
$tfp_{it} = f_t^{-1}(k_{it}, l_{it}, m_{it})$.
Therefore, equation *(4)* becomes :

$$y_{it} = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it}) + \mu_{it} = \Phi_t(k_{it}, l_{it}, m_{it}) + \mu_{it} \tag{12}$$

where $\Phi_t(k_{it}, l_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + \gamma l_{it} + \beta_m m_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it})$
Using the first stage moment condition, we have:

$$E[\mu_{it}|I_{it}] = E[y_{it} - \Phi_t(k_{it}, l_{it}, m_{it})|I_{it}] = 0 \tag{13}$$

Where I_{it} represents a set of information. We note that unlike the OP and LP methods, no coefficient is identified in the first step. In short, all the coefficients are estimated in the second step using the following second stage moment condition:

$$\begin{aligned} E[\xi_{it} + \mu_{it}|I_{it-1}] = E[& y_{it} - \beta_0 - \gamma l_{it} - \beta_k k_{it} - \beta_m m_{it} - g(\Phi_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) \\ & - \beta_0 - \beta_k k_{it-1} - \gamma l_{it-1} - \beta_m m_{it-1})|I_{it-1}] = 0 \end{aligned} \tag{14}$$

Where Φ_{t-1} is replaced by its estimate from the first stage. The coefficients γ, β_k, β_m are estimated through a first order Markov process.

Using a Cobb Douglass production function with neutral productivity differences in the sense of Hicks does not allow for the identification of factor biases in technological change, especially since value added or sales are generally used as a measure of output. Non-parametric techniques - Elasticity calibration and Data Envelopment Analysis (DEA) - solve this problem.

[7] There is a technology characterized by a linear relationship between intermediate inputs and output, and these intermediate inputs are proportional to output (Gandhi, Navarro, and Rivers (2011a)).

2.7 Calibration method

Introduced by Solow (1956, 1957), the basic idea of this method is to calibrate the different elasticities of the factors of production using two crucial assumptions[8]:

- The constant returns to scale assumption
- The perfect market assumption in both the labor and goods and services markets

Thus, with these two assumptions and as long as we have information on the different factors of production as well as on output, we are able to calculate equation *(5)*.
By calibration, the elasticities of labor (β_l) and intermediate inputs (β_m) are obtained by the ratio between the labor cost and the production value, the ratio between intermediate inputs and the production value respectively. Assuming constant returns, the elasticity of the capital stock becomes $\beta_k = 1 - \beta_l - \beta_m$.
Finally, we use the average of each elasticity over time.

The calibration method is particularly interesting in that it calculates TFP in the simplest possible way (no estimation is required). However, the assumption of perfect competition in product and input markets assumption is a strong assumption, especially in the construction sector, which tends towards an imperfect market structure.

2.8 Data Envelopment Analysis approach (DEA)

The DEA approach, also called non-parametric frontier estimation, constructs for each observation a linear combination of all other observations (normalized by production) for explicit comparison. Here, no particular production function is assumed.
Efficiency (productivity) is defined as a linear combination of output over a linear combination of factors of production. The weights on the factors (u_l, u_k) and output (v_q) are chosen directly by maximizing the efficiency (productivity) denoted by θ. Observations that are not dominated are labeled 100% efficient. Dominance occurs when another firm, or a linear combination of other firms, produces more of all output with the same aggregate of factors, using the same weights to aggregate the factors.
A linear maximization program is solved separately for each observation. For unit 1 (firm-year), in the case of a single production, the problem is :

$$\begin{aligned} \max_{v_q, u_l, u_k} \quad & \theta_1 = \left(\frac{v_q Q_1 + v^*}{u_l L_1 + u_k K_1} \right) \\ \text{subject to :} \quad & \left(\frac{v_q Q_i + v^*}{u_l L_i + u_k K_i} \right) \leq 1;\ i = 1, \ldots, N. \\ & v_q, u_l + u_k > 0; u_l, u_k > 0 \\ & v^* \geq 0 \end{aligned} \tag{15}$$

Where $v^* = 0$ when returns to scale are constant.

The efficiency of all firms cannot exceed 100% when the same weights are applied. A normalization is necessary to properly define the problem: $u_l L_1 + u_k K_1 = 1$ and v^* is an additional game variable to allow for varying returns to scale. When we have constant returns to scale ($v^* = 0$), the production frontier is a ray through the origin in the aggregate production-factor space. Under varying returns to scale, this frontier is the piece-wise linear envelope of all production plans.

The efficiency measure θ_i can be interpreted as the productivity difference between unit i and the most productive unit. The efficiency score ranges from 0 to 1. An efficiency score of 1 means that the firm is fully efficient. The level and growth rate productivity estimates are defined as follows, respectively:

[8] These assumptions can be relaxed. However, we will need other assumptions to calculate a user cost of capital.

$$logA_{it}^{DEA} - \overline{logA_t}^{DEA} = log\theta_{it} - \frac{1}{N_t}\sum_{j=1}^{N} log\theta_{jt} \tag{16}$$

$$logA_{it}^{DEA} - logA_{it-1}^{DEA} = log\theta_{it} - log\theta_{it-1} \tag{17}$$

The main advantage of the DEA method is the absence of functional form or behavioral assumptions of firms and no distributional assumption is imposed on the inefficiency term. The underlying technology is completely indeterminate and allowed to vary across firms. However, it has serious limitations in that, first, it interprets any deviation from the set production potential as inefficiency. This means that all factors that affect the firm's performance are considered to be under the firm's control. The consequence is that the estimated inefficiency is biased due to exogenous factors such as measurement error (Kumbhakar and Lovell, 2003). The second limitation is that it measures inefficiency relative to the best performing unit among observations, making its result susceptible to outlier bias (Coelli et al., 2005).
Based on this review of the TFP, we present a summary result of each method.

Methods	Advantages	Drawbacks
Fixed effects	Resolution of the simultaneity bias when we assume strict exogeneity of the inputs.	Underestimation of the capital elasticity.
Instrumental variables and GMM (Blundell and Bond, 1999)	Resolution of the simultaneity bias and these methods do not impose strict exogeneity.	The challenge is to find valid instruments.
Olley and Pakes (1996)	Resolution of simultaneity and attrition biases.	The monotonicity condition no longer holds when the investment value is zero and the functional dependence problem.
Levinsohn and Petrin (2003)	Resolution of simultaneity and attrition biases	The functional dependence problem.
Wooldridge (2009)	Resolution of simultaneity and attrition biases by performing a one-step system GMM.	This method is more time-consuming and probably more error-prone than the two-step approach.
Ackerberg, Caves and Frazer (2015)	Resolve simultaneity and attrition biases by addressing the functional dependence problem suffered by the Olley and Pakes and Levinsohn and Petrin methods.	Like the previous methods, the problem of this method is the specification of a functional form of the production function.
Elasticity calibration	It calculates the TFP in the simplest way possible.	The condition of a perfect market, both in the labor market and in the market for goods and services, is a strong assumption.
Data Envelopment Analysis (DEA)	The main advantage of the DEA method is the absence of functional form or behavioral assumptions of firms and no distributional assumption is imposed on the inefficiency term.	As it is a deterministic method, it is vulnerable to measurement errors.

Table 1: Summary of methods

The following section will empirically test these different methods.

3 Empirical Application

3.1 Data

Our data come from the Esane approximate results file (FARE), which contains accounting information from tax returns that are consistent with information from the Annual Sector Survey. The FARE system aims to build up a coherent set of business statistics. It combines administrative data (obtained from the annual profit declarations that companies make to the tax authorities, and from annual social data that provide information on employees) and data obtained from a sample of companies surveyed by a specific questionnaire to produce structural business statistics.

We mobilize data from the construction sector in metropolitan France from 2009 to 2018. Of the three sub-sectors in the construction sector (building construction including real estate development, civil engineering, and specialized construction), we focus only on the building construction. Specifically, this is the "Residential and Non-Residential Building Construction (Sector 4120)" sub-sector, which includes general construction or "all crafts" firms with overall responsibility for the construction of a building. It also includes the conversion or renovation of existing residential structures. This sub-sector has the advantage of being highly focused on on-site work and is also representative of the construction sector as shown in the estimation results in the Appendix. Moreover, unlike specialized construction, which is a highly atomized sector, building construction contains more large companies. Firms are reported for at least 3 years and our data do not show a temporal break for the same firm. Furthermore, to overcome selection bias, we need to work on an unbalanced panel data set. Indeed, the attrition bias is lower the more the panel is unbalanced with a large number of samples (Ackerberg et al., 2015).

We use total gross production as the output variable. Three inputs have been mobilized: labor, capital and intermediate inputs. The labor input is measured by the number of full-time equivalent employees. The capital stock is approximated by gross tangible fixed assets. The gross investment is also given by FARE. We measure intermediate inputs by the difference between total gross production and value added at factor costs. Because we have nominal values, we deflate them using price indices in the French construction sector obtained from the STAN 2020 edition database (constant price 2015) to obtain real values. These deflators cover added value, output, investment, capital and intermediate inputs. Table 2 presents summary statistics for the production function from 2009 to 2018.

Variables	N	Mean	Sd	Min	Max
Real output (€$\times 1,000$)	27,544	59.067	404.247	0.13	21,164.158
Real added value (€$\times 1,000$)	27,544	13.432	84.313	0.006	4,343.993
Employment	27,544	19.195	101.866	1	4,635
Real Investment (€$\times 1,000$)	27,544	0.646	4.641	0.001	437.016
Real Capital stock (€$\times 1,000$)	27,544	5.575	41.342	0.002	2,179.456
Real Materials (€$\times 1,000$)	27,544	0.457	3.225	0.001	166.266

Table 2: Summary statistics of production variables

The table clearly shows that the average number of employees in the sample does not reach 20 (average employment amounts to 19.195 employees). This situation is specific to the sector, which has many small firms. Second, these data reveal strong heterogeneity across firms. While the minimum number of employees is 1, the maximum number is 4,635. For all these reasons, it is interesting to examine the structure of our sample through the size of the firms.

The INSEE size classification (classification according to the number of employees, turnover and balance sheet) will be difficult to apply to a sector composed essentially of micro-enterprises. Thus, we assume that firm size depends only on the number of employees. To do this, we follow Baldwin et al. (2002) classification that considers a firm to be :

- Micro if the number of employees is < 20;
- Small if the number of employees is between 20 and 99;

- Medium if the number of employees is between 100 and 499;
- Big if the number of employees is ≥ 500.

The following table shows that 84.10% of our sample has less than 20 employees and 13.23% are small businesses. Medium-sized companies represent 2.23%, while large companies are very few (0.44%).

Firm size	Number of observations	Proportion (%)
Micro	23,164	84.10
Small	3,645	13.23
Medium	615	2.23
Big	120	0.44

Table 3: Proportion by size

However, although they are few in number, the large companies in the French construction sector have enormous weight. This is explained by the following graph on average total labor productivity[9]:

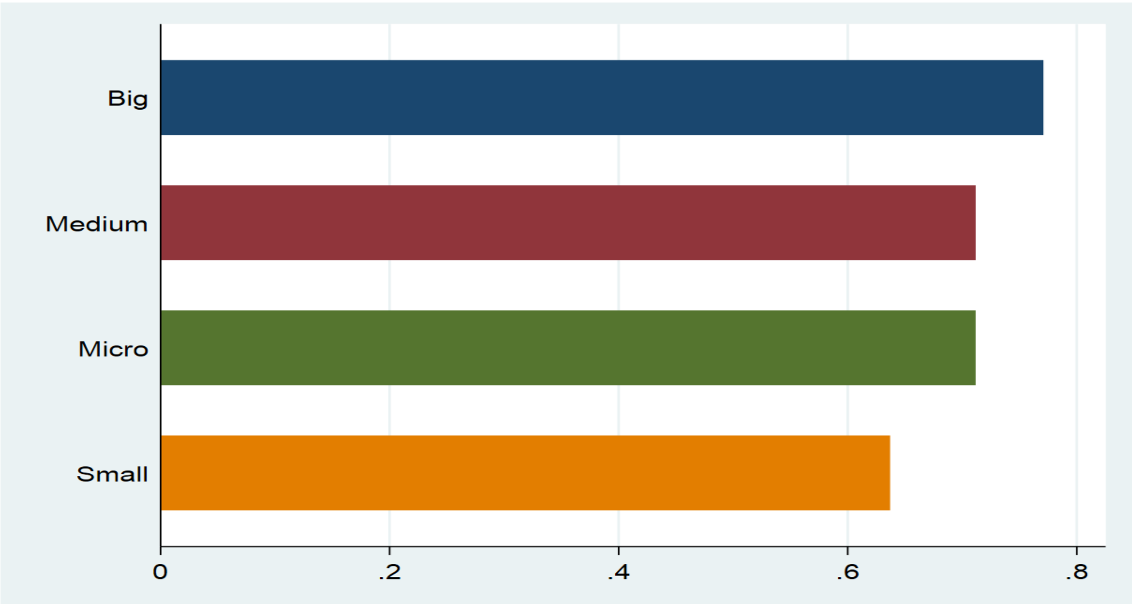

Figure 1: Average labor productivity by firm size

Figure 1 shows that, on average, the labor productivity of large firms is the highest, followed by intermediate-sized firms. However, the gap between medium-sized firms and micro-firms is not large. Average labor productivity is lowest at the small firm level.

3.2 Estimation results

In the following table we present the estimation results of the production function based on the methods presented in section 2. All reported estimates are obtained through an unbalanced panel of firms (allowing for implicit entry and exit). Wooldridge's (2009) estimator is used as the GMM method. Its advantage is that it provides consistent and efficient parameter estimation using the one-step system-GMM approach. It solves potential serial correlation, heteroscedasticity as well as the endogeneity due to simultaneity and

[9] Labor productivity is measured as the ratio of real value added to the number of full-time equivalent employees.

attrition using lagged values as instrument (Ackerberg et al. (2015)).
The Olley and Pakes (1996), Levinsohn and Petrin (2003), Wooldridge (2009), and Ackerberg, Caves, and Frazer (2015) methods are estimated using the "prodest" command[10] developed by Mollisi and Rovigatti (2017), which has the advantage of correcting for attrition and simultaneity bias and adding control variables. We denote the Fixed Effects, Olley and Pakes, Levinsohn and Petrin, Wooldridge and Ackerberg, Caves and Frazer estimators by (FE), (OP), (LP), (WRDG) and (ACF) respectively. The calibration method is designated by (CM).

Dependent variable: *ln(Y/L)*	FE	WRDG	OP	LP	ACF
γ	-0.0420***	-0.0200***	-0.0249***	-0.0201***	-0.0146***
	(0.00443)	(0.000712)	(0.00168)	(0.00123)	(1.06e-06)
β_k	0.0395***	0.0636***	0.0518***	0.0578***	0.0455***
	(0.00345)	(0.00226)	(0.00546)	(0.00369)	(1.05e-06)
β_m	0.819***	0.776***	0.783***	0.796***	0.797***
	(0.00579)	(0.000899)	(0.00293)	(0.00438)	(5.80e-06)
N	27,544	22,760	27,544	27,544	27,544
ID	4,784	4,784	4,784	4,784	4,784
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Production function estimates

Where γ=Scale effect=$(\beta_l + \beta_k + \beta_m - 1)$; *ln(Y/L)*=Production per worker in logarithm; β_k is the elasticity of capital per worker; β_m is the elasticity of intermediate inputs per worker; N is the number of observations; ID is the Number of legal units; Firm FE and Year FE are the individual and time fixed effects respectively.

To better interpret these results, we will make some general comments and comparisons between the different methods. With respect to the estimated elasticities, there are two important points in the general comments. First, whatever the estimator, the value of γ is negative, which means that returns to scale are slightly decreasing in the French construction. In other words, output varies less than the variation in the inputs used. This result shows that the more we produce, the more expensive it is to produce an additional unit in French construction. Moreover, we can link this dis-economy of scale to the organizational structure of the construction market.
Indeed, in microeconomics, diseconomies of scale are the cost disadvantages that firms accumulate as a result of increasing their size or output, leading to the production of goods and services at higher unit costs. For this reason, large companies, although few in number, are imposing themselves on the French construction market. We are therefore in a market where a few companies (large or medium-sized) have a certain amount of market power and where a large number of micro or small companies are under the weight of these giants: the French construction market is similar to an oligopoly situation. This result is supported by Lowe's (1987) conclusion that the construction industry deviates significantly from the perfect competition model.

Second, the elasticity of capital per worker in the production process is very low ranging from 0.0395 (FE method) to 0.0636 (Wooldridge method). This low elasticity of capital is not a surprise insofar as construction, retaining its manual character, is very labor intensive. Physical assets will inevitably be less

[10] 50 replicates are performed for the semi-parametric and Wooldridge (2009) methods. We use investment as a proxy in Wooldridge (2009).

noticeable in a sector composed mainly of micro-firms. The weakness of investment spending in the French construction industry has been noted through the study of Ferrand (2021). The author shows that, since the beginning of the 2000s, the rate of investment in French construction, on average, does not exceed 10% of GDP. This rate seems to be the norm when an international comparison is made (comparison with the United States or with the euro zone). The elasticity of per capita intermediate goods in per capita output is more than 0.77% in each of the estimates.

Comparing the estimation methods, we find that the elasticity of capital intensity provided by the FE method is always lower, while the elasticity of intermediate inputs per worker and the absolute value of the scale effect (γ) are still high. These results are consistent with the findings of Van Beveren (2012) who showed that the FE method leads to a low capital coefficient. We add that the fixed effects estimator overestimates the effect of intermediate inputs and the scale effect in our sample.

The estimation results of the other 4 methods (WRDG, OP, LP and ACF) are quite similar especially between the OP and LP methods. However, the absolute value of the scale effect obtained by the ACF method is the lowest (0.0146%). This result is probably linked to the problem of functional dependence (non-dynamic labor input) that is blamed on the OP and LP methods. The ACF method can be a good estimator of TFP in the construction sector. Not only does it correct for the OP and LP methods by making the labor coefficient dynamic, but it also highlights the crucial role of intermediate inputs in the production process. Using intermediate goods as a proxy for unobserved productivity is particularly important for the construction industry, where expenditures on equipment and machinery rentals are very large. Moreover, no amount of labor can replace the concrete, asphalt, wood and other materials needed to build.

The Wooldridge (2009) estimator provides the highest coefficient of capital per worker (0.0636%). The author states that this method has the advantage of easily obtaining robust standard deviations and makes effective use of the moment conditions implied by the OP and LP assumptions. Wooldridge (2009) argues that two-step estimators (OP and LP in this case) are inefficient for two reasons. First, they ignore the contemporaneous correlation of errors between two equations. Second, they do not effectively account for serial correlation or error heteroscedasticity. However, we have a loss of information (4,784 fewer observations) with the Wooldridge (2009) estimator due to the use of lagged instruments which is not the case with the ACF method which estimates elasticities in one step as well (second step).

Elasticities can also be obtained by calibration or by the DEA method. (sections 2.6 and 2.7). This means that no estimation is required to obtain them.

3.3 Non-parametric results

In this subsection, we present the results related to the calculation of TFP by the elasticity calibration method and by the DEA method. The following Table 5 presents the results of the calibration of the different elasticities.

Observation (N)	Labor (β_l)	Capital (β_k)	Materials (β_m)
27,544	0.21	0.14	0.65

Table 5: TFP calculation using calibration method

The calibration method has the advantage of showing that labor (21%) explains output better than capital (14%) in the construction sector. This result is consistent with the idea that the construction industry in general and the French construction industry in particular is labor-intensive. Materials (65%) contribute more to output, consistent with estimation methods. Although the calibration method requires strong assumptions, in this case the assumption of perfect markets, it provides results consistent with the reality of the sector. Based on these results, we represent the relationship between output and each factor of production (labor, capital and materials).
The most striking observation is that the capital stock explains 50.37% of output, while labor and materials explain 70.55% and 97.86% of output respectively.

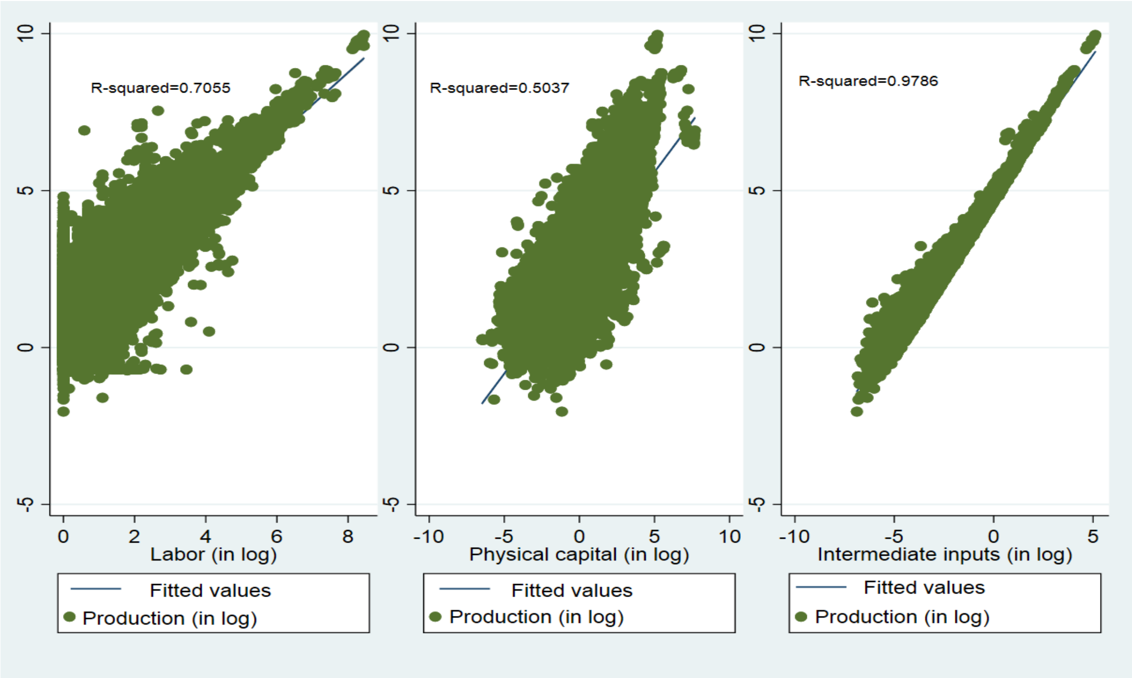

Figure 2: Relationship between output and inputs

The DEA method is used through the BCC model (Banker et al. (1984)) which assumes variable returns to scale (VRS model). Indeed, previous estimation results have shown that the constant returns assumption cannot be applied to the French construction sector. We adopt the input-oriented DEA approach because it allows us to determine the extent to which a firm's input use could be reduced if they were used efficiently to achieve the same output level. The effective decision units (DMUs) are represented by the different legal units. Since it is not possible to display the efficiency score or productivity (θ) of each of the 4,784 legal units in a given year, we will present the average efficiency by size.
The average efficiency scores are given in the following table :

Size \ Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Micro	0.49	0.49	0.50	0.49	0.49	0.48	0.47	0.50	0.51	0.50
Small	0.55	0.55	0.56	0.55	0.54	0.53	0.53	0.53	0.54	0.55
Medium	0.78	0.76	0.78	0.77	0.77	0.75	0.76	0.76	0.76	0.78
Big	0.90	0.89	0.90	0.91	0.90	0.89	0.89	0.88	0.88	0.89

Table 6: Average efficiency by size

According to Table 6, only large companies have an average score equal to about unity. In other words, all big firms have total technical efficiency. It also means that, on average, all big firms in the French construction sector operate on or very close to the production frontier from 2009 to 2018. This is a sign that big firms are using resources to their full potential. The efficiency score decreases with firm size. Medium-sized and small firms have higher average scores than micro-firms. These results confirm once again that the large companies, despite their paltry numbers, have total influence in the sector.

3.4 Comparison of TFP methods

In this subsection, we compare the different firm-level TFP calculation methods that have been obtained and focus on the correlation between them. In order to make the methods comparable, we normalize the parametric, stochastic and calibration methods: $z_{it} = \left(\frac{x_{it}-min(x)}{max(x)-min(x)}\right)$

Where $x = (x_{1t}, ..., x_{nt})$ and z_{it} is the i^{th} normalized value during the period t.
The following table provides a descriptive statistic of the TFP methods.

Methods	N	Mean	Sd	Min	Max
Fixed effects	27,544	0.103	0.035	0	1
Wooldridge (2009)	27,544	0.131	0.04	0	1
Olley and Pakes (1996)	27,544	0.125	0.038	0	1
Levinsohn and Petrin (2003)	27,544	0.117	0.038	0	1
Ackerberg, Caves and Frazer (2015)	27,544	0.115	0.037	0	1
Calibration method	27,544	0.165	0.092	0	1
Data Analysis Envelopment	27,544	0.507	0.2	0.132	1

Table 7: Summary statistics of TFP methods

Table 7 clearly shows that the TFPs obtained by the TFP estimation methods (FE, WRDG, OP, LP, and ACF) are very similar to each other. The average TFP for these 5 estimators is between 0.103 (FE estimator) and 0.131 (WRDG estimator). The different standard deviations are also very close, about 0.04 in each case. The FE, stochastic and calibration methods range from 0 to 1 because they have been normalized. The non-parametric methods (calibration and DEA methods) have slightly higher means. The calibration method provides a average TFP equal to 0.165 and the DEA method provides a mean equal to 0.507.

Finally, using the productivity levels obtained, it is possible to calculate the overall average industry productivity for each year based on each estimator.

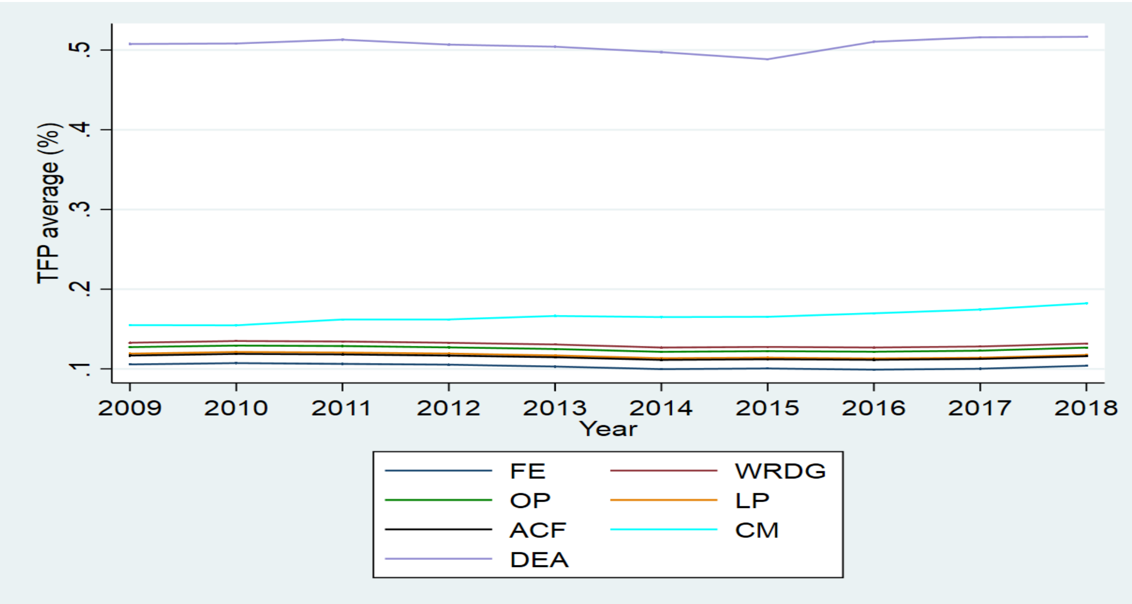

Figure 3: Total average evolution of total factor productivity

Figure 3 shows the average change in industry productivity between 2009 and 2018. The figure clearly shows that the TFP of the "residential and nonresidential building construction" industry displays a rather stable trend regardless of the method. We also add that the estimation methods are quite similar, especially the semi-parametric methods (OP, LP and ACF). The DEA method has a higher average evolution than the others. The calibration method is slightly above the rest. A similar graph by firm size for each method is also available in the Appendix.
But what about the correlation between these methods?
Table 8 below shows the Spearman rank correlation between the TFP calculation methods. Since the relationship between the TFPs is not necessarily linear, it is appropriate to use the Spearman correlation.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) TFP_{FE}	1.0000						
(2) TFP_{WRDG}	0.9283	1.0000					
(3) TFP_{OP}	0.9564	0.9942	1.0000				
(4) TFP_{LP}	0.9595	0.9909	0.9930	1.0000			
(5) TFP_{ACF}	0.9563	0.9806	0.9891	0.9937	1.0000		
(6) TFP_{CM}	-0.3569	-0.1165	-0.1641	-0.2314	-0.2605	1.0000	
(7) TFP_{DEA}	0.2379	0.3965	0.3773	0.3283	0.3232	0.4190	1.0000

Table 8: TFP correlation

Not surprisingly, there is a strong positive relationship between TFP estimation methods (more than 92%), which is consistent with the work of Van Beveren (2012). The strongest is the correlation between the OP and WRDG methods (0.9942) followed by the correlation between LP and ACF (0.9937). The correlation between the OP and LP methods is 0.9930.
However, the correlations between the non-parametric methods (DEA and calibration methods) and the other methods are relatively low, even negative with the calibration method. We have a negative correlation between the calibration method and the other methods. However, this negative correlation is less strong. The highest and lowest correlation (in absolute value) is between the CM and FE methods (0.3569) and between the CM and WRDG methods (0.1165), respectively.
The DEA method is positively correlated with the WRDG, OP, LP, and ACF methods at over 32%. The correlation between the DEA and FE methods is less significant (0.2379). The strongest correlation - positive - (0.4190) is between the two non-parametric methods (DEA and CM).

4 Conclusion

This paper reviewed some methods for calculating TFP. The first part of the paper theoretically explored the strengths and weaknesses of each TFP estimation method, ranging from parametric, semi-parametric to non-parametric techniques. In the second part, we mobilized firm-level data from 2009 to 2018 from the FARE database (Statistique structurelle annuelle d'entreprises issue du dispositif ESANE) in the French construction sector, particularly in the construction of residential and non-residential buildings. Our estimation results (especially the semi-parametric methods and the Wooldridge method) reveal few differences in input elasticities among the methods used. However, some remarks can be made.

In each of the estimates (FE, Wooldridge (2009) or semi-parametric methods), returns to scale are decreasing. This result may be related to the organizational structure of the sector in which big firms dominate the market. We are close to an oligopolistic market. Wooldridge's (2009) estimator produces a statistically significant estimate of capital relative to other approaches. Consistent with the literature, the fixed-effects estimator provides a lower coefficient on capital than the other methods. On the other hand, we have evidence that the same estimator overestimates both the value of the scale effect (in absolute value) and the coefficient on intermediate inputs.

Because the Ackerberg et al. (2015) methodology treats labor as a dynamic input whose choice affects future profits, the scale effect (in absolute value) is lower. Also, the capital elasticity obtained by the ACF method is lower than that obtained by the OP, LP and WRDG methods. However, it is likely to provide the most plausible estimates among the TFP estimation methods because it corrects for the OP and LP methods and does not lose information compared to the WRDG method.

The non-parametric TFP methods reveal two major results. First, based on the calibration method, the building sector in France is labor-intensive. Second, using the DEA method, we show that on average, all big firms operate on or very near the production frontier from 2009 to 2018. The TFP estimation methods are highly correlated with each other, except for the correlations between the non-parametric methods (DEA and calibration methods) and the other methods. They are very weak, even negative with the calibration method.

Our contribution to the literature on productivity measurement is obvious. Using several methods, we measure TFP in an economically important sector of the French economy that has faced a productivity gap in recent years. We show that returns to scale are decreasing in the sector and that the semi-parametric estimation methods and the Wooldridge (2009) method used on firm-level data are quite similar. However, as the choice of the appropriate method is strongly conditioned by the research question, the ACF method, to our knowledge, can be a good estimator of TFP in the French construction sector.

References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. *Handbook of econometrics*, 6:4171–4276. Publisher: Elsevier.

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural identification of production functions. *Munich Personal RePEc Archive.*

Ackerberg, D. A. (2016). Timing assumptions and efficiency: Empirical evidence in a production function context. *University of Michigan mimeo.*

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

Alzyadat, J. A., Alotaibi, A. F., Almubdel, A. H., and Alrshaid, N. A. (2021). Estimating total factor productivity in the Saudi Arabia construction sector. *Amazonia Investiga*, 10(40):45–53.

Baldwin, J. R., Jarmin, R. S., and Tang, J. (2002). The trend to smaller producers in manufacturing: A Canada/US comparison. *Statistics Canada, Analytical Studies-Economic Analysis, Series 1F0027MIE.*

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9):1078–1092. Publisher: INFORMS.

Banker, R. D., Charnes, A., Cooper, W. W., and Maindiratta, A. (1988). A comparison of DEA and translog estimates of production frontiers using simulated observations from a known technology. In *Applications of modern production theory: Efficiency and productivity*, pages 33–55. Springer.

Banker, R. D. and Natarajan, R. (2008). Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations research*, 56(1):48–58. Publisher: INFORMS.

Bergeaud, A., Cette, G., and Lecat, R. (2016). Productivity trends in advanced countries between 1890 and 2012. *Review of Income and Wealth*, 62(3):420–444. Publisher: Wiley Online Library.

Bjurek, H. (1996). The Malmquist total factor productivity index. *The Scandinavian Journal of Economics*, pages 303–313. Publisher: JSTOR.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric reviews*, 19(3):321–340. Publisher: Taylor & Francis.

Botrić, V., Božić, L., and Broz, T. (2017). Explaining firm-level total factor productivity in post-transition: manufacturing vs. services sector. *Journal of International Studies*, 10(3). Publisher: Centre of Sociological Research/Tsentr Sotsiolohichnykh Doslidzhen.

Boussofiane, A., Dyson, R. G., and Thanassoulis, E. (1991). Applied data envelopment analysis. *European journal of operational research*, 52(1):1–15. Publisher: Elsevier.

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

Cette, G., Fernald, J., and Mojon, B. (2016). The pre-Great Recession slowdown in productivity. *European Economic Review*, 88:3–20. Publisher: Elsevier.

Charnes, A., Cooper, W., Lewin, A. Y., and Seiford, L. M. (1997). Data envelopment analysis theory, methodology and applications. *Journal of the Operational Research society*, 48(3):332–333. Publisher: Taylor & Francis.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978a). A Data Envelopment Analysis Approach to Evaluation of the Program Follow through Experiment in US Public School Education. Technical report, Carnegie-Mellon Univ Pittsburgh Pa Management Sciences Research Group.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978b). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444. Publisher: Elsevier.

Charnes, A., Cooper, W. W., Wei, Q. L., and Huang, Z. M. (1989). Cone ratio data envelopment analysis and multi-objective programming. *International journal of systems science*, 20(7):1099–1118. Publisher: Taylor & Francis.

Chau, K. W. and Walker, A. (1988). The measurement of total factor productivity of the Hong Kong construction industry. *Construction management and economics*, 6(3):209–224. Publisher: Taylor & Francis.

Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J., and Battese, G. E. (2005a). Additional topics on data envelopment analysis. *An introduction to efficiency and productivity analysis*, pages 183–208. Publisher: Springer.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2005b). *An introduction to efficiency and productivity analysis.* springer science & business media.

Comin, D. (2010). Total factor productivity. In *Economic growth*, pages 260–263. Springer.

Cooper, W. W., Seiford, L. M., and Tone, K. (2000). Data envelopment analysis. *Handbook on data envelopment analysis*, pages 1–40. Publisher: Springer.

Crawford, P. and Vogl, B. (2006). Measuring productivity in the construction industry. *Building Research & Information*, 34(3):208–219. Publisher: Taylor & Francis.

Diewert, W. E. (1980). Capital and the theory of productivity measurement. *The American Economic Review*, 70(2):260–267. Publisher: JSTOR.

Diewert, W. E. (1988). *The early history of price index research.* National Bureau of Economic Research Cambridge, Mass., USA.

Diewert, W. E. (1995). *On the stochastic approach to index numbers*, volume 1. Department of Economics, University of British Columbia.

Ding, S., Guariglia, A., and Harris, R. (2016). The determinants of productivity in Chinese large and medium-sized industrial firms, 1998–2007. *Journal of Productivity Analysis*, 45(2):131–155. Publisher: Springer.

Dobbelaere, S. and Mairesse, J. (2007). Panel data estimates of the production function and product and labor market imperfections. *Journal of Applied Econometrics.* Publisher: ECB Working Paper.

Dvouletỳ, O. and Blažková, I. (2021). Exploring firm-level and sectoral variation in total factor productivity (TFP). *International Journal of Entrepreneurial Behavior & Research.* Publisher: Emerald Publishing Limited.

Eberhardt, M. and Helmers, C. (2010). Untested assumptions and data slicing: A critical review of firm-level production function estimators. *Department of Economics, University of Oxford.*

Felipe, J. (1999). Total factor productivity growth in East Asia: A critical survey. *The Journal of Development Studies*, 35(4):1–41. Publisher: Taylor & Francis.

Ferrand, D. (2021). Les mutations de l'investissement dans le bâtiment. *Constructif*, (2):62–67. Publisher: Fédération Française du Bâtiment.

Fischer, M. M., Scherngell, T., and Reismann, M. (2009). Knowledge spillovers and total factor productivity: evidence using a spatial panel data model. *Geographical Analysis*, 41(2):204–220. Publisher: Wiley Online Library.

Gandhi, A., Navarro, S., and Rivers, D. (2008). Estimating Production Functions with Heterogeneous Firms. Technical report, Working Paper, University of Wisconsin–Madison.

Gandhi, A., Navarro, S., and Rivers, D. (2009). Identifying production functions using restrictions from economic theory. *Unpublished, Mimeo. University of Wisconsin-Madison, Madison, USA.*

Gandhi, A., Navarro, S., and Rivers, D. (2011a). Does value added overstate productivity dispersion? Identification and estimation of the gross output production function. Technical report, Mimeo, University of Wisconsin-Madison.

Gandhi, A., Navarro, S., and Rivers, D. (2011b). On the identification of production functions: How heterogeneous is productivity? Technical report, CIBC Working Paper.

Gandhi, A., Navarro, S., and Rivers, D. (2017). How heterogeneous is productivity? A comparison of gross output and value added. *Journal of Political Economy, forthcoming.*

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016. Publisher: The University of Chicago Press Chicago, IL.

Gonçalves, D. and Martins, A. (2016). The determinants of tfp growth in the portuguese manufacturing sector. *GEE Papers*, 62. Publisher: Gabinete de Estratégia e Estudos, Ministério da Economia.

Griffith, R., Huergo, E., Mairesse, J., and Peters, B. (2006). Innovation and productivity across four European countries. *Oxford review of economic policy*, 22(4):483–498. Publisher: Oxford University Press.

Griliches, Z. (1998a). Introduction to" R&D and Productivity: The Econometric Evidence". In *R&D and productivity: The econometric evidence*, pages 1–14. University of Chicago Press.

Griliches, Z. (1998b). Productivity, R&D, and the data constraint. In *R&D and productivity: The econometric evidence*, pages 347–374. University of Chicago Press.

Griliches, Z. (2007). 12. R&D and Productivity: The Unfinished Business. In *R&D and Productivity*, pages 269–284. University of Chicago Press.

Griliches, Z. and Mairesse, J. (1983). Comparing productivity growth: An exploration of French and US industrial and firm data. *European Economic Review*, 21(1-2):89–119. Publisher: Elsevier.

Griliches, Z. and Mairesse, J. (1995). Production functions: the search for identification. *NBER working paper.*

Hammad, M. S., Omran, A., and Pakir, A. H. K. (2011). Identifying ways to improve productivity at the construction industry. *Acta Technica Corviniensis-Bulletin of Engineering*, 4(4):47. Publisher: Faculty of Engineering Hunedoara.

Harhoff, D. (1998). R&D and productivity in German manufacturing firms. *Economics of Innovation and New Technology*, 6(1):29–50. Publisher: Taylor & Francis.

Hoch, I. (1962). Estimation of production function parameters combining time-series and cross-section data. *Econometrica: journal of the Econometric Society*, pages 34–53. Publisher: JSTOR.

Hu, X. and Liu, C. (2016). Energy productivity and total-factor productivity in the Australian construction industry. *Architectural science review*, 59(5):432–444. Publisher: Taylor & Francis.

Hulten, C. R. (2007). *1. Total Factor Productivity: A Short Biography.* University of Chicago Press.

Inklaar, R. and Timmer, M. P. (2008). GGDC productivity level database: International comparisons of output, inputs and productivity at the industry level. Technical report, Groningen Growth and Development Centre, University of Groningen.

Inklaar, R. and Timmer, M. P. (2013). Capital, Labor and TFP in PWT8. 0. *University of Groningen (unpublished)*, pages 23–24.

Inklaar, R. and Timmer, M. P. (2014). The relative price of services. *Review of Income and Wealth*, 60(4):727–746. Publisher: Wiley Online Library.

Inklaar, R., Timmer, M. P., and van Ark, B. (2008a). Data for Productivity Measurement in Market Services: An International Comparison. *International Productivity Monitor*, 16.

Inklaar, R., Timmer, M. P., and Van Ark, B. (2008b). Market services productivity across Europe and the US. *Economic Policy*, 23(53):140–194. Publisher: Oxford University Press.

Ji, Y.-b. and Lee, C. (2010). Data envelopment analysis. *The Stata Journal*, 10(2):267–280. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Joshi, R. N. and Singh, S. P. (2010). Estimation of total factor productivity in the Indian garment industry. *Journal of Fashion Marketing and Management: An International Journal*. Publisher: Emerald Group Publishing Limited.

Kreuser, C. F. and Newman, C. (2018). Total factor productivity in South African manufacturing firms. *South African Journal of Economics*, 86:40–78. Publisher: Wiley Online Library.

Kumbhakar, S. C. and Lovell, C. K. (2003). *Stochastic frontier analysis*. Cambridge university press.

Lalinsky, T. (2013). Firm competitiveness determinants: results of a panel data analysis. *Available at SSRN 2548947*.

Larson, D. F., Butzer, R., Mundlak, Y., and Crego, A. (2000). A cross-country database for sector investment and capital. *The World Bank Economic Review*, 14(2):371–391. Publisher: Oxford University Press.

Lee, C. and Ji, Y.-b. (2009). Data envelopment analysis in Stata. In *Stata Conference DC*. Citeseer.

Lee, K.-R., Leem, B., Lee, C. W., and Lee, C. (2011). Malmquist productivity index using DEA frontier in Stata. *Stata Journal*, 2(2):1–9.

Levinsohn, J. and Petrin, A. (1999). When industries become more productive, do firms? Technical report, National Bureau of Economic Research.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341. Publisher: Wiley-Blackwell.

Li, Y. and Liu, C. (2010). Malmquist indices of total factor productivity changes in the Australian construction industry. *Construction management and economics*, 28(9):933–945. Publisher: Taylor & Francis.

Lipsey, R. G. and Carlaw, K. (2000). What does total factor productivity measure? *International Productivity Monitor*, 1:31–40. Publisher: Centre for the Study of Living Standards.

Lipsey, R. G. and Carlaw, K. I. (2004). Total factor productivity and the measurement of technological change. *Canadian Journal of Economics/Revue canadienne d'économique*, 37(4):1118–1150. Publisher: Wiley Online Library.

Lowe, J. G. (1987). The measurement of productivity in the construction industry. *Construction management and economics*, 5(2):101–113. Publisher: Taylor & Francis.

Lynde, C. and Richmond, J. (1993). Public capital and total factor productivity. *International economic review*, pages 401–414. Publisher: JSTOR.

Manjón, M. and Mañez, J. (2016). Production function estimation in Stata using the Ackerberg–Caves–Frazer method. *The Stata Journal*, 16(4):900–916. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Marschak, J. and Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica, Journal of the Econometric Society*, pages 143–205. Publisher: JSTOR.

Nehru, V. and Dhareshwar, A. (1994). New estimates of total factor productivity growth for developing and industrial countries. Technical report, The World Bank.

Oh, D., Heshmati, A., and Lööf, H. (2014). Total factor productivity of Korean manufacturing industries: Comparison of competing models with firm-level data. *Japan and the World Economy*, 30:25–36. Publisher: Elsevier.

Olley, G. S. and Pakes, A. (1992). The dynamics of productivity in the telecommunications equipment industry. Technical report, National Bureau of Economic Research.

Ozyurt, S. (2009). Total factor productivity growth in Chinese industry: 1952–2005. *Oxford Development Studies*, 37(1):1–17. Publisher: Taylor & Francis.

Petrin, A. and Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics*, 43(4):705–725. Publisher: Wiley Online Library.

Petrin, A., Poi, B. P., and Levinsohn, J. (2004). Production function estimation in Stata using inputs to control for unobservables. *The Stata Journal*, 4(2):113–123. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Po-Chi, C., Ming-Miin, Y. U., Chang, C.-C., and Shih-Hsun, H. S. U. (2008). Total factor productivity growth in China's agricultural sector. *China Economic Review*, 19(4):580–593. Publisher: Elsevier.

Qingwang, G. and Junxue, J. (2005). Estimating Total Factor Productivity in China [J]. *Economic Research Journal*, 6(1):51–60.

Richardson, D. (2014). *Productivity in the construction industry*. Australia Institute Canberra, ACT.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The stata journal*, 9(1):86–136. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Rovigatti, G. and Mollisi, V. (2018). Theory and practice of total-factor productivity estimation: The control function approach using Stata. *The Stata Journal*, 18(3):618–662. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Rymes, T. K. (1983). More on the measurement of total factor productivity. *Review of Income and Wealth*, 29(3):297–316. Publisher: Wiley Online Library.

Tan, W. (2000). Total factor productivity in Singapore construction. *Engineering, construction and architectural management*. Publisher: MCB UP Ltd.

Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of economic surveys*, 26(1):98–128. Publisher: Wiley Online Library.

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *The Journal of Industrial Economics*, 55(3):529–569. Publisher: Wiley Online Library.

Van Biesebroeck, J. (2008). The Sensitivity of Productivity Estimates. *Journal of Business & Economic Statistics*, 26:311–328. Publisher: American Statistical Association.

Wen, G. J. (1993). Total factor productivity change in China's farming sector: 1952-1989. *Economic Development and Cultural Change*, 42(1):1–41. Publisher: University of Chicago Press.

Wooldridge, J. M. (2009). Econometrics: Panel Data Methods.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2015a). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2):420–445. Publisher: University of Wisconsin Press.

Wooldridge, J. M. (2015b). *Introductory econometrics: A modern approach*. Cengage learning.

Xerfi|MCI (2019). Une étude exclusive sur la productivité et les principaux gisements actuels dans la construction. *BATIMAT OBSERVATOIRE DE LA CONSTRUCTION TECH*, 3.

Yasar, M., Raciborski, R., and Poi, B. (2008). Production function estimation in Stata using the Olley and Pakes method. *The Stata Journal*, 8(2):221–231. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Yismaw Ayelign, L. S. (2019). Comparison of Recent Developments in Productivity Estimation: Application on Ethiopian Manufacturing Sector. *Academic Journal of Economic Studies*, 5(3):20–31.

Zhi, M., Hua, G. B., Wang, S. Q., and Ofori, G. (2003). Total factor productivity growth accounting in the construction industry of Singapore. *Construction management and economics*, 21(7):707–718. Publisher: Taylor & Francis.

Şeker, M. and Saliola, F. (2018). A cross-country analysis of total factor productivity using micro-level data. *Central Bank Review*, 18(1):13–27. Publisher: Elsevier.

Appendix

Production function estimation using total construction

Dependent variable: *ln(Y/L)*	FE	WRDG	OP	LP	ACF
γ	-0.0520***	-0.0239***	-0.0296***	-0.0281***	-0.0107***
	(0.00143)	(0.000192)	(0.000527)	(0.000474)	(7.20e-06)
ln(K/L)	0.0578***	0.0757***	0.0707***	0.0691***	0.0560***
	(0.00105)	(0.000686)	(0.00258)	(0.00364)	(7.20e-06)
ln(M/L)	0.773***	0.704***	0.723***	0.730***	0.747***
	(0.00193)	(0.000322)	(0.00115)	(0.00452)	(7.20e-06)
N	481,348	403,994	481,348	481,348	481,348
ID	77,354	77,354	77,354	77,354	77,354
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Table 9: Production function estimates

The estimation results in Table 8 are in perfect agreement with those obtained in the body of the paper. All of these results emphasize that our sample is in harmony with the full sample.

Average change in total factor productivity by size

The following graph shows the average evolution of the TFP obtained by each method according to the size of the firm. Regardless of the method used, large firms consistently show a high average variation in TFP. For FE, CM and DEA methods, the larger the size of the company, the higher the average variation in TFP. However, with both the LP and ACF methods, micro-firms have a higher average TFP change than medium-sized firms in recent years, and even the same as large firms in 2018 with the ACF method.

Figure 4: Average change in total factor productivity by size