Firm level productivity under imperfect competition in output and labor markets*

S. Amoroso[1], B. Melenberg[1], J. Plasmans[1,2], and M. Vancauteren[3]

[1]CentER, Tilburg University
[2]University of Antwerp
[3]Hasselt University, Statistics Netherlands

October 14, 2011

PROVISIONAL DRAFT

Abstract

This paper examines the interaction between product market and labor market imperfections and discuss the impact on total factor productivity (TFP). By merging the empirical and theoretical Industrial Organization (IO) literatures on production functions, we contribute to the existing literature in two ways.

First, following the idea of McDonald and Solow (1981), later contextualized in a IO framework by Bughin (1993), we provide a simple way to model how firms deal with labor market rigidities.

Second, correcting for both simultaneity and omitted output price, we analyze to what extent the estimate of unobserved productivity is sensitive to the omission of labor market imperfection.

Using a firm-level dataset of 21 Dutch manufacturing industries over the period 1989–2008, we show that, neglecting a "wage markup", might lead to an underestimation of the true value of the price–cost margin at the aggregate level and leads to an optimistic scenario of growing firm–level productivity.

*Corresponding author: s.amoroso@uvt.nl. We thank the National Science Foundation of the Netherlands (NWO) for financial support (within the framework of the Dynamism of Innovation research project). We also thank Bert Balk (Erasmus University Rotterdam and Statistics Netherlands), Alexandra van Geen (Harvard University), participants of the Colloquium on "Dynamics of Innovation" at Tilburg University on the 19th of December 2007 and the participants of the International Workshop on "Dynamics of Innovation" at University of Antwerp on the 26th of March 2010 for fruitful discussions.

1 Introduction

In the last years, research in empirical industrial organization has focused on the endogeneity issues concerning the estimation of production functions at the establishment level. The most recent estimation techniques rely on a two–step estimation method introduced by Olley and Pakes (1996) and later extensions/modifications (Levinsohn and Petrin (2003); Wooldridge (2009)).

Parallel to the econometrical literature, another strand of research, led by the seminal papers of Hall (Hall (1986); Hall (1988); Hall (1991)), has focused on imperfections in product or in labor markets (Dickens and Katz (1987)).

Only a few studies study the possibility of having imperfect competition in both product and labor factor markets. Among others, Bughin (1993), Bughin (1996), Crépon et al. (2002), Dobbelaere (2004), and Galí et al. (2007), consider the possibility of imperfections in both product and factor markets, by taking into account that wages are no longer exogenous.

Bughin, studying the Belgian chemical industry (Bughin, 1993) and four Belgian manufacturing sectors (Bughin, 1996), considers imperfections in product and factor markets, but does not provide insights on the unobserved productivity. Moreover, he does not consider all kinds of endogeneity issues. Also in Crépon et al. (2002) and in Dobbelaere (2004), the main focus is only on the heterogeneity in price–cost markup and workers bargaining power parameters, rather than on productivity and on endogeneity issues (they do not consider selection or omitted output price biases).

With this paper, we contribute by merging these two literature, so as to provide an informative and intuitive way to model how firms deal with rigidities that prevent them to set the marginal revenue product of labor equal to its marginal cost. The starting set–up originates from the idea of McDonald and Solow (1981). They assume that both wages and labor are bargained between firms and unions. We show how this efficient bargaining affects the firm's product (pricing) decision, hence its market power. In particular, using a firm-level dataset of 21 Dutch manufacturing industries over the period 1989–2008, we show that, neglecting a "wage markup", might lead to an underestimation of the true value of the price–cost margin at the aggregate level. The underestimation of the product markups is significant and vary between 2% and 3%.

Moreover, correcting for both simultaneity and omitted output price, we analyze to what extent the estimate of unobserved productivity is sensitive to the omission of labor market imperfection.

The paper is organized as follows. In Section 2, we formulate a measure of total factor productivity (TFP) that allows for both output and labor market power. Section 3 reviews the main estimation techniques. Section 4 describes the data and results on the relevant structural parameters are reported in Section 5. In Section 6 we discuss the possible policy implications concerning the results for the TFP measure. In the final section we conclude.

2 The model

2.1 The standard setting

Assuming a Cobb-Douglas production function, the gross output Q_{it} of firm i at time t, relates to three specific inputs as follows:

$$Q_{it} = A_{it} K_{it}^{\theta_{iKt}} L_{it}^{\theta_{iLt}} M_{it}^{\theta_{iMt}}, \tag{1}$$

where K_{it} denotes capital, L_{it} labor, and M_{it} intermediate goods, the latter consisting of materials and energy, for firm i at period t. A_{it} represents the Hicksian neutral efficiency level of firm i, and is defined as TFP.[1] Taking natural logs of (1) results in a linear production function,

$$q_{it} = \theta_0 + \theta_{iKt} k_{it} + \theta_{iLt} l_{it} + \theta_{iMt} m_{it} + a_{it} \tag{2}$$

where lower-case letters refer to natural logarithms and $\ln(A_{it}) = \theta_0 + a_{it}$. θ_0 measures the mean productivity level across firms and over time, while a_{it} is an unobservable (to the econometrician, but partly known from the producer) firm and time–specific deviation from that mean.

The time-varying, input-dependent elasticity of scale θ_{it} is equal to the sum of all output elasticities with respect to the three nonnegative factor inputs:

$$\theta_{it} \equiv \sum_{k \in K,L,M} \frac{\partial Q_{it}}{\partial X_{ikt}} \frac{X_{ikt}}{Q_{it}} = \sum_{k \in K,L,M} \theta_{ikt}.$$

It is well known that direct OLS estimation of the logarithm of (2) potentially yields biased results, as the input choices are likely correlated with the error term a_{it}.

[1] MFP (*Multi-Factor Productivity*) is sometimes used interchangeably with TFP, even if there is a slight difference between what they may include. Indeed, taking into account all the factors influencing output levels can be unrealistic, therefore MFP may be a more appropriate term to use. However, the term TFP continues to be used more widely.

2.2 Omitted prices

As we observe deflated gross output, input coefficients will be biased as firm-level price variation is correlated with input choice. To see this, we can express the deflated gross output as $Y_{it} \equiv \frac{Q_{it}(P_{it})P_{it}}{P_t^j}$, where P_{it} is the price of firm i at time t, and P_t^j is the industry j ($\equiv j(i)$) price index. In logs $y_{it} = q_{it} + (p_{it} - p_t^j)$. If we then would take equation (2) with y_{it} as dependent variable, the unobserved firm-level price deviations $(p_{it} - p_t^j)$ will enter the production function (2) as an extra error component, introducing correlation with the input choices, $E(x_{it}(p_{it} - p_t^j) \neq 0)$, where $x_{it} = (l_{it}, k_{it}, m_{it})'$, and yielding biased input coefficients (Klette and Griliches, 1996).

Therefore, in order to estimate the production function consistently without information on establishment-level prices, it is necessary to impose some structure on the demand system (Foster et al., 2008).

Following Klette and Griliches (1996) and Loecker (2007), a simple conditional Dixit-Stiglitz demand system is expressed as:

$$Q_{it} = Q_t^j (P_{it}/P_t^j)^{\eta_{it}} \exp(u_{it}^d) \tag{3}$$

where Q_t^j is the sector j production index, and u_{it}^d is an idiosyncratic firm-specific demand shock. Moreover, assuming that consumers have an unbounded taste for variety, it is reasonable to assume that every firm will produce a distinct variety, and η_{it} is the firm-specific cross price elasticity of demand for differentiated goods in the industry. Taking logarithms and deriving for p_{it}, the demand in equation (3), can be rewritten as

$$p_{it} = p_t^j + \frac{1}{\eta_{it}}(q_{it} - q_t^j) - \frac{1}{\eta_{it}} u_{it}^d. \tag{4}$$

Taking into account the demand (4), the log deflated output can be expressed as

$$y_{it} = q_{it} + \frac{1}{\eta_{it}}(q_{it} - q_t^j) - \frac{1}{\eta_{it}} u_{it}^d. \tag{5}$$

Finally, substituting q_{it} with the production function (equation (2)), the deflated gross output can be written as:

$$y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + \gamma_{iLt}l_{it} + \tilde{a}_{it} - \frac{1}{\eta_{it}} q_t^j + u_{it}, \tag{6}$$

where the error $u_{it} \equiv -u_{it}^d/\eta_{it}$, the unobserved productivity $\tilde{a}_{it} \equiv a_{it}/\mu_{it}$ $\gamma_{i0t} \equiv \theta_0/\mu_{it}, \gamma_{ikt} \equiv \theta_{ikt}/\mu_{it}, k = K, L, M$, and $\mu_{it} \equiv \eta_{it}/(1 + \eta_{it})$ is defined as the markup for $\eta_{it} < -1$, such that $-1/\eta_{it} = (\mu_{it} - 1)/\mu_{it} > 0$.

2.3 Labor market rigidities: union bargaining power

We now relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power.

Many authors have studied the influence of market power of unions, by introducing wage rigidities through efficiency wages. For instance, Hall (1991)'s model assumes that the firm wages and level of employment are jointly determined according to an efficient bargaining scheme between the firm and its workers. Following the McDonald and Solow (1981) efficient bargaining model, in which both wage and employment are bargained between firms and their workers, we show that the wage of workers is determined at a level which is higher than the firm's marginal revenue of labor. Workers in firms with some degree of market power on the output market can earn wages that are much higher than the competitive industry wage level.

Introducing the nominal input prices R_{it}, W_{it}, and Z_{it} as firm i's rental price of capital, wage rate, and unit price for intermediate goods, respectively, the efficient bargaining model can be summarized as follows.

The workers in the firm bargain with the firm over both the levels of employment L_{it} and of the wage W_{it}. According to McDonald and Solow (1981) the workers' objective in their efficient bargaining model can be specified in two alternative ways. Either as the unions aggregate gain to the workers from membership, $L_{it}(W_{it} - \overline{W}_{it})$, or, taking account of the unemployment benefits, as $L_{it}W_{it} + \overline{W}_{it}\ (N_{it} - L_{it})$, where $\overline{W}_{it}$ is the reservation wage (i.e. the theoretical wage valid on an imperfectly competitive output market and a perfectly competitive labor market), W_{it} the negotiated wage, and N_{it} is the labor supply. McDonald and Solow (1981) judge the first specification as the most appropriate one for real life. In fact, in the second specification, if $\overline{W}_{it}$ falls, the firm would have to increase its wage offer to make up for a reduction in $\overline{W}_{it}$, to keep the level of union utility unchanged. Hence, we advocate McDonald and Solow (1981)'s suggestion and take $L_{it}(W_{it} - \overline{W}_{it})$ as the union utility function.

The firm's objective is to maximize its short run profit, given by the difference between the total revenue and the total costs, i.e., as $P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it}$.

The efficient bargaining model can be written as a weighted average of the logarithms of workers' aggregate gain from union membership and the firm's short run profit:

$$\max_{W_{it},L_{it}} \left[\phi_{it} \log(L_{it}(W_{it} - \overline{W}_{it})) + (1-\phi_{it}) \log(P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it})\right],$$

where $\phi_{it} \in [0,1]$ is the degree of union bargaining power.

The best-known formal solution to this efficient bargaining model is Nash's one. Hence, maximizing with respect to employment and to wage, and then combining the two first order conditions, yields the reservation wage:[2]

$$\overline{W}_{it} = \frac{\partial P_{it}(Q_{it})Q_{it}}{\partial L_{it}} = \frac{P_{it}(Q_{it})}{\mu_{it}}\frac{\partial Q_{it}}{\partial L_{it}}. \tag{7}$$

Further, we can express the bargained wage rate as a function of the bargaining parameter, ϕ_{it}, and the ratio between profits and cost of labor (see the Appendix for the derivation):

$$\frac{W_{it} - \overline{W}_{it}}{W_{it}} = \frac{\phi_{it}}{1-\phi_{it}}\frac{(P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it})}{L_{it}W_{it}}. \tag{8}$$

Now, defining $\frac{W_{it}-\overline{W}_{it}}{W_{it}} \equiv h_{it}$ as the wage markup, one can see how this is directly depending on the union's bargaining power.

2.3.1 Discussion

From the solution of the efficient bargaining model, we can derive a simple expression for the elasticity of labor. In particular, multiplying both sides of (7) by L_{it}/Q_{it} and after some simple algebraic manipulations, we can express the labor elasticity as a function on the labor share and the wage markup:

$$\gamma_{iLt} \equiv \theta_{iLt}/\mu_{it} = s_{iLt}(1 - h_{it}). \tag{9}$$

We then make use of (9) in the estimating equation (6), to derive the labor elasticity, as well as the wage markup. To retrieve the bargaining parameter ϕ_{it}, we use equation (8). The resulting production function is:

$$y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + (1-h_{it})s_{it}l_{it} - \frac{1}{\eta_{it}}q_t^j + \tilde{a}_{it} + u_{it}, \tag{10}$$

where we can directly obtain an estimate of h_{it}, along with the estimation of the other regression parameters.[3].

---

[2]The complete derivation of the efficient bargaining model is provided in Appendix.

[3]We can alternatively specify the labor elasticity as $\gamma_{iLt} = s_{iLt}\left(1 - \frac{\phi_{it}}{1-\phi_{it}}\frac{\Pi_{it}}{L_{it}W_{it}}\right)$, where $\Pi_{it} \equiv P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it}$. Plugging this expression into the estimating equation, we can directly obtain an estimate of $\frac{\phi_{it}}{1-\phi_{it}}$. Indeed the resulting estimating equation becomes:

$$y_{it} - s_{it}l_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} - \frac{\phi_{it}}{1-\phi_{it}}s_{it}l_{it}\frac{\Pi_{it}}{L_{it}W_{it}} - \frac{1}{\eta_{it}}q_t^j + \tilde{a}_{it} + u_{it}. \tag{11}$$

We use this specification as a robustness check.

At this stage, it is intuitively clear how the exclusion of frictions of the labor market might lead to underestimating the firm's market power. As a matter of fact, when there is no imperfect competition on the labor market, firms set the wage at the lowest value possible, ultimately equal to the competitive wage $\overline{W}_{it} = W_{it}$ (and, therefore, $h_{it} = 0$). For W_{it} that tends to $\overline{W}_{it}$, the markup decreases, given that the elasticity and the share of labor are constant, which is inversely related to the output markup μ_{it}.

This apparently direct relationship between the wage and the product price markup could be mistakenly interpreted as if the larger the firm's rent, the larger the wage markup (as in Dobbelaere (2004)). However, it is just an underestimation of the true level of price–cost margins that is caused by the omission of direct effects of wage bill on marginal costs (Bughin, 1993). As a matter of fact, finding a significant estimate for the wage markup parameter h_{it} means that the workers' unions have a degree of bargaining power which erodes the existing monopoly rents. Therefore, price–cost margins and wage markups should be negatively correlated.

Estimation of equation (10) can be done following several estimation approaches. In section 4, we will discuss results from GMM estimation and other more recent approaches, such as the Levinsohn and Petrin (2003)'s two–stage (LP, 2003), and the Wooldridge (2009) GMM. To follow these two latter approaches one needs to make certain timing assumptions which might be incompatible with our model. Therefore, we report results from these other two approaches only for the whole manufacturing industry and we do not discuss them in detail at the sectoral level.

3 Review of estimation techniques

Before discussing the results, we want to review and discuss the most recent estimation techniques.

3.1 The within and GMM estimators

Assuming that the unobserved productivity is constant over time ($\tilde{a}_{it} = \tilde{a}_i$), the potential endogeneity between $\tilde{a}_i$ and the regressors is controlled by exploiting the panel structure of the data (fixed–effects model). This approach, besides having the evident disadvantage of assuming no changes of productivity over time, requires a strict exogeneity assumption, where the error term u_{it} is assumed to be uncorrelated with input choices.

The reason why this assumption does not hold is because u_{it} could contain information on exogenous and unpredictable shocks (unexpected deregulation or privatization) that may affect the inputs prices, thus the inputs choices.

If we are not willing to assume strict exogeneity and we believe that $\tilde{a}_{it}$ evolves over time instead, we need to solve the simultaneity issue by means of instrumental variables (IV). The choice of the GMM estimator, in a static framework, can be legitimized by a small exit rate. Given that the exit rate is equal to 2.7%[4], the selection bias is negligible and we can discount the potential negative correlation between capital input and the error term (see Olley and Pakes (1996); OP, for brevity). Furthermore, given our choice of static representation of a firm's production function, the simultaneity between the firms' inputs choice and their knowledge of some part of the error term $\tilde{a}_{it}$ (that can be interpreted as technology and/or management) can be controlled in two ways.

Moreover, to allow for arbitrary heteroskedasticity and clustering on firm, GMM estimation is required.

The moment condition necessary to estimate equation (10) is:

$$E\left((\tilde{a}_{it} + u_{it}) \cdot x_{it}\right) = 0$$

where $x_{it} \equiv (q_t^j, l_{it-1}, l_{it-2}, m_{it-1}, m_{it-2}, k_{it-1}, k_{it-2})$

The cost of adopting this static approach is that we do not allow for the possibility that the unobserved productivity could be correlated with past choices of inputs.

3.2 The Olley and Pakes (1996) estimation algorithm and its derivatives

The other two approaches are a two–step estimation method introduced by OP (1996) and later modified by LP (2003) and a variant of the LP estimator proposed by Wooldridge (2009) who provides sufficient orthogonality conditions to identify all the relevant parameters and to obtain efficient GMM estimates and standard errors in one step.

Both the LP and Wooldridge approaches rely on the OP specification of a dynamic model that controls for both simultaneity and selection problems.

[4] See the description of the data.

At each period, the firm decides whether leaving the market or continuing to produce. If it exits, it will receive a sell–off value; if it continues, it maximizes the current profits and the expected profits of next period, given the investment and exit decisions. Both investment and exit depend on the level of productivity and on the state variables. To select the state variables, OP make a crucial assumption on the timing and dynamic nature of inputs, that is some inputs are more "dynamic in nature" than others.

In particular, OP define the capital as a dynamic input, because its choice of the current period affects the choice of the next period. This directly implies that the choice of labor has no dynamic implications. But if labor contracts are long term, as in unionized industries, then it is evident how the assumption of non-dynamicity of the labor is void.

Therefore, if labor has to be considered as a dynamic input, the optimal static solution of the efficient bargaining clearly cannot fit into the Markov-perfect dynamic model proposed by OP.

However, if we are willing to assume that labor does not enter the profit function, and, more importantly, that the bargained level of labor today has no impact on the future bargained level of tomorrow, we can then use the LP approach.

LP propose to use intermediate input to proxy the unobserved productivity at the firm-level.[5]

Therefore, we can express the unobserved productivity $\tilde{a}_{it}$ as a function of intermediate inputs and capital. In particular, the demand function for intermediate inputs is expressed as a function of capital and productivity, $m_{it} = f_t(a_{it}, k_{it})$. f is indexed by t to allow input prices and market conditions to vary across time.

Under the assumption that intermediate input demand is monotonic in a_{it}, one can invert (as in OP):

$$a_{it} = f_t^{-1}(m_{it}, k_{it}) \equiv g_t(m_{it}, k_{it}). \tag{12}$$

Using the inverse demand function specified in equation (12), we can

[5] OP show how investment can be used as a proxy for unobserved productivity. However, when inverting the investment rule, in order to express productivity as a function of investment, one condition is that the investment variable has to be strictly positive. As investment, in the data, are often zero, LP propose the use of intermediate input such as materials and energy as a proxy for unobserved productivity, to overcome the problem of lumpy investment.

rewrite the estimating equation (10) as:

$$y_{it} = \gamma_{iLt} l_{it} + \varphi_t(m_{it}, k_{it}) - \frac{1}{\eta_{it}} q_t^j + u_{it}, \tag{13}$$

where

$$\varphi_t(m_{it}, k_{it}) \equiv g_t(m_{it}, k_{it})/\mu_{it} + \gamma_{i0t} + \gamma_{iKt} k_{it} + \gamma_{iMt} m_{it}$$

is a high order polynomial in m_{it}, k_{it}.

Under the conditional assumption that

$$E(u_{it}|l_{it}, m_{it}, k_{it}, q_t^j) = 0, \tag{14}$$

we are able to identify the polynomial $\varphi_t(m_{it}, k_{it})$ and the regression parameters γ_{iLt}[6] and η_{it}. In other words, we are able to obtain estimates of the labor elasticity (and, consequently, of the ratio between the bargained and the reservation wage $\frac{\overline{W}_{it}}{W_{it}}$, and of the bargaining parameter ϕ_{it}), and of the markup μ_{it}.

Finally, as in OP and LP, we assume that the technological progress follows a first order Markov process $a_{it} = E(a_{it}|a_{it-1}, \chi_{it}) + \xi_{it}$, where χ_{it} is a survival indicator variable that corrects for selection bias[7], ξ_{it} represents the innovative shock to productivity and is assumed to be uncorrelated with productivity and capital in the previous period $t-1$.

Moreover, if for some function $h_t(\cdot)$, $E(a_{it}|a_{it-1}, \chi_{it}) \equiv h_t[g_t(m_{it-1}, k_{it-1}; P_{it})]$ (Wooldridge, 2009), where P_{it} is the probability that firm i in sector j survives in the next period ($P_{it-1} \equiv P(\chi_{it} = 1)$).

Therefore, we can rewrite the first order Markov technological progress as $a_{it} = h_t[g_t(m_{it-1}, k_{it-1}; P_{it})] + \xi_{it}$.

To identify the capital and materials coefficients, the identifying moment condition is

$$E(\xi_{it}|k_{it}, m_{it-1}, k_{it-1}) = 0. \tag{15}$$

[6] According to Ackerberg et al. (2006), collinearity between labor and the polynomial in materials and capital in this first stage of the estimation algorithm can cause the labor coefficient to be unidentified. This collinearity arises from the fact that labor, like materials and capital, needs to be allocated in some way by the firm, at some point in time. However, as the collinearity derives from the assumptions made on the perfect variability and non–dynamicity that characterize both labor and material inputs, one can still allow equation (13) to have some identifying information on θ_{iLt}.

[7] Both theoretical and empirical literatures (cfr. Ackerberg et al. (2007) for a review) on entry and exit point at the causal relationship between exit of firms and productivity differences at the firm level, leading to the conclusion that higher productivity will lower the firm's exit probability.

Substituting this expression into equation (10) and rearranging terms, yields:

$$\begin{aligned} y_{it} &= \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + \gamma_{iLt}l_{it} + \\ & h_t[g_t(m_{it-1}, k_{it-1}; P_{it})] - \frac{1}{\eta_{it}} q_t^j + \xi_{it} + u_{it}. \end{aligned} \tag{16}$$

Estimation of equation (16) can be done in two steps (as in OP and LP) or in one single step, extending the moment conditioning set of the second stage of LP.

The LP estimation strategy consists of two stages. In the first stage, with a simple GLS estimation of equation (13), estimates of the polynomial $\varphi_t(m_{it}, k_{it})$ and the regression parameters γ_{iLt} and η_{it} are obtained.

Substituting $\hat{\gamma}_{iLt}$ and $\hat{\eta}_{it}$, and $\hat{\varphi}_t(m_{it-1}, k_{it-1})$ into (16), and approximating $h_t(\cdot)$ with a flexible polynomial, NLLS estimation yields ($\sqrt{n}$) consistent estimates of materials and capital input elasticities.

The approach of Wooldridge takes into account the potential collinearity issue in the first stage of LP (Ackerberg et al., 2006) and considers the estimation of equation (16). In particular, his identifying strategy consists in relaxing the assumption on what is included in firm's i (in sector j) information set I_{it}[8], allowing for the possibility that also the labor input is included. Therefore, the conditional assumption used in LP's second stage (equation (15)) is extended to any past outcomes of (l_{it}, m_{it}, k_{it}): $E(\xi_{it}|k_{it}, l_{it-1}, m_{it-1}, k_{it-1}, \ldots, l_{i1}, m_{i1}, k_{i1}) = 0$.

Wooldridge, then, considers estimation (16) suggesting to use polynomials of order three or less to approximate the function $h(\cdot)$.

4 Data

We extract data from Statistics Netherlands for the years 1989-2008. The output and the input variables are defined as follows. As an output measure, we use the deflated value of gross output Y_{it} ($\equiv \frac{Q_{it}P_{it}}{P_t^j}$) of each firm i in sector j in period t. Labor (L_{it}) refers to the number of employees in each

[8]In OP and LP, the technological progress a_{it} is assumed to follow a first–order Markov process, where past realizations of a_{it} constitutes the information set. In other words, $a_{it} = E(a_{it}|I_{it-1}) + \xi_{it} = E(a_{it}|a_{it-1}) + \xi_{it}$, but since a_{it-1} is expressed as an inverse function of m_{it-1} and k_{it-1}, the conditional expectation $E(\xi_{it}|I_{it-1}) = E(\xi_{it}|a_{it-1}) = E(\xi_{it}|k_{it}, m_{it-1})$, $k_{it} \in I_{it-1}$ only depends on them, that is, the firm i does not take into account the level of labor in its information set.

firm for each year,[9] collected in September of that year. The corresponding wages W_{it} include gross wages plus salaries and social contributions before taxes. The costs of intermediate inputs ($Z_{it}M_{it}$) include costs of energy, intermediate materials and services. The unit user costs R_{it} of capital stock K_{it} are calculated as the sum of the depreciation of fixed assets and the interest charges.

The nominal gross output and intermediate inputs are deflated with the appropriate price indices from the input-output tables available at the NACE rev. 1 two-digits sector classification.[10] For capital, we use a two-digit NACE deflator of fixed tangible assets calculated by Statistics Netherlands.

The data extracted from the Production Survey (PS) constitutes an unbalanced panel data of 6727 firms (with a minimum of 2001 firms in 2001 and a maximum of 5607 enterprises in 2006 and 1997) with 65866 observations spanning over 20 years and over 21 industries. Firms with missing data on one of the variables used in the empirical analysis are omitted. We exclude firms exhibiting inputs growth of more than 200 percent or less than -50 percent (3822 observations dropped). We also exclude from the sample firms with an output growth of more than 300 percent or less than -90 percent (1372 observations).

Finally, the resulting sample consists of 60672 observations (6718 firms).

Throughout our sample period, the PS surveys included some changes in their population designs resulting in an unbalanced panel of the entire population. As a result, we cannot distinguish whether the entry or exit rates of firms resulted from survey response behavior or real economic structural behavior. The number of firms (N) for each NACE rev. 1 industry is calculated by Statistics Netherlands. Table **??** reports the sectors that were chosen with a corresponding NACE two-digit code and the corresponding number of firms.

Table 4.1 reports the means, medians, standard deviations, and first and third quartiles of the included data for our main variables. In particular, a summary of the logarithms of deflated revenues and of the inputs, along with

[9]For each enterprise, jobs are added and adjusted for part-time and duration factors, resulting in number of men/years expressed as Full Time Equivalents (FTEs)(*Source:* Statistics Netherlands)

[10]NACE Rev. 1 is a 2-digit activity classification which was drawn up in 1989. It is a revision of the General Industrial Classification of Economic Activities within the European Communities, known by the acronym NACE and originally published by Eurostat in 1970.

Table 4.1: Descriptive Statistics

variable	mean	sd	median	p25	p75	N
y_{it}	8.896	1.401	7.971	8.758	9.721	60835
l_{it}	4.048	1.061	3.367	3.932	4.663	"
m_{it}	8.361	1.533	7.361	8.265	9.297	"
k_{it}	5.561	1.708	4.570	5.516	6.584	"
s_{iLt}	0.270	0.127	0.179	0.258	0.345	"
s_{iMt}	0.615	0.147	0.520	0.620	0.716	"
s_{iKt}	0.045	0.044	0.019	0.035	0.058	"
age_{it}	7.055	4.863	3	6	10	"
$exit_{it}$	0.027	0.162	–	–	–	"
q_t^j	0.940	0.153	0.853	0.966	1.039	"

Note: p25 and p75 are, respectively, the 25th and the 75th percentile.

input shares in revenue, is presented. As one can see, the dispersion of the logarithms of deflated output and inputs is considerably large.[11]

During 1989-2008, the capital input constitutes almost 11 percent of gross output on average. The mean share of labor is almost 28 percent, and intermediate inputs constitute more than half of gross output (61.1 percent). Moreover, the relative dispersion of all these variables is considerably large, especially for the share of capital. The exit rate is quite small (8.2 percent) and 75 percent of the firms have been active on the market for 3 to 9 years.

5 Empirical results for the complete sample of Dutch firms

In this section we present results for the entire manufacturing industry over the period 1989-2008, without looking at the potential heterogeneity in the structural parameters among sectors and over time. In other words, we impose that the price elasticities and the labor market friction parameters, along with the input factor elasticities, are constant and independent of the number of produced goods (Loecker, 2007), i.e. $\eta_{it} = \eta$, $h_{it} = h$, $\gamma_{iMt} = \gamma_M$, and $\gamma_{iKt} = \gamma_K$. The estimating equation (10) becomes:

$$y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1-h) s_{it} l_{it} - \frac{1}{\eta} q_t^j + \tilde{a}_{it} + u_{it} \qquad (17)$$

[11]Averages over time and standard deviations for each sector are reported in Table ?? in Appendix C.

Table 5.1: GMM: Results for the whole manufacturing industry

	a	*b*	*c*
$\hat{\theta}_L$	0.069	0.235	0.211
	0.002	*0.008*	*0.006*
$\hat{\theta}_M$	0.983	0.811	0.732
	0.016	*0.012*	*0.004*
$\hat{\theta}_K$	0.078	0.068	0.062
	0.004	*0.003*	*0.003*
$\hat{\theta}$	1.131	1.113	1.005
	0.018	*0.016*	*0.002*
$\hat{\mu}$	1.136	1.108	–
	0.018	*0.016*	
h	0.779	–	–
	0.006		
$\hat{\phi}$	0.640	–	–
	0.002		
Time dummies	Yes	Yes	Yes
Sectoral dummies	Yes	Yes	Yes

Note: Sample period 1989-2005; dependent variable: log. gross deflated output y_{it}
a: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1-h)s_{it}l_{it} - \frac{1}{\eta_i}q_t^j + \tilde{a}_{it} + u_{it}$, where $1 - h = \frac{\overline{W}_{it}}{W_{it}}$
b: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} - \frac{1}{\eta_i}q_t^j + \tilde{a}_{it} + u_{it}$
c: $y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + \gamma_L l_{it} + a_{it}$.

These constraints are clearly restrictive, because they imply no differences in cross price elasticities or in wage markup across firms and/or sector. Therefore, section 4.2 explores the empirical results where the relevant parameters are allowed to differ among product segments.

Table 5.1 reports estimation results for all the relevant parameters of equation (17). The Table is divided in three columns. Column *a* reports estimates of all the relevant structural parameters of the estimating equation. The second column, *b*, contains the regression coefficients of the model correcting only for output imperfection and omitted price. Results without corrections for omitted price bias and labor imperfect competition are reported in the last column, *c*.

Table 5.2: LP&Wooldridge GMM: Results for the whole manufacturing industry

	Levinsohn&Petrin			Wooldridge–LP GMM		
	a	*b*	*c*	*a*	*b*	*c*
$\hat{\theta}_L$	0.066	0.238	0.220	0.058	0.212	0.196
	0.002	*0.006*	*0.005*	*0.002*	*0.008*	*0.006*
$\hat{\theta}_M$	0.949	0.934	0.699	0.856	0.735	0.677
	0.055	*0.055*	*0.050*	*0.025*	*0.020*	*0.017*
$\hat{\theta}_K$	0.075	0.036	0.068	0.076	0.060	0.057
	0.026	*0.018*	*0.016*	*0.004*	*0.004*	*0.004*
$\hat{\theta}$	1.090	1.207	0.987	0.989	1.007	0.930
	0.030	*0.042*	*0.036*	*0.021*	*0.023*	*0.016*
$\hat{\mu}$	1.100	1.080	–	1.118	1.079	–
	0.014	*0.012*		*0.017*	*0.016*	
h	0.778	–	–	0.811	–	–
	0.005			*0.006*		
$\hat{\phi}$	0.640	–	–	0.647	–	–
	0.001			*0.001*		
Time dummies	*Yes*	"	"	"	"	"
Sectoral dummies	*Yes*	"	"	"	"	"

Note: Sample period 1989-2005; dependent variable: log. gross deflated output y_{it}

a: $y_{it} = (1-h)s_{iLt}l_{it} + \varphi_t(m_{it}, k_{it}) - \frac{1}{\eta}q_t^j + u_{it}$, where $1 - h = \frac{\overline{W}_{it}}{W_{it}}$

b: $y_{it} = \gamma_L l_{it} + \varphi_t(m_{it}, k_{it}) - \frac{1}{\eta}q_t^j + u_{it}$

d: $y_{it} = \theta_L l_{it} + \varphi_t(m_{it}, k_{it}) + a_{it}$.

The first result we are going to discuss is the evidence of imperfect competition on the labor market.

... ...

Results for OP-derived estimation strategies are reported in Table 5.2. The Table reports results for the entire manufacturing market and is organized in two parts. One displays estimation results of equation (13), and the other reports the ones of equation (16). Each part (Levinsohn & Petrin and Wooldridge–LP) consists of three columns. Columns *a* report estimates of all the relevant structural parameters of the two estimating equations. The second columns, *b*, contain the regression coefficients of the model correcting only for output imperfection and omitted price. [12] Results of LP's estimating equation (without corrections for omitted price bias and imperfect competition) are reported in the last columns, *c*.

[12]Note that this model has been first proposed by Klette and Griliches (1996), but Loecker (2007) was the first to implement correction for output market imperfection into the semiparametric estimation framework introduced by OP (1996).

5.1 Across-Industry Estimates

Given the evidence of sectoral specificity of capital and labor[13] (Dosi (1999); Ramey and Shapiro (2001)), we investigate the heterogeneity of the manufacturing industry, by studying across-industry firms' production behavior.

For each of the 21 industries, we estimate equation (10) with and without the extension to labor imperfections, by means of the GMM estimator, as it is a robust estimation technique. As instruments we take an appropriate number of lagged inputs levels (up to two lags, depending on the size of the sector). Year dummies are always included. The estimated parameters are reported in Table 5.3. Testing the hypothesis of heterogeneity across sectors yields the conclusion that all the structural parameters statistically differ from sector to sector for the the majority of cases.

This confirms the assumption of sectoral specificities. Each sector has its own functioning, and the firm belonging to a specific sector adopts a different production strategy compared to a firm in another sector.

Quite consistently with what we found for the whole manufacturing industry, excluding imperfections on the labor market leads to an underestimation of the markups. Indeed, 60% of the sector (15, 20, 21, sectors from 23 to 29, and sectors 34 and 35) show underestimated markups and the underestimation is significant at the 5%, except for sector 20.

The estimated price–cost margins of the rest of the sectors either do not change or significantly increase.

When assuming imperfect competition on the labor market and on the output market, the 75 percent of the cases (ranging from 0.795 in the wearing apparel sector, to 1.324 in the chemicals sector) firms set their prices above the marginal costs (markups significantly larger than one). Compared with other studies, we find relativity small markups.

Indeed, Dobbelaere (2004) finds estimates ranging from 1 to 2.268 and finds constant and/or decreasing returns to scale in 60 % of the Belgian sectors. Our results are more in line with Bughin (1996) findings. He reports estimates of profit margins ranging from 0.14 to 0.27, corresponding to markups of 1.16 and 1.37, respectively.

However, none of these studies corrected for omitted output price and

[13] The "social embeddedness" of firms' routines and strategies is likely to be driven by socially specific factors, such as the nature of the local labor markets, workforce training institutions, financial institutions. Furthermore, Ramey and Shapiro (2001) suggest significant sectoral specificity of physical capital and substantial costs of redeploying the capital.

Table 5.3: GMM estimates of μ_{it} and ϕ_{it} for 21 industries

NACE			Hp:$\phi = 0$	
j	$\widehat{\mu}^j$	$\hat{\phi}$	$\widehat{\mu}^j$	N.obs
15	1.063 (0.012)	0.585 (0.002)	1.015 (0.011)	7750
17	0.992 (0.016)	0.636 (0.002)	1.047 (0.016)	1581
18	0.795 (0.016)	0.630 (0.001)	0.991 (0.018)	396
19	0.949 (0.024)	0.623 (0.001)	1.014 (0.016)	347
20	0.953 (0.010)	0.692 (0.001)	0.940 (0.009)	1969
21	1.093 (0.018)	0.594 (0.002)	1.022 (0.014)	2500
22	1.002 (0.012)	0.634 (0.002)	1.059 (0.012)	6255
23	1.101 (0.044)	0.454 (0.001)	1.026 (0.049)	193
24	1.324 (0.023)	0.504 (0.002)	1.208 (0.019)	3885
25	1.114 (0.015)	0.593 (0.002)	1.048 (0.013)	3202
26	1.168 (0.023)	0.534 (0.001)	1.113 (0.018)	2704
27	1.196 (0.027)	0.606 (0.001)	1.090 (0.020)	1240
28	1.120 (0.014)	0.656 (0.002)	1.053 (0.011)	9885
29	1.133 (0.015)	0.666 (0.001)	1.076 (0.012)	8554
30	1.011 (0.057)	0.613 (0.002)	1.124 (0.049)	180
31	1.079 (0.020)	0.629 (0.001)	1.112 (0.020)	1808
32	1.136 (0.041)	0.619 (0.002)	1.143 (0.032)	398
33	1.158 (0.027)	0.653 (0.001)	1.204 (0.023)	2062
34	1.093 (0.018)	0.682 (0.001)	1.018 (0.015)	1574
35	1.133 (0.022)	0.670 (0.001)	1.086 (0.016)	1523

Note: Estimating equation: $y_{it} = \gamma_0 + \gamma_K^j k_{it} + \gamma_M^t m_{it} + (1-h)^j s_{it} l_{it} - \frac{1}{\eta^j} q_t^j + \tilde{a}_{it} + u_{it}$. Results for sector 36 have not been reported as statistically non significant and not interpretable. Standard errors in parentheses; sample period 1989-2008.

imperfect competition, simultaneously. Unfortunately, we do not know any other study that provide empirical evidence of the rent-sharing in the labor market for the Netherlands, so as to compare.

Comparing the parameters μ^j and ϕ^j, we see that firms share their monopoly rents with labor unions, which have quite large bargaining elasticities (ranging from the lowest 0.454 of sector 23, to the highest 0.692 of sector 20, on average).

Moreover, both the wage markup and the bargaining elasticity are negatively correlated with the price–cost margins. In particular the correlation coefficient between ϕ^j and μ^j is -0.369.

This confirms the idea that strong unions erode firms' rents.

This result is in contrast with the findings of Crépon et al. (2002) and Dobbelaere (2004), who find positive correlation. Dobbelaere (2004) interprets this positive correlation between labor bargaining and output market power as the effect of the exit of firms. In particular, she guesses that strong unions reduce the firms' share of rents, forcing some of the firms to exit the market, therefore decreasing the degree of market power.

Another explanation she provides deals with the fact that stronger unions are attracted by those sectors where rents can be extracted.

To our advise, both these two interpretations have underlining problems. The first interpretation does not mesh well with the static setting adopted in the paper and implied in the price–cost margin definition, which does not allow for the dynamic aspects of competition (such as the implications of selection and reallocation effects).

The second interpretation concerns more the profitability of the firm, rather than its level of price–cost margin. A more profitable firm can attract workers that are able to extract some of the surplus. But a higher markup does not necessarily mean that the enterprise is profitable, as it does not take into account relative cost efficiencies (see Boone and van der Wiel (2007); Boone (2008); Griffith et al. (2008) for a discussion on relative profits and relative cost efficiencies). Therefore, we tested the correlation between the wage markup and the relative profits measure (computed as in Boone and van der Wiel (2007)). We find that indeed these two measures, profit elasticity and union power, are positively correlated ($\rho = 0.46$, significant at the 5% level. Results of the profit elasticities per sector are not reported, but available upon request.)

6 Policy Conclusions: Impact on TFP

In this paper we propose a measure of TFP derived from estimating a production function which accounts for both imperfect competition on the labor market and on the output market.

Assuming a Cobb-Douglas production technology, the production function derived in Section 2 is

$$y_{it} = \gamma_{i0t} + \gamma_{iKt}k_{it} + \gamma_{iMt}m_{it} + (1 - h_{it})s_{it}l_{it} - \frac{1}{\eta_{it}}q_t^j + \tilde{a}_{it} + u_{it},$$

where $\gamma_{i0t} \equiv \theta_0/\mu_{it}, \gamma_{iKt} \equiv \theta_{iKt}/\mu_{it}, k = K, L, M$, and $\tilde{a}_{it} \equiv a_{it}/\mu_{it}$. a_{it} represents the productivity that is not directly observable, i.e. the TFP.

Therefore, we consider a measure of TFP computed as the following:

$$\hat{a}_{it} \equiv \hat{\mu}_{it}(y_{it} - (\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (\widehat{1-h})s_{it}l_{it} - \frac{1}{\hat{\eta}}q_t^j)). \tag{18}$$

To compute the aggregate TFP index, we follow De Loecker and Konings (2006) and taking a share weighted sum of the firm-level TFP computed on the entire sample of firms.

The firm–specific weights consist of output based market shares. We also standardized the TFP index to 1 in 1989, so as to compare the evolution over time of the different structural assumptions and estimators.

As a matter of fact, one interesting aspect of studying the firm's behavior under imperfect competition is comparing the levels of TFP under different market structure assumptions.

Namely, imperfect competition on both output and labor markets, imperfect competition on the output market, and perfect competition.

As we adopted three different econometric approaches, we are also able to compare the TFP levels among estimation strategies.

Figure 1 shows three panels, each representing the standardized weighted average TFP level over the period 1989–2008 obtained from three estimation approaches. For each estimation technique, we also reported TFP indexes for the three different assumptions on competition.

The solid lines represent the firms' average productivity when assuming imperfect competition on both markets ($tfp(a)$). The dashed lines correspond to the case of output market imperfect competition ($tfp(b)$). The dotted lines describe the evolution of the TFP index when neither form of imperfect competition is assumed, but the returns to scale are variable ($tfp(c)$).

The first panel displays the pattern of the TFP index obtained using the GMM estimator (therefore, controlling only for endogeneity and not for a possible selection bias). The second one shows the productivity index deriving from LP estimation. The third graph reports results from the Wooldridge–GMM estimator.

As one can see, one common feature to all three graphs is that the TFP index obtained by assuming both product and labor market frictions ($tfp(a)$) always lies below the other two indexes and is not increasing over time. The other two types of productivity indexes, $tfp(b)$ and $tfp(c)$, which do not consider labor market imperfections are increasing over time.

This is an interesting result, as some of the statistical sources (OECD is an example), as well as other empirical work, that do not consider labor market imperfect competition find increasing TFP and/or a positive change throughout the time period. However, these studies compute TFP changes using the T'ornqvist index number formula or following a structured estimation approach that does not correct for labor frictions.

In the first two graphs, there is still an appreciable difference between $tfp(b)$ and $tfp(c)$. In particular, assuming imperfect competition on the output market yields a lower TFP index. On the other hand, in the third graph the predicted productivity indexes of the two models (b and c) seem to coincide.

Overall, we conclude that neglecting a “wage markup”, might lead to an underestimation of the true value of the price–cost margin at the aggregate level and leads to an optimistic scenario of growing firm–level productivity.

References

Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes. *Handbook of Econometrics*, 6.

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural Identification of Production Functions. Technical report, UCLA mimeo.

Bloch, H. and Olive, M. (2001). Pricing over the cycle. *Review of Industrial Organization*, 19(1):99–108.

Boone, J. (2008). A new way to measure competition. *The Economic Journal*, 118(531):1245–1261.

Boone, J. and van der Wiel, H. (2007). How (not) to measure competition. CPB Discussion Paper 91, CPB Netherlands Bureau for Economic Policy Analysis.

Bughin, J. (1993). Union-firm efficient bargaining and test of oligopolistic conduct. *The Review of Economics and Statistics*, 75(3):pp. 563–567.

Bughin, J. (1996). Trade unions and firms' product market power. *The Journal of Industrial Economics*, 44(3):pp. 289–307.

Chirinko, R. S. and Fazzari, S. (1994). Economic fluctuations, market power, and returns to scale: Evidence from firm-level data. *Journal of Applied Econometrics*, 9(1):47–69.

Crépon, B., Desplatz, R., and Mairesse, J. (2002). Price-cost margins and rent sharing: Evidence from a panel of french manufacturing firms. *Revised version of CREST Working Paper No. G9917*.

De Loecker, J. and Konings, J. (2006). Job reallocation and productivity growth in a post-socialist economy: Evidence from slovenian manufacturing. *European Journal of Political Economy*, 22(2):388–408.

Dickens, W. T. and Katz, L. F. (1987). Inter-industry wage differences and theories of wage determination. NBER Working Papers 2271, National Bureau of Economic Research, Inc.

Diewert, W. (1993). Duality approaches to microeconomic theory. In Arrow, K. J. and Intriligator, M., editors, *Handbook of Mathematical Economics*, volume 2 of *Handbook of Mathematical Economics*, chapter 12, pages 535–599. Elsevier.

Diewert, W. E. and Fox, K. J. (2008). On the estimation of returns to scale, technical progress and monopolistic markups. *Journal of Econometrics*, 145(1-2):174–193.

Dobbelaere, S. (2004). Estimation of price-cost margins and union bargaining power for belgian manufacturing. *International Journal of Industrial Organization*, 22:1381–1398.

Dosi, G. (1999). *Innovation Policy in a Global Economy*, chapter Some Notes on National Systems of Innovation and Production, and their Implications for Economic Analysis. Cambridge University Press.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425.

Galí, J., Gertler, M., and Lpez-Salido, J. D. (2007). Markups, gaps, and the welfare costs of business fluctuations. *The Review of Economics and Statistics*, 89(1):44–59.

Griffith, R., Boone, J., and Harrison, R. (2008). Measuring competition. *Social Science Research Network Working Paper Series*.

Hall, R. E. (1986). Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity*, 17(2):285–338.

Hall, R. E. (1988). The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*, 96(5):921–47.

Hall, R. E. (1991). Invariance properties of solow's productivity residual. Working Paper 3034, National Bureau of Economic Research.

Klette, T. J. and Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4):343–61.

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

Loecker, J. D. (2007). Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity. NBER Working Papers 13155, National Bureau of Economic Research, Inc.

McDonald, I. M. and Solow, R. M. (1981). Wage bargaining and employment. *American Economic Review*, 71(5):896–908.

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

Ramey, V. A. and Shapiro, M. D. (2001). Displaced capital: A study of aerospace plant closings. *Journal of Political Economy*, 109(5):pp. 958–992.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Wu, Y. and Zhang, J. (2000). Endogenous markups and the effects of income taxation:: Theory and evidence from oecd countries. *Journal of Public Economics*, 77(3):383–406.

A Varying markups and returns to scale in a general setting

This Appendix presents a detailed analysis of the production function and the TFP set-up allowing for market imperfections and scale economies. The derivation of markups, scale elasticities, and TFP is based on a single-product production technology.

In particular, we let each firm $i \in \{1, ...N)$ in sector j face the following production function for period t:

$$Q_{it} = A_{it} F_i(\mathbf{X}_{it}) \qquad i = 1, 2, \ldots N \quad t = 1, ..., T, \tag{19}$$

where Q_{it} measures firm i's gross output, $\mathbf{X}_{it} \equiv (X_{ij1t}, X_{ij2t}, \ldots, X_{ijrt})'$ denotes the vector of r nonnegative factor inputs, $F_i(.)$ is the core of the (differentiable) production function, and A_{it} is TFP measured as the rate of a Hicks-neutral disembodied technology.

How does imperfect competition enter (19)? It is a well-known fact that firms with market power do not set their value of the marginal product equal to their corresponding factor price. It is assumed that each firm i faces an inverse demand function, $P_{it}(Q_{it}) = \left(\frac{Q_{it}}{Q_{it}}\right)^{\frac{1}{\eta_{it}}} P_{it}$, which represents the market price as a function of the output, Q_{it}, the sectoral production (index) Q_{it}, the industry price index P_{it}, and the time-varying, firm-specific cross price elasticity of demand for differentiated goods in the industry, η_{it}.

Firm i's optimization problem can be written as:

$$\max_{Q_{it}, \mathbf{X}_{it}} \quad \left\{ P_{it}\left(Q_{it}\right) Q_{it} - \mathbf{V}'_{it} \mathbf{X}_{it} \mid A_{it} F_i(\mathbf{X}_{it}) \geq Q_{it} \right\}, \tag{20}$$

where $\mathbf{V}_{it} \equiv (V_{ij1t}, V_{ij2t}, \ldots, V_{irt})'$ is firm i's vector of r input prices. Assuming, in the first instance, that there is imperfect competition on the output market and perfect competition on the input markets (a monopolistic firm acting as a price-setter on its output market and a price-taker on its input markets), the first order conditions (FOCs) implied by the solution of (20) yield the following equations for the Lagrange multiplier and the nominal input prices:

$$\begin{aligned} P_{it}(Q_{it}) \left(\frac{1}{1 + \frac{1}{\eta_{it}}} \right) &= P^* \quad \text{and} \\ \left(\frac{1}{1 + \frac{1}{\eta_{it}}} \right) P_{it}(Q_{it}) \frac{\partial Q_{it}}{\partial \mathbf{X}_{it}} &= \mathbf{V}_{it}, \end{aligned} \tag{21}$$

where, according to Diewert (1993) and Diewert and Fox (2008), the Lagrange multiplier P^* is firm i's shadow (or marginal) price of output. Under profit maximization and market power, this price enables firm i to set each input's marginal product, $\partial Q_{it}/\partial X_{ijkt}$, above the respective factor cost: $\partial Q_{it}/\partial X_{ijkt} = ((\eta_{it}+1)/\eta_{it})V_{ijkt}/P_{it}(Q_{it})$ for $k = 1, \ldots, r$, where the term between brackets is firm i's *markup.* Note that in case of perfect competition, the cross price elasticity tends to infinite and the markup goes to one.

Multiplying both terms of the second FOC in (21) by $\mathbf{X}_{it}/Q_{it}$ and rearranging terms, this can be rewritten as:

$$\frac{\partial Q_{it}}{\partial \mathbf{X}_{it}} \frac{\mathbf{X}_{it}}{Q_{it}} \equiv \theta_{it} = \left(\frac{\eta_{it}}{\eta_{it}+1}\right) \frac{\mathbf{V}_{it}\mathbf{X}_{it}}{P_{it}(Q_{it})} \equiv \mu_{it} s_{it}, \tag{22}$$

where $\frac{\eta_{it}}{\eta_{it}+1} \equiv \mu_{it}$ is firm's i markup and s_{it} denotes the firm i's total factor input share. It can also be written as the sum of the cost shares of the k inputs of the total production as $s_{it} = \sum_{k=1}^{r} s_{ikt} = \sum_{k=1}^{r} V_{ikt}X_{ikt}/(Q_{it}P_{it}(Q_{it}))$. Therefore, following (22), the output elasticity of input k can be expressed as:

$$\theta_{ikt} \equiv \mu_{it} s_{ikt}, \qquad k = 1, \ldots J_i. \tag{23}$$

Equation (23) says that the output elasticity of any individual input k equals the markup times the share of input k in the total production value of firm i

An important feature of the markup given in (22). It allows for time-varying returns to scale. Under constant (non–variable) returns to scale ($\theta_{it} = \theta$) and constant markups ($\mu_{it} = \mu$), equation (22) is equivalent to the measure proposed by Hall (1988). Chirinko and Fazzari (1994) find a strong correlation between economies of scale and markups over time, which is obviously also implied by (22) and (23).

Several studies also relate markups to business cycles (e.g., Wu and Zhang (2000) and Bloch and Olive (2001)). Although it is realistic to allow for dynamic effects in competition, one should be wary of the effect of business cycles on the markup since it may not reflect changes in competition (Boone and van der Wiel (2007); Boone (2008); Griffith et al. (2008)).

B Bargaining model

This Appendix provides the solution of Nash bargaining model and the derivation of wage markup and union's bargaining power.

We define the negotiation of wages and labor as a bargaining problem:

$$\max_{W_{it},L_{it}} \left[\phi_{it} \log(L_{it}(W_{it} - \overline{W}_{it})) + (1-\phi_{it}) \log(P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it})\right].$$

The solution to the efficient bargaining problem is obtained by maximizing with respect to employment and to wage, and then combining the two FOCs. Maximizing with respect to labor yields:

$$\frac{\phi_{it}}{L_{it}} = (1-\phi_{it})\frac{\left(W_{it} - \frac{\partial P_{it}(Q_{it})Q_{it}}{\partial L_{it}}\right)}{[P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it}]}. \tag{24}$$

Symmetrically, maximizing with respect to wage leads to:

$$\frac{\phi_{it}}{L_{it}} = (1-\phi_{it})\frac{(W_{it} - \overline{W}_{it})}{(P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it})}. \tag{25}$$

Combining equations (24) and (25), and taking account the definition of the markup in the first section of this Appendix, we can write the reservation wage $\overline{W}_{it}$ as:

$$\overline{W}_{it} = \frac{\partial P_{it}(Q_{it})Q_{it}}{\partial L_{it}} = \frac{\partial \left(Q_{it}/Q_{it}\right)^{1/\eta_{it}} P_t^I Q_{it}}{\partial L_{it}} = \frac{P_{it}(Q_{it})}{\mu_{it}}\frac{\partial Q_{it}}{\partial L_{it}}. \tag{26}$$

Rewriting equation (25), after some algebra, we can express the wage as the following:

$$\frac{W_{it} - \overline{W}_{it}}{W_{it}} = \frac{\phi_{it}}{1-\phi_{it}}\frac{(P_{it}(Q_{it})Q_{it} - W_{it}L_{it} - R_{it}K_{it} - Z_{it}M_{it})}{L_{it}W_{it}}. \tag{27}$$

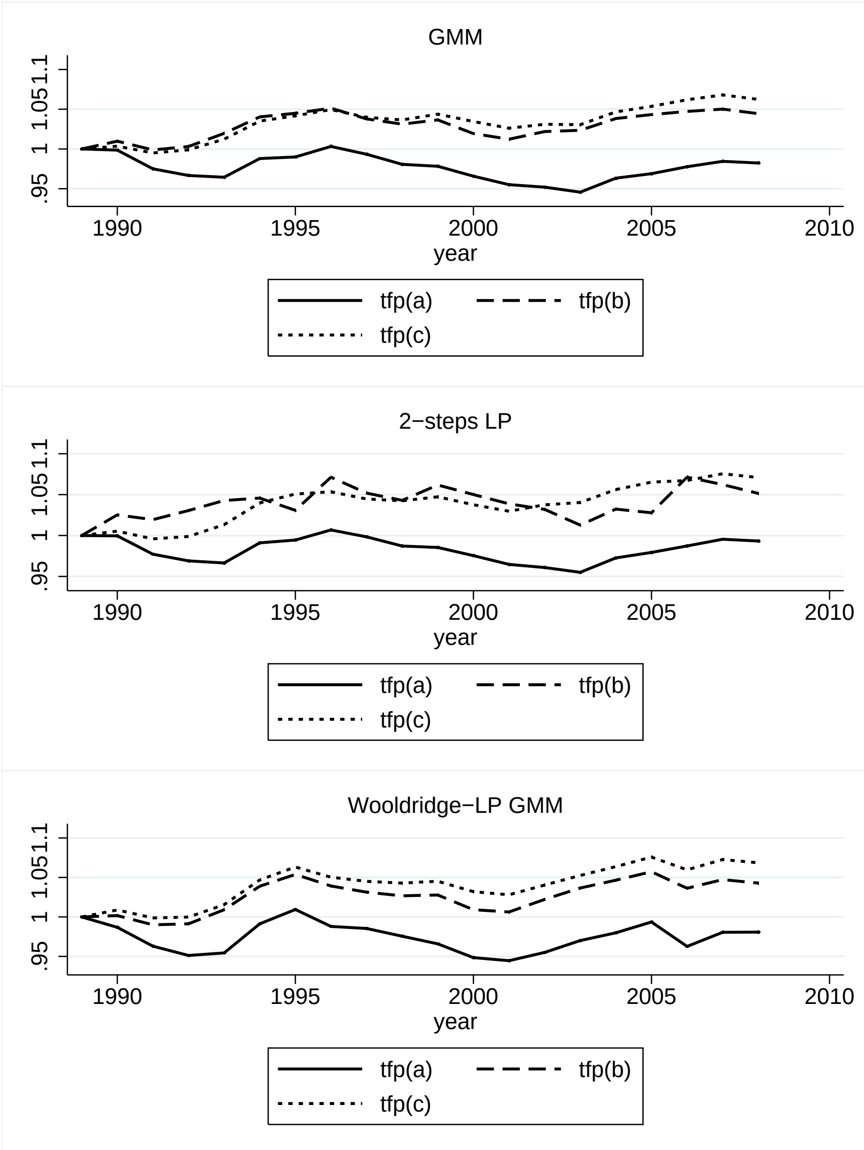
Defining $\frac{W_{it}-\overline{W}_{it}}{W_{it}} \equiv h_{it}$ as the wage markup, one can see how this is directly depending on the union's bargaining power. Moreover the expression for the wage markup in equation (27), reveals the intrinsic link between product and labor market distortions. As a matter of fact, when there is no imperfect competition on the labor market, firms set the wage at the lowest value possible, the competitive wage $\overline{W}_{it} = W_{it}$, which is inversely related to the output markup μ_{it}. But in a fully imperfect competitive setting, the marginal revenue product of labor $(\frac{P_{it}(Q_{it})}{\mu_{it}}\frac{\partial Q_{it}}{\partial L_{it}})$ is equal to the reservation wage (see equation (26)). Therefore, the larger the firms' rent, the larger the wage mark-up h_{it}, and, consequently, the union's bargaining power.

From the solution of the efficient bargaining model, we can derive a simple expression for the elasticity of labor. Multiplying both sides of (26) by L_{it}/Q_{it} and after some simple algebraic manipulations, we can express the labor elasticity as a function on the labor share and the wage markup:

$$\gamma_{iLt} \equiv \theta_{iLt}/\mu_{it} = s_{iLt}\frac{\overline{W}_{it}}{W_{it}} = s_{iLt}(1-h_{it}) \tag{28}$$

We make use of (28) in the estimating equation (6), to derive the labor elasticity, as well as the wage markup. To retrieve the bargaining parameter ϕ_{it}, we use the equation (27).

Figure 1: TFP Index over time



Note: Sample period 1989-2008; dependent variable: log. gross deflated output y_{it}

$TFP(a) \equiv \frac{\hat{\eta}+1}{\hat{\eta}}(y_{it} - (\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (1 \hat{-} h) s_{it} l_{it} - \frac{1}{\hat{\eta}} q_t^j))$, where $1 - h = \frac{\overline{W}_{it}}{W_{it}}$
$TFP(b) \equiv \frac{\hat{\eta}+1}{\hat{\eta}}(y_{it} - (\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + \hat{\gamma}_L l_{it} - \frac{1}{\hat{\eta}} q_t^j))$
$TFP(c) \equiv y_{it} - (\hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + \hat{\gamma}_L l_{it})$.