Changes in Competition of Small vs. Large Firms resulting resulting from International Trade[*]

CHANGWOO NAM
JIYOON OH[†]
Korea Development Institute

Using Korean plant-level manufacturing data, this paper examines the effect of lowering trade barriers on changes in markups of small and large firms, exporters and non-exporters. We find that large firms decide on higher markups in each sector as they have greater market power in integrated markets. Also, exporters set higher markups through higher observable productivity than non-exporters. Even after controlling for productivity and other firm characteristics, markups are proportional to market share, which can be interpreted to mean that market power only influences firm price strategy. Interestingly, the markup distribution, which is more closely related to the competition from globalisation, has been decreasing over time while the performance gap measured as sales has been stable over time. We caution that even if the performance gap measured in absolute terms may be widening, this does not imply that the level of competition between large and small firms is weakening.

[*] This working paper is funded by Economic Research Institute for ASEAN and East Asia (ERIA).
[†] Correspondent Author's Contact: Dr. Oh, Jiyoon: jiyoon.oh@kdi.re.kr

1. Introduction

Globalisation has been regarded as one of the main driving forces behind changes in market environments, such as degree of competitiveness between firms. Intuitively, more integrated markets confer a benefit on more productive firms as they can sell more products in a bigger market. The firm selling its product in the domestic market can grow as a global company. This leads to the exit of less productive firms in the market, thereby polarising firm performance. On the other hand, the influx of foreign products from world markets makes the market environment more competitive, so firms with monopolistic power due to market frictions can lose their market power in the domestic market. It creates a level playing field for all firms, so "second movers" with small market share can enjoy more equal benefits. It alleviates inequality between firms, especially in terms of firm size.

Since globalisation has two opposite effects on inequality between firms, it is a natural question whether more integrated markets benefit all firms equally. There is a wide range of academic literature on globalisation and its effect on aggregate output or firm performance, but relatively little work has been done on the different impacts of globalisation on firm performance.

For policy administration, firm size is a convenient measure. In many countries, firm policy has been implemented discriminately according to firm size. Tax benefits accrue to a greater extent to small firms, as regulations for large firms are stricter. Even if firm size contains many characteristics suggesting productivity, firm age, market power, size itself is actually obscure property. For example, firm size is not directly linked to productivity. There are ongoing debates about why large firms are large. Are they big because of their advanced technology or did they just benefit from early market entry, giving them a "first mover" advantage. Economists who had thought that small firms are the engine of growth and the entity of creative destruction have come to realize that many small firms are actually at the low level of innovation. (Hurst and Pugsley, 2011) However, there is a general consensus that market share is the obvious determinant of market power. It seems appropriate, therefore, to study the effect of globalisation on changes in the relative market power of firms. Markups are commonly investigated as a proxy variable to capture market power.

In this paper, we investigate whether more integrated markets through globalisation expand or shrink the market power gap between large and small firms. Using plant-level data of the Korean annual survey of manufacturing, we rigorously estimate plant markups and assess markup trends over time. We then empirically examine the effect of lowering trade barriers on changes in markups of small and large plants. Through this exercise, we can test our educated guesses of markup variations in small and large firms in international models. Furthermore, we can directly observe the market power gap between small and large firms measured by markups, and investigate whether this gap converges when markets become more open through trade liberalization.

In terms of the theoretical literature, our paper is closely associated with the recent development of the heterogeneous model of international trade. Markups have received a lot of attention from economists and policymakers as they measure the effect of various competition and trade policies on market power. Recently, the theoretical study of firm heterogeneity in terms of productivity or size has been combined with heterogeneity in markups. Melitz and Ottaviano (2008) suggest a monopolistically competitive model of trade with firm heterogeneity. In their model, market size and trade affect the strength of competition. Larger and more integrated markets through trade tend to have lower markups. However, this paper does not point out the difference between small and large firms. Decreases in markups when market size is bigger through trade, is linear in terms of productivity.

Other types of theoretical models of endogenous markups, such as Atkeson and Burstein (2008), Oh (2013), and Edmund, Midrigan and Xu (2013) emphasize the increasing schedule of optimal markup with respect to a firm's market share. These types of models rely on a similar setting of monopolistically competitive markets, except that the number of competitors in a particular industry or for a particular product is small. In such settings, firms take into account the effect of their pricing decisions on the equilibrium of the prices of industry goods. Price elasticity of demand decreases with a firm's market share. Thus an optimal markup, which is the inverse of price elasticity, increases with firm size. Large firms assign higher markups than small firms. A reduction in trade barriers reduces the industry share of domestic producers, thus reducing their markups. Interestingly, the optimal markup is convex-increasing in a firm size. Therefore, the adjustment in markup of large firms is larger than that of small firms with the same reduction in market share due to international trade.

This notion of differences in markups between small and large firms has hardly been investigated empirically. Roberts and Supina (1996) show that plant-specific markups of price over marginal cost vary according to the size of producers. For three products, markups decline in size and in two cases they increase. Edmund, Midrigan and Xu (2013) accurately calibrate their model with Taiwanese manufacturing plant-level data and argue that endogenous markup setting shows much larger gains from trade than Ricardian models. Bigger welfare gains in their model are driven by a significant reduction in large firms' markups. They imply that import competition reduces the gap between large and small in terms of firm markup. However, they never show any empirical evidence that plant-specific markups decrease after trade barriers are lower.

Our main empirical findings confirm many theoretical predictions. First, the level of markup is obviously higher in more productive firms, as predicted by Melitz and Ottaviano (2008). Second, markup increases as market share rises. This reinforces the increasing relationship between markup and market share. As Atkeson and Burstein (2008) and Oh (2013) predicted, larger market share leads to higher markup because large firms can enjoy more market power due to lower demand elasticity. Third, markups of exporters are higher on average than markups of non-exporters. This makes sense, that exporters are mostly more productive and larger firms, which can afford to pay fixed costs for exporting, as Melitz (2003) predicts. Fourth, we create distributions of firm markups at a given point every year, and compare them. Interestingly, the mean of markups has decreased over time, and so has the dispersion. Even though we cannot identify the main factors behind for convergence of markups, a competition effect from globalisation is definitely one of the plausible factors. In order to identify and quantify the effect of import competition on markup dispersions, we regress industry markup dispersions on industry import penetrations. Generally speaking, import competition reduces markup dispersion. Lastly, although the overall picture of markup distribution has been increasingly condensed over time, we find that when individual firms expand their market share in overseas markets, their markups increase.

The remainder of this paper is organized as follows. We introduce our theoretical model and briefly provide theoretical predictions about empirical results in Section 2. Section 3 introduces our empirical framework and our estimation routine. Section 4 provides our main empirical results and a discussion of them. The final section concludes.

2. Theoretical Background

In this section, we illustrate how variations in firm size are theoretically related to differing levels of firm markup. We first lay out the market structure in the model to examine the mechanisms underlying market share and markup. This model is based on the monopolistic competition suggested by Dixit and Stiglitz (1977), except that it has a few competitors rather than a continuum of firms. The goods market features differentiated oligopoly competition within a quantity-setting game.

We construct a model of imperfect competition in which final goods consist of a continuum of industry goods and each industry goods market consists of N_j firms. The final good is produced using a constant returns to scale production function, which aggregates a continuum of industry goods.

$$Y = \left(\int_0^1 y_j^{1-1/\eta} dj \right)^{\frac{\eta}{\eta-1}} \tag{1}$$

where y_j denotes the output of industry j. The elasticity of substitution between any two different industry goods is constant and equals η. Final goods producers behave competitively.

In each industry, there are N_j firms producing differentiated goods that are aggregated into industry goods through an aggregate constant elasticity of substitution (CES) production function. The output of goods in industry j[1] is given by

[1] The term N1-11P implies that there is no variety effect in the model.

$$y_j = N^{\frac{1}{1-\theta}} \left(\sum_{i=1}^{N_j} y_{ij}^{1-1/\theta} \right)^{\frac{\theta}{\theta-1}} \tag{2}$$

where y_{ij} is the output of firm *i* in industry *j*. Within each industry of N_j firms, a firm sets its quantity. The elasticity of substitution between any two intra-industry goods is constant and equals θ. It is assumed that the elasticity of substitution between any two goods within an industry is higher than the elasticity of substitution across industries, $1 < \eta < \theta$.

The final good producer solves a static optimization problem that results in the usual conditional demand for each industry good,

$$y_j = \left(\frac{P_j}{P} \right)^{-\eta} Y,$$

where p_j is the industry *j* price and *P* is the price of final goods,

$$P = \left(\int_0^1 P_j^{1-\eta} dj \right)^{\frac{1}{1-\eta}}, \tag{3}$$

Denoting the price of good *i* in industry *j* by p_{ij} ,

$$P_j = N_j^{\frac{1}{\theta-1}} \left(\sum_{i=1}^{N_j} p_{ij}^{1-\theta} \right)^{\frac{1}{1-\theta}}, \tag{4}$$

the inverse demand functions for goods within an industry are given by:

$$\left(\frac{p_{ij}}{P}\right) = \left(\frac{y_{ij}}{y_j/N_j}\right)^{-1/\theta} \left(\frac{y_j}{Y}\right)^{-1/\eta}.$$

Dixit and Stiglitz (1977) assume that each firm is small relative to the economy, and therefore does not influence the equilibrium price and quantity. In this model, the assumption of a small number, *N*j, of firms in each industry implies that a firm's quantity choice affects the industry price. Within a given industry, each firm takes into account the effect that the pricing and production decisions of other firms has on the demand for its own goods. Therefore, the price elasticity of demand ϵ (S_{ij}) of firm (*i*) is a decreasing function of the firm's when the substitutability of within-industry goods is higher than that of between-industry goods ($\eta < \theta$). In equation (6), the demand elasticity is a market share weighted average of two values [η, θ]: when yij is close to zero, the perceived demand elasticity of firm *i* in industry *j* is equal to θ, which is the same as in Dixit and Stiglitz (1977). On the other hand, if yij is close to one, the demand elasticity of firm i is the same as that of the monopoly firm in industry j[2].

$$\epsilon\left(S_{ij}\right) = \left[\frac{1}{\theta}\left(1 - S_{ij}\right) + \frac{1}{\eta}S_{ij}\right]^{-1} \qquad (5)$$

From eq (4) and eq (5), these market shares can be written as a function of prices in equation (7)

$$S_{ij} = \frac{p_{ij}y_{ij}}{P_jY_j} = \frac{p_{ij}^{1-\theta}}{\sum_{i=1}^{N_j} p_{ij}^{1-\theta}} \qquad (6)$$

Derived directly from the demand elasticity in equation (6), firm markup is an increasing function of its market share from (8).

---

[2] If firms compete in a price-setting game (Bertrand competition) within an industry, the demand elasticity would be $\epsilon_{y,p} = \theta\left(1 - s_i\right) + \chi s_i$.

$$\mu_{ij}(S_{ij}) = \frac{\epsilon(S_{ij})}{\epsilon(S_{ij}) - 1} = \frac{1}{1 - \frac{1}{\theta}(1 - S_{ij}) - \frac{1}{\eta}S_{ij}} \tag{7}$$

Firm markups are combined into aggregate industry markup ($\bar{\mu}_j$). Aggregate markup can be expressed in two ways: the input-share weighted average of firm markup, which is equal to the revenue-share weighted harmonic average of firm markup.

$$\bar{\mu}_j = \sum_{i=1}^{N_j} x_{ij}\mu_{ij} = \left(\sum_{i=1}^{N_j} \frac{S_{ij}}{\mu_{ij}}\right)^{-1} \tag{8}$$

where $x_{ij} = \frac{Input_{ij}}{Input_j}$ is the input share[3] of firm i in industry *j*.

In a symmetric industry equilibrium, aggregate industry markup $\bar{\mu}_j$ is equal to aggregate markup $\bar{\mu}$. Going forward, we will restrict our attention to symmetric industry equilibrium.

The assumption of ($\theta > \eta$) implies that each firm's markup of its price over marginal costs is an increasing function of that firm's market share within an industry. At one extreme, if the firm has a market share S_i approaching zero, it faces only the industry elasticity of demand 0 and chooses a markup equal to $\theta / (\theta — 1)$. At the other extreme, if the firm has a market share approaching one, it has lower elasticity of demand across industries η and sets a higher markup equal to $\eta / (\eta — 1)$. The difference $\theta — \eta$ actually determines how much the demand elasticity changes in response to shifts in market share. As $\theta — \eta$ gets bigger, the effect of market share on demand elasticity and markup becomes increasingly significant.

$\Gamma(s)$ refers to the elasticity of the markup with respect to market share. Note that

[3] In the case that input prices are common to all firms, input shares of any input are equal within firms. For instance, the labour input share $\frac{h_{ij}}{H_j}$ of firm *i* in industry *j* is the same as the capital input share $\frac{k_{ij}}{K_j}$, if firms face the same wage rates and capital rental prices.

$\Gamma(s)$ is an increasing and convex function of s. In the constant markup model, $\Gamma(s) = 0$,

$$\Gamma(s) = \frac{s}{1 - \frac{1}{\theta}(1-s) - \frac{1}{\eta}s}\left(\frac{1}{\eta} - \frac{1}{\theta}\right)$$

This convexity plays an important role in the dynamics of aggregate markup. Due to this convexity, aggregate markup increases as market shares across firms become more dispersed or unequal.

In addition to convexity, the level of aggregate markup is influenced by a composition effect. Since aggregate markup is the input-share weighted average of firm markups, a large firm's high markup weighted by its high input-share contributes significantly to raising aggregate markup, and vice versa. This composition effect implies that the pricing behaviours of large firms play a dominant role in the dynamics of aggregate markup.

It is worth mentioning that a firm's markup does not change unless its market share changes. When there are uniform changes such as cost reductions for all firms, relative prices do not change between firms; therefore, market share stays constant. This is an important departure from a generic sticky price model in which an exogenous price-setting friction causes variations in markup for the representative firm.[4] In our model, aggregate fluctuations cannot change aggregate markup. Only changes in relative productivity between firms matter in determining aggregate markup.

The model described above can apply to how globalisation can influence firm decisions in terms of markups. In terms of increases in importing, trade liberalization and the resulting greater influx of imported goods make domestic markets more competitive. Rises in import penetration in an industry naturally reduces the market share of domestic firms. Based on the theoretical framework above, this effect lowers the level of domestic markups. Furthermore, the speed of lowering markups is accelerated in large firms rather than small firms due to the convex schedule of optimal markup. In this sense, we can say that globalisation generates more competition and

[4] It follows that our model can explain why large firms are reluctant to cut prices in recessions—due to the low demand elasticity they face.

reduces market power inequality between large and small firms.

When it comes to globalisation through exporting, applying this modified imperfect competition framework is ambiguous. It is obvious domestic firms lose domestic market share to foreign competitors. For exporting producers, domestic market share may not change at all after participating in exporting, but entry to exporting may change the firm distribution through the selection process.

In terms of the economic literature, we can lean on the endogenous markup model suggested by Melitz and Ottaviano (2008). Even if the details are different, the basic mechanism is closely related to our model above. Competition from entry lowers the level of markups. Melitz and Ottaviano (2008) prove theoretically that the mean of markups decreases and the average level of productivity of firms increases as markets become more integrated through trade liberalization. This makes sense, that the selection effect pushes the least productive firms out of the market. A more competitive environment makes firms reduce prices and markups as well. Interestingly, Melitz and Ottaviano also predict the dispersion of firm performance measures such as price, markup, and firm size: the variance of cost, prices, and markups are lower in bigger markets because the selection effect decreases support for these distributions. On the other hand, the variance of firm size (in terms of either output or revenue) is larger in bigger markets due to the direct magnifying effect of market size on these variables.

Regarding the dispersion of firm performances, the two different directions of price and quantity are very interesting. Even if the degree of competition increases, firm size distribution will be more unequal. To determine if globalisation actually increases the level of competition or is more beneficial for large firms, markup distribution is a better measure than firm size distribution. In Section 4 we will present our empirical results regarding the dispersion of markups.

However, the effect of increases in firms' exports on firms' markups is ambiguous in the sense that variations in markups across firms in cross-sectional analysis may not show the same pattern in time series analysis. Thus, the real effect of international trade on differing markups according to size should be measured empirically. De Loecker and Warzynski (2012) showed that exporting makes firms increase markups in time series analysis as well as in cross-sectional analysis.

3. Estimation

3.1. Production Function Estimation

The problem of estimating the production function has been an important issue since the beginning of economics as production functions are a fundamental component of all economics. In fact, the econometric subject is the possibility that the major determinants of firm's production decision might be unobservable to econometricians. Thus, this measurement error induces the endogeneity problem due to the relation between observed inputs and unobserved productivity shocks. Olley and Pakes (1996, hereafter referred to as the OP model), Levinsohn and Petrin (2003, hereafter the LP model), Ackerberg, Caves and Frazer (2006, hereafter the ACF model), and De Loecker and Warzynski (2012, hereafter the DLW model) are seminal papers leading to the introduction of new techniques for identification of production functions. OP model and LP model cannot avoid the multicollinearity issue when they estimate the labour coefficient of production function in the estimation scheme. DLW model owes ACF model in terms of the full identification in the second stage of structural estimation. In addition, these papers are somewhat more structural in nature-using observed input decisions to control for unobserved productivity shocks (De Loecker and Warzynski (2012)). These techniques have been used in a large number of recent empirical papers including Pavcnik (2002), Fernandes (2007), Criscuola and Martin (2009), Topalova and Khandelwal(2011), Blalock and Gertler (2004), and Alvarez and Lopez (2005).

3.2. Markup Estimation

Estimating markups has a long tradition in industrial organization and international trade. Re-searchers in industrial organization are interested in measuring the effect of various competition and trade policies on market power through estimating unobservable markups. In this paper, we use a simple empirical framework in the DLW model to estimate markups. Our approach of following the DLW model nests the price-setting model used in applied industrial organization and international trade and relies on optimal input demand conditions obtained from standard cost minimisation and the ability to identify the output elasticity of a variable input. This framework eliminates issues related to input adjustment costs. Also, this methodology suggests that the output elasticity of a variable factor of production is exactly equal to its expenditure share in total revenue as price equals marginal cost of production, solving the cost minimisation problem.

Markup estimates are obtained using production data where we observe output, total expenditure on variable inputs, and plant-level revenue datasets. The DLW model in particular requires a measure of output that disregards price differences across firms. Therefore, we use real output value from the Korean dataset. Recent literature makes empirical approach very suitable to those types of datasets from several countries. (Foster, Haltiwanger, and Syverson [2008]; Goldberg et al. [2010]; Kugler and Verhoogen [2008]; De Loecker and Warzynski [2012]).

Some assumptions are introduced following the DLW model. First, constant returns to scale is not imposed, and second, the user costs of capital do not need to be observed or measured in our model. This relaxation results in a flexible methodology and reliable estimates such as the DLW model. We then use our empirical model to verify whether exporters, on average, charge higher markups than their domestic counterparts in the same industry, and how markups change according to firm size, i.e., how the market share changes. This framework is well suited to relating markups to any observed plant-level activity potentially correlated with plant-level productivity.

3.3. Local Constant Kernel Model

In recent decades, the literature on nonparametric econometric methods has offered solutions to the problems related to parametric misspecification of econometric regression models. This misspecification problem can be generically generated in production or markup estimations because the functional form of production is wholly determined by the researcher's arbitrary decision. However, nonparametric regression techniques basically do not oblige the researcher to assume and specify a functional form of production for the relationship between the firm's decision variables and the production variable (output production or value-added production). Fully nonparametric models are most often applied to cross-sectional data, while they are seldom applied to panel data sets (Czekaj and Henningsen [2013][5]).

There still exists a possibility that the DLW model has a multicollinearity problem because it uses nth order nonparametric series regression with inter-variable components in the first stage of structural estimation even though it fully estimates coefficients necessary to compute the markups in the second stage formed by the generalized method of moments (GMM) structure. Therefore, we use a local constant kernel model (hereafter, the LCK model) with unordered discrete data at the first stage of structural estimation. The LCK model is a fully nonparametric model that uses a time variable and an individual identifier as additional (categorical) explanatory variables (Racine and Li [2004]). In this formation we do not need to consider separately the production part of labour and capital, and the productivity shock observable to firm managers before their input decisions (on labour, investment, materials so on), but unobservable to econometricians. The fully nonparametric regression, that is, LCK, model only focuses on how to accurately estimate data. At the same time, the LCK model captures non-linear individual and time effects that do not need to be assumed to be additive and separable.

In our analysis, we use a fully nonparametric and nonseparable panel data model (LCK model) that has been suggested by Henderson and Simar (2005), Racine (2008), and Gyimah-Brempong and Racine (2010). They estimate an undefined function as a fully nonparametric two-ways effects panel data model with individual effect and time effect as categorical explanatory variables using the nonparametric regression method proposed by Li and Racine (2004) and Racine and Li (2004). Those papers use both continuous and categorical explanatory variables for fully nonparametric specification. This estimator does not require any data transformation with a loss of observations. In addition, the intercept of the dependent variable and the slopes of the explanatory variables on the dependent variable are not fixed according to the interaction between time periods and individuals in the fully nonparametric model. Hence, this estimator does not imply any restrictions on the most general specification of panel data models. Furthermore, the bandwidths of the explanatory variables can be selected using data

[5] Czekaj and Henningsen (2013) only compare the fittability of ordinary least squares (OLS), semiparametric and fully nonparametric regressions. It is not their purpose to solve unbiased estimators for unobserved productivity shocks in firm decisions.

driven cross-validation methods. The overall shape of the relationship between the dependent variable and the covariates, the individual, and time is entirely determined by the data.

Finally, we compare the empirical results with LCK, DLW and conventional ordinary least squares (OLS) models. It is found that the LCK model fits better with the production data and is more consistent with the economic theory compared with the DLW and OLS models. This means that the LCK model captures the non-linear individual and time effects by using a discrete smoothing parameter, and the fitted value-added is determined by the local weighted average rather than by labour, capital, and material variables.

3.4. Structure to Estimate Markups

We explain the structural model to obtain plant-level markups relying on standard cost minimisation conditions for variable inputs following the DLW model. These conditions imply that the markup is the output elasticity of an input in relation to the share of that input's expenditure in total sales and the firm's markup (DLW model). To obtain output elasticities, we need estimates of the production function, for which we rely on proxy methods developed by the DLW model. We follow the restrictions that DLW imposes, and we discuss our model in detail in below given DLW model.

3.4.1. Deriving Markups

A firm i produces output at time t with the implicit production technology:

$$Q_{it} = Q_{it}\left(X_{it}^{1}, \ldots, X_{it}^{N}, K_{it}, z_{it}\right),$$

in which it relies on N variable inputs such as labour, intermediate inputs, and electricity. In addition, the firm relies on a capital stock, K_{it}, which is treated as a dynamic input in production, which means the amount of investment at t is determined given the information at t — 1. The productivity shock z_{it} evolves exogenously following a first order Markov process, and the labour in production is a non-dynamic input, which means the amount of labour at t is related to the observed productivity shock z_{it}. However, the only restriction we impose on Q_{it} to derive an expression of the markup is that Q_{it} is continuous and twice differentiable with respect to its arguments.

Producers have a cost-minimisation problem, such as the associated Lagrangian function:

$$\mathcal{L}\left(X_{it}^{1},\ldots,X_{it}^{N},K_{it},z_{it}\right)=\sum_{j=1}^{N}P_{it}^{X^{j}}X_{it}^{j}+r_{it}K_{it}+\chi_{it}\left(Q_{it}-Q_{it}\left(\cdot\right)\right),$$

in which $P_{it}^{X^j}$ and r_{it} show a firm's input price for a variable input *j* and capital, respectively. The first-order condition for any variable input is

$$\frac{\partial \mathcal{L}_{it}}{\partial X_{it}^{j}}=P_{it}^{X^{j}}-\chi_{it}\frac{\partial Q_{it}\left(\cdot\right)}{\partial X_{it}^{j}}=0,$$

in which χ_{it} is the marginal cost of production at a given level of output as $\partial \mathcal{L}_{it}/\partial Q_{it}=\chi_{it}$. Then we can generate the following expression after some calculus:

$$\frac{\partial Q_{it}\left(\cdot\right)}{\partial X_{it}^{j}}\frac{X_{it}^{j}}{Q_{it}}=\frac{1}{\chi_{it}}\frac{P_{it}^{X^{j}}X_{it}^{j}}{Q_{it}}. \tag{9}$$

The equation (9) can be rewritten following DLW (2012) such that

$$\mu_{it}\frac{P_{it}^{X}X_{it}}{P_{it}Q_{it}}=\epsilon_{it}^{X},$$

in which the output elasticity on an input X is denoted by ϵ. This expression shows that the markup is the measure for the output elasticity on an input divided by the share of an input's expenditure in total sales such that

$$\mu_{it}=\frac{\epsilon_{it}^{X}}{\sigma_{it}^{X}}, \tag{10}$$

where σ_{it}^{X} is the share of expenditure on input X_{it} in total sales $P_{it}Q_{it}$. This means that an estimate of the output elasticity of one variable input in production and data on the expenditure share are enough to obtain a measure of plant-level markups using production data. The expenditure share can be directly obtained from observed micro data.

This derivation is standard and has been used throughout the literature, especially the DLW model (2012). DLW model's contribution is to provide consistent estimates of the output elasticities while allowing some inputs to face adjustment costs and recover firm-specific estimates of the markup related to various economic variables.

3.4.2. Output Elasticities and Markups

For estimates of the output elasticities $_{\mathrm{Eit},}$ production functions are implicitly assumed to contain a scalar Hicks-neutral productivity term and common technology parameters across the set of producers. But, when taking the log of production, the overall function can be estimated by using a fully nonparametric regression, LCK model. The latter does not imply that output elasticities of inputs across firms are constant, except for the special case of Cobb-Douglas.

The production function is

$$Q_{it} = G\left(X_{it}^{1}, \ldots, X_{it}^{N}, K_{it}, z_{it}; \beta\right) = F\left(X_{it}^{1}, \ldots, X_{it}^{N}, K_{it}; \beta\right) \exp\left(z_{it}\right),$$

in which a set of common technology parameters β govern the transformation of inputs to units of output, combined with the firm's productivity z_{it}.

This expression contains most specifications used in empirical work such as the translog production function. The main advantage of restricting production technologies of this form is proxy methods suggested by LCK, DLW, and OLS to obtain consistent estimates of the technology parameters β at the second stage. At the first stage, the total function of production G will be estimated. We consider the log version of equation (10) given that the output elasticity of a variable input j, $\epsilon_{it}^{X^j}$ is given by $\partial \ln G(\cdot) / \partial \ln X_{it}^{j}$ and is by definition independent of a firm's productivity level.

We implicitly assume that measurement errors occur in the output data and we assume unanticipated shocks to production, which we combine into v_{it}. It is assumed that the log output is given by $q_{it} = \ln Q_{it} + v_{it}$, where v_{it} are unanticipated shocks to production and i.i.d. (independent, identically distributed) shocks including measurement error. Also, the first stage of our estimation separates the overall production part and the measurement error from the data. It is important to emphasize that we explicitly count on the fact that firms do not observe v_{it} before optimal input decisions.

Therefore, the production function we estimate for each industry separately, is defined as

$$q_{it} = f\left(x_{it}; \beta\right) + z_{it} + \nu_{it},$$

in which we collect all variable inputs in x_{it}, and β contains all relevant coefficients. We consider flexible approximations to f (.), therefore we can use the LCK model, and explicitly write the production function we estimate on the data in general terms. For instance, our main empirical specification relies on any functional form that implies that f (-) is approximated by a fully non-parametric specification (LCK model), or a second order nonparametric series where all (logged) inputs, (logged) inputs squared, and interaction terms between all (logged) inputs are included (DLW model). We recover the translog production function when we drop higher-order and interaction terms. The departure from the translog production function (DLW model) is important for our purpose, to compare the empirical results.

Our fully nonparametric approach can nest various specifications of the production function, and only need the proper order of approximation of production functions at the second stage of the structural estimation framework. However, in order to obtain consistent estimates of the production function at the second stage, we need to control for unobserved productivity shocks, which are potentially correlated with input choices such as the insight of OP and LP models, and we use the DLW model approach while relying on materials to proxy for productivity. In this case, we do not need to reconsider the underlying dynamic model when considering modifications to OP setup when dealing with additional state variables. We describe the estimation framework while relying on a dynamic control for capital and discuss the additional assumptions.

We follow the DLW model (2012) and use material demand,

$$m_{it} = m_t\left(k_{it}, z_{it}, \mathbf{x}_{it}\right),$$

to proxy for productivity by inverting m(.), where we collect additional variables potentially affecting optimal material demand choice in the vector x_{it}. The inclusion of these additional control variables shows the only restriction we impose on the underlying model of competition (DLW model). Once those variables are appropriately accounted for in the estimation routine to obtain output elasticities, we can analyse how markups are different across firms and time, and how they relate to firm-level characteristics such as the globalisation or export status.

$z_{it} = m_t^{-1}\left(k_{it}, m_{it}, \mathbf{x}_{it}\right)$ is used to proxy for productivity in the production function estimation. The use of a material demand equation to proxy for productivity is important for researchers considering multicollinearity and estimating output elasticities and markups. Especially, as long as $\partial m/\partial z > 0$ conditional on the firm's capital stock and variables captured by $\mathbf{x}_{it},\ m_t^{-1}\left(k_{it}, m_{it}, \mathbf{x}_{it}\right)$ can be used to proxy for z_{it} being used to index a firm's productivity. In this setting, the DLW model (2012) finds it useful to refer to Melitz and Levinsohn (2006) who also rely on intermediate inputs to proxy for unobserved productivity while allowing for imperfect competition. Melitz and Levinsohn (2006) show that this monotonicity condition holds as long as more productive firms do not set lower markups than less productive firms. This is the main part of the DLW model's idea.

3.4.3. Steps for Estimating Markups

Basically, our analysis departs from De Loecker and Warzynski (2012) and gives up on identifying any parameter at the first stage since conditional on a nonparametric function in capital, materials, and other variables affecting input demand, identification of the labour coefficient is not plausible. Even though they use nonparametric series regression with inter-variable, high-order components, we use the fully nonparametric regression with continuous and discrete data. We are concerned with more flexible production functions and allow for an undefined functional form between the various inputs, identification of the labour coefficients at the first stage.

Our procedure consists of two steps and follows the DLW model. However, let us consider a value-added production function with the general form, which is given by

$$q_{it} = \Phi\left(k_{it}, l_{it}, m_{it}, \iota_i, \iota_t, \mathbf{x}_{it}\right) + \nu_{it}, \tag{11}$$

also for the comparison with the DLW model, given by

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{ik}^2 + \beta_{lk} l_{it} k_{it} + z_{it} + \nu_{it}, \tag{12}$$

in which lower case means the natural logarithms. k_{it} and l_{it} are log labour and log capital in firm i in period t and q_{it} denotes log value-added, and l_i and l_t in (11) are the individual and time identifiers as categorical explanatory variables.

At the first stage we run a fully nonparametic kernel regression (LCK model) of (11), then we obtain estimates of expected output ($\hat{\Phi}_{it}$) and an estimate for v_{it}. Expected output is given by

$$\hat{\Phi}_{it} = \frac{\sum_{i=1}^{n} q_{it} \mathcal{K}_\delta\left(k_{it}, l_{it}, m_{it}, \iota_i, \iota_t, \mathbf{x}_{it}\right)}{\sum_{i=1}^{n} \mathcal{K}_\delta\left(k_{it}, l_{it}, m_{it}, \iota_i, \iota_t, \mathbf{x}_{it}\right)}, \tag{13}$$

in which K is the kernel function for the vector of mixed variables[6]. For the DLW model,

$$\hat{\Phi}_{it} = \hat{\beta}_l l_{it} + \hat{\beta}_k k_{it} + \hat{\beta}_{ll} l_{it}^2 + \hat{\beta}_{kk} k_{ik}^2 + \hat{\beta}_{lk} l_{it} k_{it} + \hat{f}_t\left(k_{it}, l_{it}, m_{it}\right) + \nu_{it},$$

[6] Please refer to Racine (2008) and Racine and Li (2004) for details of fully nonparametric estimation with continuous and discrete data, and how to find optimal smoothing parameters for discrete data. Also, see Appendix A for the basics of nonparametrics.

in which $\hat{f}_t$ is estimated by high-order polynomial series of k_{it}, l_{it}, and m_{it}.

The second stage estimates coefficients for the production function through the law of motion for productivity such that

$$z_{it} = g\left(z_{it-1}\right) + \eta_{it}.$$

Following the DLW model, we allow for the potential of additional (lagged and observable) decision variables to affect current productivity outcomes (in expectation), in addition to the standard inclusion of past productivity. By allowing plant-level decisions such as export participation and investment, which directly affect a firm's future profit, the DLW model tackles concerns of De Loecker (2010), who discusses potential problems of restricting the productivity process to be completely exogenous.

After the first stage, we can compute productivity for any value of β, where $\beta = (\beta_l, \beta_k, \beta_{ll}, \beta_{kk}, \beta_{lk})$, , using $z_{it}\left(\beta\right) = \hat{\Phi}_{it} - \beta_l l_{it} - \beta_k k_{it} - \beta_{ll} l_{it}^2 - \beta_{kk} k_{ik}^2 - \beta_{lk} l_{it} k_{it}.$ The innovation to productivity given $\beta, \eta_{it}\left(\beta\right)$ is recovered by regressing $z_{it}\left(\beta\right)$ on its lag $z_{it-1}\left(\beta\right).$. Then, we use generalised moment conditions to estimate parameters of the production function such that

$$\mathbb{E}\left[\eta_{it}\left(\beta\right) \times \left(l_{it-1}, l_{t-1}^2, l_{t-1}k_t, k_{it}, k_{it}^2\right)'\right] = 0.$$

The moments above are from the DLW model and exploit the fact that the investment is assumed to be decided at a period ago and therefore should not be correlated with the innovation in productivity. We use lagged labour to identify the coefficients on labour since current labour is expected to react to shocks to productivity, and hence $\mathbb{E}\left[l_{it}\eta_{it}\right]$ is expected to be nonzero. In fact, DLW (2012) require input prices to be correlated over time while using lagged labour as a valid instrument for current labour, and they already find very strong evidence for that requirement by running various specifications that essentially relate current wages to past wages.

~

The estimated output elasticities are computed using the estimated coefficients of the production function. Under a translog value-added production function, the output elasticity for labour (1) is given by

$$\hat{\epsilon}_{it}^{l} = \hat{\beta}_l + 2\hat{\beta}_{ll} l_{it} + \hat{\beta}_{lk} k_{it}.$$

In addition, a Cobb-Douglasproduction implies that the output elasticity of labour is simply given by $\hat{\beta}_l$. . Finally, using expression (10) and our estimate of the output elasticity, we compute markups directly. However, we only observe $\tilde{Q}_{it}$, which is given by $Q_{it} \exp(\nu_{it})$. The first stage of our procedure gives us an estimate for v_{it} and we use it to compute the expenditure share such that

$$\hat{\sigma}_{it}^{X} = \frac{P_{it}^{X} X_{it}}{P_{it} \frac{\tilde{Q}_{it}}{\exp(\hat{\nu}_{it})}}.$$

This correction like the DLW model is important as it removes any variation in expenditure shares resulting from variation in output not correlated with $\Phi(k_{it}, l_{it}, m_{it}, \iota_i, \iota_t, \mathbf{x}_{it})$, or output variation not related to variables impacting input demand including input prices, productivity, technology parameters, and market characteristics, such as the elasticity of demand and income levels. These estimates for the markup as given by equation (10) for plant i at time t are computed while allowing for considerable flexibility in the production function, consumer demand, and competition (DLW [2012]).

4 Empirical Results

In this section, we use our empirical model to estimate markups for Korean manufacturing firms, and test whether exporters and non-exporters, also large and small plants have, on average, different markups. In addition, we substantially investigate how markups change with correlation with market share and export status, additionally, industry import penetration, and as such we are the first, to our knowledge, to provide robust econometric evidence of this relationship with unbalanced fixed effect regression and dynamic unbalanced panel regression.

After estimating the output elasticity of labour and materials, we can compute the implied markups from the first order conditions as described above. We use our markup estimates to discuss several major findings. First, we compare our markup estimates to the DLW model and the OLS model. Second, we look at the relationship between markups and plant-level export status and market size, and industry import penetration effect in both the cross-section and the time series. Third, we briefly discuss the relationship between markups and other economic variables.

4.1. Background and Data

We use a plant-level dataset covering firms selected in Korean manufacturing during the period 1980–2001. The data are provided by the Korean Statistical Office and contains plant-level accounts for an unbalanced panel of 91,522. We have the information about market entry and exit, as well as detailed information on plant-level export status and export sales. At every point in time t, we know whether the firm is a domestic producer, an export entrant, an export quitter, or a continuing exporter. Table 1 provides some summary statistics about numbers of observations, observation period, manufacturing industries, and plants in data. In addition, Table 2 presents basic statistics of input variables related to production, value-added, export, material cost, labour and capital. The unit of variables except monthly average employees is Mil. KRW.

Table 1: Data Statistics

This table lists numbers of observations, observation period, manufacturing industries, and plants in data.

	Value
Number of Observations	576,690
Observation Period	> 5 year
Number of Industries	69
Number of Plants	91,522

Table 2: Statistics of Input Variables

This table lists basic statistics of input variables related to production, value-added, export, material cost, labour and capital. The unit of variables except monthly average employees is Mil. KRW.

Variable	Min	Median	Max	Mean	Std. Dev.
Nominal Production	2	400	17,100,000	4,150	81,154
Nominal Export	0	0	8,466,105	1,230	44,315
Nominal Material Cost	0.2	161	9,288,284	2,271	46,490
Real Material Cost	0.0	3.1	140,137	42	819
Monthly Average Employees	2	13	33,553	45	315
Property, Plant and Equipment	0.5	141	9,041,855	2,010	43,344
Real Production	2.1	495	16,500,000	4,676	82,190
Real Value-Added	0.0	195	5,107,007	1,461	26,035

4.2. Estimated Markups

We obtain an estimate of each plant's markup and unobservable productivity shock (or total factor productivity, TFP) and compare the average or median with the DLW and the OLS approach (simple regression of the first stage without the second stage of structural estimation) in Table 3. Although our focus is not so much on the exact level of the markup and TFP, we want to highlight that the estimated markups and TFP are comparable to those obtained with different methodologies, but are different in an important way.

Our procedure generates industry-specific production function coefficients, which in turn deliver firm-specific output elasticity of variable inputs. The latter are plugged in the FOC of input demand together with data on input expenditure to compute markups. We list the median markup using a set of specifications to highlight our results in Table 3. We first present results using our standard methods using the LCK model. We present our results using value-added functions (for value-added production functions, we rely on the output elasticity of labour to compute markups), allowing for nonparametric series regression (DLW model) and a conventional OLS model (CD production).

Table 3: Statistics of TFPs and Markups

This table lists the statistics of TFPs and markups estimated by a local constant kernel (LCK) model, a De Loecker and Warzynski (DLW, 2012) model, and an OLS model. The root mean squared error (RMSE) shows the deviation of fitted value-added (VA) from real value-added. The lower panel shows correlations of LCK, DLW, and OLS markups.

Model		RMSE	1%	Median	99%	Mean	Std. Dev.
LCK							
	q	0.39					
	TFP		1.08	3.45	4.55	3.32	0.65
	Markup		0.41	1.61	8.93	2.09	2.51
DLW							
	q	0.65					
	TFP		0.72	3.45	6.32	3.33	0.98
	Markup		-0.57	1.68	9.75	2.21	2.27
OLS							
	Markup		0.62	2.05	9.35	2.57	2.39
Correlation		LCK	DLW		OLS		
LCK		1					
DLW		0.54	1				
OLS		0.87	0.48		1		

As can be seen in Table 3, the RMSE of the LCK model is much lower than that of the DLW model. This means the measurement error from the LCK model is estimated to be small as long as suitable to data compared to the DLW model. In addition, the median of the LCK model is slightly lower than that of the DLW model, but much lower than that of the OLS model. The literature argues that the simple OLS model (based on CD function) has biased estimates for coefficients so that markup estimates from the OLS model might have relatively upward-bias compared to other structural estimation. However, the interesting thing is that OLS markups are higher correlated to

LCK markups than DLW markups. The correlation between LCK and DLW markups is only 0.54, which is much lower than we expect because LCK and DLW markups basically share the estimation framework except the first stage for ϕ_{it}. Figure 1 shows distributions of markups estimated by the LCK, DLW, and OLS models, respectively, which are left-skewed sequentially by list. In addition, Figure 2 presents distributions of LCK markups over time from 1980 to 2001. As time goes by, the distributions of markups are getting denser and lower, which can be interpreted as the effect on firms' markups from changes in the competitive environment and the globalisation of the Korean economy.

Figure 1: Distributions of Markups According to Estimation Models

This figure shows distributions of markups estimated by a local constant kernel (LCK) model, a De Loecker and Warzynski (DLW: 2012) model, and an OLS model. The vertical line shows the frequency of distributions, and the horizontal line shows markups from 0 to 10 in the figure. The solid line represents the distribution of LCK markups, the dashed line is for DLW markups, and the dotted line is for the distribution of OLS markups.

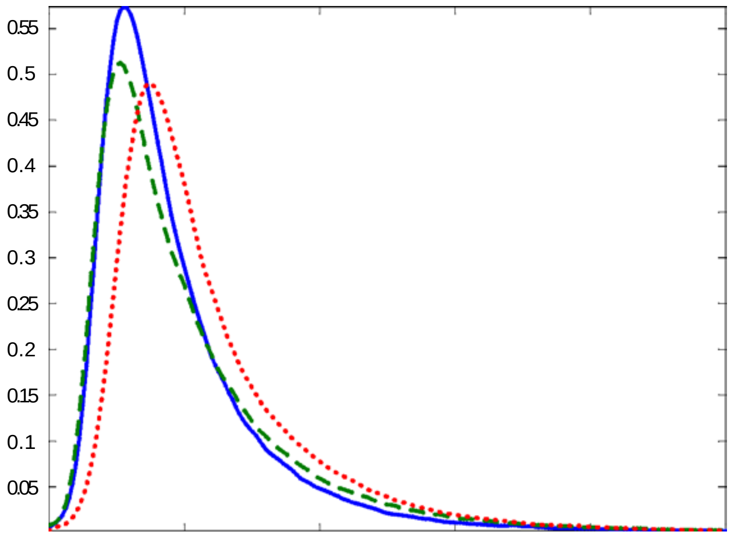


Figure 2: Distributions of Markups over Time

This figure shows distributions of LCK markups and standard deviations of LCK markups and log of sales over time from 1980 to 2001. The vertical line in the upper panel shows the frequency of distributions over time, and the horizontal line shows markups from 0 to 5 in the figure. The arrow shows the direction of medians of markups over time. The solid line in the lower panel represents standard deviations of LCK markups and the dashed line depicts standard deviations of log sales in real terms.

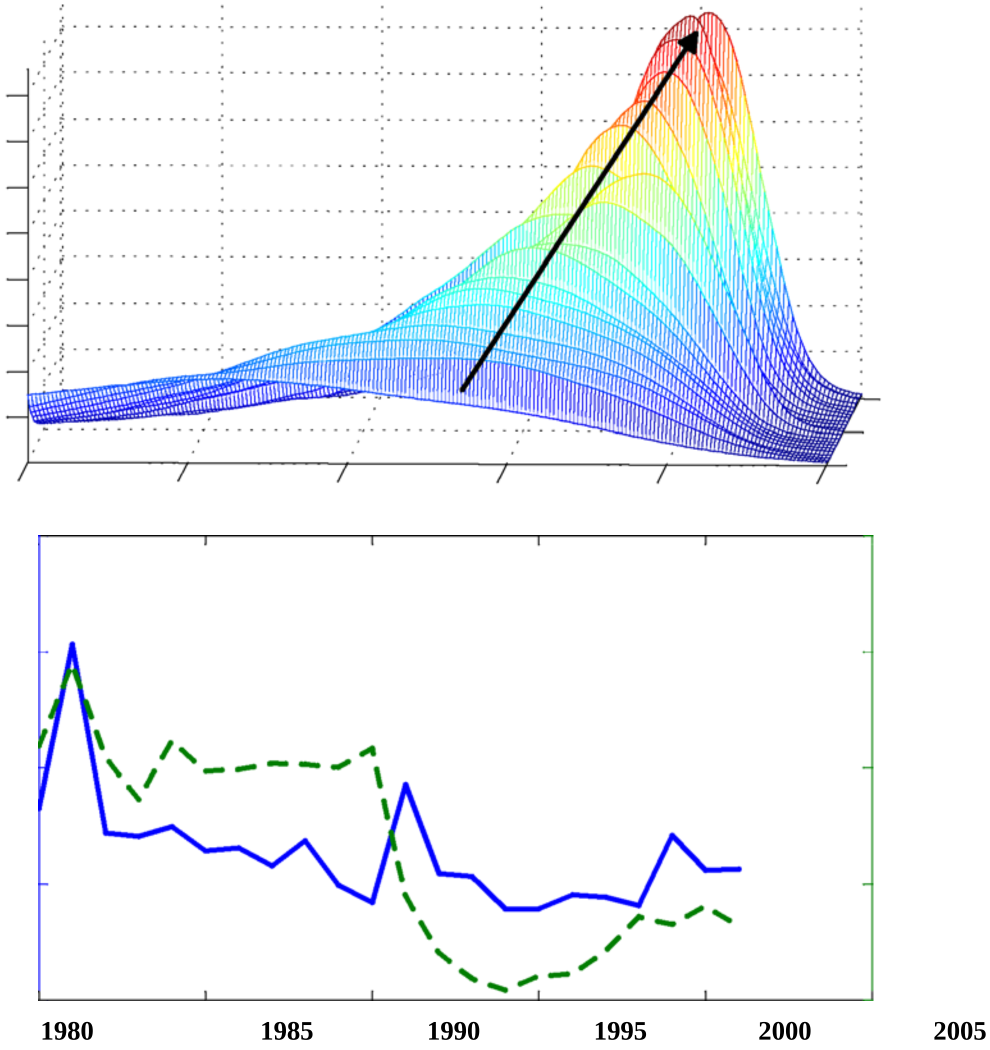


Table 4 presents means of four groups' markups (LCK model) as independent sorts of size and globalisation (export status). The small plants are in the lower 30 percent of sales in each industry at each time period, and the large plants are in upper 30 percent of sales in each industry at each time period, and the other sort is exporter or non-exporter.

As can be seen in Table 4, mean differences between large firms and small firms given export status is relatively larger than mean differences between exporters and non-exporters given firm size. We can interpret this to mean that firm markups are affected by firm size rather than by firm globalisation strategy. Figures 3–5 show distributions of exporters and non-exporters, large and small plants, and four groups' markups as independent sorts of size and globalisation. As can be seen from the lower panel in Figure 5, mean and median differences of large and small firms' markups given the export status is bigger than those of exporters and non-exporters given firm size over time. However, we can see that mean and median differences decrease over time, which is contrary to the notion that the polarisation between large and small or exporters and non-exporters would be deteriorating over time. This phenomenon might occur as a result of stronger competition in the industry, in other words, the markup gap decreases in the degree of competition intensified over time even though the innovation polarisation deteriorates over time.

Table 4: Means and Differences of Markups

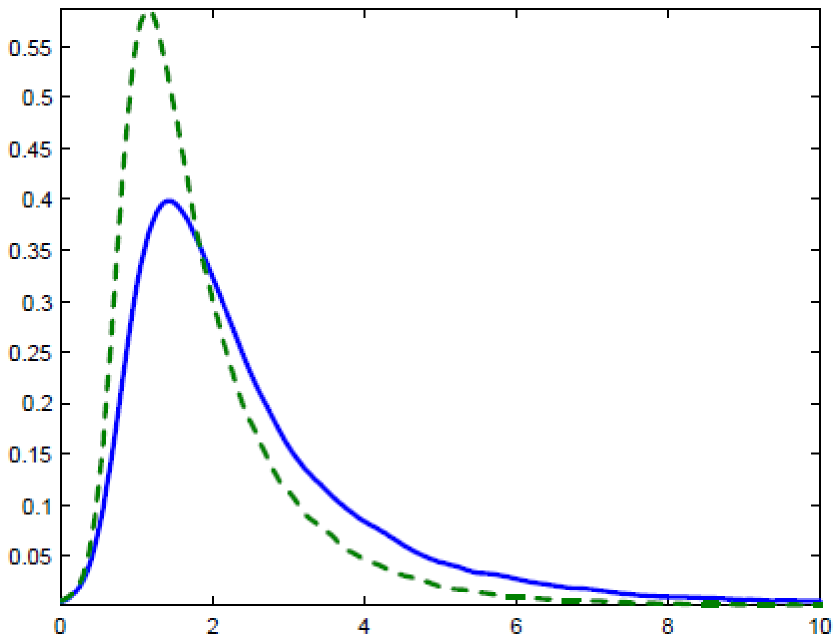
This table lists means of four groups' markups (LCK model) as independent sorts of size and globalisation. The small plants are in the lower 30 percent of sales in each industry, the large plants are in the upper 30 percent of sales in each industry, and the other sort is exporter or non-exporter. Numbers of plants are annual averages from1980 to 2001. t-statistics in parentheses are for mean differences, defined as mean difference divided by the standard error (the standard deviation of mean difference divided by the square root of number of years).

	Means of Markups		
	Exporter(A)	Non-Exporter(B)	(A)-(B)
Large Plant(C)	3.63	2.84	0.79 (12.7)
Small Plant(D)	2.05	1.97	0.08 (2.96)
(C)-(D)	1.57	0.86	
	(16.4)	(16.2)	

	Numbers of Plants		
			Exporter
			Large Plant
Small Plant	2,569	7,551	

Figure 3: Distributions of Exporters and Non-Exporters' Markups

This figure shows distributions of exporters’ and non-exporters' markups. The solid line in the upper panel represents the distribution of exporters' markups, and the dashed line is for the distribution of non-exporters' markups. The solid line in the lower panel shows the difference between medians of exporters’ and non-exporters' markups over time, and the dashed line depicts the difference between means of exporters’ and non-exporters' markups over time.



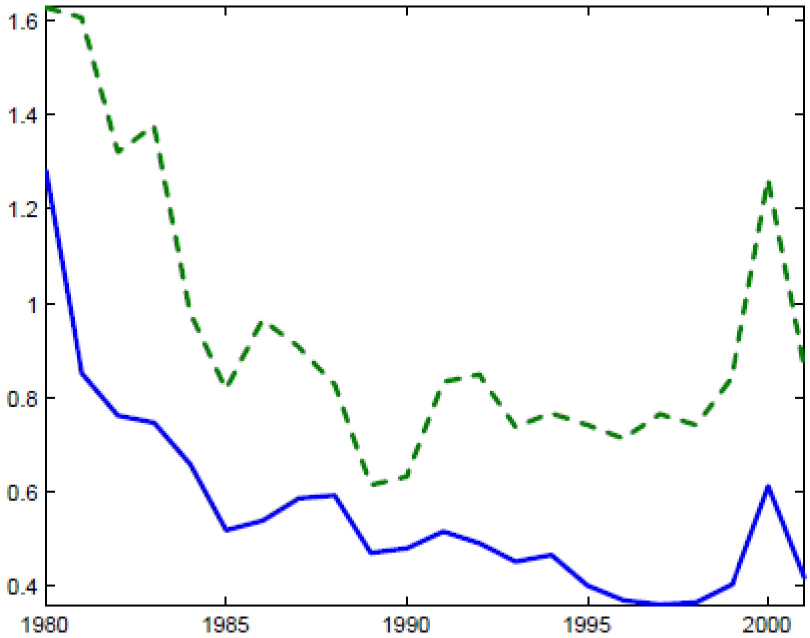
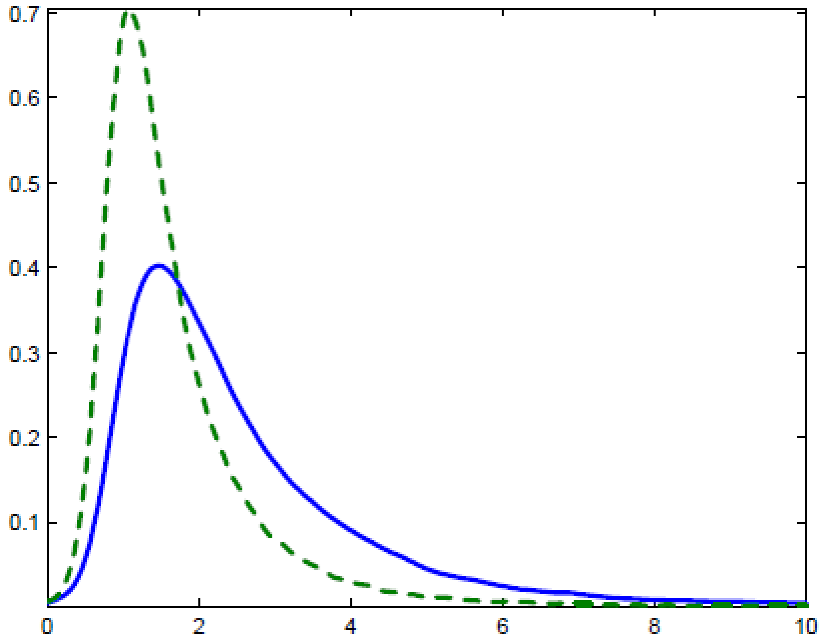


Figure 4: Distributions of Small and Large Plants' Markups

This figure shows distributions of small plants’ and large plants' markups. The small plants are in the lower 30 percent of sales, and the large plants are in the upper 30 percent of sales. The solid line in the upper panel represents the distribution of large plants' markups, and the dashed line is for the distribution of small plants' markups. The solid line in the lower panel shows the difference between medians of large and small plants' markups over time, and the dashed line depicts the difference between means of large and small plants' markups over time.


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10

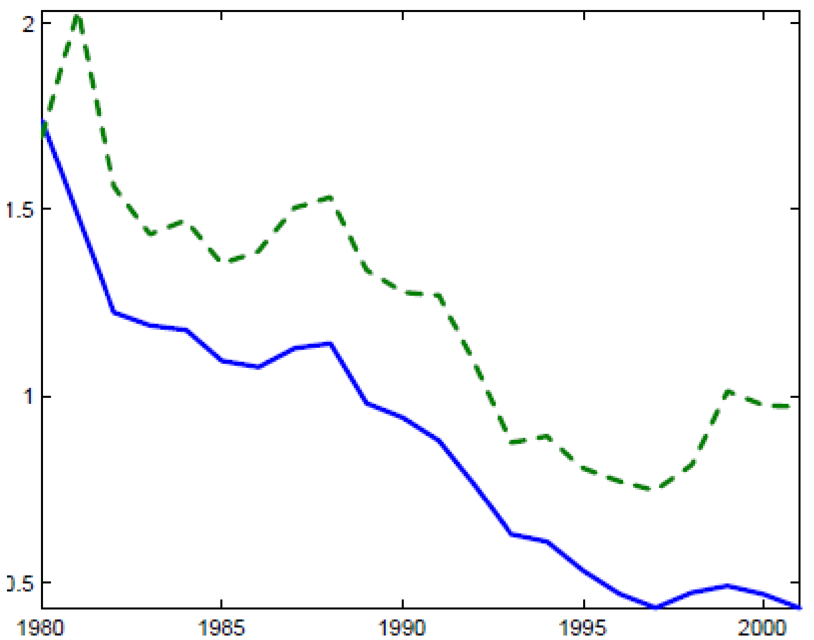
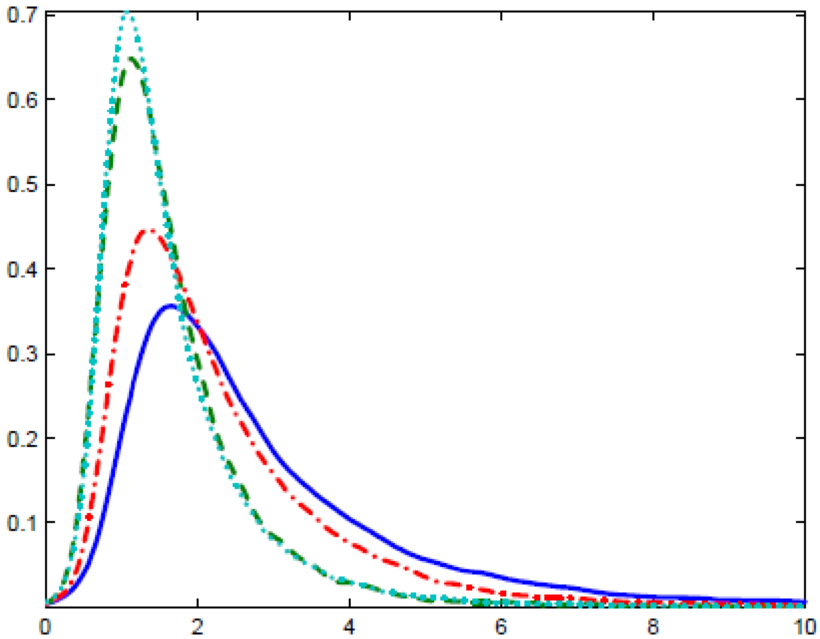
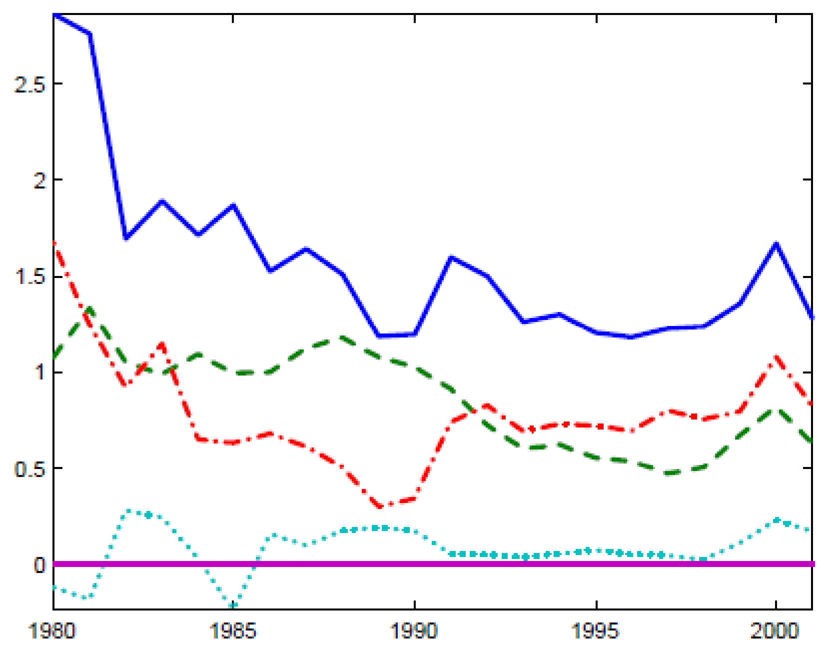

2
1.5
1
0.5
1980
1985
1990
1995
2000

Figure 5: Distributions of Markups of Plants sorted by Size and Globalisation

This figure shows distributions of four groups' markups as independent sorts of size and globalisation. The small plants are in the lower 30 percent of sales in each industry, the large plants are in the upper 30 percent of sales in each industry, and the other sort is exporter or non-exporter. The solid line in the upper panel represents the distribution of large and exporting plants' markups, the dashed line depicts the distribution of small and exporting plants' markups, the dashed-dotted line represents the distribution of large and non-exporting plants's markups, and the dotted line shows the distribution of small and non-exporting plants' markups. The solid line in the lower panel shows the difference between medians of large and small exporters' markups over time, the dashed line depicts the difference between medians of large and small non-exporters' markups, the dashed-dotted line shows the difference between medians of large exporters and non-exporters over time, and the dotted line represents the difference between medians of small exporters and non-exporters over time.





4.3. Unbalanced Panel Data Analysis for Markups

We can now turn to the main focus of our application—how markups on average are affected by size, globalisation, and productivity shock and whether markups change when the import penetration in industry increases. We discuss unbalanced panel data analysis for markups in fixed effects regression and dynamic panel regression.

The estimation framework introduced above was not explicit about firms selling in multiple markets. In light of our application we want to stress that our measure of markups for globalisation is a share-weighted average markup across the multi-markets, where the weight by market is the share of an input's expenditure used in production sold in that market. We can correctly compare markups across producers and time without requiring additional information on input allocation across production destined for different markets. To compare markups across markets within a plant, we do require either more data or more theoretical structure to pin down the input allocation by final market.

Given plant-specific markups, we can simply relate a plant's markup to its size and globalisation (export status) in a regression framework. As noted above, we are not interested in the level of the markup and instead we estimate the percentage difference in markups depending on the plant's size (market share in industry and export status).

The unbalanced panel specification we take to the data is given by

$$\ln \hat{\mu}_{ijt} = X_{ijt}\gamma + \iota_i + \iota_j + \iota_t + \xi_{ijt},$$

in which ι_i, ι_j, and ι_t are individual, industry, and time effects, respectively. X_{ijt} is an independent vector with industrial import penetration ratio, log (z_{it}), log (K_{ijt}/L_{ijt}), log(market share$_{jjt}$), export dummy, and export dummy x log(market shareijt). We control for labour and capital use, log(K_{ijt}/L_{ijt}), in order to capture differences in factor intensity, as well as full year-industry inter-actions to take out industry specific aggregate trends in markups (ι_j). We collect all the controls in a vector X_{ijt} with γ the corresponding coefficients.

We rely on our approach to test whether, on average, exporters have different markups as well as a different slope for exporters' market share. The latter, to our knowledge, has not been documented and we see this as a first important set of results. We are interested in the coefficients on the various control variables, so further below we will discuss the separate coefficients of other economic variables such as total factor productivity and industry import penetration. We estimate this fixed effect regression at the manufacturing level and include a full interaction of year and industry dummies. Once we have estimated coefficients of export dummy and export dummy x log(market share$_{ijt}$), we can compute the level of markup difference by applying the percentage difference to the constant term, which captures the domestic markup average. We denote this markup ratio between exporters' markup μ^E_{ijt} and non-exporters' markup μ^N_{ijt}, , and we compute it by applying

$$\mathbb{E}\left[\frac{\mu^E_{ijt}}{\mu^N_{ijt}} | X_{ijt} \text{ except export status}\right] = \exp\left[\gamma^E + \gamma^{E\times\log(\text{market share})} \log\left(\text{market share}_{ijt}\right)\right]$$

after estimating the relevant parameters. Table 5 presents our results.

Table 5: Market Share and Export Effects on LCK Markups in Unbalanced Panel

This table shows results of fixed effect regressions in unbalanced panel data for markups estimated by local constant kernel (LCK) model such as

$$\ln \hat{\mu}_{ijt} = X_{ijt}\gamma + \iota_i + \iota_j + \iota_t + \xi_{ijt},$$

in which l_i, l_j, and l_t are individual, industry and time effects, respectively. X_{ijt} is an independent vector with industrial import penetration ratio, log(z_{it}), log(K_{ijt}/L_{ijt}), log(market share$_{jjt}$), export dummy, and export dummy x log(market share$_{ijt}$) - ** and * refer to the statistical significance levels at 1 percent and 5 percent, respectively. Robust standard errors in brackets are clustered within plants.

	(1)	(2)	(3)	(4)	(5)
Import Penetration	0.000		0.000		
	[0.000]		[0.000]		
Log(zijt)	1.333**	1.065**			
	[0.027]	[0.018]			
Log(Kijt/Lijt)	0.071**	0.072**	0.065**	0.062**	0.062**
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Log(market sharejjt)	-0.077**	-0.039**	-0.016**	0.012**	0.016**
	[0.002]	[0.001]	[0.002]	[0.001]	[0.001]
Dummy(exporter)	0.344**	0.329**	0. 176**	0.131**	0.042**
	[0.015]	[0.013]	[0.018]	[0.013]	[0.002]
Dummy(exporter)	0.047**	0.040**	0.022**	0.011**	
xLog(market shareijt)	[0.002]	[0.001]	[0.002]	[0.001]	
Industry dummy (L~)	yes	yes	yes	yes	yes
R-sq: within	0.31	0.40	0.14	0.27	0.27
Num. of Plants	61,549	78,803	61,557	78,812	78,812
Num. of Obs	320,385	565,899	320,679	566,756	566,756

We run the fixed effect regression for the various estimates of the markups as described above. The parameter γ^E is estimated very significantly in all specifications (1) — (5) and values are between 0.042 and 0.344, which means that the exporters' markup is, on average, about 4.2 percent to 34.4 percent greater than non-exporters' markup, and values for coefficient - $\gamma^{E\times\log(\text{market share})}$ range from 0.011 to 0.047. The parameter for the log market share ranges from -0.077 to 0.016. As expected, all the results except base of market share level relying on translog technology are very similar because the variation in markups is almost identical across the various specifications. One important message from this table is that the parameter of market share does not have consistent signs, suggesting that this unbalanced fixed effects regression might have omitted variables.

Under assumptions of dynamic unbalanced panel data analysis of Arellano and Bond [1991], we take to the data is given by

$$\Delta \ln \hat{\mu}_{it} = \alpha \Delta \ln \hat{\mu}_{it-1} + \Delta X_{it}\gamma + \Delta \xi_{it},$$

in which Xit is an independent vector with industrial import penetration ratio, log(z_{it}), log(K_{it}/L_{it}), log(market share$_{it}$) export dummy, export dummy x log(market share$_{it}$), and time dummy. The second lags of log(z_{it}), log(K_{it}/L_{it}), log(market shareit), export dummy x log(market share$_{it}$), and the first differences of industrial penetration ratio, export dummy, and time dummy are used as instrument variables in difference GMM system. Table 6 presents our results. The parameter γ^E is estimated very significantly in all specifications (1) — (5) and values are between 0.054 and 0.396, which are slightly higher than values in Table 5, and values for coefficient $\gamma^{E\times\log(\text{market share})}$ range from 0.010 to 0.043, which are similar to results of fixed effects regressions. The parameters for the log market share in all specifications have robust positive signs. The significances for the import penetration ratio are weak, thus we need to consider other variables to capture the industrial characteristics. In addition, similarly to DLW (2012), TFP increases the markup on average by 16.3 percent to 24.0 percent.

Table 6: Market Share and Export Effects on LCK Markups in Dynamic Unbalanced Panel

This table shows results of difference GMMs in dynamic unbalanced panel data (Arellano and Bond [1991]) for markups estimated by using a local constant kernel

(LCK) model, such as

$$\Delta \ln \hat{\mu}_{it} = \alpha \Delta \ln \hat{\mu}_{it-1} + \Delta X_{it}\gamma + \Delta \xi_{it},$$

in which Xit is an independent vector with industrial import penetration ratio, log(z_{it}), log(K_{it}/L_{it}), log(market share$_{it}$), export dummy, export dummy x log(market share$_{it}$), and time dummy. The second lags of log(z_{it}), log(K_{it}/L_{it}), log(market share$_{it}$), export dummy x log(market share$_{it}$), and the first differences of industrial penetration ratio, export dummy, and time dummy are used as instrument variables in difference GMM system. ** and * refer to the statistical significance levels at 1 percent and 5 percent, respectively. Robust standard errors in brackets are estimated by the finite-sample corrected two-step covariance matrix.

	(1)	(2)	(3)	(4)	(5)
Log(markup$_{it-1}$)	0.169**	0.175**	0.179**	0.185**	0.184**
	[0.005]	[0.004]	[0.006]	[0.004]	[0.004]
Import Penetration	0.000		0.001*		
	[0.000]		[0.000]		
Log(z_{it})	0.240**	0.163**			
	[0.075]	[0.038]			
Log(K_{it}/L_{it})	0.079**	0.045**	0.074**	0.036**	0.034**
	[0.008]	[0.005]	[0.009]	[0.005]	[0.006]
Log(market share$_{it}$)	0.080**	0.115**	0.079**	0.109**	0.118**
	[0.015]	[0.010]	[0.016]	[0.011]	[0.012]
Dummy(exporter)	0.121	0.396**	0.054	0.383**	0.063**
	[0.134]	[0.102]	[0.152]	[0.107]	[0.006]
Dummy(exporter)	0.010	0.043**	0.002	0.042**	
x Log (market share$_{ijt}$)	[0.017]	[0.013]	[0.020]	[0.014]	
Num. of Plants	48,674	76,472	48,686	76,502	76,502
Num. of Obs	199,926	370,917	200,203	371,701	371,701

For comparison of the DLW and OLS models, Table B.1-B.4 shows the results of unbalanced fixed effects and dynamic panel regressions. Tables for the DLW model show that the DLW model still has the negative signs for market share, and weak consistent signs for the parameter of productivity shock. The OLS model has negative signs for variables related to export dummy. Therefore, the results of the LCK model are robustly consistent with industrial organization and international economic theories compared with those of the DLW and OLS models.

For the last exercise, we directly quantify how import competition can influence the dispersion of markups. Since large firms set higher markups than small firms, the dispersion of markups is closely related to the gap in markups between small and large firms. Table 7 reports industry panel fixed-effect regressions. It shows that import competition measured as import penetration reduces the differential in firm markups. In the first column, the standard deviation of industry markups decreases by about 0.07 percent with a 1 percentage point increase in import penetration. The inequality of markups between firms decreases with intensified international competition, as the theory predicts.

Table 7: Import Penetration Effect on Dispersion of Markups

This table shows results of industry panel fixed-effect regressions

$$\ln SD_{jt} = \alpha + \beta_1 \ln IMPR_{jt} + \beta_2 \ln K/L_{jt} + \beta_3 \ln \mu_{jt}$$

in which 5Djt is a standard deviation of markup in an industry *j*, IMPR is an industry import penetration, log(K_{jt}/L_{jt}) is an industry capital-labour ratio, and μ_{it} is log industry average markup. Overall import penetration can be categorized into two types. One is only imports from China, and the other is imports from the rest of the world. This classification comes from Bernard, Jensen, and Schott (2006). They emphasize that the response of industry employment to import competition from low-wage countries such as China can be different from typical import competition from the rest of the world. The dependent variable of columns (1) and (2) is the standard deviation of industry markups, and columns (3) and (4) use an inter-quartile range of industry markups as a response variable. ** and * refer to the statistical significance levels at 1 percent and 5 percent, respectively.

	(1)	(2)	(3)	(4)
Import Penetration	-0.067**		0.068	
	[0.019]		[0.039]	
Import Penetration(Other)		-0.099**		0.057
		[0.020]		[0.043]
Import Penetration(China)		1.162**		0.504
		[0.315]		[0.675]
$Log(K_{jt}/L_{jt})$	0.240**	0.211**	0.378**	0.368**
	[0.035]	[0.033]	[0.069]	[0.072]
$Log(\mu_{i_t})$	0.040**	0.056**	0.113**	0.119**
	[0.016]	[0.015]	[0.031]	[0.033]
Num. of Industry	16	16	16	16
Num. of Obs	130	130	130	130

We further estimate whether import competition from low-wage countries such as China has a stronger effect on domestic firm behaviour. Interestingly, the second column reveals positive signs for the effect of import penetration from China on markup dispersion, while import penetration from the rest of the world retains its negative effect on markup dispersion. In some sense, it is embarrassed, but it can be possible if forces of import competition are concentrated on only very small firms. Products from China are usually low quality and low price. These types of goods are commonly made by domestic small firms. It would be plausible to assume that if the good markets are segmented by high and low quality goods, and the substitution between high and low quality products is very low, import competition from low-wage countries would affect only low price goods. If domestic small firms face stronger competition from low wage countries, they have to cut prices to stay in the market, whereas large firms with high quality goods are generally able to maintain their prices. In this case, the standard deviation of markups rises with increased import penetration.

For our robustness check we use an inter-quartile range of markups as markup dispersion in each industry as a dependent variable. However, import penetration has an insignificant effect on markup dispersion in this case. It implies that changes arising from import penetration are concentrated on the lower tail or upper tail of the support of markups. If the changes in markup dispersion occur uniformly or in overall support, the result should be the same if we use the standard deviation as a dispersion measure. We can conjecture that very small firms or very large firms are more influenced by import penetration, considering the different results of the standard deviation and inter-quartile range.

5 Conclusions

In this paper, we show that large firms decide on higher markups in each industry as they have greater market powers in integrated markets. Also, exporters set higher markups through higher observable productivity than non-exporters. This is empirically consistent with the theory that firms conditional on higher observable productivity decide on higher markups. Interestingly, even after controlling for productivity and other firm characteristics, the level of markup is proportional to the market share. A one percent increase in market share leads to a 0.080.12 percent increase in markup. It draws attention since it is the evidence that the firm strategy of price is reflected by pure market power.

To answer the question whether globalisation confers unequal benefits on small and large firms, we generate markup distribution and find that the mean and the dispersion of markups have been decreasing over time. However, the average firm size and firm size distribution have been increasing. These patterns are predicted exactly by the theoretical model of trade. The main hurdle is to identify the effect of globalisation. In order to investigate the effect of globalisation, we use industry panel fixed-effect regressions. For the proxy of globalisation, the import penetration is used. It is a disadvantage that import penetration only captures the effect of importing as globalisation includes both imports and exports. It turns out that import competition reduces the markup gap between small and large firms, as predicted by the theory.

Methodologically, we develop De Loecker and Warzynski (2012) and control the endogeneity problem by using the difference GMM in dynamic unbalanced panel data suggested by Arellano and Bond (1991). Compared to De Loecker and Warzynski (2012), our estimate of markups has smaller errors and reasonable levels of average and median of markups.

This paper has an important message regarding enterprise policies. A widening of the performance gap, measured as output or sales, between large and small firms, cannot be interpreted as meaning that globalisation negatively affects the welfare of consumers. It is likely that globalisation strengthens competition between all firms, so the gap in prices or markups shrinks as a result of the selection effect, which benefits consumers. Therefore, protective policies to shield SMEs from the effects of globalisation may interfere with the selection process and harm productivity growth.

References

Ackerberg, D., K. Caves and G. Frazer (2006), *Structural Identification of Production Functions*. Unpublished.

Atkeson, A. and A. Burstein (2008), 'Pricing-to-Market, Trade Costs and International Relative Prices', *American Economic Review*, 98(5), pp.1998–2031.

Alvarez, R. and R. Lopez (2005), 'Exporting and Performance: Evidence from Chilean Plants', *Canadian Journal of Economics*, 38(4), pp.1384–400.

Arellano, M., and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58, pp.277–97.

Blalock, G. and P. Gertler (2004), 'Learning from Exporting: Revisited in a Less Developed Setting', *Journal of Development Economics*, 75(2), pp.397–416.

Choi, Y.-S. and C.H. Hahn (2010). The Effects of Imported Intermediate Varieties on Plant Total Factor Productivity and Product Switching: Evidence from Korean Manufacturing. ERIA Research Project Report 2010.

Criscuolo, C. and R. Martin (2009), 'Multinationals and U.S. Productivity Leadership: Evi-dence from Great Britain', *Review of Economics and Statistics*, 91(2), pp.263–81.

Czekaj, T. and A. Henningsen (2013), 'Panel Data Specifications in Nonparametric Kernel Regression: An Application to Production Functions', *IFRO Working Paper*, Department of Food and Resource Economics, University of Copenhagen.

De Loecker, J. (2010), 'A Note on Detecting Learning by Exporting', *NBER Working*

Paper Series, No. 16548, Cambridge, MA: NBER.

De Loecker, J. and F. Warzynski (2012), 'Markups and Firm-level Export Status', *American Economic Review*, 102(6), pp.2437–71.

Dixit, A.K., and J.E. Stiglitz (1977), 'Monopolistic competition and Optimum Product Diversity', *American Economic Review* 67 (June): 297–308.

Edmund, C., V. Midrigan and D. Yi Xu 2013. Competition, Markups, and the Gains from International Trade. Mimeo.

Fernandes, A. (2007), 'Trade Policy, Trade Volumes and Plant-Level Productivity in Colombian Manufacturing Industries', *Journal of International Economics* 71, pp.52–71.

Foster, L.S., J. Haltiwanger and C. Syverson (2008), 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?', *American Economic Review* 98 (1), pp.394–425.

Goldberg, P.K., A.K. Khandelwal, N. Pavcnik and P. Topalova (2010), Imported Intermediate Inputs and Domestic Product Growth: Evidence from India', *Quarterly Journal of Economics* 125 (4), pp.1727–67.

Gyimah-Brempong, K. and J.S. Racine (2010), 'Aid and Investment in LDCs: A Robust Approach', *The Journal of International Trade & Economic Development: An International and Comparative Review* 19, pp.319–49.

Henderson, D.J., and L. Simar (2005)' 'A Fully Nonparametric Stochastic Frontier Model for Panel Data', *Working Paper 0519*, Department of Economics, State University of New York at Binghamton.

Hurst, E., and B. Pugsley (2011). What Do Small Businesses Do? Brookings Papers on Economic Activity, pp. 73–142.

Kugler, M., and E. Verhoogen (2008). The Quality-Complementarity Hypothesis: Theory and Evidence from Colombia. *NBER Working Paper Series*, No. 14418, Cambridge, MA: NBER..

Levinsohn, J., and A. Petrin (2003), 'Estimating Production Functions Using Inputs to Control for Unobservables', *Review of Economic Studies*, 70(2), pp.317–41.

Li, Q. and J.S. Racine (2004), 'Cross-validated Local Linear Nonparametric Regression', *Statistica Sinica* 14, pp.485–512.

Melitz, M., and J.A. Levinsohn (2006). *Productivity in a Differentiated Products Market Equilibrium*. Unpublished

Melitz, M.J., and G.I.P. Ottaviano 2008, 'Market Size, Trade, and Productivity', *Review of Economic Studies* 75, pp.295–316.

Oh, J. (2013). The Cyclicality of Firm Size Distribution and its Effect on Aggregate Fluctuations. Mimeo.

Olley, S.G., and A. Pakes (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica*, 64(6), pp.1263–97.

Pavcnik, N. (2002), 'Trade Liberalization Exit and Productivity Improvements:

Evidence from Chilean Plants', *Review of Economic Studies* 69, pp.245–76.

Racine, J.S. (2008), 'Nonparametric Econometrics: A primer', *Foundations and Trends in Econometrics* 3, pp.1–88.

Racine, J.S., and Q. Li (2004), 'Nonparametric Estimation of Regression Functions with both Categorical and Continuous Data', *Journal of Econometrics* 119, pp.99–130.

Roberts, M.J., and D. Supina (1996), 'Output Price, Markups, and Producer Size', *European Economic Review* 40, pp.909–21.

Topalova, P., and A. Khandelwal (2011), 'Trade Liberalization and Firm Productivity: The Case of India', *The Review of Economics and Statistics*, 93(3), pp.995–1009

Wooldridge, J.M. (2009), 'On Estimating Firm-Level Production Functions using Proxy Variables to Control for Unobservables', *Economic Letters* 104(3), pp.112–14.

Appendix

A. Local Constant Kernel

Regression

A.1 Basics of LCK Regression

The nonparametric model is as follows:

$$y_i = g(x_i) + u_i, \; i = 1, \ldots, n,$$

in which the functional form g (•) is unknown. If g (•) is a smooth function, we can estimate g (•) nonparametrically using kernel methods so that we consider g (•) as the conditional mean of y given x such that

$$g(x) = \mathbb{E}[y_i | x_i = x],$$

due to the general result of nonparametric theory. We note that $\mathbb{E}[y_i | x_i = x] = \int y f_{y,x}(x, y)\, dy$ can be replaced with $\int y \hat{f}_{y,x}(x, y)\, dy$ with the unknown probability density function $f_{y,x}(x, y)$ estimated by the kernel method such that

$$\hat{f}_{y,x}(x, y) = \frac{1}{nh_0 \ldots h_q} \sum_{i=1}^{n} \mathcal{K}\left(\frac{x_i - x}{h}\right) k\left(\frac{y - y_i}{h_0}\right), \qquad (14)$$

in which $\mathcal{K}\left(\frac{x_i - x}{h}\right) = k\left(\frac{x_{i1} - x_1}{h_1}\right) \times \ldots \times k\left(\frac{x_{iq} - x_q}{h_q}\right)$ and where k is a kernel function satisfying basic conditions of nonparametrics, h_0 is the smoothing parameter associated with y, and $h_0 \ldots h_q$ are bandwidths for x_i. From equation (14), we obtain the estimate

$$\hat{g}(x) = \frac{\sum_{i=1}^{n} y_i \mathcal{K}\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^{n} \mathcal{K}\left(\frac{x_i - x}{h}\right)}, \tag{15}$$

which is simply a weighted average of yi, because we can rewrite (15) as

$$\hat{g}(x) = \sum_{i=1}^{n} y_i w_i,$$

in which $w_i = \mathcal{K}\left(\frac{x_i - x}{h}\right) / \sum_{j=1}^{n} \mathcal{K}\left(\frac{x_j - x}{h}\right)$ is the weight attached to y_i

A.2 Cross-Validation Method for Bandwidth

Once we have the continuous explanatory variables xi, the optimal bandwidth h is determined by the cross-validation method, minimizing

$$CV(h_1, \ldots, h_q) = \min_{h} \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{f}_{-i}(x_i)\right)^2 m(x_i), \tag{16}$$

in which $\hat{f}_{-i}(x_i) = \frac{\sum_{l \neq i}^{n} y_l \mathcal{K}\left(\frac{x_i - x_l}{h}\right)}{\sum_{l \neq i}^{n} \mathcal{K}\left(\frac{x_i - x_l}{h}\right)}$ is the leave-one-out kernel estimator of $f(x_i)$ and $m(x_i)$ is a weight function that rules out boundary observations and $\leq m(\cdot) \leq 1$. Then, the asymptotic result of optimal bandwidth is

$$n^{1/(q+4)} \hat{h} = \hat{a} \rightarrow_p a,$$

in which a is uniquely defined, positive, and finite to asymptotically minimize the first leading term of *CV* (*h*).

A.3. LCK Regression with Mixed Data

We now turn to a nonparametric approach with continuous and discrete variables. From a statistical point of view, smoothing discrete variables may introduce some bias, but it is also known that it reduces the finite-sample variance resulting in a reduction in the finite-sample mean squared error of the nonparametric estimator.

Coming back to a nonparametric regression model given by

$$y_i = g\left(x_i^c, x_i^d\right) + u_i,\ i = 1, \ldots, n,$$

in which x^c and x^d are continuous and discrete variables, respectively. Then we define the estimate of unknown PDF as

$$\begin{aligned}\hat{f}_{y,x}(x,y) &= \frac{1}{nh_0 \ldots h_q}\sum_{i=1}^{n}\mathcal{K}_\delta\left(\frac{x_i - x}{h}\right)k\left(\frac{y-y_i}{h_0}\right)\\ &= \frac{1}{nh_0 \ldots h_q}\sum_{i=1}^{n}\mathcal{W}_h\left(\frac{x_i^c - x^c}{h}\right)L\left(x^d, x_i^d, \lambda\right)k\left(\frac{y-y_i}{h_0}\right),\end{aligned}$$

in which $\delta = (h, \lambda)$, $\mathcal{W}$ is a symmetric, nonnegative univariate kernel function, and $L\left(x^d, x_i^d, \lambda\right) = \prod_{s=1}^{r}\lambda_s^{1\left(x_{is}^d \neq x_s^d\right)}$ where $1\left(x_{is}^d \neq x_s^d\right)$ is an indicator function which equals one when $x_{is}^d \neq x_s^d$ and zero otherwise. The smoothing parameter for x^d is assumed to be $0 \leq \lambda \leq 1$. Then,

$$\hat{g}\left(x_i^c, x_i^d\right) = \frac{\sum_{i=1}^{m} y_i \mathcal{K}_\delta\left(x, x_i^c, x_i^d\right)}{\sum_{i=1}^{n}\mathcal{K}_\delta\left(x, x_i^c, x_i^d\right)},$$

which is analogous to equation (13).

Least squares cross-validation selects 9 = (h, A) to minimize the following function:

$$CV(h,\lambda) = \min_{h,\lambda}\sum_{i=1}^{n}\left(y_i - \hat{g}_{-i}\left(x_i^c, x_i^d\right)\right)^2 m\left(x_i^c, x_i^d\right),$$

in which $\hat{g}_{-i}\left(x_i^c, x_i^d\right)$ and $m\left(x_i^c, x_i^d\right)$ are the same as (16). Note that when $\lambda = 1,\ L\left(x^d, x_i^d, \lambda\right)$ becomes unrelated to $\left(x^d, x_i^d\right)$. Finally, the asymptotic results of smoothing parameter δ is

$$n^{1/(q+4)}\hat{h} = \hat{a} \rightarrow_p a,$$
$$n^{2/(q+4)}\hat{\lambda} = \hat{b} \rightarrow_p b,$$

in which a and b are uniquely defined, positive, and finite to asymptotically minimize the first leading term of $CV(h, \lambda)$.

B. Additional Tables

Table B.1: Market Share and Export Effects on DLW Markups in Unbalanced Panel

This table shows results of fixed effect regressions in unbalanced panel data for markups estimated by a De Loecker and Warzynski (DLW, 2012) model such as

$$\ln \hat{\mu}_{ijt} = X_{ijt}\gamma + \iota_i + \iota_j + \iota_t + \xi_{ijt}.$$

Other descriptions remain the same as in Table 5.

	(1)	(2)	(3)	(4)	(5)
Import Penetration	0.000		0.000		
	[0.000]		[0.000]		
Log(zzjt)	0.656**	0.238**			
	[0.043]	[0.013]			
Log(Kzjt/Lzjt)	0.030**	0.016**	0.027**	0.016**	0.015**
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Log(market sharezjt)	-0.086**	-0.042**	-0.075**	-0.036**	-0.027**
	[0.002]	[0.001]	[0.002]	[0.001]	[0.001]
Dummy(exporter)	0.316**	0.302**	0.263**	0.262**	0.029**
	[0.016]	[0.013]	[0.015]	[0.012]	[0.002]
Dummy(exporter)	0.043**	0.035**	0.036**	0.030**	
xLog(market sharezjt)	[0.002]	[0.001]	[0.002]	[0.001]	
R-sq: within	0.14	0.27	0.12	0.27	0.27
Num. of Plants	60,843	78,231	60,872	78,289	78,289

Num. of Obs	313,937	557,260	314,374	559,753	559,753

Table B.2: Market Share and Export Effects on DLW Markups in Dynamic Unbalanced Panel

This table shows results of difference GMMs in dynamic unbalanced panel data (Arellano and Bond (1991)) for markups estimated by a De Loecker and Warzynski (DLW: 2012) model such as

$$\Delta \ln \hat{\mu}_{it} = \alpha \Delta \ln \hat{\mu}_{it-1} + \Delta X_{it}\gamma + \Delta \xi_{it}.$$

Other descriptions remain the same as Table 6

	(1)	(2)	(3)	(4)	(5)
Log(markupit-1)	0. 199**	0.201**	0.197**	0.205**	0.205**
	[0.006]	[0.004]	[0.006]	[0.004]	[0.004]
Import Penetration	-0.000		-0.000		
	[0.000]		[0.000]		
Log(zit)	0.147**	-0.083**			
	[0.051]	[0.017]			
Log(Kit/Lit)	0.068**	0.018**	0.067**	0.026**	0.027**
	[0.009]	[0.006]	[0.009]	[0.006]	[0.006]
Log(market shareit)	0.082**	0.051**	0.073**	0.068**	0.083**
	[0.017]	[0.011]	[0.016]	[0.011]	[0.012]
Dummy(exporter)	0.286*	0.627**	0.292*	0.594**	0.042**
	[0.130]	[0.097]	[0.128]	[0.099]	[0.006]
Dummy(exporter)	0.032	0.077**	0.033	0.072**	
xLog(market shareijt)	[0.017]	[0.012]	[0.017]	[0.012]	
Num. of Plants	47,608	75,658	47,648	75,752	75,752
Num. of Obs	194,394	363,654	194,777	365,723	365,723

Table B.3: Market Share and Export Effects on OLS Markups in Unbalanced Panel

This table shows results of fixed effect regressions in unbalanced panel data for markups estimated by an OLS model such as

$$\ln \hat{\mu}_{ijt} = X_{ijt}\gamma + \iota_i + \iota_j + \iota_t + \xi_{ijt}.$$

Other descriptions remain the same as Table 5.

	(1)	(2)	(3)
Import Penetration	0.001		
	[0.000]		
Log(Kijt/Lijt)	0.077**	0.089**	0.088**
	[0.001]	[0.001]	[0.001]
Log(market shareijt)	-0.084**	-0.068**	-0.056**
	[0.002]	[0.001]	[0.001]
Dummy(exporter)	0.289**	0.278**	-0.024**
	[0.018]	[0.013]	[0.002]
Dummy(exporter)	0.043**	0.039**	
xLog(market shareijt)	[0.002]	[0.001]	
Industry dummy (tj)	yes	yes	yes
R-sq: within	0.13	0.29	0.29
Num. of Plants	61,579	78,834	78,834
Num. of Obs	321,010	567,279	567,279

Table B.4: Market Share and Export Effects on OLS Markups in Dynamic Unbalanced Panel

This table shows results of difference GMMs in dynamic unbalanced panel data (Arellano and Bond (1991)) for markups estimated by an OLS model such as

$$\Delta \ln \hat{\mu}_{it} = \alpha \Delta \ln \hat{\mu}_{it-1} + \Delta X_{it} \gamma + \Delta \xi_{it}.$$

Other descriptions remain the same as Table 6

	(1)	(2)	(3)
Log(markupit-1)	0.176**	0.181**	0.181**
	[0.006]	[0.004]	[0.004]
Import Penetration	0.001**		
	[0.000]		
Log(Kit/Lit)	0.095**	0.054**	0.052**
	[0.009]	[0.005]	[0.005]
Log(market shareit)	0.038**	0.079**	0.091**
	[0.016]	[0.010]	[0.011]
Dummy(exporter)	-0.009	0.292**	0.040**
	[0.155]	[0.104]	[0.005]
Dummy(exporter)	-0.002	0.033*	
xLog(market shareijt)	[0.020]	[0.013]	
Num. of Plants	48,744	76,573	76,573
Num. of Obs	200,519	372,262	372,262