Lund University

Master Thesis

On the Link Between Innovation, Markups and Exporting

Author:
Hannes Jägerstedt

Supervisor:
Joakim Gullstrand
Assistant Supervisor:
Patrik Tingvall

August 29, 2016

Abstract The relationship between innovation, markup and exporting is complex and any empirical analysis is likely to contain both self-selection and simultaneity biases. I use recently developed GMM estimation techniques to recover time-varying markups at the firm level. I then cast these into a simultaneous-equation framework to estimate the influence of R&D investments on markups and export behavior of Swedish manufacturing firms. My analysis shows that firms can significantly improve their domestic market power through innovative investments. Moreover, a one percentage point increase in the markup gives rise to a 0.013 per cent export surge per employee on average. This confirms the hypothesis that, through innovation, firms establish domestic market power before turning to foreign markets.

Contents

1 Introduction

The decision in to engage exporting has played an increasingly important role in the field of international trade during the last three decades. In a study of the U.S. manufacturing industry over the period 1976-1987, Bernard and Jensen (1995) find substantial variation in export participation rates across industries. For example, only 18 per cent of U.S. manufacturing firms ever export. The authors also find that trading firms are larger, more productive, employ a higher fraction of skilled workers and pay higher wages. These stylized facts have been confirmed for a wide range of different countries, suggesting that comparative advantages is perhaps not the foremost driver of trade flows (Bernard et al., 2012). Instead, theorists must turn to models of heterogeneous firms for an explanation.

Why some firms end up supplying foreign markets while others keep servicing only domestic customers is the central question raised by Melitz (2003) in one of the currently most influential papers in the trade literature. Melitz rests his theoretical case on two central notions - that there is a fixed cost associated with exporting and that firms differ in their productivity. Taken together, these two notions lead to a situation where only the most productive firms are able to venture outside the domestic borders while less productive firms refrain from doing so. One can easily think of possible fixed costs that might occur for a firm that wants to export its goods. Such costs may include gaining an understanding of regulatory frameworks, establishing a distribution network and overcoming cultural barriers. Melitz' argument that fixed costs constitute a watershed that separates exporters from non-exporters has indeed been confirmed in numerous papers (Roberts and Tybout 1997 and Clerides et al. 1998).

An alternative explanation of why we observe that exporters are more productive than non-exporters is formulated in the 'learning by exporting' hypothesis. The idea here is that firms become more efficient following past exporting experiences through the gain of knowledge and technical expertise not found domestically. Some evidence has emerged on the hypothesis, suggesting an issue of reverse causality in the relationship between productive and trading firms (De Loecker, 2007). Both the self-selection process of engaging in foreign trade and the 'learning by exporting' hypothesis have paved way for much subsequent research and the matter of which theory most accurately describes reality is still not settled.

Despite its impact, two features of the Melitz (2003) model are particularly unsatisfying. First, firms have no way of actively improving their own market positions. Final goods producers are entirely in the hands of consumers, who demand all available product varieties in equal proportions. Market power is divided equally across all firms and depends only on the elasticity of demand and the number of firms active on the market. The second, and related, feature concerns how productivity is modeled. Firm level productivity is randomly drawn from a probability distribution and so entirely exogenous to the firm. While several papers have made attempts to model heterogeneous market power, these models remain silent on the question of what determines productivity differences (Melitz and Ottaviano 2008 and Bernard et al. 2003). Perhaps the closest explanation so far is found in the endogenous growth literature. Segerstrom and Stepanok (2015) augment a standard quality ladders endogenous growth model with the assumption that it takes time to learn how to export. The quality ladders model constitutes a theoretical backbone for the relationship between innovation, market power and export behavior. The most common measure of market power is the markup ratio, also referred to as the price-cost margin (Hall, 1988). In this paper, I study the impact of R&D investments on markups and export behavior of Swedish manufacturing firms during the period 1996-2006. I find that R&D investments enable firms to charge higher markups on average, which provides evidence that firms can positively affect their market power through innovating. Moreover, I find that higher-markup firms export more on both the extensive and intensive margin. This finding suggests that firms first establish a solid market position domestically and then turn their eyes to foreign markets.

Crepon et al. (1998) provide an econometric framework to evaluate the result of innovation activities on productivity. Productivity is traditionally measured as the residual of some production function, i.e. the variation in sales not attributable to input use. The problem with this approach is that this residual captures not only efficiency gains, but also changes in input costs and markups (De Loecker and Van Biesebroeck, 2016). In this paper, I use a method recently proposed by De Loecker and Warzynski (2012) to isolate the markup effect. De Loecker and Warzynski (2012)'s econometric methodology allows me to estimate time-specific markups at the firm level in order to study firm market power, free of cost and efficiency variations. When combined, these methods constitute a framework well-suited to study the mechanisms of the quality ladders model.

The rest of the paper is structured as follows. Section 2 provides a brief overview of quality ladders model, which constitutes the theoretical foundation for the analysis. A description of the data set and variable definitions is given in section 3. In section 4, the econometric methodology is discussed, followed by the results in section 5. Section 6 concludes.

2 Quality ladders and market power

The quality ladders model provides an explanation both to how productivity differences across firms arise and to how these differences translate into firms' market positions (Grossman and Helpman, 1991). The model emphasizes the role of innovative R&D and horizontal product differentiation. Productivity heterogeneity is achieved through stochastic innovation success. For each product available on the market, there is an associated quality ladder. At any given point in time, the current state-of-the-art version of a product is on a step with all lower-quality versions below it. Further up the ladder are higher quality versions, yet to be discovered. The topmost firm is referred to as 'quality leader', whereas the ones at lower steps are simply 'followers'. Any follower can improve the quality of its own product by investing in R&D. If successful, the firm pushes its product to a step above the incumbent and becomes a new quality leader. Consumers demand only the highest quality version available of each product variety. Thus, the current quality leader temporarily captures the entire market for its product. For as long as it remains on top, a leader acts as a monopolist. Followers act on a perfectly competitive market, earning zero profits.

To capture the market power of firms, it is customary to study their price-cost margins (Hall, 1988). A profit-maximizing firm will operate at an output at which marginal revenue equals marginal cost. Defining the elasticity of demand as $\eta \equiv -\frac{dQ}{dP}\frac{P}{Q}$, we can write this condition as

$$MR = P(1 - \frac{1}{\eta}) = MC \Rightarrow P = \frac{\eta}{\eta - 1} MC \tag{1}$$

It can readily be seen that a perfectly elastic demand ($\eta = \infty$) will result in a perfectly competitive equilibrium where P = MC. This is where followers operate in the quality ladders model. A quality-leading monopolist operates where demand is elastic, but not perfectly elastic such that $1 < \eta < \infty$. The resulting price is some fraction ($\frac{\eta}{\eta-1}$) above marginal cost. This fraction is the definition of the *markup*. For a given level of marginal costs, the markup will depend on the (elasticity of) demand facing the firm. It is easily seen that a higher elasticity will lead to a smaller markup. Firms are thus able to charge higher prices for their products, the less price sensitive their costumers are. It is in this sense that the markup is a measure of a firm's market power. In the quality ladders endogenous growth model, the elasticity of demand depends negatively on the degree of quality improvement. The implication here is that the markup is endogenous to firms' decision to innovate.

In addition to demand factors, the markup also depends on factors more closely related to firm decisions. Fixed costs that need to be covered is expected to push the price up above marginal cost. Scale of production, on the other hand, will be negatively correlated to the markup. This is because a monopolist faces a trade-off between price and quantity produced. It will need to lower the price on the entire volume of sales if it wants to increase the quantity sold.

In a recent contribution, Segerstrom and Stepanok (2015) incorporate a new assumption into the quality ladders framework - that it takes time for firms to learn how to export. This assumption leads to an incentive structure in which two types of R&D activities are possible. Firms can either do quality-improving investments or invest in learning how to export an existing product. Only followers have the incentive to engage in the former activity. The return on an investment aimed at quality-improvement is strictly higher for a follower than a leader. The reason for this is that leaders already make profits, which would only marginally improve following a new quality-improvement whereas followers would get a yield corresponding to the entire profit. Therefore, a leader cannot successfully compete for financial capital with the goal of upgrading the quality of its own product. Instead, an incumbent leader directs its R&D efforts to reach new markets. The model is laid out schematically in figure 1.

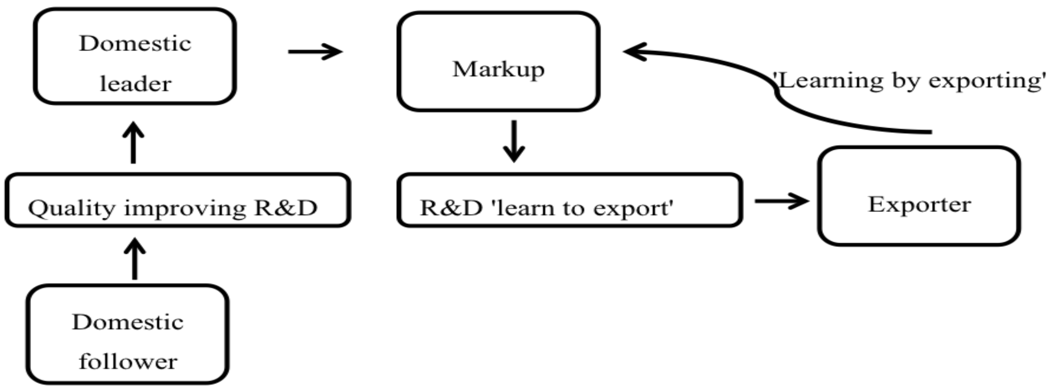


Figure 1: The relationship between innovation, markups and exporting in the quality ladders framework

There are a few reasons why a quality leader might be in a better position to enter export market than non-leader. First, having produced the best-quality version of its product domestically, a leader is in a position to engage in foreign quality-competition in a way that an inferior producer never could. This includes getting access to financial capital. A number of studies confirm that exporters sell higher quality products (Hallak and Sivadasan 2009 and Kugler and Verhoogen 2008). Second, the stream of monopoly profits enables the leader to incur the fixed cost associated with entering foreign markets such as setting up a distributional network. This argument is the one made in most heterogeneous firms models, most notably Melitz (2003). Third, a firm making profits is able to pay higher wages and thus attract workers with certain expertise. Empirical findings suggest that firms invest substantially in hiring workers with past exporting experience prior to reaching new markets (Molina and Muendler, 2013).

According to Segerstrom and Stepanok (2015) we should be able to observe two things when studying the relationship between innovation, markups and exporting: Innovators charge higher markups than non-innovators and, similarly, exporters charge higher markups than non-exporters. The second part is indeed found in recent studies, but these findings are not tied explicitly to the innovation success of firms (De Loecker and Warzynski 2012 and Peters 2013). In fact, to this author's knowledge, there are no previous studies that looks at how innovation affects markup ratios. In what follows, I study these links simultaneously to see whether export participation is related to the decision to innovate via a strengthened market power.

3 Data

I obtain data on innovative activities from the Community Innovation Survey (CIS), a survey that covers three-year periods and is conducted by Statistics Sweden (SCB) every two years. The dataset contains the CIS from 1996, 2000, 2004 and 2006. I consider both innovation input and innovation output. Input measures how much a company invests in R&D and output is what those investments result in. Innovation input is the sum of expenditures on seven research categories: intramural and extramural R&D, purchasing of machinery and other equipment, purchasing of other external knowledge, training, marketing and activities associated with the introduction of an innovation onto the market. I then define R&D intensity as innovation input per employee.

What measure to use for innovation output is somewhat more debated. Crepon et al. (1998) propose to use either the number of patents or innovative sales, i.e the turnover (per employee) that can be attributed to a new successful innovation. For several reasons I will opt for the latter. Innovative sales capture more directly the market success of an innovation as well as its profitability. Also, measures of innovation are based on self-reported data which may be affected by measurement error. Such measurement errors in explanatory variables introduce attenuation bias to parameter estimates. As noted by Heshmati and Lööf (2006), innovative sales are usually less prone to measurement errors than patents. Lastly, data on patents are not available in the CIS dataset whereas innovative sales are.

I define innovators as in Heshmati and Lööf (2006), namely as those firms who have positive innovation investment and also have positive innovative sales. Three indicators capture the objective with R&D investments - improving quality on an existing product, expanding market share or widening the product range. These are measured on a four-step scale from one to three. From these indicators, I construct two other binary variables which show whether quality improvements or increased market share are the most most highly prioritized goals. Quality takes value one if quality improvement is the most important objective and zero otherwise. Similarly, if the main objective is expanding to new markets, Gainshare is one (and zero otherwise).

I collect firm level characteristics and trade data from SCB. Data on exports and imports are not available before 1997, so in effect there are three waves of surveys available. All variables are deflated using CPI with base year 1995. In addition to innovation measures, I include the following variables in the analysis: Firm size is the yearly average number of employees, Human capital is the share of high-skilled workers employed. Capital stock is the value of physical capital used in production. Value added is firm sales less intermediate goods expenditure. Exports is the total value of exported goods and I define an exporter as a firm having positive export volumes in the period studied. Physical capital, total sales and exports are reported in thousands of Swedish krona (SEK). Some firms have multiple export destinations. This I deal with by taking an average over all export destinations to obtain a single measure for each firm and time period.

3.1 Data censoring and outliers

I remove all observations with zero or negative value added from the analysis due to the logarithmic transformation of the variable. To reduce the effect of outliers, I drop firms with markups outside of the 1th and 99th percentiles. Since the CIS is only answered by firms with ten or more employees, I restrict attention to those companies. I further restrict the analysis to firms within the Swedish manufacturing industry. After censoring there are 4,337 firms in the sample, of which 1,782 (41%) are innovative and 2,555 (59%) are not. Due to a large turnover of firms in every new CIS wave, I am unable to utilize the panel structure of the data. Only 270 out of 4,337 firms are present througout the entire sample period (see table 3). The data are therefore considered as pooled cross-sectional data.

3.2 Descriptives

Table 1 shows summary statistics for key variables by innovative status. First note that innovative firms generate higher sales and value added per employee than non-innovators. An intuitively appealing explanation for this observation is that innovative investments lead to higher-quality products or better processes, which enable an improved market position and higher sales for the firm, holding the number of employees constant. At this point, however, it is not possible to distinguish causality from correlation. It may well be that more productive firms with higher per-labor sales are more inclined to - or in a better position to - engage in R&D investments.

Among both types of firms there are those with negative value added. These are most likely newly started firms which are not yet profitable. The fact that innovators are larger than non-innovators is consistent with earlier empirical research, which suggests that firm size is highly correlated to the decision to innovate (Klette and Kortum, 2004). On the average, innovators employ a higher fraction of well educated workers than non-innovators. This is is not very surprising as, in most cases, research requires skilled labor.

Perhaps more interesting is the relationship between innovative status and exporting. At the extensive margin, 91 per cent of innovators export compared to 75 per cent of non-innovators. This is indicative of a relationship between innovation and exporting, either confirming the 'learning by exporting' hypothesis or the notion that innovation is a prerequisite for exporting. Innovating firms trade more also on the intensive margin. Median exports per employee is more than three times as large for innovators than non-innovators, at nearly 15,000 SEK against 4,300 SEK. Interestingly, non-innovating firms have higher mean exports per employee than do innovators. An explanation for this fact is that some firms in the Swedish exporting sector that do not innovate are exporters of raw materials with large trade volumes. Indeed, the non-innovating sample has higher variance and the highest maximum value of per-employee trade volumes.

In table 2, statistics are presented by export status of the firm. Just as found in Bernard and Jensen (1995), exporters are larger, more capital-intensive and are able to generate more value per employee. Contrary to earlier findings, exporters employ a smaller fraction of high-skilled workers on the average. When comparing medians, however, an exporting firm has more highly educated workers than a non-exporter. This is consistent with the explanation that a few, large exporters in less skill- and R&D-intensive industries account for a large proportion of total export volumes.

The R&D propensity and per-employee measures reveal some interesting facts. Exporters are more inlined to do innovative R&D than non-exporters, with propensitites at 61 per cent and 32 per cent, respectively. Most firms that do invest in R&D spend only small amounts, so that median levels of per-employee R&D are close to zero. In the sample at hand, the largest per-employee sums are invested by non-exporting firms. One possible explanation for this is that exporters are larger, necessitating lower per-employee measures. Another explanation is that a few R&D intensive firms are responsible for the bulk of innovation investments among non-exporters. This is supported by the high maximum value and standard deviation within the group. Innovation output, measured as the share of total sales attributable to a new innovation, is more than twice as high among exporters than non-expoters. This suggests that, in the sample at hand, exporters are more capable of capitalizing on their innovative investments.

Lastly, non-exporters seem inclined to do quality-improving research, whereas exporters are more likely to set market expansion as their primary goal. This consistent with incentive structure in the quality ladders model proposed by Segerstrom and Stepanok (2015).

Table 1: Summary Statistics by innovative status.

	Mean	Median	SD	Min	Max
Non-innovators (n = 2555)					
Sales per employee	1,542	1,157	1,292	78.45	20,319
Value added per employee	483.40	434.89	235.33	-425.23	5,587
Employees	118	27	591	10	22,582
Capital per labor	251.38	145.15	349.40	0.25	3,598
High skill share	0.18	0.12	0.20	0.00	1.00
Exporter	0.75	1.00	0.43	0.00	1.00
Exports per employee	25.57	4.30	108.37	0.00	4,392
R&D per employee	-	-	-	-	-
Innovative sales	-	-	-	-	-
Innovators (n = 1782)					
Sales per employee	1,787	1,427	1,732	156.59	33,543
Value added per employee	551.60	499.74	250.73	-476.39	3,272
Employees	269	53	953	10	18,899
Capital per labor	270.74	168.87	376.03	1.21	6,573
High skill share	0.26	0.18	0.22	0.00	1.00
Exporter	0.91	1.00	0.29	0.00	1.00
Exports per employee	24.94	14.72	51.28	0.00	1,340
R&D per employee	36.82	0.17	150.03	0.00	4,930
Innovative sales	0.23	0.15	0.23	0.01	1.58
Total (N = 4337)					
Sales per employee	1,643	1,271	1,494	78.45	33,543
Value added per employee	511.42	459.92	244.06	-476.39	5,587
Employees	180	33	765	10	22,582
Capital per labor	259.33	153.33	360.66	0.25	6,573
High skill share	0.22	0.14	0.21	0.00	1.00
Exporter	0.82	1.00	0.39	0.00	1.00
Exports per employee	25.31	8.94	89.43	0.00	4,392
R&D per employee	20.79	0.01	197.57	0.00	11,234
Innovative sales	0.10	0.00	0.19	0.00	1.58

Quantities are expressed in SEK 1,000.

Table 2: Summary Statistics by export status.

	Mean	Median	SD	Min	Max
Non-exporters (n = 794)					
Sales per employee	1,075	850.47	989.04	135.83	17,539
Value added per employee	455.92	419.33	255.89	143.64	5,587
Employees	34	18	72	10	972
Capital per labor	149.52	59.70	292.44	0.55	3,062
High skill share	0.26	0.11	0.28	0.00	1.00
Gainshare	0.04	0.00	0.20	0.00	1.00
Quality	0.29	0.00	0.46	0.00	1.00
R&D propensity	0.32	0.00	0.47	0.00	1.00
R&D per employee	21.15	0.00	400.09	0.00	11,234
Innovative sales	0.05	0.00	0.13	0.00	1.00
Exporters (n = 3543)					
Sales per employee	1,770	1,391	1,557	78.45	33,543
Value added per employee	523.86	471.77	239.61	-476.39	3,438
Employees	213	42	842	10	22,582
Capital per labor	283.94	179.58	369.84	0.25	6,573
High skill share	0.21	0.14	0.19	0.00	1.00
Gainshare	0.06	0.00	0.23	0.00	1.00
Quality	0.19	0.00	0.39	0.00	1.00
R&D propensity	0.61	1.00	0.49	0.00	1.00
R&D per employee	20.71	0.02	109.29	0.00	4,930
Innovative sales	0.11	0.00	0.20	0.00	1.58
Total (N = 4337)					
Sales per employee	1,642	1,271	1,494	78.45	33,543
Value added per employee	511.42	459.92	244.06	-476.39	5,586.66
Employees	180	33	765	10	22,582
Capital per labor	259.33	153.33	360.66	0.25	6,573
High skill share	0.22	0.14	0.21	0.00	1.00
Gainshare	0.05	0.00	0.23	0.00	1.00
Quality	0.20	0.00	0.40	0.00	1.00
R&D propensity	0.56	1.00	0.50	0.00	1.00
R&D per employee	20.79	0.01	197.57	0.00	11,234
Innovative sales	0.10	0.00	0.19	0.00	1.58

Quantities are expressed in SEK 1,000.

4 Empirical Method

The empirical strategy of this paper can be divided into three parts. First, I obtain a reliable measure of markup at the firm level. Second, I use that measure to estimate the effect of engaging in innovative R&D on markups. Third, I examine how such an effect translates into changes in export behavior of Swedish firms. While the second and third steps could in principle be addressed separately, it will be shown that it is far superior to consider them jointly using well-established econometric methods as suggested by Crepon et al. (1998).

4.1 Econometric approach

Given the objective to examine the effect of innovative R&D on markups and export status, the problem at hand can be formulated as

$$m = x_1'\beta_1 + \beta_i i + \epsilon_1 \tag{2}$$

$$e = x_2'\beta_2 + \beta_m m + \epsilon_2 \tag{3}$$

where m is markup, i is innovation output and e is a dummy indicating export status. x_1 and x_2 are vectors of control variables.

Equation-by-equation estimation using OLS gives rise to two major econometric concerns. First, a key issue with studies involving innovation measures is that the sample of innovative firms cannot generally be considered representative for the entire population. This is because not all firms engage in innovative activities. If the decision to do so is influenced by certain factors in a non-random fashion, it is hard to argue that the sample is drawn randomly from a population. In effect, inference based on the sample will be biased. To account for such self-selection bias, Heckman (1979) proposes a simple generalized Tobit model to model firms' innovative behavior. It consists of two equations - one governing the decision to innovate and another governing the extent of the innovative investments. The error terms in the two equations are correlated because the last equation is conditioned on having positive R&D expenditures. The Heckman selection estimator elegantly models this correlation and reduces the impact of selection bias[1]. Second, the assumption of homogeneity of regressors is violated. In equation 2 there is possible reverse causality running from markup to innovation output in the first equation. Given

[1]For a technical exposition of the generalized Tobit, see Heckman (1979) or Verbeek (2012).

that innovators are predominately non-leading firms, having lower market power, the resulting bias is likely to be negative. On the other hand, it is possible that the market power of a firm affects the market success rate of its new innovation positively, in which case the bias would be positive. A similar problem is encountered in the export equation 3. According to the theory by Segerstrom and Stepanok (2015) exporters should, on average, have higher markups due to their dominant market status. But we also expect a higher markup resulting from a successful innovation to enable export participation. Since both effects are positive, this bias should be positive.

To alleviate biases of the type presented here, Crepon et al. (1998) introduced an approach using simultaneous equations. The model is generally referred to as the Crepon-Duguet-Mairessec (CDM) model and it has been widely used for the last two decades. In their original approach, the authors set up a system of four equations: Two governing the selection into and extent of innovative activities (the Heckman selection equations), one describing how such activities translate into innovation output and one examining how innovation output affects productivity. In essence, the first two are used to address the selectivity issue and the second part models the simultaneity. In what follows, a very similar approach is used but with two notable changes: (i) Instead of the productivity response, I am concerned here with the response of the markup and (ii) a fifth equation on export status is added. The entire system of equations is as follows,

$$g = x_0'\beta_0 + \epsilon_0 \tag{4}$$

$$k = x_1'\beta_1 + \epsilon_1 \tag{5}$$

$$i = x_2'\beta_2 + \beta_k k^* + \beta_{imr} imr + \epsilon_2 \tag{6}$$

$$m = x_3'\beta_3 + \beta_i i + \epsilon_3 \tag{7}$$

$$e = x_4'\beta_4 + \beta_m m + \epsilon_4 \tag{8}$$

where the first equation governs the selection into innovation activities and k measures the intensity of such activities conditional on being selected. Equation 6 is innovation output (i) as a function of R&D (k). Equation 7 is the response of the markup (m) to innovation output and finally, equation 8 shows the export status (e) as a function of the markup.

4.2 Estimation and variable selection

In the original CDM-model the reduced form version of the entire system of equations was modeled simultaneously using Asymptotic Least Squares and then the structural parameters was recovered. This method relies on the assumption that the error terms in all equations are correlated. In this paper, this assumption is relaxed somewhat: The error terms ϵ_0 and ϵ_1 are correlated to each other but not to the remaining error terms. On the other hand, ϵ_2, ϵ_3 and ϵ_4 are jointly correlated. Estimation will be performed in two steps. First, the Heckman selection estimator is used to model the selection process into R&D activities (equations 4 and 5). For efficiency, the two equations are estimated jointly using full information maximum likelihood. The remaining equations (6 - 8) are then estimated jointly using three-stage least squares (3SLS). This is a system estimator and an extension of the 2SLS instrumental variable estimator. Basically, it adds an estimation step step to compute generalized least squares (GLS) estimates for efficiency. In each of the equations, the endogenous variables k, i and m are instrumented using their predicted values. To tie the system together, the inverted Mill's Ratio is included in the first of the second set of equations. It is defined as the ratio of the probability density and cumulative density functions of a standard normal distribution.[2] As argued by Heshmati and Lööf (2006), this is equivalent to estimating the entire system in one step, but with a small efficiency loss. The ease of computation and interpretation of the results favors the 3SLS method. Due to the use of an estimated (predicted) innovation input measure, standard errors for the 3SLS estimator are bootstrapped.

The vector x_0 contains employment, physical capital, capital-labor ratio, and percentage share of high-skilled workers. Employment, capital and capital-labor ratios are expressed in logarithms. x_1 is identical to x_0 with one exception: The number of employees are not included. This is because empirical research suggests that while firm size is strongly correlated to the decision to engage in R&D, it does not affect the intensive margin of investments (Klette and Kortum, 2004). This difference constitutes the exclusion restriction required when using a Heckman Selection estimator. Turning to equation 6, the vector x_2 includes firm size, physical capital, capital-labor ratio all expressed in logarithms; and the share of high-skilled labor. Imr is the inverse of the Mill's ratio from the probit estimation of equation 4.

In the markup equation covariates x_3 is identical to x_2. In the last equation, the log of export intensity is explained by the markup and covariates x_4, which are firm size,

[2] For a discussion of the inverse Mill's ratio, see for instance Verbeek (2012).

physical capital, capital-labor ratio all expressed in logarithms. Two-digit industry and time dummies are included in all equations.

5 Results

This section shows the results from estimation of the models in the previous section. To investigate the properties of potential simultaneity bias, I present both single-equation and simultaneous equation estimates.

5.1 Estimated Markups in Swedish manufacturing

I use two different measures of the markup. The main measure is estimated using a method proposed by De Loecker and Warzynski (2012), discussed at length in Appendix A. The method essentially boils down to estimating a production function using GMM. The other measure is called the Lerner index or price-cost margin. It requires no econometric approach, but can be calculated directly from observables using the formula (value added - payroll)/(value added + material costs).

Figure 2 compares markups obtained using the De Loecker & Warzynski method to those implied by the calculated price-cost margins. Medians values are 1.19 and 1.11, which are interpreted as prices of 19 and 11 per cent above marginal cost, respectively. It is comforting that the distributions are centered not too far apart. The higher standard deviation of the De Loecker & Warzynski measure is somewhat strange, since the estimation routine should take into account variations present in the data such as unobserved productivity differences. Strange as it may seem, it is potentially a benefit to the analysis, since increased variation improves efficiency of the estimates. Given the low degree of correlation between the measures (0.17) it is plausible that the measures yield very different results. To this end, in the end of this chapter I examine the use of the alternative markup. In what follows, however, the main measure is used exclusively.

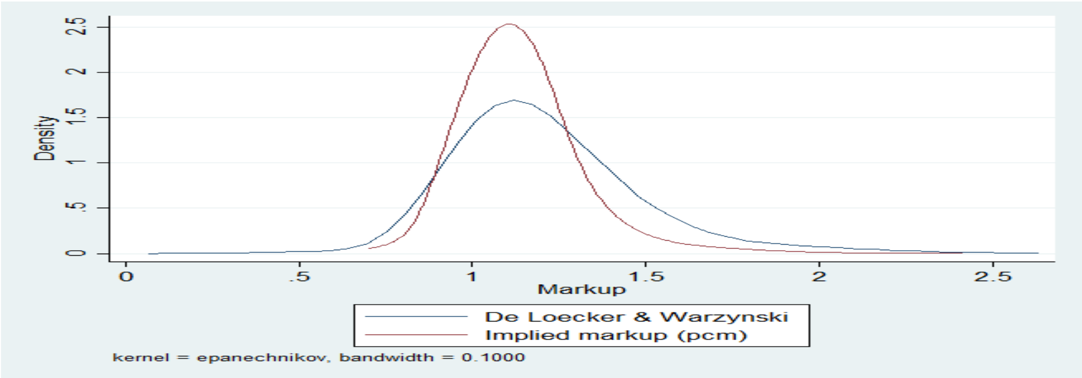
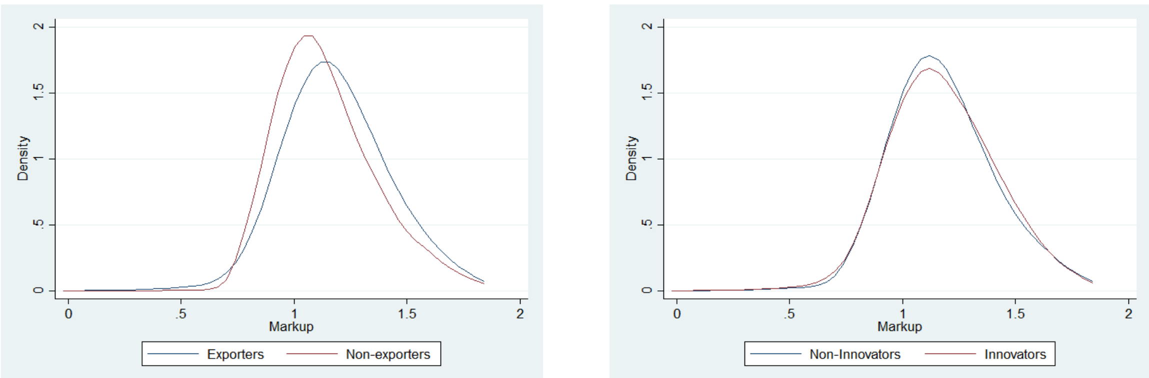


Figure 2: Kernel distributions of markups computed using DLW and PCM

Panel (a) of figure 3 compares kernel distributions of markups for exporters and non-exporters. The distribution for exporters is shifted slightly to the right. Median markups are 1.10 for non-exporters and 1.17 for exporters. This confirms the finding of De Loecker and Warzynski (2012), that exporters have higher markups than non-exporters.

A comparison by innovative status is made in panel (b) if figure 3. Oddly, there seems to be almost perfect overlap between the two distributions. On innovative status alone, no differences in markups emerge between the groups in the sample. Bear in mind, however, that this is a purely descriptive analysis which may suffer from the endogeneity issues described earlier.



(a) Exporters vs non-exporters

(b) innovators vs non-innovators

Figure 3: Kernel distribution of markups by innovative and export status

Table 4 shows mean and median values of estimated markups per industry. There is some variation across different industries. For instance, tobacco products have average markups of around 37 per cent, compared to 14 per cent on leather products. This illustrates the relationship between markups and the price elasticity of demand quite nicely.

5.2 Single equation results

Table 5 reports the OLS-estimates of the set of equations 7 - 8. Here, I estimate each equation separately using OLS and *not* using the instrumental variable (3SLS) approach. In other words, I account for neither selectivity nor potential simultaneity bias at this stage. I do this to get a feel for the sign and magnitude of such bias. I measure both the markup, share of innovative sales and the high skill share in levels so changes are to be interpreted as percentage points.

The first column holds the markup equation, in which the parameter of main interest is innovative sales. An increased share of innovative sales by one percentage point will, ceteris paribus, increase the markup by 0.07 percentage points. This suggests that following a successful innovation, firms can on average charge higher prices relative to marginal costs. This is in line with the predictions of the quality ladders model. The number of employees is negatively related to the markup. This is also reasonable if thought of as a measure of size (or scale) of production. According to the theoretical price setting behavior of a monopolist, a larger scale of production squeezes the markup. Similarly, the capital stock can be taken as a proxy for fixed costs of production, covered by setting a price above marginal cost. It is therefore not surprising to see it positively correlated to the markup. It is somewhat perplexing, however, to see a negative impact of human capital on the markup. It seems reasonable that firms employing well educated workers should be more cutting edge and thus able to diversify their products. On the other hand, this model is controlling for innovation success. So the parameter on high skill labor should be interpreted holding innovative sales fixed. It is via increased creativity and innovation that high skilled labor 'pull their weight'. Hiring more high skilled labor drives up marginal costs and when innovation output is held fixed, it does not allow for a corresponding price increase. As a consequence, the markup is negatively affected.

In column (2) I show the export equation. A one percentage point increase in the markup gives rise to a 0.004 per cent increase in the per-employee exports on average. While of small magnifude, this response is in line with what we expected. As for the control variables, the parameters all behave as expected: Exports increase in firm size, capital and skill intensity. While it provides some first suggestions, these estimates are not likely to be consistent due to the reverse causality discussed above. The direction of the bias will perhaps emerge later on, when more (theoretically) reliable estimates are displayed. For now, it is sufficient to keep in mind the shortcomings of the ordinary least squares method used to produce these results.

Table 5: OLS estimation of markup and export equations

VARIABLES	(1) Markup	(2) Export intensity
Innovative sales (%)	0.0692**	
	(0.0291)	
log of Employees	-0.0615***	0.549***
	(0.0102)	(0.0644)
log of Capital Stock	0.0661***	0.298***
	(0.00682)	(0.0482)
High skill share	-0.313***	0.552
	(0.0370)	(0.386)
Markup		0.404**
		(0.174)
Observations	1778	1612
Adjusted R-squared	0.266	0.533
Time FE	Yes	Yes
Industry FE	Yes	Yes

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

5.3 Simultaneous estimation results

Table 6 shows the results from estimation of the Heckman selection procedure. The first column corresponds to equation 4 and shows the propensity of all firms to engage in R&D activities. The estimated parameters of firm size, capital intensity and human capital are broadly in range with Heshmati and Lööf (2006). Doubling the amount of employees increases the probability of doing R&D with roughly 13.5 per cent. Sweden-based multinationals seem more inclined to invest in research than the other two categories. This could be due to a higher pressure from international competition among these firms relative to those active only in the domestic market. Since the parameters on the other ownership indicators are insignificant, we cannot say much about these other than that they seem to have little impact on the decision to do research.

Turning to equation 5 in column (2), it is evident that R&D intensity increases significantly with capital intensity as well as skill intensity. For instance, a one percentage point increase in the share of high skilled workers will increase R&D expenditure per labor by roughly seven per cent. The only ownership indicator that affects the intensive margin of investments is the Swedish local, which is negatively correlated to the size of investments.

Table 6: Heckman Estimation results

VARIABLES	(1) R&D propensity	(2) R&D intensity
log of Employees	0.135***	
	(0.0143)	
log of Capital/L	0.111***	0.742***
	(0.0169)	(0.0885)
High skill share	1.384***	7.429***
	(0.124)	(0.590)
Swemne	0.157*	0.640
	(0.0944)	(0.406)
Swelocal	-0.0390	-0.901**
	(0.101)	(0.449)
Fof	-0.0290	-0.272
	(0.100)	(0.448)
Observations	4,337	2,418
Loglikelihood	-9135	-9135
Wald	1021	1021
Prob > chi2	0.00	0.00

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

$\rho = 0.967$ and $\sigma = 5.257$

The 3SLS estimates are reported in table 7.[3] In column (1) I display the innovation output equation (6). I use the predicted values from the Heckman selection estimation as

[3]The R^2 value is negative in the markup equation. This is because I use actual values (and not instruments) to calculate the residuals. It poses no threat to the interpretation of the estimated parameters. See Sribney et al. (1999).

a proxy for innovation input. Not surprisingly, innovation input has a positive impact on the share of innovative sales. A one percent increase in per-labor R&D expenditure gives rise to a three percentage points increase in the share of innovative sales. So innovation input does have an impact on innovative output. It is also worth noting that human capital strongly and positively affects the innovation output. It seems that innovative success is less dependent on firm- and capital stock size.

Column (2) of table 7 corresponds to equation 7. The main variable of interest is innovative sales, which is positive and highly significant. A one percentage point increase in the share of innovative sales is associated with a markup increase of roughly two percentage points. This is of a substantially higher magnitude than the single-equation OLS estimate. This may be due to a downwards bias in the OLS-estimate that results from innovation output being endogenous in the markup equation. Firm size is not a significant, even though it does have the expected sign. The remaining control variables - share of high skilled workers and physical- and human capital - remain quite stable with markups rising with fixed costs and falling with more educated workers.

The last column shows equation 8. Both firm size and capital stock have roughly the same size and sign as in the equation-by-equation estimation. This is not too surprising, since the 3SLS technique mainly addresses the consistency (and efficiency) of the endogenous variables. The coefficient on markup in column (3) is around three times higher compared to the OLS case. Contrary to the predictions laid out in section 4, this suggests that the OLS estimate is downwards biased. It is possible that higher competition in export markets is responsible for this. When firms enter export markets, they are exposed to fierce competition that forces markups down towards the competitive level of price equal to marginal cost. In a simple regression of exports on markups, this reverse causality will show up as a negative - or less strong positive - relationship between markups and export volumes. In the 3SLS method, this reverse link is ruled out and markups are seen to positively affect export volumes. This suggests that, as the quality ladders model predicts, exporting firms have higher markups prior to exporting than their domestic competitors. Whether this is due to fixed costs associated with entering foreign markets or if the markup signals quality superiority is less clear cut. A combination of the two seems likely, but this issue is beyond the scope if this paper. The highly significant coefficient on high skilled labor confirms the findings of Molina and Muendler (2013), that firms invest in experienced labor prior to entering export markets.

To summarize the findings in table 7, the key parameters in the system are significant and of expected signs. When comparing results, it seems that the equation-by-equation ordinary least squares produces inconsistent estimates that are downwards biased. When taking both selectivity and simultaneity into account, the implied causal channel under consideration emerges more clearly.

Table 7: Three stage least squares (3SLS) estimates of innovation, markup and exports. Innovation sample.

	(1)	(2)	(3)
VARIABLES	Innovative sales (%)	Markup	Export
Predicted R&D	0.0318*		
	(0.0174)		
Innovative sales (%)		1.943***	
		(0.500)	
Markup			1.296**
			(0.574)
log of Employees	0.0161	-0.0315	0.602***
	(0.0190)	(0.0309)	(0.111)
log of Capital Stock	0.0239*	0.0513***	0.238***
	(0.0125)	(0.0187)	(0.0859)
High skill share	0.456***	-0.749***	0.842**
	(0.0858)	(0.190)	(0.413)
Inverse Mill's Ratio	0.664***		
	(0.199)		
Observations	1,616	1,616	1,616
R-squared	0.105	-1.794	0.537
χ^2	191.3	367.7	1900
Prob > χ^2	0.00	0.00	0.00

Time and industry dummies and constant included but not shown

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

5.4 Robustness Checks

Table 8 shows correlations among key variables used in the analysis. The correlation between labor and capital is high (0.84) and may cause collinearity problems. To see whether collinearity affects the analysis, I run an alternative 3SLS specification using total sales as a measure of size and per-employee capital instead of total capital stock. As is seen in table 9, the main results are still valid under the alternative specification. A successful innovation leads to an increased markup of a magnitude somewhat larger than in the main analysis, whereas he effect on export volumes is unchanged. An interesting thing to note is that my new measure for firm size - log of sales - now has the opposite sign in the markup equation. A larger value of total production now leads to a higher markup. An explanation for this is that total sales not only measures the volume but also price effects. All other things being equal, a price increase leads to a higher markup. To the extent that total sales reflects prices, it should be positively correlated to the markup measure.

In the exports equation, all control variables have expected signs - trade volumes increase in firm size, high skill share and capital per labor ratios.

Given the low correlation between the two different markup measures obtained, I want to examine whether my results are sensitive to which measure is being used. Table 10 shows the estimates from the main 3SLS analysis but with the price-cost margin used instead of the De Loecker and Warzynski markup. In the markup equation, innovation output still has a positive sign, albeit no longer significant. The lack of significance may partly be attributed to the lower variance in the price-cost margin. The magnitude is slashed to about a fifth of that in the main analysis. This could reflect the fact that price-cost margin computed this way does not take productivity differences at the firm level into account. A large value of total sales in relation to input costs can be either due to a high productivity or a high markup (or any combination of the two). If an innovative investment increases productivity, an innovative firm can experience increased sales volumes at a constant price over marginal cost. The result is that the positive correlation between innovation and markup disappears if we do not control for productivity changes. Note also that the number of employees has a significantly negative effect on the markup. This may be a computational effect. Recall that the price-cost margin was calculated as (value added - payroll/(value added + material costs). So a large payroll in relation to value added should have a negative impact on the calculated markup.

Turning to the exports equation, the magnitude of the estimated markup coefficient is over three times as large as in the main analysis. Unobserved productivity again offers a plausible explanation. All things being equal, a firm with a high markup is likely to be a productive firm. If I do not control for productivity, I will capture the effect of this productivity on export volumes. This effect is well documented to be strongly positive in previous research. As a consequence, the estimate will be upwards biased.

The use of the price-cost margin as a robustness check sheds some light on the benefits of the estimation method proposed by De Loecker and Warzynski (2012). The presence of unobserved productivity is an important issue, the direct link between current number of employees and markup is another.

6 Summary and conclusions

The purpose of this paper is to test some of the implications of a recently proposed theoretical model of international trade, namely an extension of the quality ladders model. The focus lies on the causal relationship between R&D, market power and export behavior of firms. The methods are applied to a pooled sample of three waves of the Swedish Community Innovation Survey coupled with firm-level registry data.

A number if key findings are highlighted. First, successful R&D efforts lead to an increase in the wedge between price and marginal cost - i.e. the markup. This implies that innovation success provides firms with at least temporary market power, just as predicted by the quality ladders model. Second, markups positively affect the intensive margin of exports. This may be taken as evidence that the gained domestic market power from innovative R&D allows the firm to compete in export markets as well. This is one of the main implications of the theoretical model.

I discuss two possible explanations for why firms require higher markups prior to exporting. One hypothesis concerns capital accumulation to cover fixed costs of exporting. An alternative explanation is that the markup is a measure of market power, reflecting that a firm is in good enough shape to compete internationally. Which one holds true has important bearings for policy purposes. If firms need to accumulate equity capital to finance their expansions into foreign markets, this may be a sign of a capital market failure. A market imperfection leading to inefficient price setting, this could warrant government intervention. On the other hand, if the markup is merely a reflection of firm's successful quality competition, no such measures are needed or even desired. Examining which explanation is more likely is an important step for further research to take.

A comparison between estimation techniques shows ordinary least squares estimates to be downwards biased. This is likely to depend on reverse causality. Because follower firms need to invest in order to gain market power, the positive influence of R&D investments on markups is partly offset. Similarly, the increased competition on export markets pushes prices down towards marginal costs, disguising a positive export effect of markup pricing.

One conclusion is that OLS is not suitable for estimation of multi-equation relationships when simultaneity is likely to be present. The results are shown to be robust when using an alternative measure of the markup and also when taking collinearity between regressors into account.

To further shed some light on the quality ladders model, a next step should be to test its dynamics. This requires, first and foremost, a richer panel data set. It would also be interesting to test some of the other implications of the model. These include welfare effects and within-industry reallocations stemming from trade liberalization. Ultimately, the goal should be to put the quality dynamics of the model in an international context to allow for quality competition not only domestically, but globally.

References

D. Ackerberg, K. Caves, and G. Frazer. Structural identification of production functions. *Working Paper*, 2006.

A. Bernard and B. Jensen. Exporters, jobs, and wages in the u.s. manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Special Issue Microeconomics*, pages 67–119, 1995.

A. Bernard, B. Jensen, S. Redding, and P. Schott. The empirics of firm heterogeneity and international trade. *Annual Economic Review*, 4:282–313, 2012.

A. B. Bernard, J. Eaton, J. B. Jensen, and S. Kortum. Plants and productivity in international trade. *The American Economic Review*, 93(4):1268–1290, 2003.

S. Clerides, S. Lach, and J. Tybout. Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco. *The Quarterly Journal of Economics*, 113(3):903–947, 1998.

B. Crepon, E. Duguet, and J. Mairessec. Research, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technology*, 7(2):115–158, 1998. doi: 10.1080/10438599800000031.

J. De Loecker. Do exports generate higher productivity? evidence from slovenia. *Journal of International Economics*, 73(1):69–98, 2007.

J. De Loecker and J. Van Biesebroeck. Effect of international competition on firm productivity and market power. Technical report, National Bureau of Economic Research, 2016.

J. De Loecker and F. Warzynski. Markups and firm-level export status. *American Economic Review*, 102(6):2437–2471, 2012.

A. Gandhi, S. Navarro, and D. Rivers. On the identification of production functions. *Working Paper*, 2013.

G. M. Grossman and E. Helpman. Quality ladders in the theory of growth. *Review of Economic Studies*, 58:43–61, 1991.

R. Hall. The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*, 96(5):921–947, 1988.

J. C. Hallak and J. Sivadasan. Firms' exporting behavior under quality constraints. Technical report, National Bureau of Economic Research, 2009.

J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47:153–161, 1979.

A. Heshmati and H. Lööf. On the relationshhip between innovation and performance: a sensitivity analysis. *Economics of Innovation and New Technology*, 15(4-5):317–344, 2006.

J. Klette and S. Kortum. Innovating firms and aggregate innovation. *Journal of political economy*, 115(5):986–1018, 2004.

M. Kugler and E. Verhoogen. The quality-complementarity hypothesis: Theory and evidence from colombia. Technical report, National Bureau of Economic Research, 2008.

M. J. Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725, 2003.

M. J. Melitz and G. I. Ottaviano. Market size, trade, and productivity. *The review of economic studies*, 75(1):295–316, 2008.

D. Molina and M.-A. Muendler. Preparing to export. Technical report, National Bureau of Economic Research, 2013.

M. Peters. Heterogeneous mark-ups, growth and endogenous misallocation. 2013.

M. Roberts and J. Tybout. The decision to export in colombia: an empirical model of entry with sunk costs. *American Economic Review*, 87:545–564, 1997.

P. Segerstrom and I. Stepanok. Learning how to export. *Working Paper*, 2015.

W. Sribney, V. Wiggins, and D. Drukker. Negative and missing r-squared for 2sls/iv, 1999. URL `http://www.stata.com/support/faqs/statistics/two-stage-least-squares/`.

M. Verbeek. *A Guide to Modern Econometrics*. John Wiley and Sons, The Atrium, Southern Gate, Chichester, West Sussex, United Kingdom, 2012.

7 Appendix A: Estimating the markup

This section will follow closely the methodology proposed in De Loecker and Warzynski (2012) (henceforth DLW), which relies on the insight that only when P = MC does the output elasticity of a variable factor of production (such as labor) equal that factor's expenditure share in total revenue. In effect, any difference between these should reflect a wedge between price and marginal cost and thus constitute the relevant markup. To expand on this a bit, consider a production function of the form [4]

$$Q_{it} = Q_{it}(X_{it}^{1}, ..., X_{it}^{v}, K_{it}, \omega_{it}), \tag{9}$$

where X's are variable inputs such as labor and intermediate inputs and K_{it} is firm i's capital stock at time t, combined with the assumption that all firms are cost minimizing, yield the following useful expression:

$$\frac{\partial Q_{it}(.)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{Q_{it}} = \frac{1}{\lambda_{it}} \frac{P_{it}^{X^v} X_{it}^{v}}{Q_{it}} \tag{10}$$

The left-hand side of this expression is the output elasticity of variable input v. λ_{it} is the marginal cost of production and $\frac{P_{it}^{X^v} X_{it}^{v}}{Q_{it}}$ is the nominal cost of that variable input. Defining the markup according to $\mu_{it} \equiv \frac{P_{it}}{\lambda_{it}}$ and the output elasticity of variable input v as $\theta_{it}^{V} \equiv \frac{\partial Q_{it}(.)}{\partial X_{it}^{v}} \frac{X_{it}^{v}}{Q_{it}}$ allows us to rewrite equation 10 as

$$\theta_{it}^{X} = \mu_{it} \frac{P_{it}^{X^v} X_{it}^{v}}{P_{it} Q_{it}} \tag{11}$$

which says that the output elasticity of variable input v is the markup multiplied by that input's cost share in total sales. Defining α_{it}^{X} as the cost share of input factor v in total sales, we finally arrive at

$$\mu_{it} = \theta_{it}^{X} (\alpha_{it}^{X})^{-1} \tag{12}$$

which allows for the computation of firm (and time) specific markups given (i) data on expenditure shares of variable inputs and (ii) estimated output elasticities of those inputs. When measuring output as value added, intermediate inputs are already subtracted from gross output. Consequently, the vector X includes only labor (L) in the following specification.

[4]This production function is not restricted to be of a particular form, such as a Cobb-Douglas. The only requirement is that it is continuous and twice differentiable w.r.t the arguments included

Given the algebraic construct above, the main issue will be that of estimating the output elasticities of certain variable inputs. I will use labor to define the markup. Effectively, this boils down to estimating the coefficients in a production function. There is an extensive body of literature on the identification of production functions (see Gandhi et al. (2013) for a good exposition). Looking at equation 13, it is likely to contain unobserved factors that can potentially affect both dependent and explanatory variables. In econometrics lingo, this is the concern of omitted variable bias (OVB). Most notably, unobserved productivity is expected to be correlated both to output and the choice of variable inputs. This problem has to be addressed properly in order to obtain consistent estimates of the parameters and, indeed, numerous ways forward have been proposed. DLW use a method suggested by Ackerberg et al. (2006) (henceforth ACF)

Turning then to the estimation of output elasticities, two additional assumptions are imposed on the production function: Productivity is Hicks-neutral[5] and the production function parameters ($\beta's$) are identical across firms. This effectively means that all firms operate using a common technology. The function to be estimated is

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{lk} l_{it} k_{it} + \omega_{it} + \epsilon_{it} \tag{13}$$

where y_{it}, l_{it} and k_{it} denote logs of value-added output, labor and capital. ω_{it} is a firm-specific productivity parameter and ϵ_{it} an error term, containing unanticipated productivity shocks and potential measurement error. The unobserved productivity needs to be modeled somehow. Levinsohn & Petrin (2003) propose to use a material demand function to proxy for productivity as

$$m_{it} = f_t(k_{it}, \omega_{it}, \mathbf{z}_{it}) \tag{14}$$

where $\mathbf{z}_{it}$ is a vector of variables potentially affecting material demand. This can be used to invert out productivity using

$$\omega_{it} = f_t(m_{it}, k_{it}, \mathbf{x}_{it}) \tag{15}$$

In practice, this is modeled non-parametrically by including a polynomial in m, k and **x** and their cross products in the estimation of equation 13. Once that is estimated, the

[5]Hicks-Neutral productivity means that the composition of input factors are unchanged by changes in productivity

productivity parameter can be 'backed out' as

$$\omega_{it}(\beta) = \hat{y}_{it} - (\beta_l l_{it} + \beta_k k_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{lk} l_{it} k_{it}) \tag{16}$$

In a next step, we recognize that productivity follows an AR(1)-process according to

$$\omega_{it}(\beta) = \omega_{it-1}(\beta) + \zeta_{it}(\beta) \tag{17}$$

Here, $\zeta_{it}(\beta)$, is interpreted as an 'innovation to productivity' parameter that is assumed to be uncorrelated to current choice of capital input but not to current choice of labor. That is, we expect that $E(k_{it}\zeta_{it}) = 0$ and $E(l_{it}\zeta_{it}) \neq 0$. We rely on the insight that lagged labor must be decided without regard to future productivity innovation such that $E(l_{it-1}\zeta_{it}) = 0$. Thus, current labor can be instrumented with its lag. It is now possible to form the following moment conditions that can be used for GMM-estimation, in order to recover the production function β-parameters,

$$E[\zeta_{it}(\beta)\mathbf{z}] = 0, \mathbf{z} = (l_{it-1}, k_{it}, l_{it-1}^2, k_{it}^2, l_{it-1}k_{it}) \tag{18}$$

The output elasticity of labor can then be computed as the partial derivative[6] $\frac{\partial \hat{y}_{it}(.)}{\partial l_{it}}$.

Expenditure shares of labor can be readily computed from the data. However, observed data on output is likely to contain measurement error and unanticipated shocks to production. This means that the optimal input decisions of profit maximizing firms may differ in practice from those implied by the level of output we observe. Since the way the markups are computed is derived from an optimizing behavior, this could bias our measure of the markup. Thus, following DLW, the expenditure shares are corrected using the error term from equation 13. This should 'purge' expenditure shares from variation that is not attributable to variables used to model material demand, i.e. arguments included in equation 14. Such variation may be due to input prices, firm productivity, elasticity of demand and income levels. Finally, firm and time specific markups are computed using equation 12 with corrected expenditure shares and estimated output elasticities being used.

I estimate firm-level markups for each industry separately to allow for heterogeneity across industries. Starting values for the GMM-estimations are provided using OLS.

[6]Since all variables are in logs, estimated parameters are interpreted as elasticities.

Table 3: Firm coverage over CIS periods.

	Number of consecutive years			
CIS	1	2	3	Total
2000	1133	0	0	1133
2002	1172	446	0	1618
2006	707	609	270	1586
Total	3012	1055	270	4337

Table 4: Estimated markups by sector

Sector	Mean	Median	Observations
15 Food products and beverages	1.130595	1.094123	1826
16 Tobacco products	1.374803	1.374803	2
17 Textiles	1.011822	0.9949992	370
18 Wearing apparel	1.265209	1.252718	109
19 Leather products	1.144163	1.098462	67
20 Wood and cork products	1.189008	1.168271	1875
21 Paper an paper products	1.269238	1.253908	515
22 Publishing, printing of media	1.246094	1.231339	2122
23 Coke, refined petroleum and nuclear fuel	1.170102	1.401332	14
24 Chemicals	1.255025	1.227188	631
25 Rubber and Plastic	1.162766	1.151823	1270
26 Non-metallic mineral products	1.086497	1.059168	538
27 Basic metal	1.131776	1.101349	455
28 Fabricated metal products	1.167998	1.139173	5080
29 Machinery and equipment	1.250872	1.226805	3343
30 Office machinery, computers	1.081253	1.084678	122
31 Electrical machinery, communications	1.246791	1.229571	860
32 Radio & TV equipment	1.10566	1.057112	341
33 Medical, precision and optical instruments	1.312786	1.292222	738
34 Motor vehicles, trailers	1.175214	1.15521	848
35 Other transport equipment	1.16537	1.143101	422
36 Furniture	1.184598	1.168298	1159
50 Sale, maintenance and repair of motor vehicles	1.255288	1.239372	1488
51 Wholesaletrade and comission trade	1.305669	1.284068	4478
52 Retail trade, repair of personal and household goods	1.311595	1.302452	2960
65 Financial intermediation, except insurance and pension	1.115081	1.115081	1
70 Real estate activities	1.17617	1.127099	1122
71 Renting of machinery and equipment	0.8194612	0.77455	283
72 Computer and related activities	1.188381	1.111177	2175
73 Research and development	1.164963	1.10075	254
74 Other business activities	1.119204	1.040187	6439

Table 8: Correlation matrix of independent variables and key regressors.

	ln(L)	ln(K)	High skill	ln(Sales)	ln(K/L)	Inn. sales	Markup	ln(Export)
ln(L)	1							
ln(K)	0.839***	1						
High skill	-0.0425	-0.235***	1					
ln(Sales)	0.950***	0.842***	-0.0208	1				
ln(KL)	0.311***	0.778***	-0.362***	0.375***	1			
Inn. sales	-0.0186	-0.0450	0.141***	-0.00159	-0.0575*	1		
Markup	0.103***	0.202***	-0.0971***	0.172***	0.234***	0.0731**	1	
ln(Export)	0.581***	0.608***	-0.149***	0.635***	0.377***	0.0425	0.149***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Three stage least squares (3SLS) estimates under different specification. Innovation sample.

VARIABLES	(1) Innovative sales	(2) Markup	(3) Export
Predicted R&D	0.0201***		
	(0.00705)		
Innovative sales (%)		2.487***	
		(0.109)	
Markup			1.320***
			(0.443)
log of Saless	0.0307***	0.0246***	0.814***
	(0.00475)	(0.00677)	(0.0154)
log of Capital/L	0.0220***	0.0433***	0.133**
	(0.00311)	(0.0139)	(0.0543)
High skill share	0.437***	-0.891***	0.199
	(0.0688)	(0.0412)	(0.559)
IMR	0.544***		
	(0.0418)		
Observations	1,616	1,616	1,616
R-squared	0.106	-3.190	0.567
χ^2	193.1	399.4	2175
Prob > χ^2	0.00	0.00	0.00

Time and industry dummies and constant included but not shown.
Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, *p<0.1

Table 10: Three stage least squares (3SLS) using PCM markup. Innovation sample.

VARIABLES	(1) Innovative sales	(2) Markup	(3) Export
Predicted R&D	0.0543		
	(0.0390)		
Innovative sales (%)		0.351	
		(0.269)	
Markup (pcm)			3.980*
			(2.381)
log of Employees	0.0269	-0.0354***	0.683***
	(0.0286)	(0.00468)	(0.121)
log of Capital Stock	0.0123	0.0429***	0.144
	(0.0200)	(0.00400)	(0.114)
High skill share	0.350*	-0.0619*	0.362
	(0.184)	(0.0351)	(0.242)
IMR	0.737***		
	(0.233)		
Observations	1,599	1,599	1,599
R-squared	0.102	-0.005	0.493
χ^2	189.8	307.4	1812
Prob > χ^2	0.00	0.00	0.00

Time and industry dummies and constant included but not shown.
Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, *p<0.1