The Aggregate
Le Chatelier Samuelson Principle
with Cournot Competition

Bertrand Koebel and François Laisney

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 10-009

The Aggregate
Le Chatelier Samuelson Principle
with Cournot Competition

Bertrand Koebel and François Laisney

Non-technical summary

In a market in which firms have market power, the impact of a change in wages on labour demand is not easily determined. Indeed, if after a decrease in wages one firm chooses to increase labour demand for producing more, this reduces the output price; this in turn may lead other firms to decrease their production and their demand for labour. We show that this ambiguity at the firm level is likely to be resolved at the aggregate level of the industry: under some empirically testable restrictions an increase in wages is likely to trigger a decrease in labour demand.

The Le Chatelier principle states that the sensitivity of input demands with respect to own price variations is smaller when the output level is held constant than when it is adjusted. It holds whether competition on the output market is perfect or imperfect, provided the production level of competitors is held constant. A first aim of this paper is to extend the Le Chatelier principle to the case where the production levels of competitors are allowed to vary.

For a given level of production, a cost minimizing firm has an incentive to use more intensively the factors of production whose price decreases and to substitute the other factors by the cheaper one (substitution effect). When the firm is able to set its production level in order to maximize profit, it benefits from the factor price reduction even further (expansion effect). With a competitive output market this expansion effect leads to a further increase in the demand for the factor that becomes cheaper. With imperfect competitive output markets the direction of the effect is less clear: if all competing firms increase production to exploit the reduction in factor price, the output price must fall, and this reduces each firm's incentives to expand its level of production and factor demand.

We study whether aggregate factor demands are decreasing in markets with a particular type of imperfect competition, termed Cournot competition. Despite the ambiguous result at the firm level, we show that under some conditions the aggregate expansion effect is likely to be negative. We complete this paper with an empirical investigation of its theoretical results. In order to identify the Le Chatelier principle with aggregate data, and for decomposing the impact of input price changes on input demands into a substitution and an expansion effect, it is necessary to propose an adequate empirical specification compatible with heterogeneous firms. Instead of relying on a representative firm setup which implies estimation biases when firms are too heterogeneous, we define the aggregate cost function as the conditional expectation of the microeconomic cost function, given the available aggregate information. In this context, we show that it is possible to identify the conditional expectation of the substitution and expansion effects using aggregate data only. We propose a simple test for the validity of the representative firm model.

The empirical application relies on a panel for 18 two-digit US manufacturing industries over the period 1949 to 2001. We obtain several results relative to the rate of returns to scale, the price-cost margin, the short and long run adjustment of input demand to input price change, and the impact of input price changes on output adjustment and inflation. The empirical findings confirm the validity of the Le Chatelier principle. We also find evidence for both increasing and constant returns to scale.

Das Wichtigste in Kürze (German summary)

In einem Markt, wo Unternehmen Marktmacht besitzen, lässt sich der Einfluss einer Veränderung der Löhne auf die Arbeitsnachfrage nicht leicht bestimmen. Falls infolge einer Reduzierung der Löhne ein Unternehmen seine Arbeitsnachfrage erhöht, um mehr zu produzieren, so nimmt der Preis des Produkts ab; dieses wiederum kann Wettbewerber dazu führen, ihre Produktion und somit auch ihre Arbeitsnachfrage zu senken. Wir zeigen, dass diese Ambiguität auf Firmenebene im Aggregat auf Industrieebene verschwinden kann: unter Restriktionen, die empirisch widerlegbar sind, führt eine Zunahme der Löhne zu einer Abnahme der Beschäftigung.

Das Le Chatelier Prinzip besagt, dass die Sensitivität der Nachfrage nach einem Produktionsfaktor zu Veränderungen in seinem eigenen Preis kleiner ist, wenn die Produktion konstant gehalten wird, als wenn sie angepasst wird. Diese Aussage ist gültig bei perfektem sowie imperfektem Wettbewerb, solange das Produktionsniveau der anderen Wettbewerber konstant bleibt. Ein erstes Ziel dieses Papiers ist die Erweiterung des Le Chatelier Prinzips auf den Fall, wo das Produktionsniveau der anderen Wettbewerber variieren darf.

Bei einem festen Produktionsniveau wird ein Unternehmen den Faktor, dessen Preis sinkt, intensiver benutzen, zuungunsten der anderen Faktoren (Substitutionseffekt). Kann die Produktion angepasst werden, um den Profit zu maximieren, wird das Unternehmen weiter von der Preissenkung profitieren können (Expansionseffekt). Unter perfektem Wettbewerb führt dieser Expansionseffekt zu einer weiteren Zunahme der Nachfrage nach diesem Faktor. Bei imperfektem Wettbewerb ist die Richtung des Effekts weniger klar: falls alle Wettbewerber ihre Produktion erhöhen, sinkt der Preis des Produkts und somit auch der Anreiz Produktion und Faktornachfrage zu erhöhen.

Wir untersuchen die aggregierten Nachfragenreaktionen in Märkten, die durch einen spezifischen Typ des unvollständiges Wettbewerb, nämlich Cournot Wettbewerb, gekennzeichnet sind. Wir zeigen dass, trotz der Unklarheit der Reaktion der Unternehmen, der aggregierte Expansionseffekt negativ sein dürfte.

Diese Implikation aus der Modellanalyse wird auf ihre empirische Validität untersucht. Die Zerlegung der Auswirkungen von Faktorpreisveränderungen in Substitutions- und Expansionseffekte erfordert die Verwendung unternehmensspezifischer Kostenfunktionen als Ausdruck der Heterogenität der Unternehmen. Wir definieren die aggregierte Kostenfunktion als den bedingten Erwartungswert der mikroökonomischen Kostenfunktion, gegeben die verfügbare Information. Wir zeigen, dass es in diesem Rahmen möglich ist, den bedingten Erwartungswert des Substitutions- und des Expansionseffekts anhand von aggregierten Daten zu ermitteln.

Die empirische Anwendung beruht auf einem Panel für 18 verarbeitende Industrien der USA für den Zeitraum 1949-2001. Für diese Industrien ermitteln wir die Höhe der Skalenerträge, den Umfang der aus Marktmacht resultierenden Gewinne sowie die kurze und die langfristige Reaktion der Faktornachfragen auf Preisveränderungen und die Auswirkungen von Kostenveränderungen auf die Wertschöpfung. Die empirischen Ergebnisse bestätigen die Gültigkeit des Le Chatelier Prinzips.

The aggregate
Le Chatelier Samuelson principle
with Cournot competition

Bertrand Koebel* and François Laisney**

January 2010

ABSTRACT. This paper studies the aggregate substitution and expansion effects triggered by changes in input prices, in a context where firms supply a homogenous commodity and compete in quantities à la Cournot. We derive a sufficient condition for the existence of a Cournot equilibrium and show that this condition also ensures that the Le Chatelier-Samuelson principle is likely to be satisfied in the aggregate at the Cournot equilibrium, although it may not be satisfied at the firm level. These results are confirmed by the empirical findings obtained for two-digit US manufacturing industries, which also highlight the importance of imperfect competition for understanding aggregate growth, investment and employment.

Keywords: Aggregation, returns to scale, market power, markup, own-price elasticity.

JEL Classification: C33, D24.

* Corresponding author: Beta, Université de Strasbourg, 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex (France). Tel (+33) 368 852 190. Fax (+33) 368 852 071. Email: koebel@unistra.fr.

** Beta, Université de Strasbourg, and ZEW Mannheim.

1. Introduction

This paper investigates the consequences of input price changes on input demands when the output market is imperfectly competitive. The impact of input price changes on input adjustment is described by the Le Chatelier principle, introduced in economics by Samuelson (1947). This principle states that the sensitivity of input demands with respect to own price variations is smaller when the output level is held constant than when it is adjusted. It is apparently not widely known, that in the context of imperfect competition, however, that Samuelson (1947, p.45-46) in fact showed that Le Chatelier principle is satisfied whether competition on the output market is perfect *or* imperfect, provided the production level of competitors is held constant. At the firm level, the Le Chatelier principle focused the attention of many researchers who derived it by weakening or changing underlying assumptions: see Eichhorn and Oettli (1972), Diewert (1981) and Milgrom and Roberts (1996) to name just a few. However, these authors have not considered whether the principle is still satisfied when *negative* externalities between firms affect their behaviour.

A first aim of this paper is to extend the Le Chatelier-Samuelson (LCS) principle to the case of endogenous levels of competitors' output. For a given level of output, a cost minimizing firm has an incentive to use more intensively the input whose price decreased and to substitute the other inputs by the cheaper one (the substitution effect). When the firm is less constrained and becomes able to set its output level in order to maximize its profit, it will choose the optimal output level in order to benefit from the input price reduction even further. This adjustment corresponds to an expansion effect. With a competitive output market this expansion effect is always negative. With imperfect competitive output markets à la Cournot, things are not that clear, because the externality provides incentives to reduce input demand: if all competing firms increase their output level in order to exploit the reduction in input price, the output price must fall, and this reduces each firm's incentives to expand its level of output supply and input demand. So firm level comparative statics is undetermined, only further restrictive assumptions on firm technologies or inverse demand, as these discussed by Roy and Sabarwal (2008, 2009), allow to obtain well determined results.

In the context of perfect competitive markets, Heiner (1982) and Braulke (1984) have shown that aggregation can be helpful for resolving ambiguity at the firm level. Even if it is not necessarily true at the firm level, we study whether aggregate input demands are decreasing in markets with Cournot competition. Despite the ambiguous result at the firm level, we show that under Novshek's (1985) type of conditions, which ensure the existence of a Cournot equilibrium, the aggregate expansion effect is likely to be negative, and the LCS principle is likely to be valid in the aggregate Cournot model.

We complete this paper with an empirical investigation of its theoretical results. In order to identify the LCS principle with aggregate data, and for decomposing the impact of input price changes on input demands into a substitution and an expansion effect, it is necessary to propose an adequate empirical specification compatible with heterogeneous firms. Instead of relying on a representative firm setup which implies estimation biases when firms are too heterogeneous, we proceed to aggregation using the stochastic aggregation theory developed by Lewbel (1996) and Koebel (2002). The aggregate cost function is defined as the conditional expectation of the microeconomic

cost function, given the available aggregate information. In this context, the aggregate cost function does not generally inherit the properties of microeconomic cost functions. However, we show that it is possible to derive the conditional expectation of the input demand system from the cost function by amending Shephard's lemma, and to identify the conditional expectation of the substitution and expansion effects using aggregate data only. We propose a simple test for the validity of the representative firm model.

There is a further methodological focus in this paper: we present a method for dealing with empirical issues raised by the adjustment of inputs in the long run, in the context of increasing or decreasing returns to scale. When the rate of returns and the type of competition on the output market are a priori unknown, empirical contributions (reviewed by Bresnahan, 1989) have often appended a pricing rule – compatible with both perfect and imperfect competition – to the cost and input demand system. However, most of these studies do not derive the optimal output level, nor do they report unrestricted elasticities. This shortcoming is surprising because, when output is endogenous, that is to say, optimally chosen by the production unit, output-restricted elasticities are only of limited interest for deriving policy implications. A further objective of this paper is thus to estimate elasticities that are not restricted by an output level artificially held constant. Since a closed form solution for the optimal output level can seldom be obtained from the price-margin equation, we rely on the implicit function theorem for deriving these elasticities empirically. This method has been developed and applied by Kulati-laka (1987) to model the adjustment of capital to its optimal long run level. We extend this method by deriving moment conditions which are consistent with the theoretical model and identify the substitution and expansion effects of input price changes. This method allows us to derive consistent estimates of the substitution and the expansion matrices and to test the validity of the LCS principle in the aggregate.

The empirical application relies on a panel for 18 two-digit US manufacturing industries over the period 1949 to 2001. We obtain several results relative to the rate of returns to scale, the markup, the short and long run adjustment of input demand to input price change, and the impact of input price changes on output adjustment and inflation. The empirical findings confirm the validity of the LCS principle. Whether returns to scale are increasing or not in U.S. manufacturing industries is an important but controversial empirical issue. Whereas many researchers have argued for the increasing returns to scale hypothesis (Hall, 1988, Shapiro, 1988, Diewert and Fox, 2008), just as many researchers have found evidence for the contrary (Burnside, 1996, Bartelsmann, 1995, Basu and Fernald, 1997). We find evidence for both the increasing and constant return to scale assumptions.

The next section outlines the microeconomic model and derives the LCS principle at the firm level, when the output market is imperfectly competitive. Section 3 exposes Novshek's (1985) and Amir's (1996) sufficient conditions for the existence of a Cournot equilibrium. Section 4 extends the LCS principle to the case of Cournot competition at the aggregate level, it also describes the aggregate consequences of Cournot competition in terms of input adjustment. Section 5 presents the aggregate statistical model compatible with available aggregate data. Section 6 sets up the empirical model specification and highlights how it departs from usual models of producer behavior. The empirical results are described in Section 7, and Section 8 concludes.

2. Input demands and the type of competition

The model is developed at the microeconomic level of the production unit. Vector $x \in \mathbb{R}^J_+$ denotes inputs and w the corresponding price vector. The production unit's output level is denoted by $y \in \mathbb{R}_+$. Under suitable regularity conditions the technology of a cost minimizing production unit is fully described by a twice continuously differentiable cost function c. By definition, $c(w, y) = w^\intercal x^*(w, y)$, where x^* denotes the cost minimizing input vector. The aim of this section is to describe how input demands react to input prices, at the level of the firm.

In an imperfectly competitive product market, the production unit knows the inverse product demand function $p : (y, v) \mapsto p(y, v)$ it faces, where the vector v comprises explanatory variables shifting output demand and which the production unit considers to be out of its control (as the production level of competitors, aggregate variables including the population, the unemployment rate, the level of value added tax and so on). In this subsection, we assume a representative production unit, which simplifies the notations, but this assumption is relaxed later.

The profit function π is given by

$$\pi(w, v) = \max_y \{p(y, v) y - c(w, y)\} \tag{1}$$

$$= p(y^o, v) y^o - c(w, y^o), \tag{2}$$

where $y^o(w, v)$ denotes the optimal solution to (1) and represents the output supply correspondence: one difficulty with y^o is that it is not necessarily a function; for some values of (w, v) there might be several profit maximizing output supplies. In the following we assume that the solution y^o is locally unique. The first order condition for an interior optimum is given by

$$p(y^o, v) + \frac{\partial p(y^o, v)}{\partial y} y^o = \frac{\partial c(w, y^o)}{\partial y}. \tag{3}$$

Output supply changes when the demand function shifts (variation in v) or when the cost parameters w change.

Some authors reviewed by Appelbaum (1982) and Bresnahan (1989) consider that this simple framework encompasses a variety of non-competitive pricing behaviors. The case of an oligopoly with H firms producing heterogeneous products is obtained when the inverse demand function p is specified as:

$$p(y_h, v) = P(y_1^*(y_h), \ldots y_h, \ldots, y_H^*(y_h), v),$$

where y_j^* denotes firm j's reaction function to the production choice y_h of the firm h under consideration. When products are homogenous, the inverse demand function simplifies into

$$p(y_h, v) = P\left(y_h + \sum_{j \neq h}^{H} y_j^*(y_h), v\right),$$

which is compatible with Cournot-Nash, Stackelberg or competitive fringe types of competition. There is a vast literature using the concept of conjectural variation (see Bresnahan, 1989, for a survey), which splits $\partial p / \partial y_h$ into a product of price sensitivity to aggregate output $\partial P / \partial Y$ and a conjectural variation term $1 + \sum_{j \neq h}^{H} \partial y_j^* / \partial y_h$. We do not try to separately identify these terms, as we are just interested into the total impact $\partial p / \partial y_h$, aggregated at the level of an industry. In this section, we follow the Cournot-

Nash conjecture and consider the production level of competitors as fixed while firm h is choosing its optimal production level. In our situation, v_h is specific to firm h, because it comprises the exogenous aggregate production level Y_{-h} of all other firms, beside the macroeconomic variables z which are identical for all firms: we write $v_h = (Y_{-h}, z)$.

A sufficient condition for an interior maximum is that, in addition to (3),

$$a^o(w, v) < 0 \tag{4}$$

with

$$a^o(w, v) \equiv \left[2 \frac{\partial p}{\partial y}(y^o, v) + \frac{\partial^2 p}{\partial y^2}(y^o, v) y^o - \frac{\partial^2 c}{\partial y^2}(w, y^o) \right]^{-1}. \tag{5}$$

This requirement can be fulfilled even in the case of decreasing marginal costs ($\partial^2 c/\partial y^2 < 0$), provided that the inverse demand function has the adequate shape and curvature.

By Hotelling's Lemma the input demand functions are given by:

$$x^o(w, v) = -\frac{\partial \pi}{\partial w}(w, v) = \frac{\partial c}{\partial w}(w, y^o(w, v)) = x^*(w, y^o(w, v)),$$

where the second equality follows from (3). Thus, just as in the perfect competitive case, the constant-output and the unrestricted input demand functions coincide at the optimal output level. Concerning comparative statics,

$$\begin{aligned}
\frac{\partial x^o}{\partial w^\mathsf{T}}(w, v) &= \frac{\partial x^*}{\partial w^\mathsf{T}}(w, y^o) + \frac{\partial x^*}{\partial y}(w, y^o(w, v)) \frac{\partial y^o}{\partial w^\mathsf{T}}(w, v) \\
&= \frac{\partial x^*}{\partial w^\mathsf{T}}(w, y^o) + a^o(w, v) \frac{\partial x^*}{\partial y}(w, y^o) \frac{\partial x^{*\mathsf{T}}}{\partial y}(w, y^o),
\end{aligned} \tag{6}$$

where the second equality follows from the differentiation of (3) with respect to w, yielding:

$$\frac{\partial y^o}{\partial w}(w, v) = a^o(w, v) \frac{\partial x^*}{\partial y}(w, y^o). \tag{7}$$

This allows to obtain the LCS principle in imperfect competition.

Result 1. Assuming $a^o(w, v) < 0$,
(i) the LCS result is satisfied:

$$\frac{\partial x_j^o}{\partial w_j}(w, v) \leq \frac{\partial x_j^*}{\partial w_j}(w, y^o) < 0, \tag{8}$$

(ii) an increase in input price w_j decreases the output level iff input demand x_j^* is normal:

$$\partial y^o/\partial w_j < 0 \Leftrightarrow \partial x_j^*/\partial y > 0, \tag{9}$$

(iii) an increase in input price w_j increases the output price $p^o(w, v) \equiv p(y^o(w, v), v)$ if output demand is decreasing and x_j^* is normal:

$$\{\partial p/\partial y < 0 \wedge \partial x_j^*/\partial y > 0\} \Rightarrow \partial p^o/\partial w_j > 0. \tag{10}$$

Statement (i) directly follows from (6), (ii) from (7) and (iii) from the inverse demand function and (7). This result shows how increases in input prices reduce input demand, which in turn decreases production and creates inflation. The normality requirement of input demand x_j^* is in fact equivalent to the statement that marginal cost is increasing in w_j. The conditions ($a^o(w, v) < 0$, $\partial x_j^*/\partial y > 0$ and $\partial p/\partial y < 0$) necessary for obtaining the different statements of Result 1 can be investigated empirically. In Section 5 we

4

study whether they are satisfied in the aggregate. It is useful to notice that in terms of elasticities (8) becomes

$$\varepsilon\left(x_j^o; w_j\right) \leq \varepsilon\left(x_j^*; w_j\right) < 0, \tag{11}$$

where

$$\varepsilon\left(x_j^o; w_j\right) \equiv \frac{\partial x_j^o\left(w, v\right)}{\partial w_j} \frac{w_j}{x_j^o\left(w, v\right)}.$$

The own-price elasticities of profit maximizing input demands are smaller than those derived from cost minimizing input demands. The economic intuition behind this result is that when the output level can be adjusted after a decrease in input price w_j, this change in scale is made in such a way to fully benefit from the input price reduction which is achieved by increasing x_j (and y if x_j is normal). Notice that input demands are not required to be normal (that is, increasing in the level of output) for obtaining the LCS principle.

Samuelson (1947, p.45-46) derived this principle using a revenue function noted $R\left(x\right)$, which is compatible with a perfectly competitive output market, when $R\left(x\right) = pf\left(x\right)$, but also with imperfect competition for $R\left(x\right) = p\left(f\left(x\right), v\right) f\left(x\right)$. A more general formulation of the Le Chatelier principle, yielding Result 1(i) as a special case, was provided by Eichhorn and Oettli (1972). In comparison to Samuelson's result, the above derivation of the LCS principle has the advantage of relying on the dual: it yields thereby equation (6) which resembles the Slutsky decomposition in consumer theory.
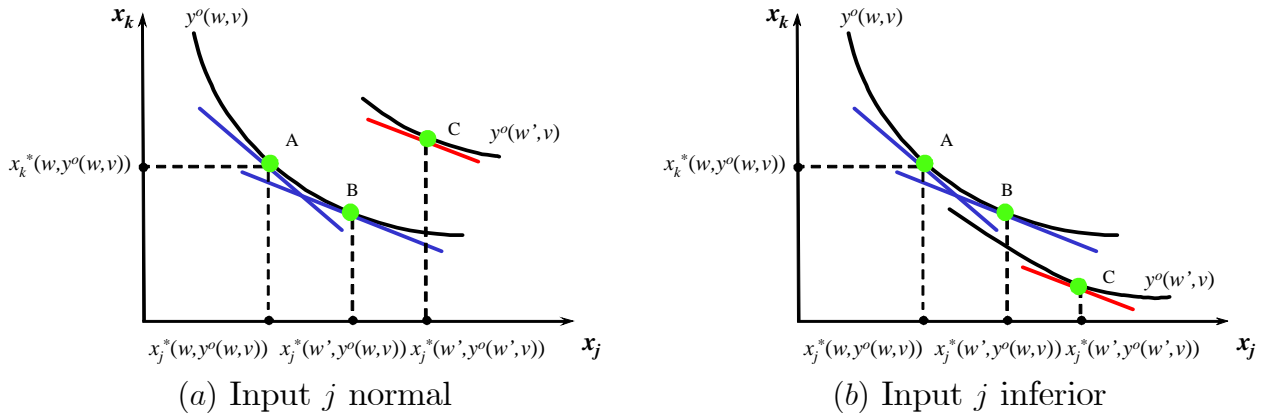

Figure 1: Substitution and expansion effects and input adjustment

Figure 1 illustrates the optimal adjustment of output and its implication for the inputs. This figure, presented by Sakai (1973) in the competitive setup, is also valid when production functions are not concave and production units have market power *as long as v is constant*. The shift from point A to point B along the isoquant corresponding to production level $y^o\left(w, v\right)$ represents input substitution caused by a decrease in the price of input j from w_j to w_j'. The shift from B to C arises when the production unit chooses the profit maximizing output level, and it depicts the expansion (or scale) effect. For normal inputs, this expansion effect is positive and by (7) it turns out that in this case the production unit increases output to its optimal level $y^o\left(w', v\right)$. When input j is inferior the converse applies (see Figure 1b): profit is maximized when the firm decreases output after the decrease of w_j (see 9). Figure 1 illustrates that in both cases unrestricted input demand $x_j^o\left(w, v\right) = x_j^*\left(w, y^o\left(w, v\right)\right)$ reacts more strongly to changes

in w_j than restricted demand $x_j^*(w, y^o)$. The LCS principle differs from the Slutsky decomposition, because production units do not want to reach the highest possible isoquant: in the situation of Figure 1b, profits are maximized by reducing the level of production.

There exist alternative sets of (weaker) assumptions which yield the conclusions of Result 1 (see Milgrom and Roberts, 1996). However, as our objective is to identify both substitution and scale effects of (6) we rely mainly on duality theory.

3. Existence of a Cournot-Nash equilibrium

Changes in input prices affect in general all firms simultaneously which in turn affects the inverse output demand function through changes in v. So, the result of former section (derived for constant v) only partly describe the consequences of changes in input prices. We consider an industry that can be relatively well described as a market with competition à la Cournot. Firms produce a similar product and are heterogeneous with respect to their cost function, their market power and their market share measured by y_h/Y. In contrast to the contestable market literature, we do not require that all (potential) firms have access to the same technology.

When products within an industry are perfectly substituable, all active firms charge or face the same price at equilibrium. The number of incumbent firms H is exogenous. In this setup, some firms make positive profits because they are able to produce cheaper than others; their technology is more efficient. The (inverse) market demand is given by

$$p(y_h, v_h) = P\left(y_h + \sum_{j \neq h}^{H} y_j, z\right), \tag{12}$$

where $v_h = \left(\sum_{j \neq h}^{H} y_j, z\right)$. Let us use the same notations as before and denote the optimal input and output levels by $x_h^o(w, v_h)$ and $y_h^o(w, v_h)$ and the price by $p^o(w, v_h)$.

In this section we consider strategic interactions between firms and describe their influence on input demand adjustments. A look at the reaction functions $y_h^o(w, Y_{-h}, z)$, $x_h^o(w, Y_{-h}, z)$ suffices to see that strategic interactions have an important impact on the output and input demand choices. It can be verified (using (3) and (5)), that the sign of $\partial y_h^o/\partial Y_{-h}$ is the same as the sign of $\partial P/\partial y + y_h \partial^2 P/\partial y^2$ and for this reason, the Cournot game can either be a game with strategic substitutes $(\partial y_h^o/\partial Y_{-h} \leq 0)$ or complements $(\partial y_h^o/\partial Y_{-h} \geq 0)$.

For simplicity, we restrict our analysis to interior points. A *Cournot equilibrium* is any H-tuple $y_h^N(w, z)$ and $x_h^N(w, z)$ such that for any active firm, (3), (5) and (12) are satisfied at $y_h^N(w, z)$ for each $h = 1, \ldots, H$. So, at a Cournot equilibrium,

$$y_h^N(w, z) = y_h^o\left(w, Y_{-h}^N(w, z), z\right), \quad x_h^N(w, z) = x_h^o\left(w, Y_{-h}^N(w, z), z\right). \tag{13}$$

If we assume that y_h^o is a continuous function in Y_{-h} for every w, z, then Brouwer's fixed point theorem can usually be applied to show that Cournot's equilibrium exists. However, in Cournot oligopolistic markets, it is restrictive to assume that y_h^o is a continuous function, because $Py_h - c_h$ is not necessarily concave in y_h for all values of (y_h, Y_{-h}, z, w). Several economists have tried to go around the assumption of concave profits for ob-

taining the existence of a Cournot equilibrium.

Novshek (1985) has shown that a H-firms Cournot equilibrium exists provided that a "firm's marginal revenue be everywhere a declining function of the aggregate output of others" that is:

$$\frac{\partial P}{\partial Y}(y_h + Y_{-h}, z) + y_h \frac{\partial^2 P}{\partial Y^2}(y_h + Y_{-h}, z) \leq 0. \tag{14}$$

This condition also implies that firms' reaction functions $y_h^o(w, Y_{-h}, z)$ are nonincreasing in Y_{-h}. Inequality (14) is satisfied if the (nonincreasing) inverse demand function is linear or concave in y, in which case the existence of a Cournot equilibrium is guaranteed. Since this condition has to be satisfied for any value of y_h, Y_{-h} and z, it can equivalently be written as

$$\frac{\partial P}{\partial Y}(Y, z) + Y \frac{\partial^2 P}{\partial Y^2}(Y, z) \leq 0, \tag{15}$$

for any Y. This formulation depends on aggregate data only and implies that condition (14) is fulfilled for *any* firm.

Amir (1996) provided a different sufficient condition ensuring the existence of a Cournot equilibrium. He showed that (Theorem 3.1) log-concavity of P w.r.t. Y, that is,

$$P(Y, z) \frac{\partial^2 P}{\partial Y^2}(Y, z) - \left[\frac{\partial P}{\partial Y}(Y, z)\right]^2 \leq 0 \tag{16}$$

for any Y, ensures that each firm's reaction correspondence is nonincreasing in Y_{-h} and that a Cournot equilibrium exists. Amir (2005) compares his requirement with Novshek's and shows that both conditions are not nested: (16) does not imply (15), nor is it implied by (15).

There are two difficulties with these aggregate conditions. On the one side, (16) is sufficient for the existence of a Cournot equilibrium, but not necessary, and so it is not the weakest possible condition for achieving existence. On the other side, the fact that (14) has to be satisfied for any value of y_h, Y_{-h} and z, is very demanding. It must even be satisfied for the case in which one firm produces the total output, which could reasonably be excluded if there is a competition law enforcing an upper bound for the market share, or alternatively, if the firms' cost functions lead them to always choose an output level smaller than Y. So we try to derive a weaker condition which ensures the existence of a Cournot equilibrium.

Result 2. Assume that for any firm, there is a maximal capacity $\overline{y}$, so that no firm chooses $y_h > \overline{y}$ for any Y and z. If

$$\frac{\partial P}{\partial Y} + \frac{\partial^2 P}{\partial Y^2} y_h \leq 0, \tag{17}$$

for any z and $0 \leq y_h \leq \overline{y}$ then Cournot's equilibrium exists.

Instead of a maximum capacity, $\overline{y}$ can also be interpreted as firm's maximum output level that a competition commission is tolerating in this oligopoly market. The proof of this result follows from the fact that the output space $[0, \overline{y}]^H$ is a complete lattice, and imposing y_h to be included in the interval $[0, \overline{y}]$ still yields a reaction correspondence that is nonincreasing in Y_{-h} just as in Novshek's case. Condition (17) is implied by the aggregate condition

$$\frac{\partial P}{\partial Y} + \frac{\partial^2 P}{\partial Y^2} Y \sqrt{\lambda} \leq 0, \tag{18}$$

for any aggregate and elementary output levels Y and $y_h \leq \overline{y}$ compatible with the

Hirschman-Herfindahl index of concentration λ.[1] Provided the restriction on the distribution of output is valid ($y_h \leq \overline{y}$), condition (18) is weaker than (15). These conditions will be tested in Section 5.

4. Aggregate comparative statics

Firm level comparative statics has been studied by Roy and Sabarwal (2008, 2009) who derive conditions ensuring monotone comparative statics at the firm level in games with strategic substitutes. In Section 2 (Result 1) it was shown that the LCS principle is satisfied for given level of aggregate production of all competitors. At a Cournot equilibrium, the total impact of a change in input prices follows from (13):

$$\frac{\partial y_h^N}{\partial w}(w,z) = \frac{\partial y_h^o}{\partial w}\left(w, Y_{-h}^N(w,z), z\right) + \frac{\partial y_h^o}{\partial Y_{-h}}\frac{\partial Y_{-h}^N}{\partial w}(w,z)$$

$$\frac{\partial x_h^N}{\partial w}(w,z) = \frac{\partial x_h^*}{\partial w}\left(w, y_h^N(w,z), z\right) + \frac{\partial x_h^*}{\partial y_h}\frac{\partial y_h^o}{\partial w}\left(w, Y_{-h}^N(w,z), z\right) + \frac{\partial x_h^*}{\partial y_h}\frac{\partial y_h^o}{\partial Y_{-h}}\frac{\partial Y_{-h}^N}{\partial w}(w,z).$$

As $\partial y_h^o / \partial w \leq 0$, it follows that $\partial y_h^N / \partial w \leq 0$ if the last term, corresponding to a change in firm h's output triggered by the strategic interaction with all other firms, does not outweigh the direct impact of an increase in w. As this last term can be positive or negative, the overall sign of $\partial y_h^N / \partial w$ is undetermined. The same remark applies to $\partial x_h^N / \partial w$: a further and indeterminate "externality induced input adjustment" is added to the substitution and expansion effects of (6) and which explains why the LCS principle is not necessarily satisfied at the firm level. It is now interesting to analyze whether the LCS principle holds for the aggregate industry.

4.1 Cournot equilibrium

We show that under condition (18), the LCS principle is likely to be satisfied in the aggregate. Let us define the aggregate input demand functions X^* for fixed levels of individual production as:

$$X^*\left(w, \{y_h\}_{h=1}^H\right) \equiv \sum_{h=1}^H x_h^*(w, y_h).$$

Similarly, the aggregate demands and supply obtained for given levels of competitors' output are defined as:

$$X^o\left(w, \{y_h\}_{h=1}^H\right) \equiv \sum_{h=1}^H x_h^o(w, Y_{-h}, z),$$

$$Y^o\left(w, \{y_h\}_{h=1}^H\right) \equiv \sum_{h=1}^H y_h^o(w, Y_{-h}, z).$$

Presumptions 1. Under the assumption that (4) and (18) are satisfied, it is expected that at the Cournot equilibrium:

---

[1] That (18) implies (17) is obvious if $\partial^2 P / \partial Y^2 \leq 0$ since $\partial P / \partial Y \leq 0$. If $\partial^2 P / \partial Y^2 \geq 0$, notice that the highest market share $\overline{y}/Y$ satisfies $\overline{y}/Y \leq \sqrt{\lambda}$, and so, $y_h \leq \overline{y} \leq Y\sqrt{\lambda}$.

(i) the aggregate (reaction) functions Y^o and X_j^o are nonincreasing in w_j;

(ii) the aggregate Nash equilibrium quantities Y^N and X_j^N are nonincreasing in w_j;

(iii) the aggregate Nash equilibrium quantities Y^N and X_j^N are less own-price elastic than Y^o and X_j^o respectively; they satisfy the aggregate LCS principle:

$$\varepsilon\left(Y^o; w_j\right) \ \leq \ \varepsilon\left(Y^N; w_j\right) \leq 0, \tag{19}$$

$$\varepsilon\left(X_j^o; w_j\right) \ \leq \ \varepsilon\left(X_j^N; w_j\right) \leq \varepsilon\left(X_j^*; w_j\right) \leq 0; \tag{20}$$

(iv) the output price is less reactive to an exogenous change in input prices at Cournot-equilibrium than at the firm optimal output level:

$$0 \leq \frac{\partial P^N}{\partial w}\left(w, z\right) \leq \frac{\partial P^o}{\partial w}\left(w, v\right). \tag{21}$$

Appendix A discusses why Presumption 1 is likely to be satisfied. Presumptions 1(i) and (ii) mean that aggregate input j and output and are likely to decrease when price of input j increases. A direct corollary of $\partial Y^N/\partial w \leq 0$ is that the individual output reactions to changes in input prices, $\partial y_h^N/\partial w$, are negative in average. Presumption 1(iii) states that the response to input price changes is less pronounced at a Nash equilibrium $(\partial Y^o/\partial w \leq \partial Y^N/\partial w)$ than in the case of the Cournot-Nash conjecture of constant output level for all competitors. At a Nash equilibrium, indeed, $\partial Y^N/\partial w$ incorporates the reduction of output quantities of all competitors triggered by the increase in input prices. The second series of inequalities (20) is an aggregate version of the LCS inequality. It means that a rule analog to Result 1 is likely to be valid in the aggregate. When all firms are able to choose their output optimally, aggregate input demand is more reactive to input prices than for fixed levels of output. This result is not straightforward in an imperfectly competitive context, because a firm has an incentive to increase its own output and input levels in reaction to a decreases in its competitor's output and input levels consecutively to an increase in input prices: $\partial y_h^o/\partial Y_{-h} \leq 0$ and $\partial x_h^o/\partial Y_{-h} \leq 0$. Claim (21) follows from the definition of the output price:

$$P^N\left(w, z\right) = P\left(Y^N\left(w, z\right), z\right) = P\left(Y^o\left(w, \{y_h\}_{h=1}^H, z\right), z\right). \tag{22}$$

There are two reasons why claims (19)-(21) can be violated at a Cournot equilibrium. First, a Cournot equilibrium can exist even if (14) or (18) is violated, in which case $\varepsilon\left(Y^N; w_j\right)$ may become positive. However, the validity of (18) can be investigated empirically. In the case where (18) cannot be rejected, this provides evidence both for the existence of a Cournot equilibrium and for the validity of the aggregate LCS principle (Presumption 1). The second source of violation of (19)-(21) is firm heterogeneity. As shown in Appendix A, the sign of $\partial Y^N/\partial w$ depends upon the size and sign of $\text{cov}\left(\frac{\partial^2 c}{\partial y^2}\frac{y^N}{Y^N}, \frac{\partial y^N}{\partial w}\right)$. This term is identically zero when firms are identical (as in many symmetric Cournot equilibria), or when they have constant returns to scale. All in all, there are several reasons to believe that Presumption 1 is empirically satisfied. We conclude this section with an example which illustrates why the LCS can be valid in the aggregate without being necessarily satisfied at the firm level.

Example 1. In a standard Cournot duopoly with linear inverse demand:

$$p = a\left(z\right) - b\left(z\right)\left(y_1 + y_2\right),$$

and cost function $c_h(w, y_h) = d_h(w) y_h$, the reaction functions are given by:

$$y_h^o(w, y_{-h}, z) = \frac{a(z) - b(z) y_{-h} - d_h(w)}{2b(z)},$$

and the Cournot equilibrium is:

$$y_h^N = \frac{a(z) - 2d_h(w) + d_{-h}(w)}{3b(z)}.$$

What happens when w increases? Whereas at the firm level the impact on y_h^N is undetermined, at the aggregate level, the impact is negative, because:

$$Y^N(w, z) = \frac{2a(z) - d_1(w) - d_2(w)}{3b(z)},$$

which decreases when w increases. The reaction curves and Nash equilibria are depicted in Figure 2 below. This figure also includes the iso-output line $y_1 + y_2 = Y$ going through the aggregate Nash equilibrium Y^N. Any point below this line corresponds to a smaller aggregate output level than Y^N. When w increases to w', the reaction functions are shifted downwards (dotted lines) because $d_h(w)$ increases in w. The new Cournot equilibrium is reached at the intersection of the dotted reaction curves, somewhere in one out of three areas A, B, C. In triangle A the output level of firm 2 increases and the one of firm 1 decrease, in rectangle B the output levels of both firm decrease and in triangle C the output level of firm 1 increase and this of firm 2 decreases. In all three cases, however, the total output level $y_1^N + y_2^N$ decreases after an increase in w to w'.[2]
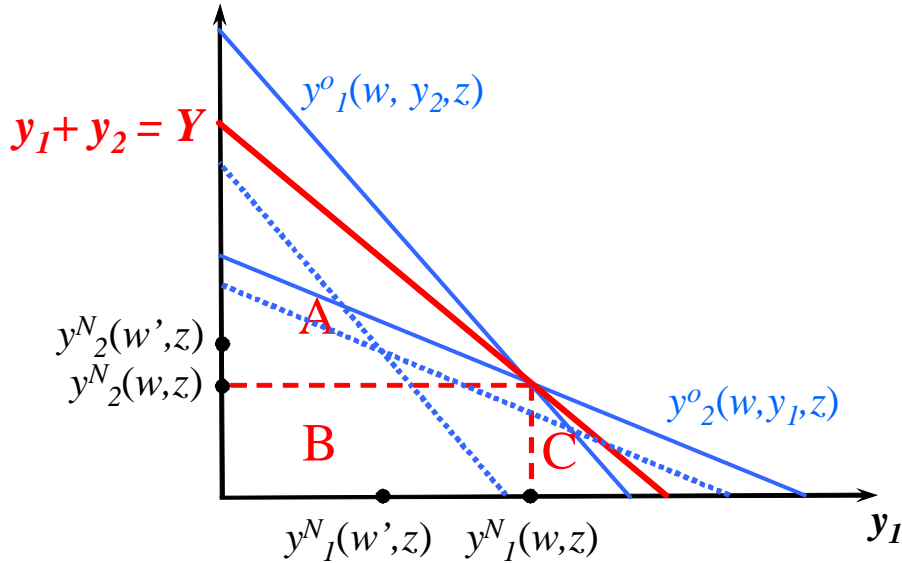

Figure 2: Comparative statics at the firm level and in the aggregate

Notice that the result depicted in Figure 2 does not decisively depend upon the slope of the reaction function, such a figure can also be obtained for both y_h^o increasing in Y_{-h} or when one reaction function is increasing and the other decreasing in Y_{-h}. The important ingredient for obtaining the aggregate comparative statics result is that at

[2] Figure 2 is also useful for illustrating claim (iv). On this figure, it can be seen that at the new input price $w' \geq w$, the aggregate output level $Y^o = y_1^o(w', y_2, z) + y_2^o(w', y_1, z)$ is smaller than $Y^N(w', z)$.

10

least one reaction function is shifted downwards after an increase in w_j, which is ensured if x_j^* is normal (see Result 1(ii)).[3] Aggregate comparative statics, however, becomes tricky when $\partial^2 P/\partial Y^2 \neq 0$ and $\partial^2 c_h/\partial y_h^2 \neq 0$ (see the discussion of Presumption 1 in Appendix A), and Figure 3 below illustrates that the claim of Presumption 1 can be violated in these nonlinear cases. This counter-example works because both reaction curves y_1^o and y_2^o cross the iso-aggregate output line $y_1 + y_2 = Y$ after w increases to w'. In this couterexample the market shares are drastically changed by a marginal change in w, which is empirically not very likely.
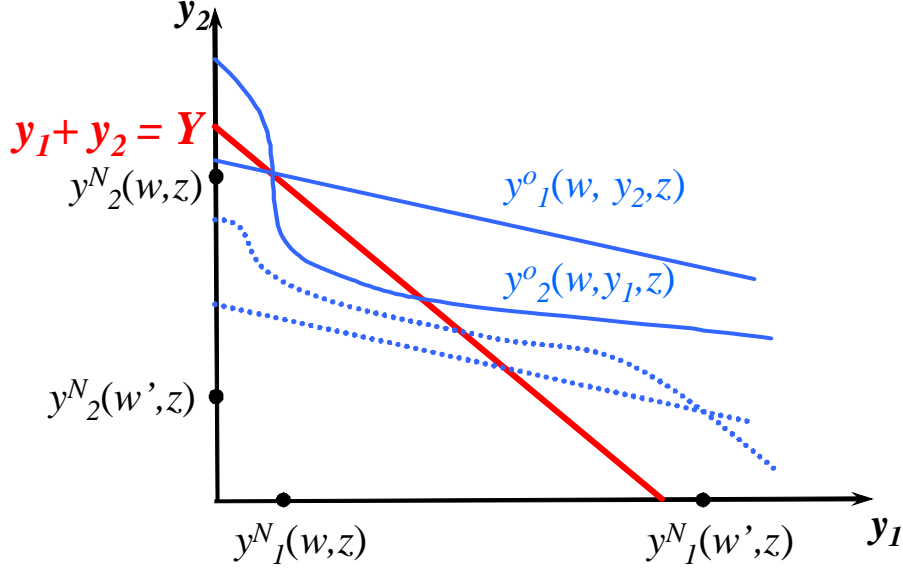
Figure 3. A counter-example to Presumption 1

4.2 Input demand reactivity and degree of competition

In order to better understand the role played by imperfect competition for obtaining the results above, let us compare the Cournot outcome with the benchmark of a market where all firms are in perfect competition. This case has been studied by Heiner (1982) and Braulke (1984). Let $Y^d : (\overline{p}, z) \mapsto Y^d(\overline{p}, z)$ denote the aggregate output demand function and define the competitive output price level $p^c(w, z)$ as the solution in $\overline{p}$ to the market clearing equation:

$$\sum_{h=1}^{H} \overline{y}_h(\overline{p}, w) = Y^d(\overline{p}, z), \tag{23}$$

where h denotes a production unit, and H is the (exogenous) number of active firms in the market. The corresponding aggregate output level is denoted by $Y^c(w, z) = Y^d(p^c(w, z), z)$. The aggregate input quantities are given by $X^c(w, z) \equiv \sum_h^H \overline{x}_h(w, p^c(w, z))$.

Whereas at the microeconomic level it is not possible to say how $\overline{x}_h(w, p^c(w, z))$ or $\overline{y}_h(w, p^c(w, z))$ vary with w, because the output-price response effect is indeterminate, Heiner (1982) has shown that this effect is well determined in the aggregate. "This reassuring effect (...) represents one of the few cases where an ambiguity at the micro

---

[3] The result can still be satisfied if only one of the reaction curves is shifted upwards, but not when all reaction functions shift upwards when w_j increases. This example shows that the claim of Presumption 1 is satisfied in more general contexts than those of supermodular or submodular games.

level is resolved at the macro level by aggregation" (Braulke, 1984, p.75). From the definition of aggregate input demand, it follows that

$$
\frac{\partial X^c(w,z)}{\partial w^\mathsf{T}} = \sum_h^H \frac{\partial \overline{x}_h}{\partial w^\mathsf{T}}(w,p^c) + \sum_h^H \frac{\partial \overline{x}_h}{\partial p}(w,p^c)\frac{\partial p^c}{\partial w^\mathsf{T}}(w,z)
$$

$$
= \frac{\partial \overline{X}}{\partial w^\mathsf{T}}(w,p^c) - \sum_h^H \frac{\partial \overline{y}_h}{\partial w}(w,p^c)\frac{\partial p^c}{\partial w^\mathsf{T}}(w,z), \tag{24}
$$

where the last equality is a consequence of the symmetry of individual demand responses. In order to conclude, notice that (23) implies that

$$
\frac{\partial p^c}{\partial w^\mathsf{T}} = \left(\frac{\partial Y^d}{\partial p} - \sum_{h=1}^H \frac{\partial \overline{y}_h}{\partial p}\right)^{-1} \sum_{h=1}^H \frac{\partial \overline{y}_h}{\partial w^\mathsf{T}},
$$

with $\partial Y^d/\partial p - \sum_{h=1}^H \partial \overline{y}_h/\partial p < 0$. This proves the positive semidefiniteness of the last term of (24) and shows that

$$
\frac{\partial \overline{X}}{\partial w^\mathsf{T}}(w,p^c) \ll \frac{\partial X^c}{\partial w^\mathsf{T}}(w,z) \ll \frac{\partial X^*}{\partial w^\mathsf{T}}\left(w,\{y_h\}_{h=1}^H\right) \ll 0. \tag{25}
$$

See Heiner (1982) and Braulke (1984) for a detailed derivation of these inequalities. Equation (25) means that in the aggregate, output price changes absorb the shock in input prices, and the reactions in input quantities then become less important than for constant output price.

How do the non-competitive aggregate equilibrium input quantities behave and compare to (25)? Some authors, like Cahuc and Zylberberg (2004, p.186) rely on (6) in order to argue that the extension effect "diminishes in absolute terms" when the market power rises. This claim is true *ceteris paribus*, that is, when the technology is independent from market power, but it is not necessarily satisfied otherwise. As a consequence this conjecture will not necessarily be satisfied at the aggregate level of an industry, where the link between the degree of competition and the size of the expansion effect, theoretically unravelled, becomes an empirical issue. For convenience we state our conjecture as follow:

Presumption 2. At the aggregate level of a heterogeneous industry, when prices adjust,
(i) the aggregate input demand functions satisfy:

$$
\varepsilon\left(\overline{X}_j; w_j\right) < \varepsilon\left(X_j^c; w_j\right) \lessgtr \varepsilon\left(X_j^N; w_j\right) < \varepsilon\left(X_j^*; w_j\right) < 0;
$$

(ii) the aggregate output supply function satisfies:

$$
\varepsilon\left(\overline{Y}; w_j\right) < \varepsilon\left(Y^c; w_j\right) \lessgtr \varepsilon\left(Y^N; w_j\right) < 0;
$$

(iii) the output price satisfies:

$$
0 < \varepsilon\left(p^c; w_j\right) \lessgtr \varepsilon\left(p^N; w_j\right).
$$

Part (i) states that the competitive p-constant input demand elasticities $\varepsilon\left(\overline{X}_j; w_j\right)$ are smaller than the p-adjusted input elasticities $\varepsilon\left(X_j^c; w_j\right)$ and $\varepsilon\left(X_j^N; w_j\right)$ whose comparison is an empirical issue. The response of $\overline{X}_j$ following an increase in w_j is greater

than the response of X_j^c and X_j^N because, in the former case, the output price $\overline{p}$ is held constant, whereas in the latter cases, the output price decreases in order to clear the output market, which calls for a smaller adjustment in input demands. However, it seems difficult to say much about how $\varepsilon\left(X_j^c; w_j\right)$ and $\varepsilon\left(X_j^N; w_j\right)$ compare using theoretical arguments only. This limitation also provides a motivation for the empirical part of this paper. In summary, this subsection shows that competitive markets do not necessarily exhibit more variability than less competitive markets.

5. Aggregation when the distribution of market shares is unobserved

In this section we explicitly derive the aggregate relationships from the disaggregate ones and show that they depend upon the distribution of market shares (and how it changes with w, Y and t). We also discuss identification of the effect of unobserved shifts in market shares on the aggregate cost and demand functions. We now include a variable t denoting time as an argument of the cost and demand functions.

5.1 Aggregate cost and input demand functions

Whereas it is natural to define total cost at the level of an industry by $C_{nt} = \sum_{h=1}^{H_n} w^\top x_{ht}$, the total cost function is given by $C_n\left(w, y_{1t}, \ldots, y_{H_n t}, t\right) = \sum_{h=1}^{H_n} c_h\left(w, y_h, t\right)$ and depends upon the whole distribution of output within the nth industry, a piece of information which is difficult to obtain and not always possible to consider explicitly. In order to set up an aggregate model, we follow Lewbel (1996) and Koebel (2002) and reparameterize $(y_1, \ldots, y_H) = \beta Y$ with $\beta = (y_1/Y, \ldots, y_H/Y)$. Then, it is always possible to define the aggregate cost and demand functions as the conditional expectations of the true but unobserved functions:

$$\mathbf{C}\left(w, Y, t\right) = \mathrm{E}_\beta\left[C\left(w, \beta Y, t\right) | w, Y, t\right] \tag{26}$$

$$\mathbf{X}^*\left(w, Y, t\right) = \mathrm{E}_\beta\left[X^*\left(w, \beta Y, t\right) | w, Y, t\right]. \tag{27}$$

The conditional expectation is taken with respect to the conditional density of market shares $f\left(\beta | w, Y, t\right)$.

The properties of the aggregate cost and input demand functions have been studied by Lewbel (1993a) and Koebel (2002) who show that microeconomic properties are not necessarily inherited in the aggregate. Indeed, comparative statics for the aggregate demands depend on the way the distribution of market shares is shifted by changes in (w, Y, t):

$$\frac{\partial \mathbf{C}}{\partial w}\left(w, Y, t\right) = \mathrm{E}_\beta\left[\frac{\partial C}{\partial w}\left(w, \beta Y, t\right) | w, Y, t\right] + \int C\left(w, \beta Y, t\right) \frac{\partial f}{\partial w}\left(\beta | w, Y, t\right) d\beta \tag{28}$$

$$\frac{\partial \mathbf{C}}{\partial Y}\left(w, Y, t\right) = \mathrm{E}_\beta\left[\frac{\partial C}{\partial Y}\left(w, \beta Y, t\right) | w, Y, t\right] + \int C\left(w, \beta Y, t\right) \frac{\partial f}{\partial Y}\left(\beta | w, Y, t\right) d\beta \tag{29}$$

$$\frac{\partial \mathbf{C}}{\partial t}\left(w, Y, t\right) = \mathrm{E}_\beta\left[\frac{\partial C}{\partial t}\left(w, \beta Y, t\right) | w, Y, t\right] + \int C\left(w, \beta Y, t\right) \frac{\partial f}{\partial t}\left(\beta | w, Y, t\right) d\beta. \tag{30}$$

Whereas the first terms of the right hand side of the equalities denote the expected microeconomic marginal cost, the last terms correspond to shifts in the distribution of

market shares following changes in w, Y and t.[4] These terms correspond to aggregation biases and can be denoted respectively by $B_{Cw}(w,Y,t)$, $B_{CY}(w,Y,t)$ and $B_{Ct}(w,Y,t)$. Notice, however, that the adding up property is inherited in the aggregate in the sense that $\mathbf{C}(w,Y,t) = w^\top \mathbf{X}^*(w,Y,t)$.

Without disaggregate data on the distribution of market shares, it will be extremely difficult to identify both the expected value of the derivatives of C w.r.t. w, Y and t and the redistribution effect separately. Both terms of decomposition (28) are for instance homogenous of degree one in w if f is homogeneous of degree zero in w. How is it possible in this context to disentangle both terms without using data specific to the distribution f and $\partial f/\partial w$? Economic theory is actually useful for giving some identifying structure to (28). Indeed, by Shephard's lemma, the first term on the right hand side of equality (28) corresponds to aggregate input demands:

$$\mathrm{E}_\beta \left[\frac{\partial C}{\partial w}(w,\beta Y,t) | w,Y,t \right] = \mathrm{E}_\beta \left[X^*(w,\beta Y,t) | w,Y,t \right] = \mathbf{X}^*(w,Y,t), \qquad (31)$$

a term which can be estimated when data on input demand quantities are available, and so, it is possible to identify the aggregation biases residually, as the difference $\partial \mathbf{C}/\partial w (w,Y,t) - \mathbf{X}^*(w,Y,t)$. Notice that identification is achieved in this way even in the case where f is homogeneous of degree zero in w.

Result 2. Under the assumption that each firm is cost minimizing, and that the aggregate conditional mean cost function $\mathbf{C}$ and all the partial derivatives $\partial \mathbf{C}/\partial w$ are observed, the aggregation biases B_{Cw} are identified.

Corollary 2. The aggregate cost function $\mathbf{C}$
(i) is homogeneous of degree one in w iff for any (w,Y,t)

$$w^\top B_{Cw}(w,Y,t) = 0 \qquad (32)$$

(ii) satisfies Shephards' lemma in the sense that $\partial \mathbf{C}/\partial w = \mathbf{X}^*$ iff for any (w,Y,t)

$$B_{Cw}(w,Y,t) = 0. \qquad (33)$$

This corollary is a direct consequence of relationships (28) and (31). Function $\mathbf{C}$ is homogeneous of degree one in w iff $w^\top \partial \mathbf{C}/\partial w = \mathbf{C}(w,Y,t) = w^\top \mathbf{X}^*(w,Y,t)$ and so the claim of Corollary 2(i) directly follows from (28). Corollary 2(ii) follows from (28) and (31). This corollary could be extended to second order derivatives of $\mathbf{C}$.

5.2 Identification of the markup, marginal cost and aggregation biases

We first derive the aggregate markup pricing relationship and discuss how it relates to the exact aggregation literature. Then we study the conditions under which the aggregation bias is identified. Let us remember that within an industry firms are facing the inverse demand function $p(y_h, v_h)$, where y_h denotes the hth firm's output and v are output demand shifters the production unit considers to be out of its influence. At the industry level with H firms, aggregate output is given by $Y = \sum_{h=1}^H y_h$ and satisfies

---

[4] Notice that

$$\int C(w,\beta Y,t) \frac{\partial f}{\partial w}(\beta|w,Y,t)\, d\beta = \operatorname*{cov}_\beta \left(C, \frac{\partial \ln f}{\partial w} | w,Y,t \right).$$

$PY = p \sum_{h=1}^{H} y_h$.

In this context, the first order condition for profit maximization for firm h is given by:

$$p + \frac{\partial p}{\partial y_h} y_h = \frac{\partial c_h}{\partial y_h}(w, y_h, t). \qquad (34)$$

Aggregating (34) over firms yields:

$$p + \sum_{h=1}^{H} \frac{\partial P}{\partial Y} y_h \frac{y_h}{Y} = \sum_{h=1}^{H} \frac{\partial c_h(w, y_h, t)}{\partial y_h} \frac{y_h}{Y}. \qquad (35)$$

Two approaches lead to a relationship that can be estimated with aggregate data. The first one assumes that microeconomic relationships exhibit properties leading to simplifications in equation (35). When all firms have the same inverse demand elasticity and the same marginal cost function $\partial c_h(w, y_h)/\partial y_h \equiv d(w)$, then (35) can be written in terms of aggregate variables only. Such an assumption, however, is incompatible with our purpose of estimating flexible returns to scale.

The second approximate aggregation approach developed by Lewbel (1996) explicitly defines aggregate functions as conditional expectation of the disaggregate relationships, thereby avoiding restrictions on individual technologies. Doing this, we obtain from (35):

$$p + \frac{\partial P}{\partial Y} Y \mathrm{E}_\beta \left[\sum_{h=1}^{H} \beta_h^2 | w, Y, t \right] = \mathrm{E}_\beta \left[\sum_{h=1}^{H} \frac{\partial c_h(w, y_h, t)}{\partial y_h} \frac{y_h}{Y} | w, Y, t \right].$$

So, the conditional expectation of the marginal revenue depends upon

$$\lambda^a(w, Y, t) \equiv \mathrm{E}_\beta \left[\sum_{h=1}^{H} \beta_h^2 | w, Y, t \right],$$

which corresponds to the conditional expectation of the Hirschman-Herfindahl index of industry concentration. Using (29), we can rewrite this equation as:

$$p + \lambda^a(w, Y, t) Y \frac{\partial P}{\partial Y} = \frac{\partial \mathbf{C}}{\partial Y}(w, Y, t) - \int C(w, \beta Y, t) \frac{\partial f}{\partial Y}(\beta | w, Y, t) d\beta, \qquad (36)$$

The main advantage of the aggregate specification (36) is that if y_h^N is the Cournot equilibrium production level of firm h, then the aggregate solution $\widehat{Y}^N(w, z, t)$ to (36) is compatible with (34), in the sense that it is as close as possible to $\sum_h y_h^N(w, z, t)$ in L^2 norm. Notice that specification (36) obtained by aggregating (34) looks similar to the usual microeconomic specifications for which $p + \lambda^a Y \partial P/\partial Y$ is interpreted as the marginal revenue as perceived by the firm: $\lambda^a = 1$ corresponds to a cartel, $\lambda^a = 1/H$ to a (symmetric) Cournot oligopoly and $\lambda^a = 0$ to perfect competition. See Bresnahan (1989) and Reiss and Wolak (2007) for a survey of this literature. All terms of (36) are identified when data on the Herfindahl index is available: P and $\partial P/\partial Y$ are identified from estimating an inverse demand function, λ^a is identified as the conditional expectation of the Herfindahl index; and $\partial \mathbf{C}/\partial Y$ is identified from the estimation of the cost function, and so the aggregation bias $B_{CY}(w, Y, t)$ can be computed residually. Without such information, however, it seems difficult to identify λ^a separately from B_{CY}. Bresnahan (1982) and Lau (1982) discussed the conditions under which a constant parameter λ^a can be uniquely identified under the assumption that a representative firm exists $(B_{CY} \equiv 0)$. Lau's result can be extended to our framework:

Result 3. Under the above assumptions,
(i) the functions λ^a and B_{CY} are nonparametrically identifiable iff the inverse demand function P is not such that $P(Y, z) = t(Y) r(z) + s(Y)$ in the case where z is a scalar, or $P(Y, z) = \mathbf{P}(Y, r(z))$ in the case where z is a vector;
(ii) for parameterized shifts in the distribution of market shares, the model is identified under broader circumstances than those mentioned in (i).

Result 3 (i) (proven in Appendix A for the sake of completeness) extends Lau's (1982) identification theorem in a context where λ^a is a function of (w, Y, t). (There is a subtlety due to the number of variables included in z.) This identifiability condition can be relatively easily tested from the estimation of an inverse demand function. When the null of separability in z cannot be rejected, Result 3(ii) recommends to specify a functional form for B_{CY}.

The main difference between Lau's and our result is that in our case, information on the cost function and on the marginal cost functions ($\mathbf{C}$ and $\partial \mathbf{C} / \partial Y$) is available. Supplementing cost data to the analysis allows us to overcome Lau's identification problem of the marginal cost function and to extend his result to the problem of identifying the aggregation bias. Indeed, aggregate cost and marginal cost depend upon the way aggregate output is distributed over firms, and this information, which was missing in the Bresnahan-Lau approach, could then provide the information we need for identifying both the average market power of an industry and the aggregation bias. For instance, when the firms' market shares are constant, $B_{CY} = 0$, and (36) allows to link the unobserved marginal cost function $\partial C / \partial Y$ to the estimable marginal cost function $\partial \mathbf{C} / \partial Y$:

$$\frac{\partial \mathbf{C}}{\partial Y}(w, Y, t) = \mathrm{E}_\beta \left[\frac{\partial C}{\partial Y}(w, \beta Y, t) | w, Y, t \right],$$

which can be treated as if it were observed. In this case, the estimation of the conditional mean cost function $\mathbf{C}$ is informative about the otherwise unobserved aggregate marginal cost C and this information allows to identify the average market power λ^a as the ratio between $\partial \mathbf{C} / \partial Y - P$ and $Y \partial P / \partial Y$. In the case where B_{CY} and λ^a are constant, (36) can be rewritten as

$$P + \frac{\partial P}{\partial Y} Y \lambda^a + B_{CY} = \frac{\partial \mathbf{C}}{\partial Y}(w, Y, t).$$

When $\partial \mathbf{C} / \partial Y$ is observed, it is possible to identify B_{CY} and λ^a iff the "regressors" P and $Y \partial P / \partial Y$ are not proportional, which is the case iff $P(Y, z) \neq r(z) Y^\theta$. This example illustrates that with more information on costs identification is achieved under broader circumstances than those given in Result 3(i).

6. Empirical specification

In the empirical application, we shall use a panel data set of N industries indexed by $n = 1, \dots, N$ with observations on T time periods indexed by $t = 1, \dots, T$. In a first subsection we present the model specification, in the second subsection we discuss the empirical content of the aggregate LCS decomposition.

6.1 The empirical model

The whole model consists of a system with a price setting rule, an inverse output demand

function, five input demand equations and a cost function:

$$p_{nt}\left(1 + \lambda_n^a\left(w_{nt}, Y_{nt}, t\right)\frac{\partial P}{\partial Y}\frac{Y}{P}\right) = \frac{\partial \mathbf{C}_n}{\partial Y}\left(w_{nt}, Y_{nt}, t\right) - B_{CY}\left(w_{nt}, Y_{nt}, t\right) + u_{nt}^Y, \quad (37)$$

$$p_{nt} = P\left(Y_{nt}, z_{nt}\right) + u_{nt}^P, \quad (38)$$

$$X_{nt} = \mathbf{X}_n^*\left(w_{nt}, Y_{nt}, t\right) + u_{nt}^X \quad (39)$$

$$C_{nt} = \mathbf{C}_n\left(w_{nt}, Y_{nt}, t\right) + u_{nt}^C \quad (40)$$

Let $u_{nt} \equiv \left(u_{nt}^Y, u_{nt}^P, u_{nt}^X, u_{nt}^C\right)^\top$. We assume that $\mathrm{E}\left[u_{nt}|w_{nt}, z_{nt}\right] = 0$. Several assumptions on $\mathrm{E}\left[u_{nt}, u_{ms}^\top\right]$ compatible with different forms of heteroskedasticity and correlation over n and t are considered. In order to be able to identify the different parameters of (37), notably the parameters of $\partial P_n/\partial Y$ and $\partial \mathbf{C}_n/\partial Y$, we append the inverse output demand (38), the set of input demand functions (39) and the cost function (40) to the system and estimate their parameters simultaneously. As $C_{nt} = w_{nt}^\top X_{nt}$, and $u_{nt}^C = w_{nt}^\top u_{nt}^X$, the cost function does not add any new information not already entailed in the input demands when the whole system is estimated in levels. However, when estimation is in first differences things are different, because for any given n and t, the error terms $u_{nt}^C - u_{n,t-1}^C$ and $u_{nt}^X - u_{n,t-1}^X$ are linearly independent when w_{nt} is not proportional to $w_{n,t-1}$.

The industry inverse output demand function is assumed to be given by:

$$P\left(Y_{nt}, z_t\right) = \exp\left(\eta_Y \ln Y_{nt} + \frac{1}{2}\eta_{YY}\left(\ln Y_{nt}\right)^2 + \eta_{Yt}\ln Y_{nt}t + f_n^P + f_t^P\right). \quad (41)$$

Several variables are candidates for inclusion in the output demand shifting vector z_t in the expression of P_n. Most of them are macroeconomic variables like GDP, total population, number of unemployed, government expenditures, exports and imports, interest rate or GDP inflation rate. Specifying z_t explicitly may yield a wrong specification of the inverse demand function, so we prefer instead including unobserved time specific variables f_t^P as well as a time trend t in (41) for representing omitted z_t variables. It is therefore not possible to explicitly test for whether P_{nt} is separable in z_t or not (the identification requirement of Result 3i). This is not such a drawback, as we believe that identification is more likely to be achieved when some ad-hoc assumption on the distributional shift functions B_{CY} and λ are made, in the spirit of Result 3(ii). Though being restrictive, such parametric specifications are still far more general than those adopted in representative firms models. The specification also includes industry specific parameters f_n^P. There are $N + T + 2$ free parameters in (38).

From the definition (26) it is clear that most microeconomic properties of c are lost in the aggregate if the conditional distribution of market shares depends upon (w, Y, t). How should we specify the aggregate model in this case? For answering this question, let us write

$$\mathbf{C}\left(w, Y, t\right) = \int C\left(w, \beta Y, t\right) f\left(\beta|w, Y, t\right) d\beta$$

$$= \int C\left(w, \beta Y, t\right) g\left(\beta\right) d\beta + \int C\left(w, \beta Y, t\right) \left(f\left(\beta|w, Y, t\right) - g\left(\beta\right)\right) d\beta$$

$$\equiv \mathbf{C}_0\left(w, Y, t\right) + \mathbf{C}_1\left(w, Y, t\right) \quad (42)$$

where $g\left(\beta\right)$ denotes the marginal (joint) density of market shares. This decomposition shows that the aggregate cost function can always be additively decomposed into a

17

function $\mathbf{C}_0$ satisfying some microeconomic properties (especially linear homogeneity in w) and a perturbation function $\mathbf{C}_1$ which depends on the gap between the conditional and marginal distribution of the market shares. This decomposition is useful for the empirical specification of the cost function. Under Lewbel's (1996) approximate proportionality assumption, this last term is identically zero and only a well behaved aggregate cost function $\mathbf{C}_0$ remains.

The microeconomic properties of $\mathbf{C}_0$ are actually helpful for identifying $\mathbf{C}_1$: any departure of $\mathbf{C}$ from linear homogeneity, Shephard's lemma and concavity in prices can be attributed to distributional shifts captured by $\mathbf{C}_1$. However, in most cases, only parts of this function can be identified separately from $\mathbf{C}_0$.[5] This is not really a problem, since we are not interested in identifying $\mathbf{C}_1$ but in modelling of the aggregate cost function and its partial derivatives.

According to Corollary 2(i), the input demand system $\mathbf{X}^*$ cannot be obtained from $\mathbf{C}$ by applying Shephard's lemma. From (28), however, we know that:

$$\mathbf{X}^* (w, Y, t) = \frac{\partial \mathbf{C}}{\partial w} (w, Y, t) - B_{Cw} (w, Y, t), \tag{43}$$

which can be plugged into (39). Although the aggregate cost function loses most microeconomic properties, the adding-up property is still satisfied by $\mathbf{C}$, and this allows us to relate the aggregation bias B_{Cw} on input demands to the function $\mathbf{C}_1$. Indeed,

$$\mathbf{C} (w, Y, t) = w^\top \mathbf{X}^* (w, Y, t)$$

$$\Leftrightarrow \mathbf{C}_0 (w, Y, t) + \mathbf{C}_1 (w, Y, t) = w^\top \frac{\partial \mathbf{C}}{\partial w} (w, Y, t) - w^\top B_{Cw} (w, Y, t)$$

$$\Leftrightarrow w^\top B_{Cw} (w, Y, t) = w^\top \frac{\partial \mathbf{C}_1}{\partial w} (w, Y, t) - \mathbf{C}_1 (w, Y, t), \tag{44}$$

where the last line follows from the definition of $\mathbf{C}$ and linear homogeneity of $\mathbf{C}_0$ in w.

We assume that B_{Cwj} is specified as:

$$B_{Cwj} (w, Y, t) \equiv \frac{1}{w_j} \left(\beta_j + \beta_{wj}^\top \ln w + \beta_{Yj} \ln Y + \beta_{tj} t \right), \tag{45}$$

and thus $J (J + 3)$ free β-parameters are involved in (45). We then impose the adding-up conditions (44) in order to obtain:

$$\mathbf{C}_1 (w, Y, t) = \gamma_C + \gamma_{\ln w}^\top \ln w + \gamma_Y \ln Y + \gamma_t t + h (w, Y, t).$$

Function h represents an arbitrary function which is homogeneous of degree one in w, but remember (footnote 6) that it is not always possible to separately identify h from $\mathbf{C}_0$, so for simplicity, we set $h (w, Y, t) \equiv \gamma_w^\top w$. The parameters γ of $\mathbf{C}_1$ are linked to the β-parameters of B_{Cwj} by

$$\gamma_C = - \sum_{j=1}^{J} \beta_j - \sum_{j=1}^{J} \iota_J^\top \beta_{wj}, \quad \gamma_{\ln w} = - \sum_{j=1}^{J} \beta_{wj}, \quad \gamma_Y = - \sum_{j=1}^{J} \beta_{Yj}, \quad \gamma_t = - \sum_{j=1}^{J} \beta_{tj}.$$

This restriction shows that only J free parameters γ_w enter $\mathbf{C}_1$ in addition of those comprised in B_{Cwj}. According to Corollary 2, the representative firm assumption can

---

[5] Consider for instance

$$\mathbf{C}_0 = \alpha_w^\top w + w^\top (A_{wy} y + A_{wt} t)$$

which is linearly homogeneous in w and

$$\mathbf{C}_1 = \beta_0 + \beta_w^\top w + \beta_y y + \beta_t t.$$

Then the parameters α_w and β_w cannot be separately identified in the expression of the aggregate cost function $\mathbf{C}_0 + \mathbf{C}_1$.

also be tested using the estimates of the B_{Cwj} and B_{CY}.

The more traditional part of the cost function is specified as a translog functional form. This specification is flexible in (w, Y, t), so that both increasing and decreasing returns to scale are a priori possible:

$$
\begin{aligned}
\mathbf{C}_{0n}(w, Y, t) = \exp \Big[&\alpha_{0,n} + \alpha_{w,n}^\top \ln w + \alpha_{Y,n} \ln Y + \alpha_{t,n} t \\
&+ \frac{1}{2} (\ln w)^\top A_{ww} \ln w + \alpha_{wY}^\top \ln w \ln Y + \alpha_{wt}^\top (\ln w) t \\
&+ \frac{1}{2} \alpha_{YY} (\ln Y)^2 + \alpha_{Yt} \ln Y t + \frac{1}{2} \alpha_{tt} t^2 \Big].
\end{aligned} \tag{46}
$$

By convention $\ln w = (\ln w_1, \dots, \ln w_J)^\top$. The parameters α_{0n}, α_{wn}, α_{Yn}, α_{tn} are industry specific, hence the subscript n. We know from (42) that function $\mathbf{C}_{0n}$ is linearly homogenous in w, so the α-parameters satisfy the usual restrictions:

$$
\iota_J^\top \alpha_{wn} = 1, \quad \iota_J^\top A_{ww} = 0, \quad \iota_J^\top \alpha_{wY} = \iota_J^\top \alpha_{wt} = 0, \tag{47}
$$

where ι_J denotes a J-vector of ones. Matrix A_{ww} is symmetric. There are $N(J+2) + J(J+3)/2 + 1$ free α-parameters involved in (46).

Departure of $\mathbf{C}$ from linear homogeneity in w can be measured by:

$$
\mathbf{C} - w_{nt}^\top \frac{\partial \mathbf{C}}{\partial w} = \mathbf{C}_1 - w_{nt}^\top \frac{\partial \mathbf{C}_1}{\partial w} = \gamma_C + \gamma_{\ln w}^\top \ln w + \gamma_Y \ln Y + \gamma_t t - \iota_J^\top \gamma_{\ln w}.
$$

This term is identically zero iff

$$
\gamma_{\ln w} = 0, \gamma_C = \gamma_Y = \gamma_t = 0. \tag{48}
$$

The aggregation bias affecting the aggregate marginal cost is specified as

$$
B_{CY}(w, Y, t) \equiv \beta_Y + \beta_{wY}^\top \ln w + \beta_{YY} \ln Y + \beta_{tY} t, \tag{49}
$$

which includes $J + 3$ further parameters.

For simplicity, we specify the conditional mean of the Hirschman-Herfindahl index as:

$$
\lambda_n^a(w_{nt}, Y_{nt}, t) = \lambda_{0n} + \Lambda_w^\top \ln w_{nt} + \lambda_Y \ln Y_{nt} + \lambda_t t, \tag{50}
$$

which has the advantage of allowing industry specific markups through the inclusion of fixed effects λ_{0n}. For estimation, (49) and (50) are substituted into the regression equation (37).

6.2 Empirical expansion and substitution matrices

Our purpose is to identify the terms of the aggregate LCS decomposition:

$$
\frac{\partial X^N}{\partial w^\top}(w, z, t) = \frac{\partial X^*}{\partial w^\top}\left(w, \{y_h^N\}_{h=1}^H, t\right) + \sum_{h=1}^H \frac{\partial x_h^*}{\partial y_h}\left(w, y_h^N, t\right) \frac{\partial y_h^N}{\partial w^\top}(w, z, t), \tag{51}
$$

which was obtained from the theoretical relationship

$$
X^N(w, z, t) = X^*\left(w, \{y_h^N\}_{h=1}^H, t\right) = \sum_{h=1}^H x_h^*\left(w, y_h^N, t\right).
$$

When output quantities are optimally allocated, they are driven by (w, z, t), i.e. $y_h = y_h^N(w, z, t)$ and $Y = Y^N(w, z, t)$, and the distribution of market shares becomes

endogenous: $\beta = \beta^N(w, z, t)$ and we can write:

$$X^N(w, z, t) = X^*\left(w, \beta^N\left(w, Y^N, t\right) Y^N, t\right) = \mathbf{X}^*\left(w, Y^N, t\right). \tag{52}$$

This equation implies that:

$$\frac{\partial X^N}{\partial w^\top}(w, z, t) = \frac{\partial \mathbf{X}^*}{\partial w^\top}\left(w, Y^N, t\right) + \frac{\partial \mathbf{X}^*}{\partial Y}\left(w, Y^N, t\right)\frac{\partial Y^N}{\partial w^\top}(w, z, t). \tag{53}$$

All the terms involved in this expression can be obtained from the estimation of $\mathbf{X}^*$ and the aggregate first order condition (36). In general, there will be no one to one correspondence between the substitution and expansion matrices of (51) and (53), but this seems to be the best approximation we can think of. Applying the implicit function theorem to (36) yields:

$$\frac{\partial Y^N}{\partial w^\top}(w, z, t) = \frac{\dfrac{\partial^2 \mathbf{C}}{\partial Y \partial w} - \dfrac{\partial B_{CY}}{\partial w} - \dfrac{\partial P}{\partial Y}Y\dfrac{\partial \lambda^a}{\partial w}}{(1 + \lambda^a)\dfrac{\partial P}{\partial Y} + \dfrac{\partial P}{\partial Y}Y\dfrac{\partial \lambda^a}{\partial Y} + Y\lambda^a\dfrac{\partial^2 P}{\partial Y^2} - \dfrac{\partial^2 \mathbf{C}}{\partial Y^2} + \dfrac{\partial B_{CY}}{\partial Y}}. \tag{54}$$

This expression is then replaced into (53) in order to obtain matrix $\partial X^N / \partial w^\top$.

6.3 Econometric issues

The observed input quantities are related to the demand functions by:

$$\begin{aligned}
X_{nt} &= \mathrm{E}\left[X_{nt}|w_{nt}, z_{nt}, t\right] + u_{nt}^X \\
&= X_n^N(w_{nt}, z_{nt}, t) + u_{nt}^X \\
&= \mathbf{X}_n^*\left(w_{nt}, Y_n^N(w_{nt}, z_{nt}, t), t\right) + u_{nt}^X \\
&= \mathbf{X}_n^*\left(w_{nt}, Y_{nt} - u_{nt}^Y, t\right) + u_{nt}^X.
\end{aligned} \tag{55}$$

All equalities are equivalent representations of the relation of interest. The first equality corresponds to the definition of the conditional mean. The second equality follows from the hypothesis that the conditional mean of input quantities (given w_{nt}, z_{nt}, t) correspond to aggregate Nash equilibrium input demands X^N. The third equality is a consequence of (52), and the fourth is obtained because observed aggregate output is related to the Nash equilibrium output by $Y_{nt} = Y^N + u_{nt}^Y$. Unfortunately, these equations are not useful as such for parameter estimation: the second equation does not allow to disentangle substitution and expansion effects; the third equality is useless, because Y^N is not known; the fourth equality is not much more helpful because the error term u_{nt}^Y is not observed. Instead of that, we have to rely upon (39). However, as the regressor Y_{nt} of (39) is correlated with u_{nt}^X when both error terms u_{nt}^X and u_{nt}^Y are correlated (see (55)), the variable Y_{nt} is not a valid instrument for estimating the parameters of the system.

Imposing $\mathrm{E}\left[u_{nt}^X|w_{nt}, z_{nt}, t\right] = 0$ on the idiosyncratic error term of (39) allows to obtain consistent estimates for

$$\mathrm{E}\left[X_{nt}|w_{nt}, z_{nt}, t\right] = \mathrm{E}\left[\mathbf{X}_n^*\left(w_{nt}, Y - u_{nt}^Y, t\right)|w_{nt}, z_{nt}, t\right] \equiv \mathbf{X}_n^*\left(w_{nt}, \widehat{Y}(w_{nt}, z_{nt}, t), t\right), \tag{56}$$

which implicitly defines the function $\widehat{Y}$: the existence of the function $\widehat{Y}$ is a consequence of the mean value theorem for integration. In general, however, $\widehat{Y}(w_{nt}, z_{nt}, t) \neq \mathrm{E}[Y_{nt}|w_{nt}, z_{nt}, t]$. The following result states that function $\widehat{Y}$ coincides with the function

Y^N we are looking for:

Result 4. Under the above assumptions,
$$\widehat{Y}\left(w_{nt}, z_{nt}, t\right) = Y^N\left(w_{nt}, z_{nt}, t\right). \tag{57}$$

Indeed,
$$\mathrm{E}\left[X_{nt}|w_{nt}, z_{nt}, t\right] = \mathbf{X}_n^*\left(w_{nt}, Y_n^N\left(w_{nt}, z_{nt}, t\right), t\right) = \mathbf{X}_n^*\left(w_{nt}, \widehat{Y}_n\left(w_{nt}, z_{nt}, t\right), t\right).$$

The first equality follows from the definition of the conditional mean of X_{nt} and (55); the second equality is a consequence of the definition (56) of $\widehat{Y}$. Result 4 then follows from the (plausible) assumption that $\mathbf{X}_n^*$ is locally monotone in Y (on an interval covering Y^N and $\widehat{Y}$). Thus, the conditional mean function $\mathrm{E}\left[\mathbf{X}_n^*\left(w_{nt}, Y_{nt}, t\right)|w_{nt}, z_{nt}, t\right]$ can be used for identifying the substitution matrix $\partial \mathbf{X}^*/\partial w^\top\left(w_{nt}, Y_{nt}^N, t\right)$ and the vector of expansion effects $\partial \mathbf{X}^*/\partial Y\left(w_{nt}, Y_{nt}^N, t\right)$. The implicit function theorem can then be relied upon for obtaining estimates for $\partial Y^N/\partial w^\top\left(w, z, t\right)$. Result 4 justifies the use of the moments $\mathrm{E}\left[u_{jnt}^X Z_{nt}\right] = 0$ (contemporaneous exogeneity) with $Z_{nt} = I\left(w_{nt}, z_{nt}, t\right)$ a K-vector of instruments, and I denoting arbitrary square integrable functions of (w_{nt}, z_{nt}, t). However, Result 4 does not claim that the only valid instruments are provided by Z_{nt}, but only that these instruments should be considered in order to obtain an implicit function $\widehat{Y}$ compatible with the theoretical model.

7. Empirical results

The data we use are described in Appendix B.[6] We present the empirical results in five stages; firstly those concerning the (inverse) output demand function; second, those relative to the rate of return to scale and markups; third, all results describing shifts along the isoquant (from point A to B on Figure 1), and fourth, results corresponding to the total impact of input price changes on input demand and output supply (from point A to C on Figure 1).

When the regressions are expressed in level as in (37)-(40), the Durbin Watson statistics for first order serial correlation of u_{nt} are found to be extremely low (near to 0.3 in average), suggesting that $\mathrm{Cov}\left[u_{nt}u_{n,t-1}|Z_{nt}\right] \neq 0$. In order to try to avoid spurious regression problems, the parameters are estimated after expressing all regressions in first differences. So, the NJ^* observations for the first year are lost, where $J^* = J + 3$ denotes the number of regressions. The vector of differenced error terms is denoted by $v_{nt} \equiv u_{nt} - u_{n,t-1}$ and the moment conditions are given by $\mathrm{E}\left[v_{nt}Z_{n,t-1}^\top\right] = 0$, where Z_{nt} denotes the instruments, and comprises w_{nt}, $\ln w_{nt}$, $t\ln w_{nt}$, $(\ln w_{j,nt})^2$ as well as N industry and $T - 2$ time dummies. The discussion leading to Result 4 suggests $\mathrm{E}\left[u_{n,t-1}Z_{n,t-1}^\top\right] = 0$, and so the moment conditions $\mathrm{E}\left[v_{nt}Z_{n,t-1}^\top\right] = 0$ are likely to be fulfilled if $\mathrm{E}\left[u_{nt}Z_{n,t-1}^\top\right] = 0$.[7] The same instruments are used for determining all GMM

---

[6] The results are obtained using TSP 5.0.

[7] We also experiment with the moment conditions $\mathrm{E}\left[\Delta u_{nt}\Delta Z_{n,t}^\top\right] = 0$, however, the overidentification test rejected the validity of the instruments at the 1% threshold. Given the fact that $\mathrm{E}\left[\Delta u_{n,t}Z_{n,t-1}^\top\right] = 0$ was not rejected, the rejection of $\mathrm{E}\left[\Delta u_{nt}\Delta Z_{nt}^\top\right] = 0$ implies that $\mathrm{E}\left[\Delta u_{n,t}Z_{nt}^\top\right] \neq 0$, which may be due to violation of strict exogeneity.

estimators, they generate $99 \times J^*$ orthogonality conditions.

7.1 The inverse demand function

Table 1 reports estimates of the inverse output demand function (38), using different estimations methods: the regression is either in levels as in (38), in first difference (FD) or with AR1 residuals (AR1). There are $N(T-1) = 936$ observations used in each single equation regression for estimating 72 parameters (among which 69 dummy variables). The AR1 regression assumes that $u_{nt}^P = \rho u_{n,t-1}^P + \epsilon_{nt}$ with ϵ_{nt} iid and comprises one additional parameter ρ. The upper panel of Table 1 reports statistics obtained from estimating the inverse output demand's parameters from the single regression (38). The lower panel reports these estimates when all J^* regressions are simultaneously run.

The estimated values of the parameters exhibit some differences from one estimation method to the other. The sensitivity of output price to output quantity is found to be quite important: the inverse demand elasticity with respect to output, $\varepsilon(P;Y)$, ranges between -0.20 and -0.86 and is significantly different from zero. We also tested for industry specific heterogeneity in the elasticities by including heterogeneous parameters η_Y for each industry in the inverse demand specification (41). In all cases (FD, AR1), the test could not reject the homogeneity hypothesis ($\eta_{Y,n} = \eta_Y$).

There is relatively little difference between the regressions in levels: the nonlinear least squares which neglects endogeneity of Y and GMM using $Z_{n,t-1}$ as instruments (Table 1 lines 2 and 4) yield quite similar results. In the GMM regression in levels, the orthogonality between u_{nt}^P and $Z_{n,t-1}$ is rejected, which casts doubts on the validity of the identifying restrictions $\mathrm{E}\left[u_{nt} Z_{n,t-1}^\top\right] = 0$.

The main difference is between the estimates in levels and in FD: both the estimates of the parameters and of their t-value are different. In the nonlinear regression, the estimated value of the inverse demand elasticity $\varepsilon(P;Y)$ is divided by three when the estimation is in FD (Table 1, line 2). This results has dramatic implications for the estimated markup as discussed below.

The results for the inverse output demand obtained from the system SUR and GMM regressions are reported in the lower panel of Table 1. These results are relatively similar to the single equation results, but the t-values are often higher due to efficiency gains of system regressions. The whole system comprises 291 parameters to be estimated on the basis of $N(T-1)J^* = 7488$ observations. The same instruments are used for all equations. So there are $99 \times 8 = 792$ orthogonality conditions imposed for parameter identification. The convergence of the SUR estimator rests on the validity of the exogeneity of the output level, an assumption which is theoretically dubious and empirically rejected when it is tested using the GMM estimator. The regressions in levels, both in the SUR and GMM case, yield implausibly high t-values. This suggests that spurious regression contaminates the regressions in levels, leading to underestimated parameter variances, and also explaining the rejection in the overidentification tests in the regressions in levels. All these remarks lead us to prefer GMM estimates of the model expressed in first differences. With this estimator, the median value of the estimated inverse output demand elasticity is about -0.85 in the system regressions, somewhat below the single regression result of -0.70. There are also some differences between the estimated value of $Y \partial P^2/\partial Y^2$, which has some implications for testing the hypothesis of

the existence of a Cournot equilibrium. This gap may be due to problems in identifying an inverse output demand equation when only information on p and Y are used in conjunction with industry specific dummies. Including information on the cost function helps to identify what we are looking for. The overidentification test does not reject the instruments' validity in the FD and AR1 cases, which is conform to Result 4. We do not find contradictions between the estimates in FD and AR1, which is not surprising since $\widehat{\rho} = 0.979$ in the AR1 regression and since the FD specification is statistically not rejected.

Table 1: Estimates for the inverse output demand

Estimation method		η_Y	η_{YY}	η_{Yt}	$\varepsilon(P;Y)$	$\partial P/\partial Y$	$Y\partial P^2/\partial Y^2$	OIT
Single	level, NLSQ	−0.293 (−2.2)	−0.103 (−2.7)	0.005 (3.7)	−0.620 (−16.3)	−0.004 (−16.3)	0.006 (12.6)	−
equation	FD, NLSQ	0.031 (1.0)	0.003 (0.3)	−0.004 (−6.2)	−0.197 (−3.8)	−0.001 (−3.8)	0.001 (3.5)	−
	level, GMM	−1.019 (−6.9)	0.111 (2.4)	−0.002 (−1.1)	−0.590 (−14.8)	−0.004 (−14.8)	0.007 (12.5)	.00
	FD, GMM	−0.089 (−0.9)	−0.033 (−0.9)	−0.001 (−0.6)	−0.704 (−4.2)	−0.005 (−4.2)	0.005 (3.8)	.49
	AR1, GMM	−0.080 (−0.8)	−0.032 (0.7)	−0.005 (−1.9)	−0.749 (−5.4)	−0.005 (−5.4)	0.006 (4.7)	.58
System	level, SUR	−0.324 (−3.98)	−0.077 (−3.3)	0.003 (4.3)	−0.575 (−19.0)	−0.004 (−19.0)	0.006 (15.1)	−
	FD, SUR	−0.098 (−3.3)	0.003 (0.4)	−0.003 (−5.2)	−0.277 (−5.0)	−0.002 (−5.0)	0.002 (4.9)	−
	level, GMM	−0.740 (−11.9)	0.037 (1.6)	−0.000 (−0.6)	−0.595 (−22.7)	−0.004 (−22.7)	0.006 (21.1)	.00
	FD, GMM	−0.797 (−3.4)	−0.544 (−5.4)	0.021 (3.9)	−0.855 (−4.2)	−0.006 (−4.2)	0.020 (4.0)	.08

Columns 2 to 7 report the median value of the corresponding statistic over all observations as well as the median Student statistic in parentheses.

The last column reports the p-value of the overidentification test.

Turning to the sufficient conditions for the existence of a Nash equilibrium satisfied, in all cases, the inverse output demand is decreasing and $\partial P^2/\partial Y^2$ is found to be statistically significant and positive, and so the inverse demand is not concave. Columns 6 and 7 of Table 1 suggest that $\partial p/\partial Y + Y\partial p^2/\partial Y^2$ is slightly positive in most cases (unreported results show that it is significantly positive for the bulk of the observations). The weaker aggregate condition (18), which also ensures existence of the Cournot equilibrium, is numerically satisfied for values of the concentration index λ^a below 0.6 in the single regression, and λ^a below 0.1 for the parameter values obtained for the system of regressions. This finding confirms the plausibility of Presumption 1. With all specifications, the alternative sufficient condition for the existence of Cournot equilibrium, (16), is statistically rejected for the bulk of the observations.

7.2 Rate of return to scale and markup

In order to avoid that misspecification of the pricing and output demand equations (37), (38) contaminates the estimation of the rate of return to scale and the markup, we first consider only the system of 5 input demands (39) and the cost function (40) for parameter estimation. Then we consider the whole system (37)-(40). We also assess the

23

sensitivity of the empirical results with respect to the regression method used, and provide results based on seemingly unrelated regressions (SUR) which neglects endogeneity of the output level, and GMM estimates, where the choice of the instruments is guided by Result 4 above.

Table 2a: SUR estimates of the rate of returns to scale, markup, and λ^a

Estimation method	Quartile	$C_1 \equiv 0$			$C_1 \neq 0$		
		$\varepsilon(\mathbf{C};Y)$	$p/(\partial \mathbf{C}/\partial Y)$	λ^a	$\varepsilon(\mathbf{C};Y)$	$p/(\partial \mathbf{C}/\partial Y)$	λ^a
SUR	0.25	0.36 (−3.9)	1.09 (0.2)	−	0.42 (−8.8)	1.15 (0.7)	−
(39)-(40)	0.50	0.66 (−0.8)	1.46 (0.8)	−	0.62 (−3.9)	1.66 (1.8)	−
level	0.75	0.89 (−0.9)	2.20 (2.1)	−	0.89 (−0.9)	2.45 (3.6)	−
SUR	0.25	0.42 (−7.0)	1.32 (1.5)	0.49 (3.5)	0.35 (−12.0)	1.58 (2.3)	0.42 (2.6)
(37)-(40)	0.50	0.56 (−4.8)	1.84 (2.1)	0.80 (4.8)	0.50 (−8.0)	2.12 (3.3)	0.83 (5.8)
level	0.75	0.74 (−2.7)	2.36 (3.6)	1.06 (6.8)	0.68 (−4.0)	2.98 (5.1)	0.99 (7.5)
SUR	0.25	0.61 (−5.4)	1.13 (0.8)	−	0.64 (−5.1)	1.14 (0.9)	−
(39)-(40)	0.50	0.79 (−2.5)	1.29 (1.9)	−	0.81 (−2.6)	1.28 (2.1)	−
FD	0.75	0.90 (−0.9)	1.64 (3.3)	−	0.90 (−0.9)	1.58 (3.4)	−
SUR	0.25	0.56 (−8.7)	1.32 (2.1)	1.84 (1.5)	0.62 (−8.7)	1.29 (2.1)	1.20 (1.9)
(37)-(40)	0.50	0.71 (−4.3)	1.46 (3.0)	1.42 (3.0)	0.75 (−3.9)	1.37 (3.0)	1.77 (3.1)
FD	0.75	0.78 (−2.9)	1.83 (4.7)	1.91 (4.8)	0.80 (−2.4)	1.69 (4.6)	2.44 (4.9)

The table reports the 0.25, 0.50 and 0.75 quantiles (over the $N(T-1) = 936$ observations) of the estimated values of $\varepsilon(c,y)$, $p/(\partial c/\partial y)$ and λ^a. The t-statistics for the hypotheses $\varepsilon(c,y) = 1$, $p/(\partial c/\partial y) = 1$ and $\lambda^a = 0$ appear in parentheses.

The equation system is nonlinear and we find several parameter values corresponding to local minimum of the SUR and GMM objective function, depending on which residual variance matrix and which parameters' starting values are chosen. In order to compare the results from different regressions we use the same starting values in all regressions. The residual variance matrix is obtained from the residuals of a first stage regression in which the residual variance matrix was set to the identity.

Different quantiles (over years and industries) of the estimated values of the rate of return to scale and markup are reported in Tables 2a and 2b. According to the SUR-results there is strong evidence for high rates of return to scale, high markups and high Hirschman-Herfindahl indices. The median value of $\widehat{\varepsilon}(\mathbf{C};Y)$ is comprised between 0.50 and 0.81 which correspond to high returns to scale. The median gap between price and marginal cost is ranging between 28 and 66%, and the median HH-index of concentration is comprised between 0.80 and 1.77! The very high estimates for $\widehat{\lambda}^a$ occur both in the representative firm model ($C_1 \equiv 0$) and the aggregate model. This is surprising in the light of Result 3(ii), which states that λ^a is identified when $C_1 \equiv 0$. This result is actually due to the fact that λ^a and $\varepsilon(P;Y)$ play a symmetric role in the nonlinear regression

24

(37), and so, ceteris paribus, small estimates (in absolute value) of $\varepsilon(P; Y)$ tend to be compensated by large values of λ^a. For first difference, Table 1 reports that the NLSQ estimates of $\varepsilon(P; Y)$ is indeed quite small, which explains why $\widehat{\lambda}^a$ is quite large in Table 2a. There are some relatively small differences in the results obtained from estimated the limited set of regressions (39)-(40) or the whole set of regressions (37)-(40). Both representative firm ($C_1 \equiv 0$) and aggregate model ($C_1 \neq 0$) yields relatively similar results.

Result 4 suggests to use instrumental variables for the estimation, and so we turn to the GMM estimates reported in Table 2b. The same specifications as in Table 2a are now estimated using $Z_{n,t-1}$ as instruments, and we minimize the length of either $\mathrm{E}\left[u_{nt}Z_{n,t-1}^\top\right]$ or $\mathrm{E}\left[v_{nt}Z_{n,t-1}^\top\right]$ in the appropriate metric.

In levels, the overidentification tests rejects the validity of the instruments $Z_{n,t-1}$. The t-statistics are quite high, and we find a strong evidence for increasing returns to scale and markup pricing. In first differences, however, the overidentifying restrictions are not rejected at the 1% threshold. The t-statistics become much smaller in absolute value. When comparing the estimates based on the regression system (39)-(40) and those of the whole system (37)-(40), we found relatively small differences and no contradictions. Note however that the only specification which is not rejected at the 5% level by the overidentification test is for $C_1 \neq 0$, first difference, and the complete system of equations.

The inclusion of additional observations through the consideration of the further regressions (37)-(38), allows us to obtain somewhat higher t-values, without really affecting the level of the estimates.

Although there is evidence for increasing returns to scale (IRS), since most estimates of $\varepsilon(\mathbf{C}, Y)$ are smaller than one, the IRS hypothesis is rejected for more than 50% of the observations. The results also confirm the existence of markup pricing (the median value of $p/(\partial \mathbf{C}/\partial Y)$ is 1.17), but the markup is significantly greater than one for only about 25% of the observations. These results are broadly compatible with those of Diewert and Fox (2008), who work with the same dataset, and also found strong evidence for increasing returns to scale. Our SUR estimates of the model specified in levels are quite close to those of Diewert and Fox. Our GMM results for the model in FD are somewhat more contrasted and suggest that many observations (about 50%) are not incompatible with the assumption of constant returns to scale, but at the same time, there is also evidence for increasing returns to scale for at least 25% of the observations.

Table 2b also reports the 0.25, 0.5 and 0.75 quantiles of the estimated HH-index of concentration and its corresponding t-value. The median value of the estimated HH-index is 0.17 and it is significantly different from zero for more than 50% of the observations. These estimates of λ^a are more plausible than those reported in Table 2a, which is directly related to the fact that the estimate of the inverse output demand elasticity $\varepsilon(P; Y)$ obtained with GMM is more realistic. If we integrate this estimated value of λ^a into our formulation (18) of the existence of the Cournot-Nash equilibrium, we cannot reject the existence assumption. This empirical finding provides empirical support for Result 2.

Table 3 reports the empirical correlation (over 18 industries) between the average value (over time) of λ^a, $\varepsilon(C, Y)$, $p/(\partial C/\partial Y)$, $\varepsilon(P, Y)$, s_π, s_π^a. Few correlations are above 0.3. These correlations suggest that the higher the implicit HH-index λ^a, the higher

Table 2b: GMM estimates of the rate of return to scale, markup, and λ^a

Estimation method	Quartile	$C_1 \equiv 0$				$C_1 \neq 0$			
		$\varepsilon(\mathbf{C};Y)$	$p/(\partial\mathbf{C}/\partial Y)$	λ^a	OIT	$\varepsilon(\mathbf{C};Y)$	$p/(\partial\mathbf{C}/\partial Y)$	λ^a	OIT
GMM	0.25	0.43 (−5.3)	0.95 (−0.5)	−		0.18 (−4.5)	0.48 (−1.5)	−	
(39)-(40)	0.50	0.66 (−1.7)	1.51 (0.8)	−	0.00	0.67 (−2.6)	1.00 (0.0)	−	0.00
level	0.75	1.07 (0.8)	2.33 (2.6)	−		1.21 (0.7)	1.68 (2.1)	−	
GMM	0.25	0.73 (−6.0)	0.97 (−0.3)	0.11 (1.2)		0.29 (−8.3)	0.81 (−1.8)	−0.22 (−1.6)	
(37)-(40)	0.50	0.87 (−1.8)	1.17 (1.7)	0.20 (3.4)	0.00	0.74 (−2.7)	1.15 (0.8)	0.12 (0.9)	0.00
level	0.75	1.00 (0.0)	1.35 (4.9)	0.37 (8.1)		1.02 (0.3)	1.83 (2.7)	1.00 (5.1)	
GMM	0.25	0.59 (−2.8)	0.91 (−0.3)	−		0.59 (−2.1)	0.91 (−0.2)	−	
(39)-(40)	0.50	0.88 (−0.7)	1.19 (0.6)	−	0.03	0.92 (−0.5)	1.05 (0.3)	−	0.04
FD	0.75	1.13 (0.4)	1.78 (1.3)	−		1.11 (0.3)	1.68 (1.1)	−	
GMM	0.25	0.64 (−5.6)	0.97 (−0.2)	0.05 (0.6)		0.64 (−3.5)	1.02 (0.1)	0.07 (0.7)	
(37)-(40)	0.50	0.89 (−1.2)	1.14 (0.7)	0.11 (2.5)	0.03	0.87 (−1.1)	1.17 (0.8)	0.13 (2.5)	0.08
FD	0.75	1.03 (−0.2)	1.48 (2.1)	0.20 (4.1)		0.99 (0.0)	1.54 (1.9)	0.23 (5.1)	

The table reports the 0.25, 0.50 and 0.75 quantiles (over the $N(T-1) = 936$ observations) of the estimated values of $\varepsilon(c,y)$, $p/(\partial c/\partial y)$ and λ^a. The t-statistics for the hypotheses $\varepsilon(c,y) = 1$, $p/(\partial c/\partial y) = 1$ and $\lambda^a = 0$ appear in parentheses. The OIT column reports the p-value of the overidentification test.

Table 3: Correlation between different estimated measures of imperfect competition

	λ^a	$\varepsilon(C,Y)$	$p/(\partial C/\partial Y)$	$\varepsilon(P,Y)$	s_π
$\varepsilon(C,Y)$	−0.77				
$p/(\partial C/\partial Y)$	−0.25	0.17			
$\varepsilon(P,Y)$	0.04	−0.28	−0.04		
s_π	0.20	0.10	−0.26	−0.20	
s_π^a	0.31	0.00	−0.22	−0.11	0.86

The entries correspond to the empirical correlation between the estimated values of the variables of the respective column and line. In a first stage, the variables are averaged over time, then the correlation is calculated over 18 industries, using only within industry variations of averages for calculating the correlation.

The values used for s_π and s_π^a are those depicted in Figure B4.

is the rate of return to scale $1/\varepsilon(C, Y)$ and there is a strong statistical relationship between these measures as $\mathrm{corr}(\lambda^a, \varepsilon(C, Y)) = -0.77$. We find that industries with a higher rate of return to scale have relatively low markups. This seems surprising at first sight, just as the finding of no important correlation between the profit rate and the rate of return to scale. Similarly, the correlation between the markup and the rate of profit is found to be negative. There is, however, some rationale behind these results: industries exhibiting IRS have to adopt markup pricing behavior in order to avoid losses, but do not necessarily benefit from a high profit rate, notably because they may not be able to capture the consumers of their product (which is consistent with $\mathrm{corr}(\varepsilon(C, Y), \varepsilon(P, Y)) = -0.28$).

7.3 Test of the representative firm assumption

As discussed in Sections 5 and 6, the representative firm model is obtained for $\mathbf{C}_1(w, Y, t)$ homogeneous of degree one in w, which is the case iff (48) is satisfied, but this hypothesis is statistically rejected at the 1% threshold.

The value of the different types of aggregation bias, expressed in percentage of the corresponding input quantity (or marginal cost) are reported in Table 4 for FD-GMM. The bias affecting capital, labour and material input demands is small and insignificant. The biases related to energy and service input demand are important, but they are not significant either. Shephards' lemma ($B_{Cj} = 0$) and the hypothesis that C is homogeneous of degree zero in input prices ($w^\top B_{Cj} = 0$) are rejected for only about 20% of the observations (but these hypotheses are rejected globally). Although there is some statistical evidence for the existence of aggregation biases, neglecting them has almost no consequences on the empirical results obtained: the left and right parts of Tables 2 are very similar.

Table 4: Estimates of the aggregation biases (FD GMM)

Bias	w_k	w_ℓ	w_e	w_m	w_s	Y
B_{Cj}	-0.00	-0.07	0.60	-0.03	0.69	-0.03
	(-0.3)	(-1.4)	(1.6)	(-1.3)	(1.7)	(-0.2)

The table reports the median value of the estimated aggregation biases, expressed in percentage of the corresponding input quantity. The aggregation bias on marginal costs, B_{CY}, is expressed in percentage of marginal cost. The t-statistic for the null that $B_{Cj} = 0$ appears in parentheses.

The bias affecting marginal costs represents about 3% of marginal cost (and about 4% of the measure of the return to scale) and is also statistically insignificant. The negative estimate of B_{CY} means that market share reallocations (changes in the distribution of output shares) following an increase in aggregate output Y only contributes marginally to reduce aggregate cost. If we rewrite, the conditional expectation of (37) as

$$\frac{\partial \mathbf{C}}{\partial Y}\frac{Y}{C} = \frac{PY}{C} + \frac{PY}{C}\lambda^a \varepsilon(P; Y) + \frac{Y}{C}B_{CY},$$

and replace the different terms of this last relationship with the median value of their corresponding estimates, we obtain

$$0.87 \simeq 1.06 - 0.13 - 0.04.$$

27

So, an alternative measure of the rate of return to scale, which excludes the cost reduction through the shift in the distribution of market shares, is given by

$$\widehat{\frac{\partial \mathbf{C}}{\partial Y}}\frac{Y}{C} - \frac{Y}{C}\widehat{B}_{CY} = 0.91$$

which is now closer to constant returns to scale for most observations. Thus, this framework provides a partial explanation for a well-documented empirical contradiction between results obtained from microeconomic and more aggregated data (see for instance Basu and Fernald (1997), Morrison-Paul and Siegel (1999)).

7.4 Substitution effect

Table 5 reports the median value of the input demand elasticities over the 18 industries and over all years. These elasticities are calculated for a given aggregate output level, and correspond to the aggregate substitution effect. They give the sensitivity of the aggregate input demand $\varepsilon\left(\mathbf{X}_j^*; w_i\right)$ and differ somewhat from $\varepsilon\left(X_j^*; w_i\right)$ which exclude shifts in the distribution of market shares.

Table 5: Elasticities for the aggregate substitution effect

$\varepsilon\left(\mathbf{X}_j^*; w_i\right)$	$\mathbf{X}_k$	$\mathbf{X}_\ell$	$\mathbf{X}_e$	$\mathbf{X}_m$	$\mathbf{X}_s$
w_k	−0.21 (−1.1)	0.00 (0.4)	0.14 (4.1)	0.00 (0.2)	0.08 (2.9)
w_ℓ	0.23 (2.5)	−0.27 (−4.4)	−0.15 (−0.6)	0.09 (1.3)	−0.12 (−0.7)
w_e	−0.12 (−3.1)	−0.00 (−0.2)	−0.36 (−1.2)	0.06 (4.3)	0.03 (0.7)
w_m	0.21 (2.5)	0.14 (2.3)	0.24 (1.2)	−0.28 (−6.7)	0.47 (3.4)
w_s	−0.08 (−1.0)	0.12 (1.8)	−0.06 (−0.5)	0.17 (2.7)	−0.76 (−2.3)
Y	0.40 (1.9)	0.76 (4.6)	0.49 (1.2)	0.92 (5.8)	0.70 (2.2)
t	−0.005 (−0.5)	−0.014 (−2.4)	0.012 (1.4)	−0.008 (−1.7)	0.014 (1.2)

The table reports the median value of the elasticities over all observations and the corresponding t-statistic in parentheses.

The upper panel of Table 5 gives the input price elasticities with the own-price elasticities on the main diagonal. It can be seen that the own-price elasticities are nonpositive, which is consistent with $\varepsilon\left(X_j^*; w_i\right) \leq 0$.

Most of the inputs are found to be substitutes, but substitution is rather limited as all cross price elasticities are below 0.25 in absolute value. Only two inputs are found to be significantly complements: capital and energy for a change in the energy price. Notice however, that this result is not symmetric: energy and capital are substitutes for changes in the capital price. These findings, however, are not contradictory: our aggregate model does not imply that matrix $\partial \mathbf{X}^*/\partial w^\top$ is symmetric, because increases in energy and capital input prices may have different impacts on the distribution of market shares $f(\beta|w, Y, t)$. So, according to (43), input demands can react asymmetrically to cross price variations. This empirical finding could explain why the literature reports a variety of contradictory results on the sign of this elasticity (see for instance Frondel

and Schmidt, 2002). The lower panel of Table 5 reports the impact of a marginal change in output $\varepsilon\left(\mathbf{X}_j^*; Y\right)$ and of time $\varepsilon\left(\mathbf{X}_j^*; t\right)$.[8] All conditional input demands are found to be nondecreasing in the output level. Technological change is not neutral, but labour saving and energy intensive.

7.5   Expansion effect

The amended Novshek type of assumption (18) is not rejected for the bulk of the observations. So, we expect that the Cournot equilibrium exists and that the LCS principle is satisfied at the aggregate level of the industry, in accordance with Presumption 1.

Table 6 comprises the median value of the estimates for the total impact of input prices on equilibrium output quantity Y^N and price p^N and input demands X^N. The first row of Table 6 shows that the median values of $\varepsilon\left(Y^N, w_h\right)$ are all negative, which confirms Presumption 1(ii), but they are not significantly different from zero. There is also some weak evidence for the inflationary impact of rising input prices as all $\varepsilon\left(p^N, w_h\right)$ are found to be nonnegative (but not significantly). This lack of precision is rather common in applications using GMM. Here it is due to the fact that the denominator of (54), which was never found to be significantly positive, is not always *significantly* negative. The median value of the denominator is estimated at -0.002 with an asymptotic t-value of -1.4. A careful inspection of the different terms of the denominator shows that $\partial^2 C / \partial Y^2$ is estimated to be significantly negative, and so the marginal cost is found to be decreasing. This lack of precision also contaminates the estimates of $\varepsilon\left(\mathbf{X}_j^N; w_i\right)$.

For the median values reported in Table 5, the LCS principle is satisfied: we can approximate the unrestricted demand elasticities by

$$\varepsilon\left(\mathbf{X}_\ell^N; w_\ell\right) \simeq \varepsilon\left(\mathbf{X}_\ell^*; w_\ell\right) + \varepsilon\left(\mathbf{X}_\ell^*; Y\right) \varepsilon\left(Y^N, w_\ell\right).$$

As the median values of $\varepsilon\left(\mathbf{X}_\ell^*; Y\right)$ and $\varepsilon\left(Y^N, w_\ell\right)$ are positive and nonpositive, respectively (Table 6), it turns out that $\varepsilon\left(\mathbf{X}_\ell^N; w_\ell\right) \leq \varepsilon\left(\mathbf{X}_\ell^*; w_\ell\right)$. However, the median values of the unrestricted elasticities reported in Table 6 do not always satisfy the LCS principle due to sampling variation. Notice that the input demand functions $\mathbf{X}_j^N$ are not homogeneous of degree zero in w and that most cross price elasticities $\varepsilon\left(\mathbf{X}_j^N; w_i\right)$ are found to be negative, due to the fact that the expansion effect (which is negative) is important in absolute value. So, expansion, or price induced change in the output level, matters for explaining changes in aggregate input demand.

How can the mostly negative elasticities of Table 6 be reconciled with the upward trending quantities reported on Figure 2? Table 5 showed that most of the growth in input demands is due to output growth, output growth in turn is hampered by input price increases, and rises with shifts in output demand.

What about the empirical validity of Presumption 2? According to this statement there is no systematic link between the degree of market power and the strength of aggregate input demand adjustment. Table 7 reports the empirical correlation (over sectors) between the mean (over time) of the estimated input demand elasticities and some indicators of imperfect competition. For all inputs except energy, there is some

---

[8] In fact, $\varepsilon\left(\mathbf{X}_j^*; t\right)$ is a growth rate or a semi-elasticity defined as $\varepsilon\left(\mathbf{X}_j^*; t\right) \equiv \partial \ln \mathbf{X}_j^* / \partial t$.

Table 6: Elasticities for total aggregate impact of input price changes

$\varepsilon\left(\mathbf{X}_j^N; w_i\right)$	Y	p	$\mathbf{X}_k$	$\mathbf{X}_\ell$	$\mathbf{X}_e$	$\mathbf{X}_m$	$\mathbf{X}_s$
w_k	−0.01 (−0.3)	0.01 (0.3)	−0.18 (−0.7)	−0.02 (−0.9)	0.11 (2.2)	−0.02 (−0.8)	0.04 (1.4)
w_ℓ	−0.22 (−0.8)	0.19 (0.9)	0.12 (0.6)	−0.37 (−1.8)	−0.35 (−0.7)	−0.09 (−0.3)	−0.26 (−0.6)
w_e	−0.04 (−0.7)	0.05 (0.8)	−0.12 (−1.8)	−0.03 (−0.9)	−0.44 (−1.0)	0.02 (0.7)	−0.04 (−0.5)
w_m	−0.37 (−1.5)	0.45 (1.7)	−0.01 (−0.9)	−0.20 (−0.9)	−0.06 (−0.1)	−0.68 (−1.9)	0.03 (0.7)
w_s	−0.06 (−0.6)	0.07 (0.6)	−0.12 (−0.9)	0.01 (0.1)	−0.13 (−0.5)	0.04 (0.4)	−0.67 (−1.4)

The table reports the median value of the elasticities over all observations and the corresponding Student statistic in parentheses.

Table 7: Correlation between expansion effects and indicators of imperfect competition

	X_k	X_ℓ	X_e	X_m	X_s
λ^a	0.22	0.31	−0.03	0.20	0.14
$\varepsilon\left(C, Y\right)$	−0.18	−0.42	−0.00	−0.07	0.04
$p/\left(\partial C/\partial Y\right)$	0.09	−0.05	0.19	−0.00	−0.07
$\varepsilon\left(P, Y\right)$	0.15	0.08	−0.12	0.11	0.13
s_π	−0.19	−0.07	−0.05	−0.17	0.04

The entries correspond to the empirical correlation between the estimated values of the expansion effects $\varepsilon\left(\mathbf{X}_j^*; Y\right)\varepsilon\left(Y^N, w_j\right)$ and some indicators of imperfect competition. In a first stage, the variables are averaged over time, then the correlation is calculated over 18 industries, using only within industry variations of averages for calculating the correlation.

positive statistical correlation between $\varepsilon\left(\mathbf{X}_j^*; Y\right) \varepsilon\left(Y^N, w_j\right)$ and $\varepsilon\left(P, Y\right)$, which means that on average, industries with a lower $\varepsilon\left(P, Y\right) < 0$ react more intensively to own price changes. This result is somewhat surprising, as the converse result was expected to hold at the firm level (see the quotation of Cahuc and Zylberberg in Section 4.2). We also find that expansion effects are somewhat smaller (in absolute value) in industries with high profit rates.

8. Conclusion

Output adjustments have important consequences on input demands. This impact, however, is rarely quantified in economic contributions, because with imperfect competitive output markets, increasing returns to scale and externalities disturb the usual representative firm's comparative statics. This paper makes three contributions to the literature: (i) it derives the circumstances under which the LCS principle holds in an aggregate Cournot economy with heterogeneous firms; (ii) it provides a framework amending the representative firm model for the specification of aggregate input demand and output supply functions; (iii) it shows that empirically the LCS principle is inherited at the level of two-digit US manufacturing industries.

Empirical results show that, in US manufacturing, input substitution is rather limited for a given output level. Output adjustments imply further important changes in the input mix. We find empirical support for increasing returns to scale and markup pricing of moderate size. This is important for understanding aggregate growth, investment and employment. There is, however, much heterogeneity over industries and time, and about 50% of the observations are compatible with a rate of returns to scale near to one, and a markup close to zero.

Aggregation effects arise through shifts of market shares over firms within an industry. These shifts seem to play a rather small but significant role for explaining the evolution of aggregate input demands. With only aggregate data at hand, however, these effects are difficult to identify with precision. Despite theoretical and empirical support in favor of identification, the empirical results turn out to be sensitive to the choice of specification and instruments.

9. Appendix A: Proof of the Results

Discussion of Presumption 1. Claim (i) is a direct consequence of Result 1.

Claim (ii): differentiating

$$P\left(Y^N, z\right) + \frac{\partial P\left(Y^N, z\right)}{\partial y} y_h^N = \frac{\partial c_h\left(w, y_h^N\right)}{\partial y}$$

with respect to w, we obtain

$$\frac{\partial P}{\partial y}\left(Y^N, z\right) \frac{\partial Y^N}{\partial w} + \frac{\partial^2 P}{\partial y^2} \frac{\partial Y^N}{\partial w} y_h^N + \frac{\partial P}{\partial y} \frac{\partial y_h^N}{\partial w} = \frac{\partial x_h^*}{\partial y} + \frac{\partial^2 c_h}{\partial y^2} \frac{\partial y_h^N}{\partial w}.$$

Taking market-shares as weights, the weighted sum of this relationship over the firms

yields

$$\frac{\partial P}{\partial Y}\left(Y^N, z\right)\frac{\partial Y^N}{\partial w} + \frac{\partial^2 P}{\partial Y^2}\frac{\partial Y^N}{\partial w}Y^N\sum_{h=1}^{H}\left(\frac{y_h^N}{Y^N}\right)^2 + \frac{\partial P}{\partial Y}\sum_{h=1}^{H}\frac{\partial y_h^N}{\partial w}\frac{y_h^N}{Y^N} \tag{58}$$

$$= \sum_{h=1}^{H}\frac{\partial x_h^*}{\partial y_h}\frac{y_h^N}{Y^N} + \sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y^2}\frac{\partial y_h^N}{\partial w}\frac{y_h^N}{Y^N} \tag{59}$$

Let us denote the Hirschman-Herfindahl index of concentration (at Cournot equilibrium) by

$$\lambda \equiv \sum_{h=1}^{H}\left(\frac{y_h^N}{Y^N}\right)^2.$$

Notice that

$$\sum_{h=1}^{H}\frac{\partial y_h^N}{\partial w}\frac{y_h^N}{Y^N} = Y^N\frac{1}{2}\frac{\partial\lambda}{\partial w} + \lambda\frac{\partial Y^N}{\partial w}.$$

Moreover,

$$\mathrm{cov}\left(\frac{\partial^2 c_h}{\partial y^2}\frac{y^N}{Y^N}, \frac{\partial y^N}{\partial w}\right) = \frac{1}{H}\sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y^2}\frac{\partial y_h^N}{\partial w}\frac{y_h^N}{Y^N} - \frac{1}{H}\frac{\partial Y^N}{\partial w}\frac{1}{H}\sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y_h^2}\frac{y_h^N}{Y^N}.$$

These expressions can be substituted into (58) for obtaining:

$$\frac{\partial Y^N}{\partial w} = \frac{\sum_{h=1}^{H}\frac{\partial x_h^*}{\partial y_h}\frac{y_h^N}{Y^N} - \frac{\partial P}{\partial Y}\frac{Y^N}{2}\frac{\partial\lambda}{\partial w} + H\mathrm{cov}\left(\frac{\partial^2 c}{\partial y^2}\frac{y^N}{Y^N}, \frac{\partial y^N}{\partial w}\right)}{(1+\lambda)\frac{\partial P}{\partial Y} + Y^N\lambda\frac{\partial^2 P}{\partial Y^2} - \frac{1}{H}\sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y^2}\frac{y_h^N}{Y^N}}. \tag{60}$$

The numerator is likely to be positive, because (i) most firms have normal input demands ($\partial x_h^*/\partial y_h \geq 0$ for most h), (ii) the inverse output demand function is likely to decrease in Y and the Hirschman-Herfindahl index to increase in w, and (iii) the covariance term is likely to be positive by (5) and (7). Notice that this covariance term is identically zero when firms are identical, or when they have constant returns to scale (as in Example 1). In order to show that the denominator is negative, let us aggregate the inverse of (5) to obtain:

$$-\sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y_h^2}\frac{y_h^N}{Y^N} < -2\frac{\partial P}{\partial Y} - \frac{\partial^2 P}{\partial Y^2}Y^N\lambda.$$

This expression can then be used for majoring the denominator of (60):

$$(1+\lambda)\frac{\partial P}{\partial Y} + Y^N\lambda\frac{\partial^2 P}{\partial Y^2} - \frac{1}{H}\sum_{h=1}^{H}\frac{\partial^2 c_h}{\partial y^2}\frac{y_h^N}{Y^N}$$

$$< \left(1+\lambda - \frac{2}{H}\right)\frac{\partial P}{\partial Y} + \left(1 - \frac{1}{H}\right)Y^N\lambda\frac{\partial^2 P}{\partial Y^2}$$

$$\leq \left(1 - \frac{1}{H}\right)\left(\frac{\partial P}{\partial Y} + \frac{\partial^2 P}{\partial Y^2}Y^N\lambda\right). \tag{61}$$

where the last inequality follows from $\lambda \geq 1/H \geq 0$ and $\partial P/\partial Y \leq 0$. Then, (17) implies that (61) and the denominator in (60) are negative and so $\partial Y^N/\partial w \leq 0$.

From the definition of X^N,

$$X^N(w, z) \equiv \sum_{h=1}^{H} x_h^N(w, z) = \sum_{h=1}^{H} x_h^*\left(w, y_h^N(w, z)\right) = \sum_{h=1}^{H} x_h^o\left(w, Y_{-h}^N(w, z), z\right),$$

it follows that

$$
\begin{aligned}
\frac{\partial X^N}{\partial w^\top}(w, z) &= \sum_{h=1}^{H} \frac{\partial x_h^*\left(w, y_h^N\right)}{\partial w^\top} + \sum_{h=1}^{H} \frac{\partial x_h^*\left(w, y_h^N\right)}{\partial y} \frac{\partial y_h^N(w, z)}{\partial w^\top} \\
&= \frac{\partial X^*}{\partial w^\top}\left(w, \left\{y_h^N\right\}_{h=1}^{H}\right) \\
&\quad + \left(\frac{1}{H} \sum_{h=1}^{H} \frac{\partial x_h^*\left(w, y_h^N\right)}{\partial y}\right) \frac{\partial Y^N(w, z)}{\partial w^\top} + H \operatorname{cov}\left(\frac{\partial x_h^*\left(w, y_h^N\right)}{\partial y}, \frac{\partial y_h^N(w, z)}{\partial w^\top}\right)
\end{aligned}
\tag{62}
$$

where the last equality comes from the definition of the covariance:

$$\operatorname{cov}\left(\frac{\partial x_h^*}{\partial y}, \frac{\partial y_h^N}{\partial w^\top}\right) = \frac{1}{H} \sum_{h=1}^{H} \frac{\partial x_h^*}{\partial y} \frac{\partial y_h^N}{\partial w^\top} - \left(\frac{1}{H} \sum_{h=1}^{H} \frac{\partial x_h^*}{\partial y}\right)\left(\frac{1}{H} \sum_{h=1}^{H} \frac{\partial y_h^N}{\partial w^\top}\right).$$

The first matrix $\partial X^*/\partial w^\top$ of the decomposition (62) is negative semidefinite, direct consequence of the negative semidefinitness of $\partial x_h^*/\partial w^\top$. The vector $\partial Y^N/\partial w$ is likely to be nonpositive by (ii) and when most firms have normal input demands, the matrix $\left(H^{-1} \sum_{h=1}^{H} \partial x_h^*/\partial y\right) \partial Y^N(w, z)/\partial w^\top$ will also comprise nonpositive components. The covariance matrix in (62) is likely to have negative entries on its main diagonal by (7). This shows that it is plausible that $\varepsilon\left(X_j^N; w_j\right) < \varepsilon\left(X_j^*; w_j\right) \le 0$. Notice that $\partial X^N/\partial w^\top$ is not symmetric in general.

Claim (iii): let us show that $\partial Y^o/\partial w \le \partial Y^N/\partial w$. This inequality follows from the definition of $Y^N = \sum_{h=1}^{H} y_h^N(w, z) = \sum_{h=1}^{H} y_h^o\left(w, Y_{-h}^N, z\right)$, which yields

$$\frac{\partial Y^N}{\partial w^\top} = \sum_{h=1}^{H} \frac{\partial y_h^o}{\partial w^\top} + \sum_{h=1}^{H} \frac{\partial y_h^o}{\partial Y_{-h}} \frac{\partial Y_{-h}^N}{\partial w^\top}.$$

Under condition (18) we have $\partial y_h^o/\partial Y_{-h} \le 0$ for each h, so there exists a number $N \le 0$ such that

$$\frac{\partial Y^N}{\partial w^\top} = \sum_{h=1}^{H} \frac{\partial y_h^o}{\partial w^\top} + N \sum_{h=1}^{H} \frac{\partial Y_{-h}^N}{\partial w^\top}.$$

This is equivalent to

$$\frac{\partial Y^N}{\partial w^\top}\left(1 - N(H-1)\right) = \frac{\partial Y^o}{\partial w^\top},$$

with $1 - N(H-1) \ge 1$ and so it follows that $\partial Y^o/\partial w \le \partial Y^N/\partial w \le 0$. The statement (19) in terms of elasticities then directly follows (all these inequalities hold at $y_h^N = y_h^o$ for $h = 1, \ldots, H$). Similarly, the last term

$$\frac{\partial X^N}{\partial w^\top} = \sum_{h=1}^{H} \frac{\partial x_h^o}{\partial w^\top} + \sum_{h=1}^{H} \frac{\partial x_h^o}{\partial Y_{-h}} \frac{\partial Y_{-h}^N}{\partial w^\top},$$

is likely to be a matrix with positive entries, hence the claim $\varepsilon\left(X_j^o; w_j\right) < \varepsilon\left(X_j^N; w_j\right) \le 0$.

Claim (iv) about output price reactiveness directly follows from the definition (22) of

P^N and claim (iii):

$$0 < \frac{\partial P^N}{\partial w}(w, z) = \frac{\partial P}{\partial Y}\left(Y^N, z\right)\frac{\partial Y^N}{\partial w}(w, z) \leq \frac{\partial P}{\partial Y}\left(Y^N, z\right)\frac{\partial Y^o}{\partial w}\left(w, \{y_h\}_{h=1}^H, z\right).$$

$\square$

Proof of Result 3. Rewrite (36) as:

$$\left(\frac{\partial \mathbf{C}}{\partial Y}(w, Y, t)\right)^{-1}\left(P(Y, z) + Y\frac{\partial P}{\partial Y}(Y, z)\lambda^a(w, Y, t)\right) = g(w, Y, t), \tag{63}$$

with

$$g(w, Y, t) \equiv 1 - \frac{B_{CY}}{\partial \mathbf{C}/\partial Y}(w, Y, t).$$

Nonidentification means that there exist $g^* \neq g$ and $\lambda^* \neq \lambda^a$ such that for any (w, z, t)

$$\left(\frac{\partial \mathbf{C}}{\partial Y}(w, Y, t)\right)^{-1}\left(P(Y, z) + Y\frac{\partial P}{\partial Y}(Y, z)\lambda^*(w, Y, t)\right) = g^*(w, Y, t). \tag{64}$$

Notice that in the case where the model is not identified, there exists a continuum of functions λ and g compatible with (63): any $\lambda^{**} = \kappa\lambda^a + (1 - \kappa)\lambda^*$ and $g^{**} = \kappa g + (1 - \kappa)g^*$ satisfy (63). (For $\kappa = -\lambda^*/(\lambda^a - \lambda^*)$ we obtain that $\lambda^{**} = 0$ which corresponds to perfect competition.)

The proof studies the conditions under which g and λ^a are unique. We follows Lau's (1982) demonstration (which considered the case where λ^a was a constant parameter) and we include it for the sake of completeness. For any (w, Y, t), we can write

$$g^*(w, Y, t) = F(g(w, Y, t), Y), \tag{65}$$

for some F. The model is not identified iff F is not the identity. By (63)-(65) we also have:

$$\left(\frac{\partial \mathbf{C}}{\partial Y}(w, Y, t)\right)^{-1}\left(P(Y, z) + \lambda^*(w, Y, t)\frac{\partial P}{\partial Y}(Y, z)Y\right)$$
$$= F\left[\left(\frac{\partial \mathbf{C}}{\partial Y}(w, Y, t)\right)^{-1}\left(P(Y, z) + \lambda^a(w, Y, t)\frac{\partial P}{\partial Y}(Y, z)Y\right), Y\right].$$

Differentiating this equation w.r.t. z yields

$$\frac{\partial P}{\partial z} + \lambda^*(w, Y, t)\frac{\partial^2 P}{\partial Y \partial z}(Y, z)Y = F_1 \cdot \left(\frac{\partial P}{\partial z} + \lambda^a(w, Y, t)\frac{\partial^2 P}{\partial Y \partial z}(Y, z)Y\right)$$

$$\Leftrightarrow \quad (1 - F_1)\frac{\partial P}{\partial z} = \frac{\partial^2 P}{\partial Y \partial z}(Y, z)Y\left[\lambda^a(w, Y, t)F_1 - \lambda^*(w, Y, t)\right]. \tag{66}$$

In the case where λ^a and g are not identified, the function F is not the identity and so $F_1 \neq 1$ and $\lambda^a \neq \lambda^*$.

If z is a scalar, (66) is satisfied for any (w, Y, t, z) iff $\phi \equiv (1 - F_1)/(\lambda^a F_1 - \lambda^*)$ depends only upon Y, in which case:

$$\frac{\partial^2 P/\partial Y \partial z}{\partial P/\partial z}(Y, z) = \phi(Y)$$

$$\Leftrightarrow \quad \frac{\partial}{\partial Y}\ln\left(\frac{\partial P}{\partial z}(Y, z)\right) = \phi(Y)$$

$$\Leftrightarrow \quad \frac{\partial P}{\partial z}(Y, z) = t(Y)r(z)$$

$$\Leftrightarrow \quad P(Y, z) = t(Y)r(z) + s(Y).$$

If z is not a scalar, then (66) implies that

$$\frac{\partial P/\partial z_i}{\partial P/\partial z_j}(Y,z) = \frac{\partial^2 P/\partial z_i \partial Y}{\partial^2 P/\partial z_j \partial Y}(Y,z).$$

By Leontief's (1947) theorem, there exists a real valued function r such that $P(Y,z) = \mathbf{P}(Y, r(z))$. Sufficiency is shown by Lau (1982, p.97). □

10. Appendix B: Data description

This study relies on data provided by the Bureau of Labor Statistics (BLS) for 18 two-digit U.S. manufacturing industries over the period 1949-2001. Unfortunately, this dataset based on the Standard Industrial Classification (SIC) has been superseded after the introduction of the North American Industry Classification in 2006. These SIC data series are not longer updated by the BLS.[9]

Although it would certainly be better to use data at more disaggregated level, there exist only few datasets comprising quantities and price indices at such level. This data set comprises information on the price and quantity of output (p, Y) and of five inputs (hence $J = 5$): capital (w_k, X_k), labour (w_ℓ, X_ℓ), energy (w_e, X_e), intermediate material input (w_m, X_m), and services (w_s, X_s). The evolution of these quantities over time is depicted in Figure B1 at the aggregate level over all 18 industries. From Figure B1, it seems clear that these variables (except perhaps the labour input) are nonstationary. In addition to the endogenous variables, several exogenous variables are also nonstationary.
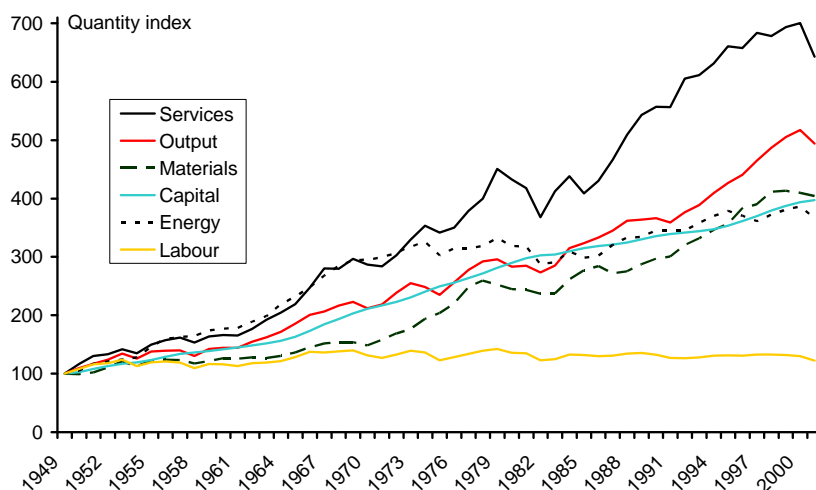

Figure B1: The evolution of input quantities. Basis 100 in 1949

The profit rate s_π defined as the ratio of (gross) profit to sales, is an important variable for assessing the relevance of imperfectly competitive behavior. Computing this variable is not an easy task because profits depend on the user costs of capital which are not observed and whose definition is not consensual. Diewert (2003) and Diewert, Harrison and Schreyer (2004) are useful references on this point. In this paper, we follow Diewert

---

[9] Data are available upon request.

(2003) and retain the user cost of capital formula,

$$w_{k,nt} \equiv w_{i,nt}\left(1 + r_t\right) - \mathrm{E}_t\left[\left(1 - \delta_{nt}\right)w_{i,n,t+1}\right]. \tag{67}$$

This equation shows that the user cost of capital w_k is increasing in the discounting rate r_t and decreasing in the expectation errors on investment goods inflation (given by $\mathrm{E}_t\left[w_{i,n,t+1}\right] - w_{i,n,t+1}$). According to Diewert (2003), we set r_t to 0.05 plus the consumer price inflation rate and assume no expectation errors on $w_{i,n,t+1}$. The few values of the user costs found to be slightly negative (5 out of 954 cases) were replaced with 0.0001.

The time series of the profit rate s_π averaged over all 18 industries is depicted in Figure B2. In U.S. manufacturing industries, the profit rate is about 5.3 percent on average over time and industries. This figure is somewhat different from that of Basu and Fernald (1997) who report an average profit rate of 3 percent. This contradiction is related to differences in the time period and the industries covered, but certainly also to the high sensitivity of the user costs of capital with respect to a priori assumptions upon (i) the expected price change of investment goods and (ii) the discounting rate of future income streams.

Given the sensitivity of the economic definition of the profit rate with respect to a priori choices, it appears helpful to complete the picture with the more robust accounting definition of the gross profit rate, which does not rely on the user cost of capital and the capital stock, but uses instead investment and investment price:

$$s_\pi^a = \frac{p_{nt}Y_{nt} - \left(w_{i,nt}X_{i,nt} + w_{\ell,nt}X_{\ell,nt} + w_{e,nt}X_{e,nt} + w_{m,nt}X_{m,nt} + w_{s,nt}X_{s,nt}\right)}{p_{nt}Y_{nt}}.$$

Notation $X_{i,nt}$ denotes gross investment and $w_{i,nt}$ its price. On average over time and industries s_π^a is equal to 8.7 percent, which seems to confirm the rather high level of profits in US manufacturing.

Average profitability over time and over industries is depicted in Figures B2 and B3, respectively. The evolution of average profitability (Figure B2) broadly reflects the business cycle, but there is also a downward trend or a break in this picture: whereas the average (economic) profit rate was about 7.8 percent over the 1949-1975 period, it declined to 2.8 percent for the post 1975 period. Such a pattern is actually observed for many industries (not reported). Figure B3 reports the average level of profitability over the period 1949-2001 for each industry and shows that profitability varies a lot across industries. Two industries exhibit a negative average rate of profit: Textile Mills Products and Primary Metal Industry (SIC numbers 22 and 33). For two other industries, average profitability exceeds 10 percent: Chemistry & Allied Products and Petroleum Refining (SIC 28 and 29). See Table B1 for the list of industry names with their corresponding SIC number.
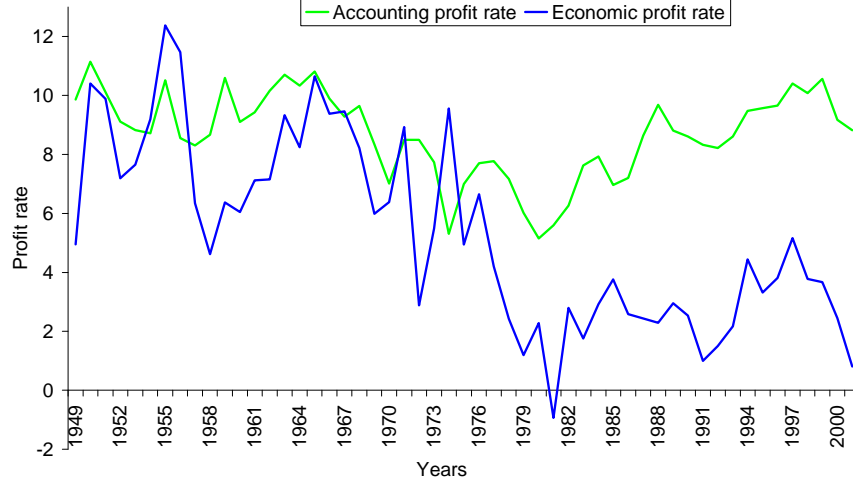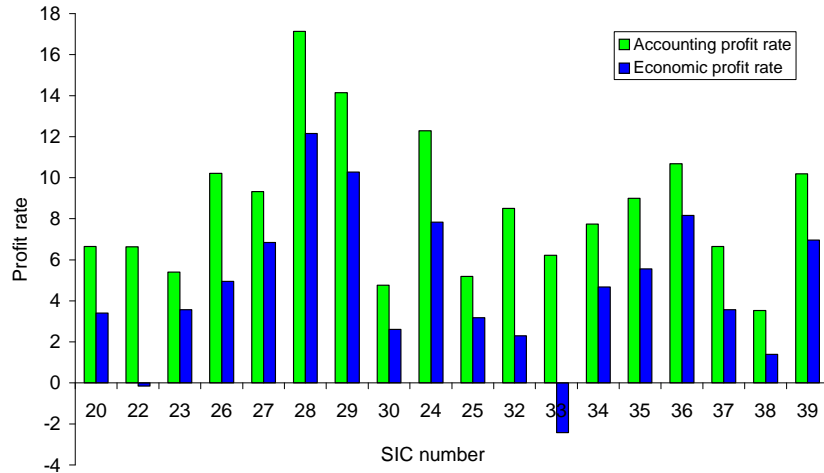
Figure B2: Profit rate over time

Figure B3: Profit rate over industries

Let us now compare the accounting and economic profit rates. In general both concepts diverge mainly because the accounting definition neglects the opportunity cost of the investment in capital goods and reduces an investment decision with uncertain returns to a static accounting exercise. From Figure B2, it can be seen that both concepts are quantitatively very close for the 1949-75 period, but the accounting profit rate becomes much larger than the economic profit rate after 1976. Similarly, Figure B3 shows that on average over the industries, the accounting profit rate is greater than the economic profit rate. There is a simple relationship between s_π^a and s_π explaining these facts. Gross investment is given by $X_{i,nt} = X_{k,nt} - (1 - \delta_{nt}) X_{k,n,t-1}$. In the steady state, $r_t = \delta_t$ and $w_{i,nt} = w_{i,n,t+1}$ (the user cost becomes $w_{k,nt} = 2\delta_t w_{i,nt}$) and the sole purpose of investment is replacement: $X_{i,nt} = \delta_t X_{k,nt}$. It follows that in the steady state $w_{k,nt} X_{k,nt} = 2w_{i,nt} X_{i,nt}$ and that $s_\pi^a - s_\pi = w_{i,nt} X_{i,nt} / (p_{nt} Y_{nt})$. In the steady state, the difference between the accounting and economic profit is given by the investment share

in turn-over. So s_π^a provides an easily calculable upper bound for (steady-state) economic profits, which helps to understand why $s_\pi^a > s_\pi$ in Figures B2 and B3. The fact that for the 1949-1975 period, we observed that $s_\pi^a \simeq s_\pi$ is mainly due to a moderate nominal interest rate combined with high inflation for the investment goods (so that $w_{i,nt} < w_{i,n,t+1}$ in (67)).[10]

Table B1. The Standard Industrial Classification for U.S. manufacturing industries

SIC No	Industry name	SIC No	Industry name
20	Food & kindred products	30	Rubber and misc. plastics products
21	Tobacco products industry	31	Leather and leather products
22	Textile mill products industry	32	Stone, clay, glass & concrete products
23	Apparel and other textile products	33	Primary metal industries
24	Lumber and wood products	34	Fabricated metal products
25	Furniture and fixtures	35	Industrial machinery and equipment
26	Paper and allied products	36	Electronic & other electric equipment
27	Printing and publishing	37	Transportation equipment
28	Chemical and allied products industry	38	Instruments, clocks, optical goods
29	Petroleum and coal products	39	Other manufacturing industries

11. References

Amir, R., 2005, "Ordinal versus cardinal complementarity: The case of Cournot oligopoly," *Games and Economic Behavior*, 53, 1-14.

Amir, R., 1996, "Cournot oligopoly and the theory of supermodular games," *Games and Economic Behavior*, 15, 132-148.

Appelbaum, E., 1982, "The Estimation of the Degree of Oligopoly Power," *Journal of Econometrics*, 19, 287-299.

Basu, S., and Fernald, J. G., 1997, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 105, 249-283.

Bartelsman, E. J., 1995, "On empty boxes: Returns to scale revisited," *Economics Letters*, 49, 59-67.

Braulke, M., 1984, "The Firm in Short-Run Industry Equilibrium: Comment," *American Economic Review*, 74, 750-753.

Bresnahan, T. F., 1982, "The Oligopoly Solution Concept is Identified," *Economics Letters*, 10, 87-92.

Bresnahan, T. F., 1989, "Empirical studies of industries with market power," in R. Schmalensee and R. D. Willig (Editors), *Handbook of Industrial Organization*, Volume 2, 1011-1057.

Burnside, C., 1996, "Production function regressions, returns to scale, and externalities," *Journal of Monetary Economics*, 37, 177-201.

Cahuc P. and A. Zylberberg, 2004, *Labor Economics*, MIT Press.

Diewert, W. E., 1981, "The Elasticity of Derived Net Supply and a Generalized Le Chatelier Principle', *Review of Economic Studies,* 48, pp.63-80.

Diewert, W. E., 2003, "Measuring Capital," NBER Working paper 9526.

Diewert, W. E., and K. Fox, 2008, "On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups," *Journal of Econometrics*, 145, 174–193.

Diewert, W. E., A. Harrison and P. Schreyer, 2004, Cost of capital services in the production account, *mimeo*.

---

[10] The commodity prices index increased by 3.3% in average over the 1949-1975 period, and by 4.6% over 1976 to 2001. For the price index of investment goods, the figures are respectively 4.3% and 3.5%.

Eichhorn, W. and W. Oettli, 1972, "A general formulation of the LeChatelier-Samuelson Principle," *Econometrica,* 40, 711-717.

Frondel, M. and C. M. Schmidt, 2002, "The Capital-Energy Controversy: An Artifact of Cost Shares?," *The Energy Journal*, 23, 53-79.

Hall, R. E., 1988, "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96, 921-947.

Heiner, R. A., 1982. "Theory of the Firm in "Short-Run" Industry Equilibrium," *American Economic Review*, 72, 555-62.

Koebel, B., M., 2002, "Can aggregation be achieved by neglecting the problem? Property inheritance and aggregation biases," *International Economic Review* 43, 223-255.

Kulatilaka, N., 1987, "The Specification of Partial Static Equilibrium Models," *Review of Economics and Statistics*, 69, 327-335.

Lau, L. J., 1982, "On identifying the degree of competitiveness from industry price and output data," *Economic Letters*, 10, 93-99.

Leontief, W., 1947, "Introduction to a Theory of the Internal Structure of Functional Relationships," *Econometrica*, 15, 361–373.

Lewbel, A., 1993, 'Distribution Movements, Macroeconomic Regularities, and the Representative Consumer', *Ricerche Economiche*, 47, 189-199.

Lewbel, A., 1996, "Aggregation Without Separability: A Generalized Composite Commodity Theorem", *American Economic Review*, 86, 524–543.

Lopez, R. E., 1984, "Estimating Substitution and Expansion Effects Using a Profit function Framework," *American Journal of Agricultural Economics*, 66, 358–367.

Milgrom, P., and J., Roberts, 1996, "The LeChatelier Principle," *American Economic Review*, 86, 173-179.

Morrison-Paul, C. J., and D. Siegel, 1999, "Estimation of Scale Economies Underlying Growth and Productivity: The Empirical Implications of Data Aggregation," *Southern Economic Journal*, 65, 739-756.

Novshek, W., 1985, "On the Existence of Cournot Equilibrium," *The Review of Economic Studies*, 52, 85-98.

Reiss P. and F. Wolak, 2007, "Structural Econometric Modeling: Rationales and Examples from Industrial Organization," *Handbook of Econometrics*, Volume 6A, Elsevier.

Roy, S. and T. Sabarwal, 2009, "Monotone Comparative Statics for Games With Strategic Substitutes," Working Paper 2008-10, Department of Economics, University of Kansas.

Roy, S. and T. Sabarwal, 2008, "On the (Non-)Lattice Structure of the Equilibrium Set in Games With Strategic Substitutes," Economic Theory, 37, 161–169.

Sakai, Y., 1973, "An Axiomatic Approach to Input Demand Theory," *International Economic Review*, 14, pp.735-752.

Samuelson, P. A., 1947, *The Foundations of Economic Analysis*, Harvard University Press.