ISSN 0819-2642
ISBN 978 0 7340 3998 9

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 1032

January 2008

Learning-by-Doing and Productivity Dynamics in Manufacturing Industries

by

Andrew Clarke

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

Learning-by-Doing and Productivity Dynamics in Manufacturing Industries

Andrew J. Clarke *

Department of Economics, University of Melbourne, VIC, 3010, Australia

Abstract

This paper estimates a structural model of learning-by-doing. Treating production experience as a state variable, this paper provides estimates of the structural parameters obtained from the first order conditions arising from the plant's maximization problem. Estimates are provided using data on 4-digit manufacturing industries and plant-level observations. Using aggregate industry data, the results indicate that estimated learning rates might be considerably lower than previous estimates. The results also reveal considerable variation in estimated learning rates, across broad industry groups, at both the plant-level and the 4-digit industry level. This implies that using results from existing studies that focus upon specific, narrowly defined industries or firms, may lead to misleading conclusions concerning the widespread importance of learning-by-doing for generating productivity dynamics within the manufacturing industry.

Key words: Production Experience, Learning-by-Doing, Structural Estimation
JEL classification: C13, D21, L23, L60

* Tel +61-3-8344-5298 Fax +61-3-8344-6899
Email address: andrew.clarke@unimelb.edu.au (Andrew J. Clarke).

[1] The financial support of Statistics Canada through the Statistic Canada Research Stipend is gratefully acknowledged. I would like to thank John Baldwin, Jeff Borland, Mick Coelli, Russell Hillberry, Alok Johri, John Rust and Lawrence Uren for their comments.

1 Introduction

At least as far back as Marshall, economists have suggested that organizations store and accumulate knowledge that affects their technology of production[1]. The returns to this production knowledge are generally associated with the concept of learning-by-doing. There is an extensive empirical literature that explores the relationship between production experience and plant productivity. Recent studies such as Bahk & Gort (1993), Irwin & Klenow (1994), Jarmin (1994), Benkard (2000), Thompson (2001), Thornton & Thompson (2001), and Cooper & Johri (2002) find that agents and organizations appear to become more productive as they gain experience at producing a particular product or service. The results from these studies suggest that a doubling of cumulative output generally leads to an approximately 20% reduction in unit costs of production[2]. In contrast to these studies, which cover a diverse range of narrowly defined industries, this paper studies the importance of learning-by-doing for generating productivity dynamics within the manufacturing industry as a whole.

Several authors have recognized that learning-by-doing introduces an inter-temporal component to output decisions made by production units[3]. Higher current production rates, which involve an increase in current costs with no immediate effect upon productivity, lead to a reduction in future costs of production through learning-by-doing. Consistent with this literature, this paper treats the stock of experience as a state variable under the control of the production unit, and explicitly models the decision of how much output to produce as a joint decision with how much production experience to accumulate.

Recent work has also emphasized the quantitative importance of 'organizational forgetting' where relatively distant production experience becomes less relevant over time. Argote et al. (1990) provide empirical evidence for this hypothesis of 'organizational forgetting' associated with the construction of Liberty Ships during World War II. Similarly, Darr et al. (1995) provide evidence for this hypothesis for pizza franchises and Benkard (2000) provides evidence for 'organizational forget-

[1] Marshall (1961) states that 'capital consists in a great part of knowledge and organization... [so that]...it is best to reckon Organization as a distinct agent of production' (p115).

[2] Argote et al. (1990), Epple et al. (1991), Bahk & Gort (1993) and Benkard (2000) provide numerous references to this historical literature.

[3] Fudenberg & Tirole (1983), Ghemawat & Spence (1985), Jarmin (1994), Irwin & Klenow (1994), and Zulehner (2003)

ting' associated with the production of commercial aircraft. Under a slightly different specification for 'organizational forgetting', Irwin & Klenow (1994) provide evidence for spillovers of experience across successive generations of chips in the semi-conductor industry[4].

The key contribution of this paper is to provide estimates of learning-by-doing parameters using the first order conditions from a structural model (of plant behavior) that allows for the accumulation of production experience, consistent with the hypothesis of organizational forgetting. Importantly, the structural model allows a separation of the parameters controlling the productivity of experience (learning-by-doing) from those controlling the accumulation of experience (organizational forgetting). The particular nature of the structural model, and the log-linear functional forms used in the estimation, have the advantage that estimation does not require data on the stock of production experience. This considerably reduces the data requirements since there is no need to track production units since birth in order to construct a series for the stock of experience. This allows estimation of learning parameters in a much broader context than previous studies[5]. Specifically, estimates of learning and forgetting rates are obtained using data on both four-digit manufacturing industries and a large sample of plant-level observations.

The results suggest that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with the 4-digit manufacturing industry data. The results also indicate that estimated learning rates might be considerably lower than existing (aggregate) estimates, such as those in Cooper & Johri (2002). The results identify variation in estimated learning rates, even at the 4-digit industry level. This is a noteworthy result that has previously not been identified in the existing literature.

Using plant-level observations, the results suggest that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with plant-level data for Science-Based industries. The estimates for this industry are consistent with previous work that has studied firms within this same broad industry group. However, the model is generally rejected for other industry

[4] Rather than applying a constant rate of depreciation to cumulative output, the specification in Irwin & Klenow (1994) allows for a constant rate of organizational forgetting across generations of specific products, with no depreciation within product generations.

[5] Existing microeconomic studies, limited by the requirement that researchers observe the entire production history since birth have tended to focus on quite specific industries or organizations.

groups suggesting that the implied productivity dynamics associated with learning-by-doing are an adequate description of the data for some, but not all, industry groups. Consequently, using results from existing studies that focus upon specific, narrowly defined industries or firms, may lead to misleading conclusions concerning the importance of learning-by-doing, as a widespread phenomena, across manufacturing industries as a whole.

An interesting result emerges when comparing the plant-level estimates to those obtained using the more aggregate 4-digit industry data. Compared to this industry data for Science-Based industries, the estimated learning rates are considerably larger at the plant level. This result is most likely driven by an 'aggregation bias' in the presence of plant-level heterogeneity in productivity dynamics associated with learning-by-doing. The results using the industry data treat each 4-digit industry as a single representative plant. The existence of considerable plant-level heterogeneity, uncorrelated with plant-level observed characteristics, together with a highly non-linear structural model, implies that the average estimated learning rates, at the plant-level, may be quite different to the learning rate of the representative (average) plant, identified in the aggregate estimates.

The structural model of proprietary learning-by-doing, presented in this paper, is quite close to that presented in Irwin & Klenow (1994), Jarmin (1994), and Zulehner (2003) but has some important distinguishing features. First, the assumption in these papers of Cournot competitors producing a homogenous product seems reasonable for focusing upon learning-by-doing in a specific industry[6]. However, in the context of manufacturing industries as a whole, this assumption is much harder to justify. Consequently, this paper treats production units as operating in monopolistically competitive output markets producing a differentiated final output.

Second, Irwin & Klenow (1994) and Benkard (2000) allow for spillovers of experience across generations of final outputs. For particular industries, such as the semi-conductor industry, where such product generations are well defined, this might be an appropriate characterization of organizational forgetting. For manufacturing industries as a whole, where product generations are less clearly defined, if measured at all, a more general specification for the accumulation of production

[6] This has the advantage that (marginal) costs of production are directly related to market shares and the elasticity of output demand. The learning parameters may then be identified from observable by specifying marginal costs of production as a function of the stock of experience. Of course, this still requires a measure of cumulative output for each firm in order to construct a measure for the stock of experience.

experience is required. This paper views the production unit as operating in an environment with an ever changing set of tasks, workers, teams, machines and information in which experience is continually accumulating and depreciating.

2 The Structural Model

A set of intermediate goods producers learn by accumulating production experience. These firms operate within a market characterized by monopolistic competition, controlling how much they wish to learn by varying the markup of price over marginal cost to maximize the present discounted value of profits[7]. These differentiated intermediate goods are used as inputs for the production of a final good labeled 'manufacturing output'.

A final good Y_t is produced in a competitive industry by an arbitrary number of identical (final goods) producers with price P_t[8]. This 'manufacturing output' is produced using intermediate goods $Q_t(i)$ as inputs to a constant returns to scale production technology

$$Y_t = \left\{ \int_0^1 \left[\psi_t(i)\, Q_t(i) \right]^{1/\mu} \, di \right\}^{\mu} \tag{1}$$

where $\psi_t(i)$ are technological parameters that affect the relative demand for input $Q_t(i)$. The conditional factor demands for each intermediate input $Q_t(i)$ that arise from the cost minimization problem, solved by these final goods producers, will represent the demand faced by each intermediate goods producer. Let $v_t(i)$ denote the output price of intermediate good $Q_t(i)$. A zero profit condition in this final goods sector ties down the equilibrium price of final output P_t as a function of the prices of the intermediate goods. The inverse demand function for each intermediate goods producer will be given by:

$$v_t(i) = f\left[Q_t(i), P_t, Y_t, \psi_t(i)\right] = \psi_t(i) \left\{ \frac{Q_t(i)}{Y_t} \right\}^{\frac{(1-\mu)}{\mu}} P_t \tag{2}$$

[7] The structural model, outlined below, uses the standard Blanchard & Kiyotaki (1987) model of monopolistic competition commonly used in aggregate general equilibrium models, modified to allow for the accumulation of production experience.

[8] This implicitly assumes that consumers have preferences defined over this good, 'manufacturing output' and other aggregate goods.

where $\psi_t(i)$ is a demand shock. The price elasticity of demand, faced by intermediate goods producer i will be given by $\mu/(1-\mu)$, the negative of the elasticity of substitution between any two intermediate inputs in final goods production.

There is a continuum of intermediate goods producers operating within a monopolistically competitive economy each producing a differentiated good $Q(i)$ with $i \in [0,1]$. Since the accumulation of production experience might be associated with a production technology that exhibits increasing returns to scale it is inappropriate to assume that production units operate within a perfectly competitive market. Monopolistic competition is one such market structure that would be consistent with an increasing returns to scale production technology. This might be motivated in several ways. First, the focus of this paper is upon the dynamics in output implied by proprietary learning-by-doing such that the strategic implications of learning-by-doing spillovers are explicitly ignored[9]. Second, treating manufacturing as a single industry, implies that alternative oligopolistic market structures, may be less appropriate. For example, the assumption of Cournot competitors producing a homogenous product seems reasonable for a specific industry. However, in the context of manufacturing industries as a whole, this assumption is much harder to justify.

Each intermediate good is produced according to the following technology:

$$Q_t(i) = A_t(i)F\left[\,K_t(i), Z_t(i), H_t(i), M_t(i)\right] \tag{3}$$

where production experience $Z(i)$ is combined with physical capital $K(i)$, labor $H(i)$, and production materials & energy $M(i)$ to produce output $Q(i)$. The term $A_t(i)$ represents a shock to total factor productivity.

This production technology differs from a standard neo-classical production technology because the firm accumulates a stock of production experience which is an input to the production technology. Production experience $Z(i)$ refers to the information accumulated by the firm, through the process of past production, regarding how best to organize its production activities. The production technology exhibits positive and diminishing returns to production experience, a feature

[9] Irwin & Klenow (1994), Thornton & Thompson (2001), and Thompson (2001) provide evidence that production units learn much more from their own output than from the output of other firms. Results from these papers suggest that as much as two-thirds of production experience may be attributed to accumulation from proprietary learning-by-doing.

often found in existing studies of learning-by-doing.

In order to represent the stock of production experience as a specific capital good that is jointly produced with output, the stock of experience will be described by an accumulation technology. This is closely related to the empirical learning-by-doing literature in which each unit of past production contributes equally to the creation of knowledge. The current specification differs in that the contribution of production in any period to the current level of experience is decreasing over time. The accumulation technology is given by:

$$Z_{t+1}(i) = Z_t(i)^{\eta}\, Q_t(i)^{\gamma} \tag{4}$$

or

$$\ln Z_t(i) = \eta^t \,\ln Z_0(i) + \gamma \sum_{k=0}^{t-2} \eta^k \,\ln Q_{t-1-k}(i)$$

where $Z_t(i)$ denotes the stock of experience available to the production unit at time t, $Z_0(i)$ denotes the initial endowment of production experience and $Q_t(i)$ denotes the current level of output[10].

The restriction $\eta < 1$ is consistent with the empirical evidence supporting the hypothesis of depreciation of production experience often referred to as organizational forgetting. This hypothesis of organizational forgetting might be motivated in several ways. It allows for the sensible idea that production knowledge may become less relevant over time as new techniques of production, new product lines and new markets emerge. In addition, the knowledge accumulated through production experience will be a function of the current vintage of physical capital. The decision to replace physical capital will imply that the existing stock of production experience will be less relevant. It also provides, in a general way, for the idea that some match specific knowledge may be lost to the firm as workers leave or get reassigned to new tasks or teams within the firm.

The hypothesis of organizational forgetting is compatible with the existence of a steady state in which the stock of experience is constant. In contrast, studies of learning-by-doing without this feature allow the stock of experience to grow unboundedly. An alternative way to bound learning is to assume that productivity increases due to learning occur for a fixed number of periods. While this

[10] There is no production in the initial period 0. In the first period, the production unit takes the initial stock of production experience as given when choosing its production level.

may be appropriate for any one task or worker within the firm, it becomes less appropriate once firms are viewed as operating within an environment with an ever changing set of tasks, workers, teams, machines, products, and information. The accumulation technology (4) is consistent with such an environment in which experience is (approximately) continually accumulating and depreciating.

The accumulation technology (4) also allows for a varying contribution of (the weighted sum) of past output to the stock of experience. Since it need not be true that each unit of last period's output produce (exactly) one unit of experience, the parameter γ controls the contribution of this previous period output to the stock of experience available to the production unit. This parameter essentially controls the productivity of investment in production experience.

The 'representative' intermediate goods producer i solves the following dynamic programming problem:

$$\begin{aligned}\mathbb{V}(A_{it}, \psi_{it}, Z_{it}) = \max_{\{H_{it}\, K_{it}\, M_{it}\, Z_{i,t+1}\}} & \{v_{it}\, Q_{it} - w_t\, H_{it} - r_t\, K_{it} - s_t\, M_{it}\} \\ & + \beta\, E_t\, \mathbb{V}(A_{i,t+1}, \psi_{i,t+1}, Z_{i,t+1} | \Omega_{it})\end{aligned}$$

subject to:

$$v_{it} = f(Q_{it}, P_t, Y_t,\, \psi_{it}) = \psi_{it}\, \left[\frac{Q_{it}}{Y_t}\right]^{\frac{1-\mu}{\mu}}\, P_t \qquad \text{inverse demand function}$$

$$Q_{it} = A_{it}\, F(K_{it},\, H_{it},\, M_{it},\, Z_{it}) \qquad \text{production technology}$$

$$Z_{i,t+1} = Z_{it}^{\eta}\, Q_{it}^{\gamma} \qquad \text{accumulation technology}$$

where v_{it} is the output price charged by the ith intermediate producer, Q_{it} is the output of this ith intermediate producer, Y_t is the quantity of final output and P_t is the price of this final output. There are three control variables: the physical capital input K_{it}, the labor input H_{it}, and the intermediate input M_{it}. There are three state variables: the endogenous state variable production experience Z_{it} and the exogenous state variables A_{it}, a stochastic disturbance to total factor productivity, and ψ_{it}, a stochastic demand disturbance, that are observed by the producer at time t. Note that all of these state variables will generally be unobservable to the econometrician.

The stochastic nature of the problem arises because the producer must choose their desired stock of experience, for period $t+1$, prior to the realization of the productivity shock $A_{i,t+1}$ and the demand shock $\psi_{i,t+1}$. Expectations are assumed to be formed rationally, conditional upon the information set Ω_{it}, which includes the current and past realizations of the exogenous variables and the endogenous state variable Z_{it}.

Each intermediate goods producer is assumed to operate within a perfectly competitive input market such that the nominal rental price of physical capital r_t, the nominal wage w_t, and the nominal price of intermediate inputs s_t are taken as given. The assumption of monopolistically competitive markets implies P_t, the price of the final good, will be taken as given. The parameter β is a discount factor satisfying $0 < \beta < 1$ that is used to discount future profits.

Consistent with previous studies of learning-by-doing, the structural model ignores all other sources of persistence in unobserved productivity, such as input costs of adjustment. The accumulation technology might reasonably be interpreted as representing costs of output adjustment making it difficult to separate the persistence implied by the accumulation of production experience from the persistence implied by input adjustment costs. This would require some quite specialized data to be able to identify the independent effect of production experience from the effect of adjustment costs and is not be possible with the data used in this paper. Of course, in reality, the situation is much more complicated with productivity dynamics determined by the interaction between learning-by-doing with costs of adjustment. This is an important area for future research.

Since the physical capital input will likely be measured with error, it is convenient to solve for the level of the physical capital input needed to accumulate a desired stock of experience $Z_{i,t+1}$, given the existing stock of experience Z_{it}, the human capital input H_{it} and materials input M_{it}:

$$K_{it} = \mathbb{K}\left(A_{it},\, H_{it},\, M_{it},\, Z_{it},\, Z_{i,t+1}\right)$$

where $\mathbb{K}$ is decreasing in H_{it}, M_{it}, and Z_{it} and increasing in $Z_{i,t+1}$. The first-order conditions associated with the dynamic programming problem become:

$$\frac{\partial \pi_{it}(\cdot)}{\partial H_{it}} = 0 \qquad \text{and} \qquad \frac{\partial \pi_{it}(\cdot)}{\partial M_{it}} = 0 \tag{5}$$

$$\frac{\partial \pi_{it}(\cdot)}{\partial Z_{i,t+1}} + \beta\, E_t \left[\frac{\partial \pi_{i,t+1}(\cdot)}{\partial Z_{i,t+1}} \,\middle|\, \Omega_{it} \right] = 0 \tag{6}$$

where:

$$\begin{aligned} \pi_{it} &= v_{it}\, Q_{it} - w_t\, H_{it} - r_t\, \mathbb{K}(\cdot)_{it} - s_t\, M_{it} \\ &= \pi(A_{it},\, \psi_{it},\, v_{it},\, P_t,\, Y_t,\, w_t,\, s_t,\, r_t,\, H_{it},\, M_{it},\, Z_{it},\, Z_{i,t+1}) \end{aligned}$$

Conditional upon a stock of desired experience $Z_{i,t+1}$, the first order conditions (5) for labor and intermediate inputs are the standard static optimality conditions. The first-order condition (6) determines the value of an additional unit of production experience for use by the producer in period $t+1$. This additional unit of experience has a (marginal) value, in terms of profits, of $\partial \pi_{it}(\cdot)/\partial Z_{i,t+1}$. The condition (6) implies experience will be accumulated up to the point where the value of an additional unit of experience this period is equal to the discounted value of this experience next period.

2.1 Some Functional Forms and Identifying Restrictions

The first order conditions of the maximization problem (5) and (6) cannot be estimated without specifying some functional form for the production technology (3). The choice of functional form is driven primarily by the nature of the data currently available. In addition, the functional form has been chosen to allow a comparison with both existing microeconomic studies and aggregate studies of production experience. Assume the production technology (3) takes the following form:[11]

$$Q_{it} = A_{it}\, K_{it}^{\theta}\, H_{it}^{\alpha}\, M_{it}^{\phi}\, Z_{it}^{\varepsilon} \qquad \alpha > 0,\ \theta > 0,\ \phi > 0,\ \varepsilon > 0. \tag{7}$$

[11] All of the previous studies of production experience use a Cobb-Douglas form for the production technology. For a given stock of experience Z_{it}, with constant returns to scale in the 'measured inputs' such that $(\alpha + \theta + \phi) = 1$, this implies the same form for the average cost function, as in Irwin & Klenow (1994).

Regardless of the form of accumulation technology, diminishing marginal productivity of production experience in the production technology leads to the result that additional units of production experience lead to relatively large reductions in average cost when the stock of experience is relatively low. The effect of the accumulation technology is to describe how changes in output levels affect the (available) stock of experience. For example, a linear technology of the form used by Benkard (2000), implies a constant marginal productivity of 'investments' in experience and a log-linear technology as in (4) implies diminishing marginal productivity of investments in experience.

Using the production technology (7) and the log-linear accumulation technology (4), the first order conditions become:

$$E_t \left[\frac{\alpha}{\mu} \frac{v_{it}\, Q_{it}}{w_t\, H_{it}} + \beta(\eta + \varepsilon\, \gamma) \frac{w_{t+1}\, H_{i,t+1}}{w_t\, H_{it}} - \beta\, \eta\, \frac{\alpha}{\mu} \frac{v_{i,t+1}\, Q_{i,t+1}}{w_t\, H_{it}} - 1 \;\middle|\; \Omega_{it} \right] = 0 \tag{8}$$

$$E_t \left[\frac{\phi}{\mu} \frac{v_{it}\, Q_{it}}{s_t\, M_{it}} + \beta(\eta + \varepsilon\, \gamma) \frac{s_{t+1}\, M_{i,t+1}}{s_t\, M_{it}} - \beta\, \eta\, \frac{\phi}{\mu} \frac{v_{i,t+1}\, Q_{i,t+1}}{s_t\, M_{it}} - 1 \;\middle|\; \Omega_{it} \right] = 0 \tag{9}$$

The structural model involves two (potentially) unobservable variables—the physical capital input K_{it} and the stock of experience Z_{it}. Both of the identifying assumptions, the Cobb-Douglas production technology (7) and the log-linear accumulation technology (4) imply that the estimating equations may be written in terms of observed variables.

It is clear that these identifying assumptions provide several advantages associated with estimation of the model. First, the system (8) and (9) may be estimated using panel data on the value of production ($v_{it}\, Q_{it}$), the total cost of the labor input ($w_t\, H_{it}$), and the total cost of intermediate inputs ($s_t\, M_{it}$). Importantly, the estimating equations do not require data on the stock of production experience Z_{it} so that the structural learning parameters might be estimated without the need to track production units from birth in order to construct a measure of the stock of production experience. Consequently, these identifying assumptions allow estimation of the learning parameters in a much broader context than previous studies.

Second the identifying assumptions imply estimation of the system (8) and (9) only requires data on the value of production, the total cost of the labor input, and the total cost of intermediate inputs. This also allows estimation of the learning parameters in a much broader context than previous studies. Typically, available data sets only contain information on the value of measured inputs and the value of output, rather than separate information on the prices and quantities of inputs and outputs. While the assumption of perfectly competitive input markets might be used to obtain input quantities by using an appropriately defined input price deflator, there may be considerable biases introduced when output is measured using deflated total revenues[12].

The identifying assumptions, implied by the Cobb-Douglas production technology and the log-linear accumulation technology may be described in terms of three conditions. First, homotheticity of the production technology, such that the elasticity of scale is independent of the scale of production, allows the (potentially) unobserved physical capital input to be expressed as a function of observed input prices and quantities. Second, the Cobb-Douglas production technology implies a constant output elasticity with respect to the stock of experience. This is a standard assumption in the learning-by-doing literature such as Bahk & Gort (1993), Irwin & Klenow (1994), or Benkard (2000) that provides for a constant learning rate, independent of the stock of experience[13]. Third, the log-linear accumulation technology implies a constant elasticity with respect to the inputs to the production of experience. This is the main point of departure from existing studies of learning-by-doing. In contrast to a linear accumulation technology which implies a constant marginal productivity associated with the inputs to the production of experience, this log-linear technology implies diminishing marginal productivity for these inputs.

It is the combination of all three of these identifying restrictions that allow the estimating equations to be written in terms of observable variables. For example, an alternative homothetic production technology, such as CES, with a log-linear accumulation technology will still require data on the stock of experience since this implies an output elasticity of production experience that depends upon the level of output and experience. Similarly, a Cobb-Douglas production technology, with a linear accumulation technology will also require data on the stock of experience.

Regardless of the estimation strategy used, an examination of the system defined by (8) and (9) reveals that only a subset of the (full) parameter vector $\mathbf{\Theta} = [\alpha, \phi, \mu, \beta, \eta, \gamma, \varepsilon]$ will be identified.

[12] Klette & Griliches (1996) discuss some of the biases associated with production function estimation when output is measured by deflated revenues.

[13] Note that this constant elasticity still implies that additional units of production experience lead to relatively large reductions in average cost when the stock of experience is relatively low

Only estimates of the ratio (ϕ/μ) or (α/μ) may be obtained so that the demand parameter μ and the production technology parameters α or ϕ cannot be separately identified. Indeed, without further restrictions only the reduced parameter vector $\mathbf{\Theta}_0 = [(\alpha/\mu), (\beta\,\varepsilon\,\gamma), (\beta\,\eta), (\phi/\mu]$ may be identified.

3 Empirical Estimation Strategy

The identifying restrictions discussed above imply that the estimating equations, defined by (8) and (9), may be estimated using panel data containing information on total revenues $(v_{it}\, Q_{it})$, the total wage bill $(w_t\, H_{it})$, and the total cost of intermediate inputs $(s_t\, M_{it})$. The driving force behind the dynamic structure of the model are the learning-by-doing parameters η, γ and ε. Using the Generalized Method of Moments (GMM) consistent estimates of the parameters may be obtained by directly estimating the dynamic equations given by (8) and (9). In contrast to estimation of the full structural model, say by maximum likelihood methods, this avoids the need to specify the (joint) distribution of the random variables A_{it} and ψ_{it}.

Using the system defined by (8) and (9) and replacing one-step ahead expected values with their realized values in time t provides the system:

$$f_1(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_0) = v_{1,it} \tag{10}$$

$$f_2(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_0) = v_{2,it} \tag{11}$$

where:

$$f_1(\mathbf{X}_{i,t-1}, \mathbf{\Theta_0}) = \frac{\alpha}{\mu}\,\frac{\mathrm{vout}_{i,t-1}}{\mathrm{pay}_{i,t-1}} + \beta(\eta + \varepsilon\,\gamma)\,\frac{\mathrm{pay}_{i,t}}{\mathrm{pay}_{i,t-1}} - \beta\,\eta\,\frac{\alpha}{\mu}\,\frac{\mathrm{vout}_{i,t}}{\mathrm{pay}_{i,t-1}} - 1$$

$$f_2(\mathbf{X}_{i,t-1}, \mathbf{\Theta_0}) = \frac{\phi}{\mu}\,\frac{\mathrm{vout}_{i,t-1}}{\mathrm{intermed}_{i,t-1}} + \beta(\eta + \varepsilon\,\gamma)\,\frac{\mathrm{intermed}_{i,t}}{\mathrm{intermed}_{i,t-1}} - \beta\,\eta\,\frac{\phi}{\mu}\,\frac{\mathrm{vout}_{i,t}}{\mathrm{intermed}_{i,t-1}} - 1$$

with $\mathbf{\Theta_0} = [(\alpha/\mu), (\beta\,\varepsilon\,\gamma), (\beta\,\eta), (\phi/\mu)]$ where vout_{it} represents the value of output $(v_{it}\, Q_{it})$ at time t, pay_{it} represents the total cost of the labor input $(w_t\, H_{it})$ at time t, and $\mathrm{intermed}_{it}$ denotes the total cost of intermediate inputs ($s_t\, M_{it}$) at time t. The hypothesis of rational expectations requires that the forecast errors $\upsilon_{1,it}$ and $\upsilon_{2,it}$ be orthogonal to any information available at time of the forecast t, including the information contained in previous forecast errors. This implies:

$$E\begin{bmatrix} f_1(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_1)\, \mathbb{W}_{i,t-1} \\ f_2(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_2)\, \mathbb{W}_{i,t-1} \end{bmatrix} = 0$$

where $\mathbb{W}_{i,t-1}$ is a P-dimensional vector whose elements are contained in the information set $\Omega_{i,t-1}$. The system defined by (10) and (11) provides the following $(2\,P \times 1)$ vector of sample moment conditions, involving four (4) parameters to be estimated

$$g(\mathbf{X},\, \mathbf{\Theta_0}) = \frac{1}{N} \sum_i \begin{bmatrix} \frac{1}{T} \sum_t \upsilon_{1,i,t}\, \mathbb{W}_{i,t-1} \\ \\ \frac{1}{T} \sum_t \upsilon_{2,i,t}\, \mathbb{W}_{i,t-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{T} \sum_t f_1(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_0)\, \mathbb{W}_{i,t-1} \\ \\ \frac{1}{T} \sum_t f_2(\mathbf{X}_{i,t-1}, \mathbf{\Theta}_0)\, \mathbb{W}_{i,t-1} \end{bmatrix}$$

for N cross-sectional observations over T time periods. For fixed T, estimators formed using these sample orthogonality conditions will yield consistent estimators, if, when averaged over production units, they converge to zero as N gets large. As shown by Chamberlain (1984), the hypothesis of rational expectations implies the sample average of the moment conditions should converge to zero as $T \to \infty$, for a given N. However, in the presence of aggregate shocks forecast errors might be correlated across production units so that the cross-sectional mean forecast error need not converge to zero as $N \to \infty$, for a given T. When the aggregate shocks have a mean of zero and are serially uncorrelated, it might be expected that the effects of these aggregate shocks 'average out' over time, provided the time dimension of the panel is sufficiently large. Relative to existing panel data sets, which generally have short time dimensions, the data used in this paper cover, at least, twenty five years. Consequently, it seems reasonable to assume that the effects of any aggregate shocks are indeed averaged out and the estimator formed by using the sample orthogonality conditions will yield consistent estimates of the parameter vector. The (implicit) assumption here is that the components of the information set $\mathbb{W}_{i,t-1}$ are of no use in forecasting future aggregate shocks[14].

The hypothesis of rational expectations requires that the forecast errors of production units be uncorrelated with any information available at the time of the forecast. However, violation of this condition need not invalidate the hypothesis of rational expectations but might simply represent the inability of the econometrician to observe all elements of the agents's information set. In particular, the forecast errors might be correlated with information contained in the (observed) information set due to some unobserved factor that is fixed over time. Following Holtz-Eakin (1988), an alternative orthogonality condition may be constructed where the forecast error contains an additive fixed effect that is correlated with $\mathbb{W}_{i,t-1}$. In this case twice-lagged values of the endogenous variables will be orthogonal to the first difference of the forecast error since 'differencing' eliminates the fixed effect. Unfortunately, with persistent annual data, twice-lagged values of the endogenous variables will represent quite weak instruments providing only weak identification of the structural parameters[15].

The hypothesis of rational expectations implies that candidates for the instruments $\mathbb{W}_{i,t-1}$ include any variables in the information set at time $(t-1)$, including lagged values of the endogenous variables. The instrument set $\mathbb{W}_{i,t-1}$ includes single lags of each of the following (six) ratios and a constant:

$$\left[\frac{\text{vout}_{i,t-1}}{\text{pay}_{i,t-1}}, \frac{\text{vout}_{i,t-1}}{\text{intermed}_{i,t-1}}, \frac{\text{pay}_{i,t}}{\text{pay}_{i,t-1}}, \frac{\text{intermed}_{i,t}}{\text{intermed}_{i,t-1}}, \frac{\text{vout}_{i,t}}{\text{pay}_{i,t-1}}, \frac{\text{vout}_{i,t}}{\text{intermed}_{i,t-1}}\right]$$

Given these instruments, there at most $(2\,P - K) = 12$ over-identifying restrictions. Alternatively, the system may be estimated by selecting at most two instruments, and a constant, from $\mathbb{W}_{i,t-1}$, with a total of 2 over-identifying restrictions.

[14] Alternatively, it is common to assume that aggregate shocks affect all cross-sectional units in the same way so their effects might be captured by time dummies. This will work provided that forecast errors may be decomposed into a period-specific component, reflecting aggregate shocks, and an idiosyncratic component. Given the non-linear nature of the models being estimated, it is unlikely that the underlying model will yield additively separable forecast errors that may be decomposed into an aggregate component and idiosyncratic error.

[15] Once again, given the non-linear nature of the models being estimated, it is unlikely that the underlying model will yield additively separable forecast errors that may be decomposed into an fixed time invariant component and a time-varying idiosyncratic component.

3.1 Data

The model is first estimated using the NBER manufacturing database which provides annual data for 459 manufacturing industries (using 1987 SIC codes) from 1958–1996. Since the estimating equations do not require data on the stock of experience in order to estimate the learning parameters, this industry data represents an obvious first choice to study learning-by-doing in the manufacturing industry. These industry data are publicly available on-line and contain a lengthy time series of annual observations. Details are provided in Bartlesman & Gray (1996). This is also the same data that has been used by Cooper & Johri (2002) and thus allows a direct comparison of the estimation results to their work. The measure of the value of output provided in the NBER manufacturing database is the value of 4-digit industry shipments, not adjusted for the change in inventories. Although it is feasible to construct a measure of the value of production using data on the change in the stock of inventories, Bartlesman & Gray (1996) note that this might introduce considerable measurement error, particularly for the period before 1982.

It is possible that using 4-digit industry data involves considerable 'canceling out' of idiosyncratic plant-level productivity shocks, thereby removing much of the variation useful in identifying the structural parameters associated with learning-by-doing. Consequently, the model is also estimated using the Canadian Annual Survey of Manufactures which contains plant-level data on the variables needed to estimate the parameters of the system defined by (10) and (11). Statistics Canada have used this Annual Survey of Manufactures (ASM) to construct a lengthy panel of annual observations for a large cross section of the population of Canadian manufacturing establishments over the period 1973-1997. This data has not been used previously to examine learning-by-doing in manufacturing establishments and may be used to provide some alternative estimates to those currently in the literature.

For the purposes of this paper, the Canadian ASM sample was restricted to establishments classified as plants that are present in the survey for all years from 1973 to 1997. This focus upon continuing plants can be motivated in several ways. First, the structural model considered in this paper is one in which there is no entry or exit. Consequently, the decision to restrict the sample to a balanced panel of continuing plants reflects this assumption. Second, the structural model presented in this paper is essentially a model of plant (or establishment) behavior, reflecting the level at which data is collected in the Annual Survey of Manufactures. However, the decision to exit will likely be (endogenously) determined at the firm level which would require a well specified structural model that explicitly accounts for the decision to exit[16].

A data appendix provides some detail on the both the NBER database and the Canadian plant-level data.

4 Estimation Results

NBER Manufacturing Database

Table 1 presents estimation results using the 4-digit NBER manufacturing industry data for the period 1960–1996, using the instrument set[17]:

$$\tilde{\mathbb{W}}_{i,t-1} = \left[1, \frac{\text{pay}_{i,t-1}}{\text{pay}_{i,t-2}}, \frac{\text{intermed}_{i,t-1}}{\text{intermed}_{i,t-2}}\right]$$

As noted above, in the absence of further identifying restrictions, only the structural parameters $(\beta\,\eta)$ and $(\beta\,\varepsilon\,\gamma)$ are identified. Since β is not identified, there is an infinite combination of β, η and $\varepsilon\,\gamma$ that are consistent with the estimates of $(\beta\,\eta)$ and $(\beta\,\varepsilon\,\gamma)$ presented in Table 1. However, at least in the macroeconomic literature, there is a consensus that $\beta \in [0.90, 1]$ implying a rate of time preference $\rho \in [11\%, 0\%]$. Consequently, Table 2 presents estimation results for η and $\varepsilon\,\gamma$ associated with values of $\beta = \{0.90,\ 0.95,\ 0.98\}$. In contrast to the parameters $(\beta\,\eta)$ and $(\beta\,\varepsilon\,\gamma)$, the structural parameters η, $\varepsilon\,\gamma$ and their sum $(\eta + \varepsilon\,\gamma)$ have an economic interpretation.

[16] Although the determinants of entry and exit in the presence of production experience accumulation might be an important research question, such a structural model will generally be outside the scope of this paper. This is further complicated by data considerations and especially the ability to aggregate the plant level observations in the data to obtain firm level variables. Since the data cover 25 years, this requires considerable specific knowledge regarding corporate restructuring, mergers, and acquisitions. Olley & Pakes (1996) discuss some potential selection biases that may be introduced by restricting the sample to continuing plants.

[17] This 'optimal' instrument set, as a subset of $\mathbb{W}_{i,t-1}$ is chosen by minimizing the *GMM_AIC*, *GMM_HQIC* and *GMM_BIC* selection criteria of Andrews (1999). These tests are performed using an optimal weighting matrix constructed using the 'de-meaned' sample moment conditions.

The production technology (7) and the log-linear accumulation technology (4) imply the following

$$\ln\, Q_{it} = \{\theta\, \ln\, K_{it} - \theta\, \eta\, \ln\, K_{i,t-1}\} + \{\alpha\, \ln\, H_{it} - \alpha\, \eta\, \ln\, H_{i,t-1}\}$$
$$+ \{\phi\, \ln\, M_{it} - \phi\, \eta\, \ln\, M_{i,t-1}\} + (\eta + \varepsilon\, \gamma)\, \ln\, Q_{i,t-1} + \{\ln\, A_{it} - \eta\, \ln\, A_{i,t-1}\}$$

where $(\eta + \varepsilon\, \gamma)$ represents the 'persistence' elasticity in output generated by learning-by-doing[18]. The results in Tables 1 & 2 provide an estimate of $(\eta + \varepsilon\, \gamma) = 0.7881$, the persistence elasticity in output generated by learning-by-doing, implying a 10% increase in output in period $(t - 1)$ leads to a 7.8% increase in current output.

Alternatively, the production technology (7) and the log-linear accumulation technology (4) imply the following relationship between 'measured' total factor productivity, the lag of measured total factor productivity, and the lagged level of output:

$$\begin{aligned} \ln \mathrm{tfp}_{i,t} &= \ln\, Q_{it} - \theta\, \ln\, K_{it} - \alpha\, \ln\, H_{it} - \phi\, \ln\, M_{it} \\ &= \eta\, \ln \mathrm{tfp}_{i,t-1} + \varepsilon\, \gamma\, \ln Q_{i,t-1} + \{\ln A_{i,t} - \eta\, \ln A_{i,t-1}\} \end{aligned} \tag{12}$$

In the absence of persistence in exogenous productivity A_{it}, the (structural) parameter η represents the persistence in (measured) total factor productivity associated with learning-by-doing. Similarly, $(\varepsilon\, \gamma)$ represents the 'scale effect' in total factor productivity, associated with learning-by-doing. It is clear from (12) that η cannot be identified from data on total factor productivity, without further identifying assumptions. Specifically, it would be impossible to identify persistence in total factor productivity associated with learning-by-doing from persistence in (exogenous) productivity[19].

For reasonable values of the discount rate $\beta \in [0.90, 1]$, the results presented in Tables 1 & 2 imply estimates of the persistence in total factor productivity (η) in the interval $[0.8703, 0.7833]$,

---

[18] It is clear that the studies of learning-by-doing which restrict $\gamma = \eta = 1$ imply an explosive series for output. Similarly, allowing for the depreciation of production experience $(\eta < 1)$, restricting $\gamma = 1$, also implies an explosive series provided $\eta > (1 - \varepsilon)$. Therefore either the learning rate (ε) or the 'retention rate' for production experience (η) must be sufficiently low in order to bound the level of output.

[19] For example, in the absence of learning-by-doing and $\ln\, A_{it} = \rho\, \ln\, A_{i,t-1} + \xi_{it}$ with $\xi_{it} \sim \mathcal{N}(0, \sigma_\xi^2)$:

$$\ln \mathrm{tfp}_{i,t} = \rho\, \ln \mathrm{tfp}_{i,t-1} + \xi_{it}$$

statistically different from both zero and unity at a 5% level of significance. For example, holding 'investment' in production experience fixed, these results provide an estimate of the 'retention' elasticity of production experience of approximately 0.8245 for $\beta = 0.95$, implying a 10% increase in last period's stock of experience produces approximately an 8.25% increase in the current stock of experience. Alternatively, these results imply a depreciation rate for production experience of approximately 17.5%.

Similarly the results in Tables 1 and 2 imply estimates of $\epsilon\,\gamma \in [0.0054, 0.0049]$. Holding the stock of experience constant, these results imply a 10% increase in last period's output increases total factor productivity by 0.05%. This parameter is estimated with considerable imprecision such that the point estimate is not significantly different from zero.

Although the 'joint effect' given by $(\eta+\varepsilon\,\gamma)$ provides evidence on the importance of learning-by-doing in generating productivity dynamics, the magnitude of the individual structural parameters controlling the productivity of experience (ε) and the parameters controlling the accumulation of experience (η and γ) is of particular interest. For example, estimates of ε may be used to calculate an implied learning rate that can be compared to existing studies. As noted above, this is not possible without further identifying assumptions. An examination of the system defined by (10) and (11) suggests that ε might be separately identified from γ when the accumulation technology exhibits constant returns to scale such that $\eta = (1-\gamma)$[20]. Table 3 presents estimates of the (structural) parameters, imposing constant returns to scale in accumulation. The estimates of ε imply a learning rate in the interval $[1.7401\%, 1.5648\%]$[21].

With or without constant returns to scale in accumulation, the J-test of the over-identifying restrictions is not rejected, at the 5% level. Since this J-test of the over-identifying restrictions will be a joint test of the moment restrictions and the structural model itself, including the functional forms, these results imply that the accumulation of production experience and the corresponding dynamic structure are not rejected by the data. Moreover, the results imply that the specific func-

[20] Cooper & Johri (2002) & Clarke (2006) contain a discussion as to why it might be appropriate to impose constant returns to scale in accumulation. An LM test fails to reject this restriction of constant returns to scale in accumulation. This LM test does reject small degrees of increasing returns to scale suggesting the test has some power to reject alternatives.

[21] Using the production technology (7), a doubling of the stock of experience reduces average costs of production by a factor of 2^{ε} so the learning rate is given by $(2^{\varepsilon}-1)$.

tional forms are not rejected by the data and that the errors are uncorrelated with the instruments.

As an approximate test of the model, it is possible to compare the predicted factor shares, implied by the structural model, to the mean factor shares in the data. These predicted factor shares are given by:

$$\text{lsh} = \frac{\alpha}{\mu}\,\frac{1-\beta\,\eta}{1-\beta\,\eta-\beta\,\varepsilon\,\gamma} \qquad \text{and} \qquad \text{msh} = \frac{\phi}{\mu}\,\frac{1-\beta\,\eta}{1-\beta\,\eta-\beta\,\varepsilon\,\gamma}$$

Table 1 reports estimates of these predicted factor shares and their standard errors, computed using the Delta method[22]. Although the predicted share of intermediate inputs in revenues is not significantly different from its sample counterpart, the labor share is significantly lower than the mean share in the data[23].

It is possible that the results using the full NBER manufacturing sample average some important variation across cross-sectional units, such as industry groups. Consequently, results are also presented for a set of broad industry groups. This industrial classification is based upon Baldwin & Raffiquzzaman (1994) that divides the manufacturing industry into five industry groups[24]. An Appendix provides some characteristics of these broad industry groups.

Tables 1 and 2 present results for the five broad industry groups. With the exception of 'Scale-Based' industries, the J-test of the over-identifying restrictions is not rejected. Although the predicted share of intermediate inputs in revenues is not significantly different from its sample counterpart for all but 'Product-Differentiated' industries, the labor share is significantly lower than the mean share in the data for all industrial groups.

Although there is some variation in the point estimates of $\beta\,\eta$ for each of the five broad industry groups, they are not significantly different from those for the full sample. The estimate of $\beta\,\varepsilon\,\gamma$ is only significantly different from zero and the full NBER sample, at the 5% level, for industry group 4 (Product-Differentiated Industries). For values of $\beta \in [0.90, 1]$, holding the stock of experience constant, these results imply a 10% increase in last period's output increases total

[22] There is little difference in these standard errors computed using bootstrapped samples.

[23] This result might be expected in the presence of measurement error in the wage bill, particularly if there is a utilization component associated with the labor input.

[24] An alternative classification based upon 2-digit (SIC) industry groups would result in too few observations to be able to adequately identify the parameters of interest.

factor productivity in the interval $[8.54\%, 7.69\%]$.

Imposing constant returns to scale in the accumulation technology produces estimates of ε that are significantly different from zero for industry group 2 (Labor-Intensive Industries) and industry group 4 (Product-Differentiated Industries). These estimates imply learning rates in the interval $[12.5839\%, 11.2573\%]$ for Labor-Intensive Industries and $[19.0582\%, 16.9993\%]$ for Product-Differentiated Industries.

The results presented in Tables 1–3 provide evidence that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with the 4-digit manufacturing industry data. Despite this, variation in the estimated structural parameters across broad industry groups suggests that conclusions regarding the importance of productivity dynamics generated by learning-by-doing, as a wide-spread phenomena in the manufacturing industry as a whole, might be problematic. The relatively low estimated learning rates for the full sample in the interval $[1.7401\%, 1.5648\%]$ average across industry groups with relatively large and significant learning rates and industry groups with relatively small learning rates.

Since the estimates of ε for the full sample imply considerably lower 'learning rates' than those provided in much of the existing empirical literature, it is useful to investigate this result further. The structural model presented in this paper differs from the typical study of learning-by-doing in several ways: (1) allowing for the depreciation of production experience; (2) allowing for a non-unit elasticity associated with 'investments' in production experience ($\gamma \neq 1$); (3) a log-linear accumulation technology, allowing for diminishing marginal productivity in accumulation; and (4) an optimizing model for the accumulation of production experience. It is possible to rely upon Cooper & Johri (2002) who present estimates, based upon a model with these first three (3) features. Using the same NBER manufacturing data and imposing constant returns to scale in accumulation, they estimate a regression such as (12), providing estimates of $\varepsilon = 0.16$ and $\eta = 0.55$, with an implied learning rate of approximately 11.75%. The conclusion from a comparison with this study is that these first three features do not appear to considerably reduce the estimated learning rates. Consequently, the lower estimated learning rates provided in this paper might reasonably be attributed to the final feature of the model—treating the stock of experience as a state variable under the control of production unit.

An alternative explanation for low estimated learning rates might be due to the particular instrument set used to identify the parameters of interest. In order to investigate the dependence of the results presented in this paper upon the instrument set, results are available using the 'aggregate instruments' used by Cooper & Johri (2002). Although these instruments produce larger estimates of $\varepsilon\,\gamma$ (and ε in the crs case) and lower estimates of η, the results are generally not significantly different from those presented in Tables 1–3. For example, using these instruments produces estimates of $\beta\,\varepsilon\,\gamma = -0.0027$ and $\beta\,\eta = 0.8048$ for the full sample. These results are consistent with the results presented in Tables 1–3.

Canadian Plant-Level Manufacturing Data

Table 4 presents estimation results using the Canadian plant-level data, imposing constant returns to scale in accumulation, and using a value of $\beta = 0.9627$, the annual equivalent of the Bank of Canada 90 day Commercial Paper rate[25]. In contrast to the more aggregated NBER 4-digit industry data, the J-test of the over-identifying restrictions is generally rejected. The exception is industry group 5 (Science-Based Industries). This J-test is a joint test of the moment restrictions and the structural model, including the functional forms and suggests that the accumulation of production experience and the corresponding dynamic structure are (generally) not consistent with the plant-level data.

The conclusions from the J-test suggest that attention should be restricted to the results for Science-Based industries. The broad industrial group 'Science-Based' Industries contains many of the industries that have been used to previously study learning-by-doing, for example the semi-conductor industry studied by Jarmin (1994), Irwin & Klenow (1994), and Zulehner (2003) and the commercial aircraft industry studied by Benkard (2000). The results presented in Table 4 provide an estimate of $\varepsilon = 0.40$, significant at the 10% level, with an estimated learning rate of 31%. These results are consistent with the results provided in Benkard (2000) for a commercial aircraft manufacturer providing a learning rate of approximately 35%-40%. The estimated learning rate in

[25] The hypothesis of constant returns to scale in accumulation is not rejected using an LM test.

Table 4 is larger than the average learning rate of approximately 20%, across different generations of DRAM chips, provided by Irwin & Klenow (1994) and Zulehner (2003). The results in Table 4 for η provide an annual retention rate for production experience of 71.86%, consistent with the results obtained by Benkard (2000) with an estimated retention rate of approximately 61%.

These results suggest that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with plant-level data for Science-Based industries. The estimates for this industry are consistent with previous work that has studied firms within this broad industry group. However, the J-test is generally rejected for other industry groups. This suggests that learning-by-doing does might not play an important role in generating productivity dynamics in these industries. Consequently, using results from existing studies that focus upon specific, narrowly defined industries or firms, may lead to misleading conclusions concerning the importance of learning-by-doing, as a widespread phenomena, across manufacturing industries as a whole.

An interesting result emerges when comparing the plant-level estimates to those obtained using the more aggregate NBER data. Compared to the NBER data for Science-Based industries, the estimated learning rates are considerably larger at the plant level. Since it is not possible to distinguish between proprietary learning and (external) spill-overs in the aggregate industry data, it is expected that, all else equal, learning rates should be higher at the industry level.

This result of lower estimated learning rates using the industry data is most likely driven by an 'aggregation bias' in the presence of plant-level heterogeneity in productivity dynamics associated with learning-by-doing. The existing microeconomic evidence suggests that productivity differences across plants are driven by considerable heterogeneity in plant-level idiosyncratic shocks, uncorrelated with quite detailed industry characteristics. Even within Science-Based industries, at a point in time, some plants might be accumulating considerable production experience while a large majority of plant might be accumulating very little experience with considerable heterogeneity in the productivity of this experience. The estimated learning rates at the plant level average across plants with relatively large and significant learning rates and plants with relatively small learning rates. In contrast, the results using the (aggregate) industry data treat each 4-digit industry as a single 'representative' plant. The complexity of plant-level heterogeneity coupled with the highly non-linear nature of the structural model implies that the average estimated learning rates (at the plant level) may be quite different to the learning rate of this representative plant, identified in the aggregate estimates. Consequently, it is not particularly meaningful to compare the average estimated learning rate obtained using the aggregate data to the estimated learning rate at the plant-level, which averages over plants with heterogenous productivity outcomes.

Of course care must be exercised when comparing aggregate industry data for the U.S. with plant level observations on Canadian manufacturing establishments. It is possible that the Canadian manufacturing industry is characterized by higher learning rates at both the plant level and the industry level. Using aggregate 2-digit Canadian manufacturing data over the period 1983-1997 provides estimates of $\varepsilon = 0.0728$, implying a learning rate of approximately 5%. Since ε is estimated relatively imprecisely, this is not significantly different from the estimated learning rate for both the full NBER manufacturing sample and the sample covering the same time period. This suggests that the lower estimated learning rates in the aggregate data are most likely a function of aggregating a series of heterogenous accumulation technologies, rather than any cross-country differences.

An alternative explanation for the finding of higher learning rates at the plant-level is that the larger plants account for a greater proportion of industry outputs and inputs and receive greater weight in the industry level data[26]. An examination of Figure 1, which provides a histogram of the size distribution (by employment) for the plant-level data reveals that medium size plants (20-49 employees) account for approximately 25% of the plant level observations, with just under 5% of plants with over 500 employees. If larger plants, on average, have lower learning rates, it is anticipated that industry level learning rates would be lower than the plant level learning rates. There is some weak evidence that larger plants, on average, have lower learning rates than smaller plants across some broad size categories. However, these results are quite sensitive to the choice of instruments used to identify the parameters.

[26] Unfortunately, the NBER data does not contain information of either the number of plants or the size of these individual plants used in the calculation of the industry totals

5 Conclusions

This paper has presented an alternative strategy for estimating the parameters associated with learning-by-doing using commonly available datasets. In contrast to existing (microeconomic) studies of learning-by-doing, which have largely focused upon production function estimation, this paper provides estimates of the learning parameters by estimating the first order condition from a structural model that allows for the accumulation of production experience. This estimation strategy is illustrated using both 4-digit industry data and observations collected at the plant level. The identifying assumptions have the advantage that estimation of the structural learning parameters do not require data on the stock of production experience, allowing a study of learning-by-doing in a much broader context than previous studies.

The results suggest that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with the 4-digit manufacturing industry data. The results also indicate that estimated learning rates might be considerably lower than existing estimates, such as those in Cooper & Johri (2002). The paper has identified variation in estimated learning rates, even at the 4-digit industry level. This is a noteworthy result that has previously not been identified in the existing literature.

These results indicate that the dynamic structure implied by the structural model of learning-by-doing is broadly consistent with plant-level data for Science-Based industries only. The estimates for this industry are consistent with previous work that has studied firms within this broad industry group. However, the model is generally rejected for other industry groups. This implies that using results from existing studies that focus upon specific, narrowly defined industries or firms, may lead to misleading conclusions concerning the importance of learning-by-doing, as a widespread phenomena, across manufacturing industries as a whole.

While the identifying assumptions allow learning to be studied in a much broader context than previously available, this comes at a cost. The plant-level dynamics associated with learning-by-doing are most likely too complicated to be able to be captured by a single model for the entire manufacturing industry, particularly in the presence of tremendous heterogeneity in plant-level productivity. These results, together with the results using the aggregate NBER data, imply the ‘continuous’ model of accumulation used in this paper may be a more appropriate model for the NBER data which aggregates across heterogenous plants. A richer specification for the accumulation of production experience might be better able to capture heterogeneity in the dynamics associated with learning-by-doing at the plant level. However, this significantly increases the data requirements, reducing the scope for studying learning across broadly defined industrial groups. The challenge of future empirical research is to build richer models of productivity dynamics that may be estimated given the constraints imposed by data availability.

References

Andrews, D. W. K. (1999). Consistent Moment Selection Procedures for Generalized Method of Moments Estimation. *Econometrica*, 67(3), 543–564.

Argote, L., Beckman, S. L., & Epple, D. (1990). The Persistence and Transfer of Learning in Industrial Settings. *Management Science*, 36(2), 140–154.

Bahk, B.-H. & Gort, M. (1993). Decomposing Learning by Doing in New Plants. *Journal of Political Economy*, 101(4), 561–583.

Baldwin, J. & Raffiquzzaman, M. (1994). Structural Change in the Canadian Manufacturing Sector 1970—1990. *Analytical Studies Branch Research Paper*, 61. Statistics Canada, Ottawa.

Baldwin, J. R. & Gu, W. (2003). Plant Turnover and Productivity Growth in Canadian Manufacturing. *Analytical Studies Branch—Research Paper Series*, 193.

Bartlesman, E. J. & Gray, W. (1996). The NBER Manufacturing Productivity Database. *NBER Technical Working Paper*, 205.

Benkard, C. L. (2000). Learning and Forgetting: The Dynamics of Aircraft Production. *American Economic Review*, 90(4), 1034–1054.

Blanchard, O. & Kiyotaki, N. (1987). Monopolistic Competition and Aggregate Demand Disturbances. *American Economic Review*, 77(4), 647–666.

Chamberlain, G. (1984). Panel Data. In Z. Griliches & M.D. Intriligator (Eds.). *Handbook of Econometrics Volume II*, (pp. 1248–1318). Elsevier Science B.V.

Clarke, A. J. (2006). Learning-by-Doing and Aggregate Fluctuations: Does the Form of the Accumulation Technology Matter? *Economics Letters*, 92, 434—439.

Cooper, R. W. & Johri, A. (2002). Learning-by-Doing and Aggregate Fluctuations. *Journal of Monetary Economics*, 49, 1539–1566.

Darr, E. D., Argote, L., & Epple, D. (1995). The Acquisition, Transfer, and Depreciation of Knowledge in Service Organizations: Productivity in Franchises. *Management Science*, 41(11), 1750–1762.

Epple, D., Argote, L., & Devadas, R. (1991). Organizational Learning Curves: A Method for Investigating Intra-Plant Transfer of Knowledge Acquired Through Learning by Doing. *Organization Science*, 2(1), 58–70.

Fudenberg, D. & Tirole, J. (1983). Learning-by-Doing and Market Performance. *The Bell Journal of Economics*, 14(2), 522–530.

Ghemawat, P. & Spence, M. A. (1985). Learning Curve Spillovers and Market Performance. *Quarterly Journal of Economics*, 100(5), 839–852.

Holtz-Eakin, D. (1988). Testing for Individual Effects in Autoregressive Models. *Journal of Econometrics*, 39, 297–307.

Irwin, D. A. & Klenow, P. J. (1994). Learning-by-Doing Spillovers in the Semiconductor Industry. *Journal of Political Economy*, 102(6), 1200–1227.

Jarmin, R. S. (1994). Learning by Doing and Competition in the Early Rayon Industry. *RAND Journal of Economics*, 25(3), 441–454.

Klette, T. J. & Griliches, Z. (1996). The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous. *Journal of Applied Econometrics*, 11, 343–361.

Marshall, A. (1961). *Principles of Economics*. MacMillan & Co Ltd, London, 8th edition.

Olley, S. G. & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6), 1263–1297.

Rapping, L. (1965). Learning and World War II Production Functions. *Review of Economics and Statistics*, 47(1), 81–86.

Rosen, S. (1972). Learning by Experience as Joint Production. *Quarterly Journal of Economics*, 86(3), 366–382.

Sheshinski, E. (1967). Tests of the Learning-by-Doing Hypothesis. *Review of Economics and Statistics*, 49(4), 568–578.

Thompson, P. (2001). How Much Did the Liberty Shipbuilders Learn?: New Evidence for an Old Case Study. *Journal of Political Economy*, 109(1), 103–137.

Thornton, R. A. & Thompson, P. (2001). Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding. *American Economic Review*, 91(5), 1350–1369.

Zulehner, C. (2003). Testing Dynamic Oligopolistic Interaction: Evidence from the Semiconductor Industry. *International Journal of Industrial Organization*, 21, 1527–1556.

	Full	Industry Groups				
		1	2	3	4	5
	Production Parameters					
α/μ	**0.1652***	**0.0901***	**0.1797***	**0.2032***	**0.1696***	**0.1906***
	(0.0130)	(0.0202)	(0.0113)	(0.0270)	(0.0133)	(0.0162)
ϕ/μ	**0.4837***	**0.4392***	**0.4122***	**0.6347***	**0.3281***	**0.3944***
	(0.0378)	(0.0943)	(0.0343)	(0.0898)	(0.0270)	(0.0304)
	Learning Parameters					
$\beta\,\varepsilon\,\gamma$	0.0049	0.0228	0.0336	-0.0112	**0.0769***	0.0105
	(0.0114)	(0.0228)	(0.0213)	(0.0131)	(0.0288)	(0.0075)
$\beta\,\eta$	**0.7833***	**0.8331***	**0.7813***	**0.8737***	**0.6604***	**0.8102***
	(0.0319)	(0.0432)	(0.0697)	(0.0442)	(0.0642)	(0.0432)
$\beta(\eta + \varepsilon\,\gamma)$	**0.7881***	**0.8559***	**0.8149***	**0.8625***	**0.7373***	**0.8207***
	(0.0271)[c]	(0.0311)[c]	(0.0498)[c]	(0.0543)[c]	(0.0408)[c]	(0.0398)[c]
	'Factor Shares'					
Labour Sh.	**0.1690***	**0.1044***	**0.2123***	**0.1866***	**0.2193***	**0.2018***
	(0.0061)[c]	(0.0113)[c]	(0.0051)[c]	(0.0126)[c]	(0.0063)[c]	(0.0131)[c]
Materials Sh.	**0.4947***	**0.5087***	**0.4871***	**0.5829***	**0.4242**	**0.4176***
	(0.0145)[c]	(0.0430)[c]	(0.0115)[c]	(0.0767)[c]	(0.0089)[c]	(0.0203)[c]
J-Statistic	0.4085	5.8460	0.0513	13.9755	1.2689	1.3287
df	2	2	2	2	2	2
p-value	0.8153	0.0538	0.9747	0.0009	0.5302	0.5146
N	16946	4810	4551	2812	2849	1924

[a] Standard errors shown in parentheses
[b] Bold entries denote estimate is significant at 5% level. Starred (*) entries denote estimate is significant at 1% level
[c] Standard errors calculated using the Delta method

Table 1: NBER Manufacturing Industry Data 1960–1996: Estimation Results by Industry Group

	Full	Industry Groups				
		1	2	3	4	5
	$\beta = \mathbf{0.90}$					
$\varepsilon\,\gamma$	0.0054	0.0253	0.0373	-0.0124	**0.0854***	0.0117
	(0.0127)	(0.0253)	(0.0237)	(0.0348)	(0.0320)	(0.0083)
η	**0.8703***	**0.9257***	**0.8681***	**0.9708***	**0.7338***	**0.9002***
	(0.0354)	(0.0480)	(0.0774)	(0.0491)	(0.0713)	(0.0480)
$(\eta + \varepsilon\,\gamma)$	**0.7881***	**0.8559***	**0.8149***	**0.8625***	**0.7373***	**0.8207***
	(0.0301)	(0.0345)	(0.0533)	(0.0603)	(0.0453)	(0.0442)
	$\beta = \mathbf{0.95}$					
$\varepsilon\,\gamma$	0.0052	0.0240	0.0354	-0.0118	**0.0809***	0.0111
	(0.0120)	(0.0240)	(0.0224)	(0.0329)	(0.0303)	(0.0079)
η	**0.8245***	**0.8769***	**0.8224***	**0.9197***	**0.6952***	**0.8528***
	(0.0336)	(0.0455)	(0.0734)	(0.0465)	(0.0676)	(0.0455)
$(\eta + \varepsilon\,\gamma)$	**0.7881***	**0.8559***	**0.8149***	**0.8625***	**0.7373***	**0.8207***
	(0.0285)	(0.0327)	(0.0524)	(0.0571)	(0.0429)	(0.0419)
	$\beta = \mathbf{0.98}$					
$\varepsilon\,\gamma$	0.0050	0.0233	0.0343	-0.0114	**0.0785***	0.0107
	(0.0116)	(0.0233)	(0.0217)	(0.0319)	(0.0294)	(0.0077)
η	**0.7993***	**0.8501***	**0.7972***	**0.8915***	**0.6739***	**0.8267***
	(0.0326)	(0.0441)	(0.0711)	(0.0451)	(0.0655)	(0.0441)
$(\eta + \varepsilon\,\gamma)$	**0.7881***	**0.8559***	**0.8149***	**0.8625***	**0.7373***	**0.8207***
	(0.0277)	(0.0317)	(0.0508)	(0.0554)	(0.0416)	(0.0406)
N	16946	4810	4551	2812	2849	1924

[a] Standard errors shown in parentheses
[b] Bold entries denote estimate is significant at 5% level. Starred (*) entries denote estimate is significant at 1% level

Table 2: NBER Manufacturing Industry Data 1960–1996: Estimation Results by Industry Group, by calibrated discount rate

	Full	Industry Groups				
		1	2	3	4	5
	Production Parameters					
α/μ	**0.1652***	**0.0901***	**0.1797***	**0.2031***	**0.1696***	**0.1906***
	(0.0130)	(0.0202)	(0.0113)	(0.0269)	(0.0133)	(0.0162)
ϕ/μ	**0.4837***	**0.4388***	**0.4121***	**0.6344***	**0.3281***	**0.3944***
	(0.0378)	(0.0942)	(0.0343)	(0.0897)	(0.0270)	(0.0304)
	Learning Parameters					
$\beta\,\varepsilon$	0.0224	0.1372	**0.1539***	-0.0887	**0.2265***	0.0556
	(0.0506)	(0.1138)	(0.0533)	(0.0842)	(0.0508)	(0.0344)
$\beta\,\eta$	**0.7833***	**0.8331***	**0.7812***	**0.8737***	**0.6602***	**0.8102***
	(0.0319)	(0.0433)	(0.0698)	(0.0442)	(0.0642)	(0.0432)
$\beta(\eta+\varepsilon\,\gamma)$	**0.7881***	**0.8560***	**0.8149***	**0.8625***	**0.7372***	**0.8207***
	(0.0271)[c]	(0.0311)[c]	(0.0498)[c]	(0.0543)[c]	(0.0408)[c]	(0.0398)[c]
$\beta(\varepsilon\,\gamma)$	0.0049	0.0229	0.0337	-0.0112	**0.0770***	0.0105
	(0.0114)[c]	(0.0228)[c]	(0.0213)[c]	(0.0130)[c]	(0.0288)[c]	(0.0075)[c]
J-Statistic	0.4084	5.8558	0.0513	13.9794	1.2689	1.3287
df	2	2	2	2	2	2
p-value	0.8153	0.0535	0.9747	0.0009	0.5302	0.5146
LM Test Stat[d]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	16946	4810	4551	2812	2849	1924

[a] Standard errors shown in parentheses
[b] Bold entries denote estimate is significant at 5% level. Starred (*) entries denote estimate is significant at 1% level
[c] Standard errors calculated using the Delta method
[d] L.M. test statistic (asympt). distributed as $\chi^2(1)$

Table 3: NBER Manufacturing Industry Data 1960–1996: Estimation Results by Industry Group Imposing CRS in Accumulation

	Full	Industry Groups				
		1	2	3	4	5
	Production Parameters					
α/μ	**0.0267***	**0.0206***	0.0257	**0.0336***	0.0008	0.0253
	(0.0052)	(0.0045)	(0.0144)	(0.0151)	(0.0213)	(0.0134)
ϕ/μ	**0.1682***	**0.1516***	0.0796	0.1965	0.0020	0.1727
	(0.0397)	(0.0377)	(0.0459)	(0.1089)	(0.0558)	(0.0924)
	Learning Parameters					
ε	**0.3913***	**0.3801***	**0.7177***	0.3232	**0.9904***	0.3999
	(0.0995)	(0.1003)	(0.1432)	(0.2464)	(0.2622)	(0.2331)
η	**0.7711***	**0.8138***	0.2964	**0.8003***	-20.0579	**0.7186***
	(0.0409)	(0.0309)	(0.3498)	(0.0761)	(575.1669)	(0.1083)
J-Statistic	45.1058	27.0959	8.2801	11.1277	12.1314	5.3657
df	2	2	2	2	2	2
p-value	0.0000	0.0000	0.0159	0.0038	0.0023	0.0684
N	141197	47349	31678	37024	16363	8783

[a] Standard errors shown in parentheses
[b] Bold entries denote estimate is significant at 5% level. Starred (*) entries denote estimate is significant at 1% level

Table 4: Canadian ASM Plant-Level Data 1975–1996: Estimation Results by Industry Group

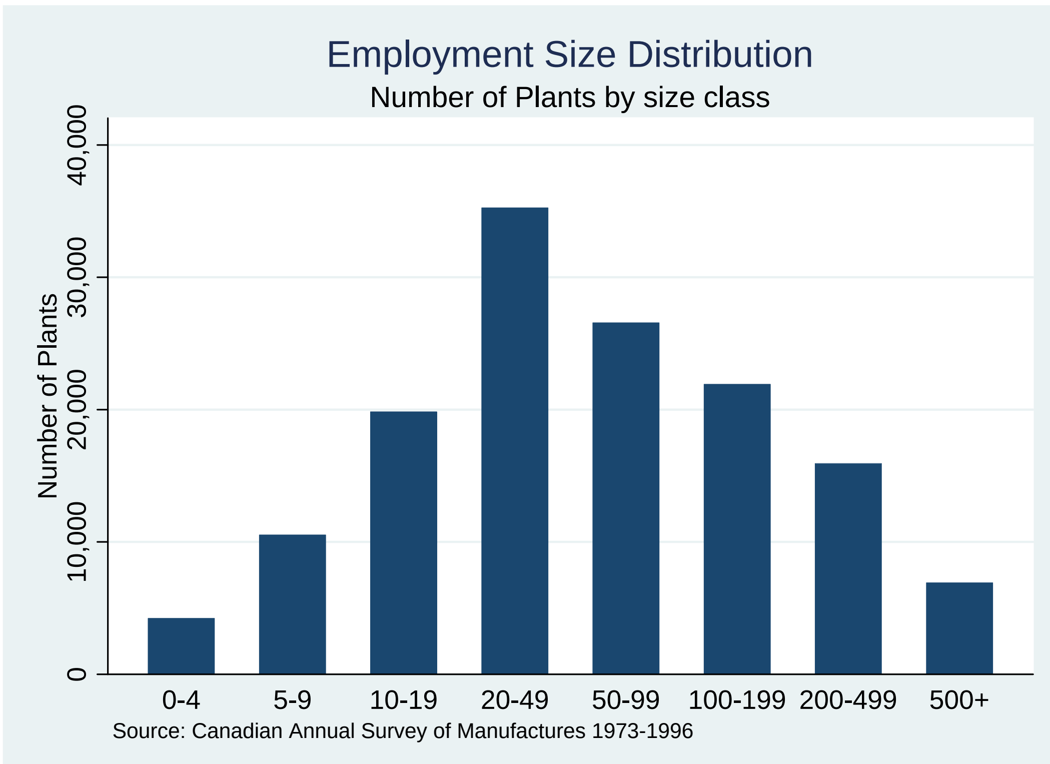

Figure 1: Employment Size Distribution

Data Appendix

OECD Industry Groups

The five broad industry groups are formed using 4-digit industry classifications rather than 2-digit codes so that it is difficult to characterize these groups according to 2-digit industrial sectors. However as a crude approximation, Natural Resource Industries include Food & Beverages, Plastic Products, some Wood Industries, some Non-Metallic Mineral Products, and Refined Petroleum & Coal Products. Labor Intensive Industries include Textiles, Clothing, Furniture & Fixtures, some Wood Industries, and some Fabricated Metal products. Scale Based Industries include Rubber & Plastic, some Paper Industries, some Printing Industries, some Primary Metals industries, some Fabricated Metals industries, and some Transport Industries. Product Differentiated Industries include some Fabricated Metals industries, Machinery Industries, and some Electrical industries. Finally Science Based Industries include some Electrical Industries, some Chemical Industries, and Instruments.

1. **Natural Resource Industries**[27]
    - primarily involved in processing of domestic raw materials
    - relatively larger materials share in output
2. **Labor Intensive Industries**
    - low physical capital to labor ratios
    - relatively larger labor share in output
    - relatively smaller plants and lower wages
    - generally protected by high tariff rates
    - relatively lower proportion of salaried employees
3. **Scale Based Industries**
    - relatively larger plants and higher wages
    - high physical capital to labor ratios
    - relatively larger labor share in output
4. **Product Differentiated Industries**
    - high advertising to sales ratios
    - producing a large number of products
    - relatively larger research and development (R & D) expenditures
    - relatively larger labor share in output
5. **Science Based Industries**
    - 'high technology' industries with high R & D expenditures
    - large percentage of workforce employed in scientific and professional occupations
    - larger plants and higher wages

[27] see Baldwin & Raffiquzzaman (1994)

NBER Productivity Database

The measure of revenue provided in the NBER data is the value of 4-digit industry shipments, in millions of dollars. Following Bartlesman & Gray (1996), the value of shipments measures the net selling value of goods and excludes discounts, returns and allowances, sales tax, excise taxes and duties and charges for outward transportation. It is not adjusted for inventory changes. The cost of intermediate inputs ($P_t s_t M_{it}$) is constructed using the cost of materials (MATCOST). This includes purchased electrical energy and fuels consumed. The total wage bill ($P_t w_t H_{it}$) is constructed using total payroll (PAY).

Canadian Annual Survey of Manufactures

An establishment is defined as the smallest unit capable of reporting certain specified input and output data. This includes materials and supplies used, goods purchased for resale, fuel and power consumed, number of employees and salaries and wages, man hours worked and paid, inventories and shipments or sales. Each establishment is identified by a unique identifier that changes if and only if the name of the establishment changes and the establishment is physically relocated and there is a change in ownership. The target population of the survey is all establishments primarily engaged in manufacturing with employees and with a value of shipments over $30,000. The survey uses two methods of data collection—direct and administrative. The direct survey method covers approximately 60% of manufacturing establishments. The long form is sent to large establishments and is highly detailed. This long form covers approximately 90% of total shipments. Unlike the long form, the short form does not collect information on raw materials or separate administrative employees from production workers. The records for the remaining 40% of establishments are extracted from administrative files.

For the purposes of the present study, the sample was restricted to establishments classified as plants that are present in the survey for all years from 1973 to 1997 and report using the long form at least once over the period 1973 to 1997. This produces a balanced panel of annual observations for 6139 plants. It excludes establishments classified as sales offices, warehouses or head offices.

There is some flexibility in the construction of the required data using the Canadian plant-level data. The value of output may be calculated using either the value of production or the value of shipments. The difference between these two measures of output reflects an adjustment for the change in inventories. Since there is no considerable difference in the results using the value of shipments, the results presented in the paper relate to the value of production. The value of output may also be calculated as the value of output attributed to either manufacturing activity (VPM) or total activity (VPT). The structural model implies that plants optimally choose total output rather than manufacturing output. In addition, the relative merits of each measure are driven by data considerations. Not all data can be decomposed into manufacturing activity and total activity. The activity for all 'small' establishments is classified as manufacturing activity. In addition, labor data are shown separately for manufacturing activity only in the case of production workers so that the number of salaried employees includes non-production workers involved in both manufacturing and non-manufacturing activity. The cost of heat and power cannot be shown separately for non-manufacturing activity. For these reasons, the results presented in this paper relate to total production (VPT) only.

The cost of intermediate inputs ($P_t s_t M_{it}$) is constructed as the sum of heat and power costs and the cost of production materials. No adjustment is made for the change in the inventory of raw materials. The total wage bill ($P_t w_t H_{it}$) includes the gross earnings of both salaried and non-salaried (production) workers before deductions for income tax. It includes payments for overtime and paid leave as well as bonuses and commissions paid. Remuneration to outside pieceworkers is included in the cost of materials.