Catalogue no. 11F0027MIE — No. 049
ISSN 1703-0404
ISBN 978-0-662-47712-9

Research Paper

Economic Analysis (EA) Research Paper Series

Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing

by John R. Baldwin, W. Mark Brown and David L. Rigby

Micro-economic Analysis Division
18-F, R.H. Coats Building, 100 Tunney's Pasture Driveway

Telephone: 1-800-263-1136

Statistics Canada Statistique Canada

Canada

Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing[1]

by
John R. Baldwin,[2] W. Mark Brown[2] and David L. Rigby[3]

11F0027MIE No. 049
ISSN 1703-0404
ISBN 978-0-662-47712-9

Micro-economic Analysis Division
18-F, R.H. Coats Building, 100 Tunney's Pasture Driveway
Statistics Canada, Ottawa K1A 0T6

February 2008

Authors' names are listed alphabetically.

We are indebted to Desmond Beckstead and Robert Gibson for research support both with regards to the longitudinal files used herein and the census and input–output calculations used to capture agglomeration effects.

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2008

All rights reserved. The content of this electronic publication may be reproduced, in whole or in part, and by any means, without further permission from Statistics Canada, subject to the following conditions: that it be done solely for the purposes of private study, research, criticism, review or newspaper summary, and/or for non-commercial purposes; and that Statistics Canada be fully acknowledged as follows: Source (or "Adapted from," if appropriate): Statistics Canada, year of publication, name of product, catalogue number, volume and issue numbers, reference period and page(s). Otherwise, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, by any means—electronic, mechanical or photocopy—or for any purposes without prior written permission of Licensing Services, Client Services Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

La version française de cette publication est disponible (n^o^ 11F0027MIF au catalogue, n^o^ 049).

---

1. This research was supported by a Canadian Studies Faculty Research grant and by matching funds from the Micro-economic Analysis Division of Statistics Canada.
2. Micro-economic Analysis Division, Statistics Canada.
3. Department of Geography, University of California, Los Angeles.

Note of appreciation:

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1-800-263-1136. The service standards are also published on www.statcan.ca under About us > Providing services to Canadians.

Economic Analysis Research Paper Series

The Economic Analysis Research Paper Series provides for the circulation of research conducted by the staff of National Accounts and Analytical Studies, visiting Fellows and academic associates. The research paper series is meant to stimulate discussion on a range of topics including the impact of the New Economy, productivity issues, firm profitability, technology usage, the effect of financing on firm growth, depreciation functions, the use of satellite accounts, savings rates, leasing, firm dynamics, hedonic estimations, diversification patterns, investment patterns, the differences in the performance of small and large, or domestic and multinational firms, and purchasing power parity estimates. Readers of the series are encouraged to contact the authors with comments, criticisms and suggestions.

The primary distribution medium for the papers is the Internet. These papers can be downloaded from the Internet at www.statcan.ca for free.

All papers in the Economic Analysis Series go through institutional and peer review to ensure that they conform to Statistics Canada's mandate as a government statistical agency and adhere to generally accepted standards of good professional practice.

The papers in the series often include results derived from multivariate analysis or other statistical techniques. It should be recognized that the results of these analyses are subject to uncertainty in the reported estimates.

The level of uncertainty will depend on several factors: the nature of the functional form used in the multivariate analysis; the type of econometric technique employed; the appropriateness of the statistical assumptions embedded in the model or technique; the comprehensiveness of the variables included in the analysis; and the accuracy of the data that are utilized. The peer group review process is meant to ensure that the papers in the series have followed accepted standards to minimize problems in each of these areas.

Publications Review Committee
Analytical Studies Branch, Statistics Canada
18th Floor, R.H. Coats Building
Ottawa, Ontario K1A 0T6

Table of contents

Abstract **6**

Executive summary **7**

1. Introduction **9**

2. Data and model **11**

2.1 Plant- and firm-specific characteristics 12
2.2 Model 13
2.3 Place-specific characteristics 15
2.4 Sample characteristics 17

3. Econometric strategy **20**

4. Panel model estimates **21**

5. Conclusion **26**

References **28**

Abstract

Productivity and wages tend to be higher in cities. This is typically explained by agglomeration economies, which increase the returns associated with urban locations. Competing arguments of specialization and diversity undergird these claims. Empirical research has long sought to confirm the existence of agglomeration economies and to adjudicate between the models of Marshall, Arrow and Romer (MAR) that suggest the benefits of proximity are largely confined to individual industries, and the claims of Jacobs (1969) that such benefits derive from a general increase in the density of economic activity in a particular place and are shared by all occupants of that location. The primary goal of this paper is to identify the main sources of urban increasing returns, after Marshall (1920). A secondary goal is to examine the geographical distance across which externalities flow between businesses in the same industry. We bring to bear on these questions plant-level data organized in the form of a panel across the years 1989 and 1999. The panel data overcome selection bias resulting from unobserved plant-level heterogeneity that is constant over time. Plant-level production functions are estimated across the Canadian manufacturing sector as a whole and for five broad industry groups, each characterized by the nature of their output. Results provide strong support for Marshall’s (1920) claims about the importance of buyer-supplier networks, labour market pooling and spillovers. The data show spillovers enhance plant productivity within industries rather than between them and that these spillovers tend to be more spatially extensive than previous studies have found.

Keywords: agglomeration economies, localization economies, productivity, urban economies

Executive summary

The growth accounting framework that is used to parse out the various determinants of economic growth focuses on changes in labour, capital, intermediate materials and an unexplained residual that is usually referred to as multifactor productivity (MFP). MFP is often attributed either to unmeasured inputs, such as managerial experience and innovative capabilities, or to externalities arising from the environment. Externalities include freely available knowledge, the social infrastructure of the economy and supporting economic structures.

Marshall’s agglomeration (localization) economies fall within the latter category. Agglomeration economies include the advantages of having the correct labour mix readily available for a firm, of having other firms that can readily supply specialized inputs and having an information flow from other firms in the same industry that reduces costs or improves the quality of the product. In all cases, distance is seen to provide cost advantages in having the right mix of labour, suppliers or information available close at hand.

Measurement of the impact of these externalities is difficult. Many early studies were hampered by their use of aggregate data and poor proxies for agglomeration. To overcome these problems, this paper makes use of detailed microdata on Canadian manufacturing plants and firms that permit both productivity and associated characteristics of the production entities to be measured. The database essentially covers the entire population, thereby reducing the selection bias associated with less comprehensive databases and it is tracked over a 10-year period.

Plants can be located precisely using constant geographic codes and tracked over time. This allows us not only to examine differences across plants at a particular point in time but, more importantly, to study how changes over time in urban characteristics have influenced productivity. Examining these changes allows us to ask whether recent growth in urban economies and changes in industrial structure have in turn fed back into changes in productivity. It also allows us to ask whether simple cross-sectional results might have been the result of selection bias—that higher productivity in the firms or plants in certain areas might arise from special characteristics of those firms (fixed effects) as opposed to the characteristics of some urban economies that give rise to agglomeration economies.

The study also links two other sources of information to the microdatabase. Census data provide information about the occupational distribution of the labour force in urban areas to test the extent to which occupational matches between firms and their urban areas are related to MFP. Input–output data are used to describe the nature of linkages that are important to different industries and then the microdata are used to ask whether suppliers in industries that the input–output tables identify as suppliers of importance are located in close proximity to each plant and whether the impact on MFP of that plant suggest that supplier links at the urban level contribute to MFP. Finally, the number of plants in the same industry located in close proximity is used to test whether there are intra-industry spillovers that arise from knowledge transfer of various sorts. This transfer could come through employees moving from one plant to another and from knowledge transfer that occurs from informal or formal contacts.

The study finds that all three sources of agglomeration economies are important. At the aggregate level, our results show that plant productivity is significantly influenced by the occupational distribution of workers, the density of the buyer-supplier network and the count of own-industry establishments within the region in which the plant is located. The labour-matching effect is empirically the largest. These results substantiate and extend earlier findings from cross-sectional investigation in the United States (Rigby and Essletzbichler 2002) and Canada (Baldwin et al. 2007).

Following Rosenthal and Strange (2001, 2003) and Henderson (2003), we explore the geographical extent of the benefits that derive from the co-location of plants. As with Rosenthal and Strange (2003), these results indicate that the benefits of own-industry co-location attenuate rapidly with distance.

The results also suggest that the impact of urban agglomeration economies with regard to labour supply, specialized suppliers and knowledge spillovers is broadly felt across industrial sectors—though the impact does differ by sector, both in terms of the size of its impact and its statistical significance.

1. Introduction

How much does geography matter to the performance of firms? Are the benefits of some locations as important to competitive advantage as the individual characteristics of business establishments themselves? What are the sources of increasing returns found in specific locations and do these location-specific economies accrue to all types of establishments or only to some? How far do the benefits of co-location extend over space?

These questions have deep roots within urban and regional economics, extending back to the work of Marshall (1920). Prompted by his evocative discussion of the local 'industrial atmosphere,' our primary concerns are the identification of the sources of agglomeration and an evaluation of their significance. In this endeavour, we follow a recent resurgence of interest in agglomeration, comprehensively reviewed by Rosenthal and Strange (2004) and Duranton (2007). This resurgence is driven, in part, by recognition of the continued importance of cities as centres of economic activity, in spite of recent innovations in transportation, information and communication technologies and the fragmentation of much production around the globe (Scott 2001). The resurgence also reflects the increasing availability of geo-referenced microdata that allow researchers to examine new arguments about urban increasing returns in new ways.

Standard theoretical claims about agglomeration hinge on geographical proximity. By locating in cities, firms have access to established pools of labour that are both specialized and deep, thus minimizing costs associated with search and training (Mincer 1984; Barron, Black and Loewenstein 1989; Becker and Murphy 1992). In addition, the creativity and diversity of talent—associated with larger cities in particular—is seen to play a critical role in the generation and incubation of new businesses (Duranton and Puga 2000, Jacobs 1969). Cities, as centres of dense economic activity, also provide individual firms with abundant opportunities for both the local sourcing of inputs and the distribution of output, thus reducing transportation costs (Chinitz 1961). This same urban concentration of firms is thought to enhance the production of knowledge and its localized spillover, either through face-to-face exchange of tacit information, or through the inter-firm mobility of human capital (Lucas 1988; Rauch 1993; Jaffe, Trajtenberg and Henderson 1993; Almeida and Kogut 1999).

Much of the early empirical analysis of agglomeration examined the general relationship between city size and productivity (Sveikauskas 1975, Moomaw 1981, Beeson and Husted 1989). Moomaw (1983) and Gerking (1993) provide reviews of this work. Subsequent efforts have become increasingly refined and have largely focused on discriminating between localization (own-industry) economies and urbanization (cross-industry) economies, or in a dynamic context between Marshall-Arrow-Romer (MAR) externalities and Jacobs' externalities (Glaeser et al. 1992). Examining industry employment growth in U.S. cities between 1956 and 1987, Glaeser et al. (1992) find that regional specialization, a proxy for MAR externalities, has little effect, while city diversity, a proxy for Jacobs' externalities, helps. They conclude that spillovers across industries are much more important at explaining employment growth than within-industry spillovers, especially in mature cities. Focusing on eight manufacturing sectors across U.S. metropolitan areas from 1970 to 1987, Henderson, Kuncoro and Turner (1995) find that MAR externalities boost growth in mature capital goods industries, but that Jacobs' externalities have no significant impact. They go on to show that in relatively immature, high-technology sectors both types of externalities are present, but that with maturity and movement of these industries out of large, diversified cities, MAR economies predominate. This is consistent with the arguments of Duranton and Puga (2000). In more sophisticated efforts to control for endogeneity, Black and Henderson (1999) and Henderson (2003) explored the nature of agglomeration economies, using panel methods on plant-level data drawn from the U.S. Census Bureau's Longitudinal Research Database (LRD). Black and Henderson (1999) reported no evidence of agglomeration economies of any kind in capital goods industries, though MAR economies operated within high-technology sectors. Henderson (2003) reported similar findings.

A series of other empirical papers has focused on identification of the sources of agglomeration economies, after Marshall (1920). Thus, Dumais, Ellison and Glaeser (1998) showed that industry employment growth in metropolitan areas is dependent upon the industry's demand for labour by occupation and on the local distribution of workers across occupations. Little evidence was found to support arguments linking employment growth with knowledge spillovers or the local buyer-supplier network. Using similar data from the LRD, a recent extension of this work across more detailed industry groupings in U.S. cities has found much stronger support for all three Marshallian agglomeration factors—the labour mix, the local density of the buyer-supplier network and spillovers (Rigby and Essletzbichler 2002). Rosenthal and Strange (2001) exploited Dunn and Bradstreet data in their exploration of the determinants of agglomeration. They regressed an index of the spatial concentration of industries at different spatial scales on proxies for knowledge spillovers, labour pooling, input sharing and other factors. They found that labour pooling was significantly linked to industry concentration across spatial scales while the influence of spillovers was limited to the zip code (local) scale. In a subsequent paper, they have shown more carefully that localization economies attentuate rapidly with distance (Rosenthal and Strange 2003).

Baldwin et al. (2007) link the arguments of a number of the papers discussed above using cross-sectional Canadian manufacturing data for 1999. Adapting measures of labour market matching and the buyer-supplier network from Rigby and Essletzbichler (2002) to the Canadian economy, they find broadly similar results. In their search for spillovers, they employ own-industry plant-counts to capture MAR economies, after Henderson (2003), and population size as a proxy for Jacobs' economies. Across the Canadian economy as a whole they find evidence of both types of externalities operating, though the localization-MAR effects are stronger. Like Rosenthal and Strange (2003), they show the geographical range of own-industry externalities is limited, extending no more than 10 kilometres.

The analysis in this paper focuses on Canadian manufacturing plants. Canadian cities are generally smaller than those in the United States and therefore the sample offers the advantage of testing for the existence of externalities in smaller markets that have not yet exhausted agglomeration economies.

Across these plants we estimate production functions that embody two sets of arguments: one captures plant characteristics, as well as those of parent firms, if applicable; the second captures place-specific characteristics originally linked to economic performance by Marshall (1920).

The production functions are estimated using panel methods to overcome omitted variable bias, which is common in cross-sectional analysis. The panel techniques require observations of individual manufacturing plants over time. We examine manufacturing establishments in 1989 and 1999. Only plants that were in business in both years comprise our sample. The longitudinal panel not only allows us to deal with fixed-effect bias but it also allows us to examine whether the relationship found in the cross-section, which is the result of developments that have taken many years, also exists in the short-run—that is, whether changes that have occurred during the 1990s in locational characteristics have affected changes in labour productivity.

At the aggregate level, our results show that plant productivity is significantly influenced by the occupational distribution of workers, the density of the buyer-supplier network—a proxy for the existence of inter-industry spillovers—and the count of own-industry establishments within the region in which the plant is located—a proxy for the existence of intra-industry spillovers. Indeed, the first two of these three agglomeration economies have elasticities that are larger than those for some plant or firm characteristics. These results are broadly consistent with earlier findings from cross-sectional investigation in the United States (Rigby and Essletzbichler 2002) and Canada (Baldwin et al. 2007).

Following Rosenthal and Strange (2001, 2003) and Henderson (2003), we explore the geographical extent of the benefits that derive from the co-location of plants. For each establishment we count the number of own-industry plants located within concentric circles of different radii. Across our full sample of Canadian manufacturing plants there is a positive relationship between own-industry plant counts, our measure of localization economies and productivity. The relationship is statistically significant for short distances (0 to 5 kilometres) but not for longer distances. This result is consistent with those of Rosenthal and Strange (2003) who find that the benefits of own-industry co-location attenuate rapidly with distance. We found no significant productivity benefit of locating in regions with a larger population, our proxy for urbanization economies (Jacobs 1969). In fact, the independent effect of city size tends to be negative, suggesting congestion effects.

Extending the analysis to five separate industry groups—characterized by output type—reveals that the different sources of agglomeration economies do not operate uniformly across the economy. The local density of buyer-supplier networks and the local labour mix exerted a significant influence on productivity in three of five industry clusters. The significance of the count of own-industry plants varied considerably across different distances depending on the industry in question.

2. *Data and model*

The variables used in our econometric models are readily separated into two groups: characteristics of individual business units or plants; and, characteristics of particular locations. Table 1 lists the variables in our models and provides brief descriptions. The plant-level information is developed from the Canadian Annual Survey of Manufactures (ASM) for 1989 and 1999. The panel techniques we employ require observations on individual establishments for at least two years.

Table 1
Variable description

Variables		Description
Abbreviation	Label	
Plant characteristics		
LP	Labour productivity	Value added divided by the number of production workers in the plant
PROFIT/VA	Profit-to-value added ratio	Value added minus wages divided by value added
PRDWRK	Production workers	Number of production workers in the plant
NPWPW	Non-production-to-production worker ratio	Number of non-production workers divided by number of production workers
Place characteristics		
LABMIX	Labour mix	Defined in Sub-section 2.2
USXLQ	Local density of upstream suppliers	Defined in Sub-section 2.2
PLANTS0-5	Plants within 5 km	Number of plants within 5 km in the same 2-digit industry
PLANTS0-10	Plants within 10 km	Number of plants within 10 km in the same 2-digit industry
PLANTS0-50	Plants within 50 km	Number of plants within 50 km in the same 2-digit industry
PLANTS0-200	Plants within 200 km	Number of plants within 200 km in the same 2-digit industry
POPLN	Population	Population of the census metropolitan area or census agglomeration where the plant is located

Sources: Statistics Canada, Annual Survey of Manufactures, Census of Population, Input–Output Accounts.

Our place-specific data are derived from the ASM, from the Census of Population in 1991 and 2001 and from the Canadian Input–output Accounts. All data were geocoded to a constant 2001 Census geography for census metropolitan areas (CMAs) and census agglomerations (CAs). In 2001, there were 141 CMAs and CAs in Canada, ranging in size from Kitimat, B.C., with a population of about 10,000, to the Toronto CMA with a population of about 4.6 million. The 141 regions contained approximately 80% of the Canadian population in 2001 and the same percentage of Canadian manufacturing establishments in 1999.

2.1 Plant- and firm-specific characteristics

The dependent variable in our analysis is labour productivity, measured as value added divided by the number of production workers. For each plant, we measure value added and production workers at their mean across three years. For 1989 these are the two adjacent years. Owing to the fact that 1999 is the last year on the longitudinal file, we take the mean level of value added and production workers for 1999 and the two previous years. Value added is measured in constant dollar terms using an industry-level deflator.

We utilize three-year means for value added and the number of production workers per plant, as well as all other plant-level characteristics, in order to reduce the year-over-year variability inherent to microdata of this kind. Plants often encounter shocks that may obscure the relationship between plant-level inputs and output—for example, because of labour hoarding. Using three-year means helps to reduce the effect of this variability on our estimates.

Labour productivity is expected to depend on several plant-level characteristics. These include plant size, capital intensity and the ratio of non-production workers to production workers. It is expected that labour productivity will be higher in plants that are larger in size because they are able to take advantage of various forms of scale economies—for example, those that result from longer production runs. Plant size is measured by the number of production workers.

The productivity of production workers is also expected to rise as the amount of machinery and equipment with which they work increases. We would like to capture the effect of mechanization with a variable measuring the capital-to-labour ratio. Unfortunately, capital stock data are unavailable at the plant level and so we use a proxy variable, the ratio of profits to value added, to represent the capital-to-labour ratio. In a competitive economy, profits just cover capital costs and therefore offer a good proxy for the relative capital embedded in different plants—especially when averaged across several years. Other studies that have used this proxy in a production function framework using the Canadian manufacturing survey microdata have generated coefficients on capital that are very close to those derived from industry-level data on capital stock (see Baldwin and Gu 2003). Profits are measured as value added minus wages. Like value added, wages are deflated to produce constant dollar estimates.

Production workers tend to generate higher levels of output if more non-production workers are contributing to the production process. For instance, more input from management and engineering functions can help to improve the organization of the production process. Hence, we expect labour productivity to be positively associated with the ratio of non-production workers to production workers.

2.2 Model

The relationships between value added, plant size and capital intensity noted above can be formally derived from a production function using Cobb-Douglas technology, where value added (VA) is expressed as:

$$VA = AK^{\alpha} L_{pw}^{\beta} L_{npw}^{\sigma}, \tag{1}$$

where K is a measure of capital input, L_{pw} is the number of production workers employed by the plant and L_{npw} is the number of non-production workers. With a little algebraic manipulation, Equation (1) may be re-written such that labour productivity (LP) is a function of capital and labour inputs:

$$LP = \frac{VA}{L_{pw}} = A\left(\frac{K}{L_{pw}}\right)^{\alpha} \left(\frac{L_{npw}}{L_{pw}}\right)^{\sigma} L_{pw}^{\beta+\alpha+\sigma-1}. \tag{2}$$

Hence, labour productivity is a positive function of the amount of capital employed per production worker, the number of non-production workers for each production worker and the size of plant, as measured by the number of production workers. The ASM provides measures of value added and the number of production and non-production workers. However, as we have already noted, it does not provide estimates of capital and therefore we need to develop a proxy $(\hat{K})$. We can estimate $\hat{K}$ from the following expression for profit (π):

$$\pi = VA - wages = r\hat{K} \tag{3}$$

where r is the rate of return on capital. The profit-to-labour ratio $r\hat{K}/L_{pw}$ can be substituted into (2), and if we assume the rate of return is equalized across plants, then

$$LP = Ar^{\alpha}\left(\frac{\hat{K}}{L_{pw}}\right)^{\alpha}\left(\frac{L_{npw}}{L_{pw}}\right)^{\sigma} L_{pw}^{\beta+\alpha+\sigma-1}. \tag{4}$$

Given this formulation, variation in profits across industries and provinces can be accounted for by industry and province fixed effects.

One of the practical issues with Equation (4) is that our proxy of the capital-to-labour ratio and our measure of productivity are very highly correlated because both contain value added in their numerator and labour in their denominator. To address this problem, we estimate a slightly different model. Multiplying (1) by VA^{α}/VA^{α} we obtain

$$VA = Ar^{\alpha}\left(\frac{\hat{K}}{VA}\right)^{\alpha} VA^{\alpha} L_{pw}^{\beta} L_{npw}^{\sigma}, \tag{5}$$

which implies

$$VA = A^{\frac{1}{1-\alpha}} r^{\frac{\alpha}{1-\alpha}} \left(\frac{\hat{K}}{VA}\right)^{\frac{\alpha}{1-\alpha}} L_{pw}^{\frac{\beta}{1-\alpha}} L_{npw}^{\frac{\sigma}{1-\alpha}}. \tag{6}$$

Labour productivity can then be defined as

$$LP = \tilde{A}\tilde{r}\left(\frac{\hat{K}}{VA}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{L_{npw}}{L_{pw}}\right)^{\frac{\sigma}{1-\alpha}} L_{pw}^{\frac{\beta+\alpha+\sigma-1}{1-\alpha}}, \tag{7}$$

where $\tilde{A} = A^{1/(1-\alpha)}$ and $\tilde{r} = r^{\alpha/(1-\alpha)}$. Equation (7) can be used to solve for the values of α, β and σ. Hence, despite the fact that we do not estimate the effect of the capital-to-labour ratio on productivity directly, we are able recover an estimate.

In order to estimate (7) we include a multiplicative error term ε and use its logarithmic transformation:

$$\ln LP_{ijk} = \ln \tilde{A} + \ln \tilde{r} + \delta_1 \ln \frac{\hat{K}_i}{VA_i} + \delta_2 \ln \frac{L_{npw,i}}{L_{pw,i}} + \delta_3 \ln L_{pw,i} + \delta_4 \ln r_i + \ln \varepsilon_i , \qquad (8)$$

where $\delta_1 = \frac{\alpha}{1-\alpha}$, $\delta_2 = \frac{\sigma}{1-\alpha}$, $\delta_3 = \frac{\beta+\alpha+\sigma-1}{1-\alpha}$ and $\delta_4 = \frac{1}{1-\alpha}$. Note also that *i* indexes plants, *j* indexes firms and *k* indexes geographic locations.

Throughout the analysis we assume that other characteristics of the firm and the characteristics of the location of the firm are transmitted through the multifactor productivity term *A*. Hence,

$$\ln \tilde{A} = \ln a + \varphi' \mathbf{lnX}_j + \theta' \mathbf{lnG}_k + \gamma_i + \eta_j + \lambda_k \qquad (9)$$

where **X** is a vector of characteristics related to the firm that controls plant *i* and **G** is a vector of characteristics that are associated with location *k*. These locational characteristics are related either to the metropolitan area associated with *k* or are calculated based on set distance from *k*, where *k* can be thought of as a point in space.

We measure two types of firm characteristics in the model. First, we identify whether the plant is part of a multi-establishment firm. This is a binary dummy variable where the reference group is single-plant firms. Our expectation is that multi-plant firms will be larger than single-plant firms. Firm size brings the benefit of firm-wide economies to the plant. For instance, larger firms may be better able to collect and analyse information that can improve management practices and thus raise productivity.

Second, we identify whether plants are foreign controlled. Foreign-controlled plants are expected to have a higher level of productivity because they have access to a broader range of experiences and technologies (Baldwin and Gu 2005). The foreign control is also a binary categorical variable where the reference group is domestically controlled plants.

2.3 Place-specific characteristics

The agglomeration variables that we develop in our productivity model, the local density of buyer-supplier networks, labour pooling and knowledge spillovers can all be traced back to Marshall (1920). We outline below the variables employed to measure these Marshallian economies, along with indicators used to capture other types of agglomeration economies.

An area's labour pool supports the needs of a particular industry if the occupational distribution of an area corresponds to that of the distribution required by that industry. The labour mix for an industry within a metropolitan area is defined after Dumais, Ellison and Glaeser (1998) as

$$LABMIX_i^m = \sum_o \left(L_{io} - \sum_{j \neq i} \frac{E_j^m}{E^m - E_i^m} L_{jo} \right)^2 ,$$

where *o* represents an occupation, *i* and *j* index industries and *m* refers to the metropolitan area. *L* measures the proportion of workers in a particular industry and occupation, while *E* measures the number of workers in a single industry or in all industries within a metropolitan area. This index is a sum of squared deviations that measures the degree to which the occupational distribution of employment of an industry is matched by the occupational distribution of the workforce in the metropolitan area as a whole, excluding the specified industry. The occupational distribution of industry workers is calculated at the national level and covers some 47 occupations at the 2-digit level using the 1991 Standard Occupational Classification, which is used for the 1991 and 2001 Censuses. We anticipate that a better match between the occupational distribution (demand) in an industry and the occupational distribution of the entire workforce of a metropolitan area (supply) will boost productivity. Improved matches reduce the value of the squared term. Thus, we expect a negative coefficient on this variable in the following regressions.

We calculate inter-industry spillovers by focusing on those industries that are closely linked via either suppliers or buyer networks. The local density of buyer-supplier networks is estimated from national input–output data and indicators of the local concentration of production within specific sectors. High correlation between estimates of the geographic concentration of upstream producers and downstream customers led us to focus on upstream activity only. To measure local variation in the density of upstream connections for each 4-digit industry and for each CMA in Canada, we identify an upstream supplier-weighted location quotient:

$$USXLQ_j^m = \sum_{i,i \neq j} w_{ij}^n \left(\frac{TVS_i^m / \sum_i TVS_i^m}{TVS_i^n / \sum_i TVS_i^n} \right).$$

The term in the parentheses is a location quotient for each industry *i* in metropolitan area *m*. The location quotients are calculated using the total value of shipments (*TVS*) of each industry and they measure the degree to which a particular city is specialized in an industry. A value of less than one would indicate an industry is under-represented, while a value greater than one would indicate the industry was over-represented. The term w_{ij} represents the weight of industry *i* as a supplier of industry *j*—that is, the proportion of all manufactured input purchases by industry *j* supplied by industry *i*. Supplier weights are estimated from inter-industry transactions and are derived from the Canadian national input–output tables. The subscripts *i* and *j* refer to each of the 236 4-digit Standard Industrial Classification (SIC) manufacturing industries, *m* refers to one of 140 or so metropolitan areas in Canada and *n* refers to the nation. Note that we also removed the influence of the own-industry in these measures, by dropping the principal diagonal from the input–output direct coefficients matrix. Metropolitan areas whose economies are specialized in industries that are significant suppliers to industry *j* will have a relatively high *USXLQ* and this is expected to have a positive effect on labour productivity in plants within industry *j*.

Note that because the labour mix and buyer-supplier network measures are defined at the metropolitan level, the values for these variables for a given industry are constant for all plants in that industry and metropolitan area. As we have noted above, this necessitates adjustment of the standard errors in our model for, as Moulton (1990) demonstrates, they can be biased when merging aggregate variables across micro units of observation.

The third agglomeration effect arises from knowledge spillovers that are generated by the close proximity of producers in the same industry in the same urban area—intra-industry spillovers. Measuring knowledge spillovers is notoriously difficult, even impossible, as Krugman (1991) claims, for they do not leave a paper trail. Jaffe, Trajtenberg and Henderson (1993) disagree. They argue patent citations can be used to track the spatial limits of knowledge flows. Nevertheless, the linking of patent information to the plant-level data that are increasingly used to study agglomeration is surprisingly underdeveloped. Rigby and Essletzbichler (2002) show that flows of knowledge embodied in intermediate goods enhance the productivity of agglomerated plants, but that sheds little light on the role of disembodied information flows. We spent some time examining the influence of local own- and cross-industry patents, in industries of use and make, on plant labour productivity, but were discouraged by the results that were broadly insignificant. Our measures all used simple counts of patents within metropolitan areas and industries linked to the patent classification rather than citations. Raw patent counts for 1999, earlier years or groups of years were not significantly related to productivity.

As a result, we follow Henderson (2003) and Rosenthal and Strange (2003) and use counts/densities of plants in specific geographical areas as a proxy for intra-industry knowledge spillovers. To define 'geographical areas' we exploited data on the latitude and longitude of individual plants to define concentric circles of varying distances around each. The concentric circles employed had radii of 0 to 5 kilometres, 0 to 10 kilometres, 0 to 50 kilometres and 0 to 200 kilometres. We admit that these distances were chosen in an ad hoc fashion, though we do not have much theory to suggest over precisely what distances particular kinds of information actually flow. For each plant, we counted establishments within the same 2-digit SIC industry. In our previous work (see Baldwin et al. 2007), counts of own-industry plant numbers within the same metropolitan area consistently generated insignificant parameters in the regression models. We anticipate that as the number of plants increases, so too does the potential flow of knowledge that is expected to boost plant productivity.

We add metropolitan population size to our model as a proxy for urbanization type economies that are not captured elsewhere in our model. The benefits of urban size are many. Large urban economies bring with them greater industrial and occupational diversity that facilitate the transfer of new innovations across industries (Jacobs 1969). Large population centres also create the demand for infrastructure that can enhance the productivity of all industries (e.g., highways, airports, ports and communications networks).

2.4 Sample characteristics

Descriptive statistics for all place-specific variables and for plant variables that are continuous are reported in Tables 2 and 3. The values in Table 2 are shown for the two years over which we have drawn our observations: 1989 and 1999. These values are not logged. Table 3 reports descriptive statistics for the differences for these variables. Along with the mean, median and standard deviation for all variables, both tables report the number of observations across which the descriptive statistics were calculated.

Table 2
Descriptive statistics for plants present in 1989 and 1999

	1989				1999			
	Mean	Median	Standard deviation	Observations	Mean	Median	Standard deviation	Observations
Plant characteristics								
Labour productivity	82,775	57,910	113,862	11,323	87,298	55,644	112,083	11,323
Profit-to-value added ratio	0.58	0.58	0.16	11,323	0.58	0.58	0.18	11,323
Production workers	53	15	230	11,323	59	19	198	11,323
Non-production-to-production worker ratio	0.46	0.37	0.52	11,323	0.42	0.33	0.53	11,323
Place characteristics								
Labour mix	5.1	4.3	2.4	3,204	5.5	4.8	2.5	3,204
Local density of upstream suppliers	6.0	1.2	24.5	3,204	6.9	1.2	29.0	3,204
Plants within 5 km	41	17	74	11,323	31	13	54	11,323
Plants within 10 km	50	20	73	11,323	41	17	57	11,323
Plants within 50 km	279	134	360	11,323	203	98	264	11,323
Plants within 200 km	359	216	446	11,323	270	163	345	11,323
Population	159,220	37,932	463,249	138	178,011	39,992	535,224	138

Source: Statistics Canada, Annual Survey of Manufactures.

Table 3
Descriptive statistics for change in variables from 1989 to 1999

	Change in variables			
	Mean	Median	Standard deviation	Number of observations
Plant Characteristics				
Labour productivity	4,523	-695	118,244	11,323
Profit-to-value added ratio	0.002	0.004	0.17	11,323
Production workers	6	1	116	11,323
Non-production-to-production worker ratio	-0.04	-0.04	0.59	11,323
Place Characteristics				
Labour mix	0.09	0.35	1.45	11,323
Local density of upstream suppliers	0.34	-0.01	9.8	11,323
Plants within 5 km	-11	-2	34	11,323
Plants within 10 km	-20	-5	52	11,323
Plants within 50 km	-96	-26	155	11,323
Plants within 200 km	-186	-98	212	11,323
Population	305,942	208,638	315,286	11,323

Note: Figures shown result from a special tabulation performed by the authors.
Source: Statistics Canada, Annual Survey of Manufactures (1989 and 1999).

Plant characteristics are measured across individual manufacturing establishments (plants). We limited our sample in several ways. By construction, plants in rural areas are excluded from the study. Furthermore, only plants with a three-year average level of employment above zero are included as labour productivity with zero employment is undefined.

The sample is also restricted to plants with positive value added and positive returns to capital. For the latter, this implies value added minus wages is greater than zero. As a practical matter these restrictions are imposed because logarithmically transformed variables with a value of zero or less are mathematically undefined. They are also imposed because plants with negative value added or negative returns to capital are likely undergoing significant economic shocks. Again, this may blur the relationship between inputs and output.

Also excluded are plants that change location and industry. This can result in very large changes in our place-specific variables that may have significant influence on our estimated coefficients. It is very difficult to disentangle the influence of the change in the place characteristics from that of other factors. For instance, a plant move will change the density of local buyer-supplier links, but a move likely also coincides with the building of a new plant that may also influence productivity.

Due to the longitudinal nature of the analysis, the most significant restriction to our set of plants is that they must have lasted at least 10 years, from 1989 to 1999. In 1999, this restriction, plus all of the others noted above, reduced the number of plants in the sample from about 29,000 to about 11,300. Omitting plants that do not remain in business for at least 10 years significantly reduces the number of observations in our sample and raises questions about sampling bias. However, the results reported below are very similar to those published earlier on a much larger cross-section of plants from 1999 and we have found that they are robust to broad changes in sample characteristics.

Turning now to the descriptive statistics, there 11,323 plants present in 1989 that were in business in 1999. The descriptive statistics for these plants, at their beginning and end points, are presented in Table 2, and the changes in the same variables are presented in Table 3

The average productivity of the plants on average increased over the period. The mean labour productivity of plants present in 1989 and 1999 increased from $82,775 to $87,298. Other plant level characteristics remained relatively stable over the period. The profit to value added ratio remained essentially constant. Average and median plant sizes increased marginally, while the ratios of non-production workers to production workers fell modestly.

Shifting to our geographical or place-specific variables for each establishment, counts of plants in the same 2-digit industry within various distances were generated. All establishments, not just those that form part of our sample, are included in these counts. Population values are reported for approximately 140 CMAs or CAs that comprise the geographical units of analysis. The labour mix and upstream location quotient are calculated at the 3- and 4-digit levels of the Canadian SIC for each CMA and CA, yielding 3,204 annual observations.

Table 4 reports correlation coefficients and associated p-values for all continuous variables in the regression models. These correlations are reported after logging all variables and then differencing them. Because of the large number of observations, the p-values are frequently significant, even though the correlation coefficients themselves are relatively low. Of course, collinearity between variables does not bias our estimators, merely rendering them inefficient.

Table 4
Correlation of variables in regression models

Change in	Labour productivity	Profit-to-value added ratio	Production workers	Non-production-to-production workers ratio	Multi-plant status	Foreign-control status	Labour mix	Local density of upstream suppliers	Plants within 5 km	Plants within 10 km	Plants within 50 km	Plants within 200 km	Population
Labour productivity	1	…	…	…	…	…	…	…	…	…	…	…	…
Profit-to-value added ratio	0.61	1	…	…	…	…	…	…	…	…	…	…	…
p-value	(0.00)	…	…	…	…	…	…	…	…	…	…	…	…
Production workers	-0.26	-0.14	1	…	…	…	…	…	…	…	…	…	…
p-value	(0.00)	(0.00)	…	…	…	…	…	…	…	…	…	…	…
Non-production-to-production workers ratio	0.35	0.27	-0.39	1	…	…	…	…	…	…	…	…	…
p-value	(0.00)	(0.00)	(0.00)	…	…	…	…	…	…	…	…	…	…
Multi-plant status	0.05	0.01	0.03	0.00	1	…	…	…	…	…	…	…	…
p-value	(0.00)	(0.25)	(0.00)	(0.97)	…	…	…	…	…	…	…	…	…
Foreign-control status	0.05	0.00	0.00	0.02	0.13	1	…	…	…	…	…	…	…
p-value	(0.00)	(0.64)	(0.66)	(0.08)	(0.00)	…	…	…	…	…	…	…	…
Labour mix	-0.21	-0.02	0.00	0.02	-0.03	-0.01	1	…	…	…	…	…	…
p-value	(0.00)	(0.09)	(0.60)	(0.04)	(0.00)	(0.55)	…	…	…	…	…	…	…
Local density of upstream suppliers	0.06	0.03	0.11	0.00	0.01	0.03	-0.01	1	…	…	…	…	…
p-value	(0.00)	(0.00)	(0.00)	(0.75)	(0.35)	(0.00)	(0.25)	…	…	…	…	…	…
Plants within 5 km	0.06	0.01	0.05	0.00	0.02	0.02	-0.11	0.00	1	…	…	…	…
p-value	(0.00)	(0.20)	(0.00)	(0.62)	(0.02)	(0.06)	(0.00)	(0.99)	…	…	…	…	…
Plants within 10 km	0.03	0.00	0.02	0.01	0.02	0.00	-0.06	0.04	0.35	1	…	…	…
p-value	(0.00)	(0.65)	(0.07)	(0.26)	(0.06)	(0.90)	(0.00)	(0.00)	(0.00)	…	…	…	…
Plants within 50 km	0.05	0.01	0.04	-0.01	0.04	0.00	-0.14	0.05	0.31	0.45	1	…	…
p-value	(0.00)	(0.39)	(0.00)	(0.41)	(0.00)	(0.72)	(0.00)	(0.00)	(0.00)	(0.00)	…	…	…
Plants within 200 km	0.05	0.01	0.04	-0.02	0.03	0.00	-0.16	0.02	0.31	0.32	0.73	1	…
p-value	(0.00)	(0.25)	(0.00)	(0.01)	(0.00)	(0.74)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	…	…
Population	-0.04	-0.01	0.06	0.02	-0.01	-0.01	-0.02	-0.01	-0.09	-0.15	-0.12	0.09	1
p-value	(0.00)	(0.28)	(0.00)	(0.03)	(0.24)	(0.34)	(0.00)	(0.40)	(0.00)	(0.08)	(0.00)	(0.00)	…

… not applicable
Notes: All variables were logged for 1989 and 1999 and then differenced.
Source: Statistics Canada, Annual Survey of Manufactures (1989 and 1999).

3. *Econometric strategy*

One of the problems associated with estimating (8) is that there are unobserved fixed effects associated with plant *i*, its related firm *j*, and location *k* that may be correlated with our vector of geographical characteristics **G**. We represent these unobserved fixed effects in (9) by γ_i, η_j, and λ_k that are associated with the plant, firm, and location, respectively.

To illustrate this problem, consider the case of a leading firm that was established by chance in a location several decades in the past (e.g., 3M). Over the ensuing decades, the firm grew because of its superior production processes and product development (an unobserved characteristic). Often successful firms generate spin-offs as employees who developed technical and management expertise started their own businesses producing related products. Hewlet Packard in Silicon Valley is an example. With the development of the firm, and its spin-offs, local input suppliers emerge and the workforce of the local community is transformed, increasingly matching that of this geographic cluster of firms. In this case, we would observe a positive association between the level of labour productivity of the firms found in this cluster—the original firm and its spin-offs—and the mix of labour, the presence of upstream suppliers and the number of firms in the same industry. This result is traceable not to localization economies—labour matching, buyer-supplier links, and knowledge spillovers—but to the special nature of the progenitor firm.

The same logic applies to geographic locations. The concentration of firms may be related to natural features—for example, access to a resource stock—rather than any form of localization economy à la Marshall.

To address these issues, we substitute (9) into (8) and take the first difference across periods:

$$\ln \Delta LP_{ijk} = \Delta \ln a + \delta_1 \Delta \ln \frac{\hat{K}_i}{VA_i} + \delta_2 \Delta \ln \frac{L_{npw,i}}{L_{pw,i}} + \delta_3 \Delta \ln L_{pw,i} + \delta_4 \Delta \ln r_i + \varphi' \mathbf{\Delta lnX}_j + \theta' \mathbf{\Delta lnG}_k + \Delta \ln \varepsilon_i, \quad (10)$$

In so doing, we eliminate the firm- and location-level fixed effects that might be correlated with our Marshallian localization economies. But, of course, in doing so we are giving ourselves a harder task in isolating these effects because there may be little change in the variables of interest and this will increase the standard errors of the estimates.

4. Panel model estimates

We estimate different forms of Equation (10). Our key results are reported in Tables 5 and 6. Table 5 presents output from two models that were estimated using ordinary least squares after differencing between years. All standard errors are robust and corrections have been made for potential correlation of errors between manufacturing establishments found in the same region (Moulton 1990). Model 1 shows the relationship between labour productivity and plant characteristics alone. As expected, labour productivity tends to be significantly higher in plants where the profit-to-value added ratio, our proxy for the capital-to-labour ratio, is high. This variable consistently displays the largest elasticity of all independent variables, typically raising productivity over 8% for every 10% increase in the profit-to-value added ratio. As the ratio of non-production workers to production workers rises across plants, so productivity also tends to increase. The elasticity of this variable is typically less than half that of the profit-to-value added ratio. The coefficients on these plant characteristic variables yield sensible estimates of the coefficients of the production function—the implied labour share is 0.51 (β = 0.29 and σ = 0.22), capital share (α = 0.43) and there are near constant returns to scale ($\alpha + \beta + \sigma = 0.94$).

Plants that belong to multi-establishment firms also display higher productivity values than single-establishment firms and foreign-owned plants tend to be more productive than domestic plants. In both cases these effects are significant and the elasticity on foreign-owned plants is about the same as that for multi-plant firms.

Model 2 adds our agglomeration measures. We proxy knowledge spillovers using own-industry plant counts and metropolitan area population size. Own-industry plant counts are commonly used to capture localization economies (see Henderson 2003 and Rosenthal and Strange 2003), while population is employed to capture urbanization economies. We attempt to estimate the distance across which localization economies operate by plant counts within circles of progressively greater radii. In an earlier paper (Baldwin et al. 2007) employing cross-section data for a single year, and therefore unable to control for omitted variable bias, we found that both localization economies and urbanization economies exerted a positive and significant impact on the labour productivity of individual plants. Furthermore, we found that the benefits of sharing a location with other plants in the same industry extended no more than 10 kilometres. We clarify these results next.

Model 2 in Table 5 shows that all three of our agglomeration measures have a significant influence on plant productivity. Of these, the impact of the local labour mix is most significant, exerting an impact on productivity that is approximately five times greater than the effects of multi-plant or foreign-ownership status. The partial regression coefficient on the labour mix variable shows that plants located in metropolitan areas where the occupational distribution of workers is closely related to the occupational distribution of their own workforce have significantly higher labour productivity. A 10% improvement in this occupational match raises plant productivity approximately 5%.

Table 5
General model results

	Model 1		Model 2	
	Coefficient	p-value	Coefficient	p-value
Change in plant/firm characteristics				
Profit-to-value added ratio	0.76	<0.001	0.75	<0.001
Production workers	-0.10	<0.001	-0.11	<0.001
Non-production-to-production worker ratio	0.38	<0.001	0.38	<0.001
Multi-plant status (reference = single plant)	0. 10	0.079	0.09	0.002
Foreign control status (reference = domestic control)	0.10	<0.001	0.09	<0.001
Change in place characteristics				
Labour mix	…	…	-0.51	<0.001
Local density of upstream suppliers	…	…	0.10	<0.001
Plants within 5 km	…	…	0.02	0.021
Plants within 200 km	…	…	0.02	0.403
Population	…	…	-0.15	0.045
Constant	0.04	<0.001	0.05	<0.001
Number of observations	11,323		11,323	
F	829		637	
Probability > F	<0.001		<0.001	
R-squared	0.42		0.47	
Root mean square error	0.45		0.43	

… not applicable
Notes: All variables are log transformed, with the exception of the binary variables, and differenced between the years 1989 and 1999. In all regressions, standard errors are corrected for heteroskedasticity and potential correlation of errors within census metropolitan areas and census agglomerations.
Source: Statistics Canada, Annual Survey of Manufactures (1989 and 1999).

The local density of the buyer-supplier network (a measure of inter-industry spillovers) also exerts a positive and significant impact on the labour productivity of individual establishments, with an elasticity typically about a fifth of that of the labour-mix effect. Finally, productivity increases with the number of plants in nearby proximity in the same industry—a measure of intra-industry spillovers. The value of the coefficient is positive and significant for the number of plants within 5 kilometres but is insignificant for plants within 200 kilometres. Other distances were substituted for the 200-kilometre radius (10 kilometres and 50 kilometres) with the same result.[1] Hence, like others, we find a strong distance gradient with respect to intra-industry spillovers.

Our measure of urbanization economies (population size) exerts a negative influence on plant productivity in our multivariate regression. This suggests there may be a congestion effect associated with increasing city size.

1. We also estimated the model with concentric circles at 5 to 10 kilometres, 10 to 50 kilometres and 50 to 200 kilometres and obtained similar results.

The results from our general model show that, on average, agglomeration economies raise the productivity of individual producers. However, there is no guarantee that the benefits of co-location are equally important for all businesses. One simple way of exploring this question is to examine how the different sources of agglomeration economies operate across manufacturing industries. This approach is problematic for two reasons. First, the number of plants within most 3- or 4-digit Standard Industrial Classification (SIC) industries in Canada is quite small and so it is difficult to obtain statistically significant results. Second, it would also be difficult to make sense of results that stretch across hundreds of sectors. To overcome this problem, we follow a different course, aggregating individual manufacturing industries together into five broad sectors and then estimating Model 2 across each of those sectors.

The five sectors are taken from the Organisation for Economic Co-operation and Development (OECD) (1987). They are defined as natural resource-based, labour-intensive, scale-based, product-differentiated and science-based. The original OECD classification was tailored for use with the Canadian manufacturing data. Baldwin and Rafiquzzaman (1994) list the 4-digit (SIC) industries assigned to each of the OECD sectors. Each sector is defined primarily on the basis of the factors that influence the process of competition. For resource-based industries, the primary determinant of competitive success is access to abundant natural resources. For the labour-intensive sector, it is labour costs. For scale-based industries, competition hinges on the length of production runs. In the product-differentiated group, competition depends on an ability to target production to the demands of various markets. Finally, competition in science-based sectors depends on the application of scientific knowledge.

Table 6 shows the results from estimating our panel model for each of the five OECD sectors. Overall, plant characteristics affect labour productivity in a consistent way across these five industry groupings, though the sizes of the partial regression coefficients are variable. Our firm measures, multi-plant status and domestic/foreign ownership status have the same positive sign across all OECD sectors, though the coefficients vary markedly in size and are not uniformly significant. Plants that are part of multi-establishment firms have higher productivity than single-establishment firms, though the productivity differential is statistically significant only in labour intensive and science-based OECD sectors. Similarly, while foreign-owned plants have higher productivity than domestically owned plants, the difference is significant in scale-based and science-based sectors.

The different sources of agglomeration economies generally have the same signs across OECD sectors as in Table 5. Our labour mix, or labour matching variable (LABMIX), has the anticipated sign in all sectors, and an improved match between the supply and demand of workers by occupation significantly improves productivity in scale-based, productivity-differentiated and science-based sectors. In the science-based industries, the labour mix variable has a very high elasticity: a 10% improvement in the occupation match raises plant productivity by almost 7%. This elasticity is larger than that for all other variables in the science-based model, save for our proxy of the capital-to-labour ratio. The influence of the labour mix is also relatively strong within scale-based and product-differentiated sectors of Canadian manufacturing. Our cross-sectional analysis produced broadly similar results, though the patterns of significance of the labour mix variable differed across OECD sectors. We have greater faith in the results reported here because omitted variables in the cross-sectional analysis lead to biased and inconsistent estimators.

Table 6
Sectoral model results

	Natural resource-based		Labour intensive		Scale-based		Product-differentiated		Science-based	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Change in plant/ firm characteristics										
Profit-to-value added ratio	0.85	<0.001	0.69	<0.001	0.65	<0.001	0.82	<0.001	0.91	<0.001
Production workers	-0.11	<0.001	-0.12	<0.001	-0.13	<0.001	-0.09	<0.001	-0.11	<0.001
Non-production-to-production worker ratio	0.31	<0.001	0.32	<0.001	0.46	<0.001	0.41	<0.001	0.49	<0.001
Multi-plant status (reference = single plant)	0.12	<0.001	0.10	0.118	0.02	0.331	0.03	0.558	0.13	0.005
Foreign control status (reference = domestic control)	0.04	0.158	0.06	0.131	0.09	0.019	0.16	0.106	0.19	0.017
Change in place characteristics										
Labour mix	-0.09	0.096	-0.17	0.112	-0.53	<0.001	-0.58	<0.001	-0.66	<0.001
Density of upstream suppliers	0.10	<0.001	0.06	0.040	0.19	<0.001	0.06	0.309	0.07	0.219
Plants within 5 km	0.01	0.455	0.001	0.927	0.03	0.014	0.01	0.393	0.03	0.040
Plants within 200 km	-0.20	0.022	0.15	0.017	0.15	0.015	-0.18	0.069	0.22	0.045
Population	-0.10	0.328	-0.11	0.171	-0.31	0.101	-0.16	0.269	0.30	0.175
Constant	0.08	0.002	0.08	0.001	0.10	0.001	0.02	0.310	0.03	0.456
Number of observations	3,028		2,933		2,545		2,012		805	
F	202		415		198		141		97	
Probability> F	<0.001		<0.001		<0.001		<0.001		<0.001	
R-squared	0.47		0.50		0.49		0.41		0.52	
Root mean square error	0.42		0.37		0.42		0.47		0.47	

Notes: All variables are log transformed, with the exception of the binary variables, and differenced between the years 1989 and 1999. In all regressions, standard errors are corrected for heteroskedasticity and potential correlation of errors within census metropolitan areas and census agglomerations.
Source: Statistics Canada, Annual Survey of Manufactures (1989 and 1999).

Across all OECD sectors, the density of the regional buyer-supplier network (USXLQ) exerts a positive influence on plant productivity. It is significant for all but the product-differentiated and science-based sectors. The elasticity on the buyer-supplier network variable obtains the highest level of statistical significance for the natural resources and scale-based sectors.

Own-industry plant counts are generally significant across sectors, but the apparent spatial extent of spillovers varied considerably. For scale-based and science-based industries, the number of plants within 5 kilometres **and** within 200 kilometres had a positive and significant effect on productivity. In fact, it was plants within 200 kilometres that had the strongest effect—the elasticities implied a 10% increase in the number of plants would increase productivity by about 2%. For the other sectors, the number of plants within 5 kilometres had no significant effect. However, the number of plants within 200 kilometres did significantly influence productivity. For labour intensive industries, the effect was positive, but curiously for national resource-based and product-differentiated industries the effect was negative and significant. For natural resource industries this may reflect diminishing returns to once highly productive resource endowments, but the cause of the negative elasticity for product-differentiated industries is less obvious.

5. Conclusion

The growth accounting framework that is used to parse out the various determinants of economic growth focuses on changes in labour, capital, intermediate materials and an unexplained residual that is usually referred to as multifactor productivity (MFP). MFP is often attributed either to unmeasured inputs, such as innovative capabilities, or to externalities arising from the environment. Externalities include freely available knowledge, the social infrastructure of the economy and supporting economic structures.

Marshall's agglomeration (localization) economies fall within the latter category. Agglomeration economies include the advantages of having the correct labour mix readily available for a firm, of having other firms that can readily supply specialized inputs and having an information flow from other firms in the same industry that reduces costs or improves the quality of the product. In all cases, distance is seen to provide cost advantages in having labour, suppliers or information available close at hand.

Measurement of the impact of these externalities is difficult. To overcome some of the problems that have prevented thorough studies of the agglomeration phenomenon, this paper makes use of detailed microdatabases on Canadian manufacturing plants and firms that permit both productivity and associated characteristics of the production entities to be measured. The database essentially covers the entire population, thereby reducing the selection bias associated with less comprehensive databases and it is tracked over a 10-year period. Plants can be located precisely using constant geographic codes over the 10-year period. This allows us not only to examine differences across plants at a particular point in time but, more importantly, to study how changes over time in urban characteristics have influenced productivity. Examining these changes allows us to ask whether recent growth in urban economies and changes in industrial structure have in turn fed back into changes in productivity. It also allows us to ask whether simple cross-sectional results might have been the result of selection bias—that higher productivity in the firms or plants in certain areas might arise from special characteristics of those firms (fixed effects) as opposed to the characteristics of some urban economies that give rise to agglomeration economies.

The study also links two other sources of information to the microdatabase. Census data provide information about the occupational distribution of the labour force in urban areas to test the extent to which occupational matches between firms and their urban areas are related to MFP. Input–output data are used to describe the nature of linkages that are important to different industries and then the microdata are used to ask whether suppliers in industries that the input–output tables identify as suppliers of importance are located in close proximity to each plant and whether the impact on MFP of that plant suggests that supplier links at the urban level contribute to MFP. Finally, the number of plants in the same industry located in close proximity is used to test whether there are intra-industry spillovers that arise from knowledge transfer of various sorts. This transfer could come through employees moving from one plant to another and from knowledge transfer that occurs from informal or formal contacts.

The study finds that all three sources of agglomeration economies are important. At the aggregate level, our results show that plant productivity is significantly influenced by the occupational distribution of workers, the density of the buyer-supplier network and the count of own-industry establishments within the region in which the plant is located. The labour-matching effect is empirically the largest. These results substantiate and extend earlier findings from cross-sectional investigation in the United States (Rigby and Essletzbichler 2002) and Canada (Baldwin et al. 2007).

Following Rosenthal and Strange (2001, 2003) and Henderson (2003), we explore the geographical extent of the benefits that derive from the co-location of plants. As with Rosenthal and Strange (2003), we find the benefits of own-industry co-location attenuate rapidly with distance, although when broken down by sector there are instances where intra-industry spillovers extend over longer distances.

The results also suggest that the impact of urban agglomeration economies with regard to labour supply and specialized suppliers is broadly felt across industrial sectors—though the impact does differ by sector, both in terms of the size of its impact and its statistical significance.

References

Almeida, Paul, and Bruce Kogut. 1999. “Localization of knowledge and the mobility of engineers in regional networks.” *Management Science.* 45, 7: 905–917.

Baldwin, John R., Desmond Beckstead, W. Mark Brown and David L. Rigby. 2007. *Urban Economies and Productivity*. Economic Analysis (EA) Research Paper Series. Catalogue no. 11F0027MIE2007045. Ottawa: Statistics Canada.

Baldwin, John R., and Wulong Gu. 2003. “Export-market participation and productivity performance in Canadian manufacturing.” *Canadian Journal of Economics.* 36, 3: 634–657.

Baldwin, John R., and Wulong Gu. 2005. *Global Links: Multinationals, Foreign Ownership and Productivity Growth in Canadian Manufacturing*. The Canadian Economy in Transition. Catalogue no. 11-622-XIE2005009. Ottawa: Statistics Canada.

Baldwin, John R., and Mohammed Rafiquzzaman. 1994. *Structural Change in the Canadian Manufacturing Sector (1970-1990).* Analytical Studies Branch Research Paper Series. Catalogue no. 11F0019MIE1994061. Ottawa: Statistics Canada.

Barron, John M., Dan A. Black and Mark A. Loewenstein. 1989. “Job matching and on-the-job training.” *Journal of Labor Economics.* 7, 1: 1–19.

Becker, Gary S., and Kevin M. Murphy. 1992. “The division of labor, coordination costs, and knowledge.” *The Quarterly Journal of Economics.* 107, 4: 1137–1160.

Beeson, Patricia E., and Steven Husted. 1989. “Patterns and determinants of productive efficiency in state manufacturing.” *Journal of Regional Science.* 29, 1: 15–28.

Black, Duncan, and Vernon Henderson. 1999. “A theory of urban growth.” *The Journal of Political Economy*. 107, 2: 252–284.

Chinitz, Benjamin. 1961. “Contrasts in agglomeration: New York and Pittsburgh.” *The American Economic Review*. 51, 2: 279–289.

Dumais, Guy, Glenn Ellison and Edward L. Glaeser. 1998. *Geographic concentration as a dynamic process.* Center for Economic Studies Working Paper 98-3. Washington, D.C.: Center for Economic Studies, U.S. Census Bureau.

Duranton, Gilles. 2007. *California dreamin’: the feeble case for cluster policies.* Unpublished manuscript. Department of Economics, University of Toronto.

Duranton, Gilles, and Diego Puga. 2000. “Nursery cities: urban diversity, process innovation, and the life cycle of products.” *The American Economic Review.* 91, 5: 1454–1477.

Gerking, Shelby. 1993. “Measuring productivity growth in U.S. regions: A survey.” *International Regional Science Review.* 16, 1-2: 155–185.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman and Andrei Shleifer. 1992. “Growth in cities.” *The Journal of Political Economy.* 100, 6: 1126–1152.

Henderson, J. Vernon. 2003. “Marshall’s scale economies.” *Journal of Urban Economics.* 53, 1: 1–28.

Henderson, J. Vernon, Ari Kuncoro and Matt Turner. 1995. “Industrial development in cities.” *The Journal of Political Economy.* 103, 5: 1067–1090.

Jacobs, Jane. 1969. *The Economy of Cities*. New York: Random House.

Jaffe, Adam B., Manuel Trajtenberg and Rebecca Henderson. 1993. “Geographic localization of knowledge spillovers as evidenced by patent citations.” *The Quarterly Journal of Economics.* 108, 3: 577–598.

Krugman, Paul. 1991. “Increasing returns and economic geography.” *The Journal of Political Economy.* 99, 3: 483–499.

Lucas, Robert E. Jr. 1988. “On the mechanics of economic development.” *Journal of Monetary Economics.* 22, 1: 3–42.

Marshall, Alfred. 1920. *Principles of Economics: An Introductory Volume.* (8th ed.). New York; London: Macmillan.

Mincer, Jacob. 1984. *Labor Mobility, Wages, and Job Training.* Mimeograph. New York: Columbia University.

Moomaw, Ronald L. 1981. “Productivity and city-size: A critique of the evidence.” *The Quarterly Journal of Economics.* 96, 4: 675–688.

Moomaw, Ronald L. 1983. “Spatial productivity variations in manufacturing: A critical survey of cross-sectional analyses.” *International Regional Science Review.* 8, 1: 1–22.

Moulton, Brent R. 1990. “An illustration of the pitfall in estimating the effects of aggregate variables on micro units.” *The Review of Economics and Statistics.* 72, 2: 334–338.

Organisation for Economic Co-operation and Development (OECD). 1987. *Structural Adjustment and Economic Performance.* Paris: Organisation for Economic Co-operation and Development.

Rauch, James E. 1993. “Productivity gains from geographic concentration of human capital: evidence from the cities.” *Journal of Urban Economics.* 34, 3: 380–400.

Rigby, David L., and Jürgen Essletzbichler. 2002. "Agglomeration economies and productivity differences in U.S. cities." *Journal of Economic Geography*. 2, 4: 407–432.

Rosenthal, Stuart S., and William C. Strange. 2001. "The determinants of agglomeration." *Journal of Urban Economics.* 50, 2: 191–229.

Rosenthal, Stuart S., and William C. Strange. 2003. "Geography, industrial organization, and agglomeration." *The Review of Economics and Statistics.* 85, 2: 377–393.

Rosenthal, Stuart S., and William C. Strange. 2004. "Evidence on the nature and sources of agglomeration economies." In *Handbook of Regional and Urban Economics*. *Vol. 4.* J. Vernon Henderson and Jacques-François Thisse (eds.). Amsterdam: North-Holland. 4, 49: 2119–2171.

Scott, Allen. 2001. *Global City-regions: Trends, Theory and Policy*. Oxford: Oxford University Press.

Sveikauskas, Leo. 1975. "The productivity of cities." *The Quarterly Journal of Economics.* 89, 3: 393–413.