IDENTIFYING THE SOURCES OF LOCAL PRODUCTIVITY GROWTH

Federico Cingano and Fabiano Schivardi[*]

1. Introduction

Since Marshall (1890) emphasized the importance of local scale economies for agglomeration (through technological spillovers, input-output linkages and labour market externalities), various alternative theories have been proposed to illustrate how the intensity and composition of productive activity affects local economic performance.[1] Early empirical works tried to determine whether differences in aggregate productivity levels among locations can be significantly explained by measures of the intensity of economic activity.[2] Findings in Jaffe *et al.* (1993) and Rosenthal and Strange (2000) suggest that information spillovers are an important source of externalities. A related strand of literature tries to assess what composition of the local industrial structure, if any, is most conducive to such externalities, focusing in particular on the role of sectoral specialization (localization economies) and product variety (urbanization externalities)[3] in determining spillovers within and between industries.

In light of the central role of technological externalities in modern growth theories, recent works focus on the possibility that externalities arising from local interactions might cause differences not only in productivity levels but also in growth rates. In a seminal paper, Glaeser *et al.* (1992) estimated the effects of other possible sources of technological spillovers at the local level. They found strong evidence that indicators of

* We thank Antonio Ciccone, Andrea Lamorgese, Ariel Pakes and seminar participants at the Bank of Italy, Universidad Carlos III, NARSA Meeting 2002 and the workshop on "Cities and Geography" hosted by CERAS in Paris for useful discussions. Marco Chiurato and Elena Genito provided excellent research assistance. We also thank the Italian social security administration (INPS) for supplying us with their firm-level data. We are responsible for any mistakes. The final version of this paper was published in the Journal of the European Economic Association, 2004, Vol. 2, pp. 720-742.

1 See, for example, Henderson (1974) for an early contribution and Eaton and Eckstein (1997) for a recent one.

2 See, for example, Sveikauskas (1975), Segal (1976), Ciccone and Hall (1996).

3 The seminal contribution for urbanization economies is in Jacobs (1969); Duranton and Puga (2001) offer an interesting modelization of such effects over the industry life cycle.

localization economies (also called MAR economies, for "Marshall-Arrow-Romer") have a negative growth effect in a cross-section of US cities, while urbanization (or "Jacobs") economies, spurred by production variety, are positively related to subsequent growth. Adopting a somewhat different approach, Henderson *et al.* (1995) found positive effects of productive specialization in the case of mature capital-goods industries, while production variety seemed to be more important for newly established high tech industries. Further studies helped to extend what is now known as the "urban growth" literature to countries other than the US (see, for example, Combes, 2000 for France; Cainelli and Leoncini, 1999 for Italy; and Bradley and Gans, 1996 for Australia). Their results tend to confirm that productive specialization has a negative impact on growth, while evidence on urbanization economies is less clear-cut. Such findings are quite puzzling, not only because they imply the absence of intra-industry technological spillovers, but also because they suggest there are dynamic *disadvantages* to spatial concentration.

We argue that these controversial findings in the literature may depend on a simple identification problem. Theories of dynamic externalities predict a relation between local structure and productivity; because of the lack of local productivity data, existing works have used employment growth regressions, on the assumption that productivity gains produce proportional employment gains through shifts in labour demand. This approach implicitly assumes that changes in *labour supply* are independent of local conditions. This is a strong assumption, however: for example, congestion externalities, such as higher rents and pollution, are likely to influence mobility choices, potentially breaking or even reversing the causal chain from agglomeration economies to productivity and employment growth.[4]

We overcome this problem by using a measure of growth that is closer to the theoretical notion of dynamic externalities. We exploit balance-sheet data on a large sample of Italian manufacturing firms to construct a measure of sectoral TFP with a high degree of geographical disaggregation, applying a production function estimation procedure that allows us to account carefully for endogeneity and selection problems (Olley and Pakes, 1996). We then regress productivity growth at the city-

4 In their study of city-wide characteristics (i.e. without sectoral breakdown) and city population growth, Glaeser *et al.* (1995) explicitly acknowledge that urban growth regressions can only capture the impact of local conditions on both productivity and quality of life.

industry level against precise employment-based beginning-of-period indicators of the local industrial structure. Our contribution to the literature is twofold. First, we construct a test of location economies that does not rely on the identification assumptions required for the employment growth regressions. To our knowledge, this is the first paper that tests agglomeration theories using TFP data with a high degree of both geographical and sectoral disaggregation. Second, given that we also have detailed information on employment growth, we can run the regressions previously used in the urban-growth literature and compare the results with those of the TFP regressions. This will give an indication of the importance of the identification issues in interpreting the employment-based regressions as evidence for agglomeration economies.

In testing location economies, our main results can be summarized as follows. Indicators of specialization-MAR economies have a sizeable effect on productivity: doubling the share of sectoral employment in a given location brings about an average increase in sectoral TFP of 0.2 percentage points per year, on a 10 per cent increase in the average rate of productivity growth. We also find evidence that city size matters: doubling initial employment in manufacturing raises TFP by 0.4 percentage points per year. These results are consistent with a broad theoretical literature on urban growth but in conflict with most of the empirical literature to date. We do not find that other possible sources of externalities, such as urban diversity, local competition or average firm size are significant determinants of TFP growth.

To address the relevance of the identification issues in employment growth regressions, we first construct a simple model of local conditions, productivity growth and employment determination, which formally shows that identification requires changes in labour supply to be independent from local conditions affecting productivity growth. We assess the empirical relevance of this assumption by running employment growth regressions, finding results that are opposite to those for TFP and in line with the previous literature. We perform several robustness checks and extensions, all pointing to the importance of the identification problem. Taken together, this evidence suggests that employment growth might be ill-suited to infer the sources of dynamic productivity growth, casting serious doubts on the interpretation of the results previously found in the literature as evidence for or against dynamic externalities.

The rest of the paper is organized as follows. In section 2 we describe the data sources and TFP estimation at the local geographical level; section 3 discusses the empirical specification and the main results obtained with TFP. Section 4 formally illustrates the identification issues affecting employment growth regressions and discusses results from such estimates, comparing them with those obtained for TFP. Section 5 concludes.

2. Measuring local TFP

2.1 *Data*

As in most of the literature, the unit of observation in our analysis is defined by sectoral activity at the local level. Our geographical units are Italy's local labour systems (LLS), defined as groups of municipalities characterized by a self-contained labour market, as determined by the National Institute for Statistics (Istat) on the basis of the degree of working-day commuting by the resident population. Using 1991 census data, the Istat procedure identified 784 LLSs covering the whole national territory.[5] Given that externalities are likely to arise mainly from direct interaction, this is the ideal geographical unit for studying local spillovers. In terms of the sectoral classification, we restrict our attention to manufacturing, given the well-known problems in estimating productivity in services. Following the territorial analysis of Istat, we use the 10-sector classification system reported in Table 1, which achieves a good compromise between the need for intra-sectoral homogeneity and that of a sufficient number of observations by sector for a statistically reliable analysis. Our unit of observation is the local labour system sector (which from here on we designate as L-S).

To obtain information on productivity and on its determinants at the L-S level, we combine data from three different sources. First, we exploit several waves of the Italian social security administration (*Istituto Nazionale Previdenza Sociale*, INPS) archives on the universe of Italian firms (1986-98) to compute precise measures of the local industrial structure. For all firms with at least one employee, the firms' archives provide information on the total number of employees working in each

[5] The average land area is 384 square kilometers, with a population density of 188 inhabitants per sq. km. Population ranges from 3,000 in the smallest LLS to 3.3 million in the largest.

year (with a breakdown between production and non-production workers), their average yearly earnings and some firm characteristics. In particular, for each firm we know the address (municipality and postal code) and the sector of activity (specified with a three-digit breakdown), which together allow us to classify each firm in the corresponding L-S. We use these data to compute the employment-based measures of the local industrial structure (such as indexes of productive specialization, variety, firm size and local competition).[6]

The INPS dataset has no information on production or capital stock, so that it cannot be used to compute TFP. We therefore resort to a sub-

Table 1

Firms' characteristics (average values)

Sector		Company Accounts				INPS	
		Value added*	Capital stock*	Employ-ment	No of obs.	Employ-ment	No of obs.
1	F	4,617	9,596	92	1,516	10	25,819
2	T&C	2,769	4,287	82	2,335	13	43,784
3	L&F	1,637	1,634	53	820	12	13,254
4	W&C	1,756	3,086	55	1,167	7	27,830
5	T&Gl	4,017	9,912	88	1,260	15	14,001
6	BM	6,393	17,065	157	711	39	4,224
7	Mach	4,441	5,935	112	5,582	14	91,606
8	Chem	7,460	14,843	128	2,013	27	13,785
9	P&P	4,325	7,375	90	992	12	15,634
10	Teq	9,692	36,191	555	489	115	2,353
	Total	4,749	8,418	113	16,885	14	261,549

Note: * = thousands of 1991 euros. Sectoral classification: F = Food, beverages and tobacco; T&C = Textiles and clothing; L&F = Leather and footwear; W&C = Wood, products of woods and cork; T&Gl = Timber, construction materials and glass; BM = Basic metals; Mach = Metal products, machinery and equipment; Chem = Rubber, plastic and chemical products; P&P = Paper, printing and publishing; TEq = Transportation equipment.

[6] The archives allowed for the computation of indicators that require firm-level information (see next section) that could not have been computed using industry census data, available only at the aggregate level.

sample of firms drawn from the Company Accounts Data Service (*Centrale dei Bilanci*), a large dataset collected by a consortium of banks interested in pooling information about their clients and containing detailed balance-sheet information. Data refer to a sample of 30,000 to 40,000 firms and have been available on an annual basis since 1982. Since the data are used by banks to extend loans, they are quality-controlled and contain actually reported (as opposed to imputed) figures.[7] Firms in the sample account for approximately one half of the total employment in manufacturing and, according to a Data Service report (1992), for an even higher share of sales. Table 1 reports industry-level averages for three variables of interest (value added, capital stock – constructed using the permanent inventory method, see the Appendix – and employment) in 1991.

The use of a sub-sample of firms entails two problems. First, not all the existing L-S will be present in the sample, as we established through a comparison with INPS data (the universe). If we consider for example 1991, the Company accounts dataset has at least one firm for 2,453 L-S out of 6,372; in terms of LSS, 539 of them are in our sample, of the total of 784. Given that the selection criterion is independent of localization, the probability that a given L-S is represented in our sample increases with the number of firms in it, so we will tend to exclude L-S with low levels of sectoral employment. In fact, the average sectoral employment in excluded L-Ss is only 75 workers, against almost 1,400 for those included. In terms of coverage, the L-Ss included account for a share of total sectoral employment that ranges from 86 per cent for wood to 98 per cent for metal products. Notice that the exclusion of L-Ss with very low sectoral employment is very much in line with the previous literature, which generally only considers metropolitan areas (Glaeser *et al.,* 1992).

The second potential problem is that firms are not randomly chosen. Though previous comparisons indicate that the Company accounts data are not too far from being representative of the whole population in terms of the frequency distribution by sector and geographical area (Guiso and Schivardi, 2000), the focus on the level of borrowing skews the sample towards larger firms. This can be noticed from the last two columns of

7 This dataset has been used, among others, by Guiso and Parigi (1999) to study the effects of uncertainty of firms' investment decisions, by Pagano *et al.* (1998) for the choices of going public, by Sapienza (2002) for the effects of bank mergers on interest rates on loans, by Guiso and Schivardi (2000) to explore the impact of information spillovers on firms' behaviour.

Table 1, comparing average employment and number of firms at the sectoral level for the Company accounts and the INPS databases in 1991: the left-hand skewness of the size distribution of Italian firms (in manufacturing, firms with 5 employees or fewer account for 60 per cent of the firm population but less than 10 per cent of total employment) explains much of the observed differences. Moreover, since banks are most interested in firms that are creditworthy, firms in default are not in the dataset, so that the sample is also biased towards higher than average quality borrowers. While we have no direct way of accounting for potential selection problems affecting our productivity growth estimates, we will show that employment growth regressions based on Company accounts data (the sub-sample) and the INPS data (the population) yield very similar results (see section 4). Therefore, we are confident that the selection criteria, based on turnover thresholds and on multiple banking relationships, are unlikely to induce any spurious correlation between the estimated local TFP growth rates and our explanatory variables. We also include detailed sectoral and geographical controls in our growth regressions to account for error in measurement that is correlated across space or product chains.

As for the precision of our productivity growth estimate, the average number of observations at the city-sector level is 8.5 (Table 4, last row) and, given that both the sectoral and the geographical classification are fairly detailed, in many cases we end up computing TFP with just a few firm-level observations. While this is likely to introduce noise, we think our measure is sufficiently precise for our purposes. First, as we have seen above, firms included in the Company accounts dataset account for a large share of output. Second, in order to account for the different precision with which TFP is computed, we will estimate our regressions using weighted least squares, with the weights determined by the number of firm-level observations available. We will also perform several additional robustness checks.

The final data source is the census, from which we obtain additional economic indicators at the local geographical level. In particular, we used the Italian population census (Censimento generale della popolazione, 1981) to calculate measures of human capital in the LLSs, obtained as average schooling of the working-age population, and the 1981 service and industry census, used as an alternative source of employment data in our robustness checks.

2.2 *TFP Estimation procedures*

We exploit our detailed firm-level dataset to measure TFP at the L-S level. We use the usual Cobb-Douglas production function $Y = AK^{\alpha_s} L^{\beta_s}$, where *K* and *L* denote the stock of capital and labour and *A* is the TFP, and where we allow for the coefficients α_s and β_s to vary across sectors. The traditional method assumes perfect competition in the input markets and constant returns to scale in production (Solow's assumptions) and calculates $\widehat{\beta}_S$ as the labour share in each sector and $\widehat{\alpha}_S$ as its complement to 1. The availability of firm-level data, however, allows us to estimate the coefficients directly. The advantages of estimating the production function with firm data is that Solow's assumptions are not required. In fact, the Italian labour market is heavily regulated, so that the perfect competition hypothesis is hard to justify. Moreover, by dismissing the assumption of constant returns to scale, we can disentangle TFP growth from scale effects internal to the firm, determined by the production technology and therefore independent from externalities at the local level. Indeed, with Solow's method any effect of the scale of production would be attributed to TFP, potentially introducing a significant measurement error.

The direct estimation of the production function faces well-known econometric problems. Since the level of productivity will affect both the firm's input choices and the participation decision, consistent estimation of the production function parameters makes it necessary to address problems of selection and simultaneity. We use a multi-step estimation algorithm proposed by Olley and Pakes (1996), which accounts both for the endogeneity and the selection problems, allowing for unbiased and unconstrained estimation of α_s and β_s. The procedure is briefly summarized in the Appendix. To obtain our measure of city-sector TFP we first calculate productivity at the firm level as a residual, accounting for the fact that the scale of individual plants matters if we do not impose CRS, and aggregate TFP to the city-sector level as the employment-weighted average of firm-level TFP.[8] To control for the reliability of the estimates, we also calculate the coefficients using Solow's assumptions, computing $\widehat{\beta}_S$ as the

[8] Alternatively, we could have used directly the growth of TFP at the firm level without aggregating at the city-sector level. The problem with this approach is that it would have restricted the sample to the surviving firms only, thus reducing the representativeness of the results.

average labour share in each sector and $\widehat{\alpha}_s$ as its complement to 1.[9] In this case, the city-sector estimates of TFP are obtained as $ln\ A_{c,s} = Y_{c,s} - \hat{\alpha}\ ln\ K_{c,s} - \widehat{\beta}_s\ ln\ L_{c,s}$, where $X_{c,s} = \sum_{i \in c,s} x_i$.

Table 2 reports the estimated values of α_s and β_s the two procedures. Production function estimates of $(\alpha_s + \beta_s)$ lie in the range 0.93-1.05, indicating that the CRS assumption is a good approximation for most sectors but that for a few of them it might not be inconsequential for TFP calculations, particularly in the face of changes in the average scale of production. In terms of single coefficients, the Olley and Pakes procedure tends to yield a higher labour and a lower capital coefficient, arguably because of deviations of the factor markets from the competitive paradigm. Apart from these differences, the two methods give broadly consistent results, an indication of the reliability of the estimates. In what follows we use the production function estimates as our preferred ones.

Table 2

Production function coefficients: factor share and direct estimates

	Factor shares		Direct estimates		
Sector	β	α	β	α	$\alpha + \beta$
1 F	0.56	0.44	0.63	0.39	1.02
2 T&C	0,60	0.40	0.58	0.37	0.95
3 L&F	0.61	0.39	0.62	0.43	1.05
4 W&C	0.63	0.37	0.70	0.35	1.05
5 T&Gl	0.58	0.42	0.67	0.37	1.04
6 BM	0.65	0.35	0.60	0.33	0.93
7 Mach	0.67	0.33	0.72	0.28	1.00
8 Chem	0.60	0.40	0.70	0.29	0.99
9 P&P	0.66	0.34	0.72	0.32	1.04
10 TEq	0.74	0.26	0.70	0.26	0.96

Note: α is the capital coefficient and β the labour one. The first estimates use the traditional Solow approach, the second the direct estimation of the production function coefficients using the Olley and Pakes (1996) procedure. See Table 1 for the sectoral labels.

9 For this method, less computationally intense, we also allowed the coefficients to differ by year and macro area; we did not find a significant trend in estimated coefficients along either dimension. We therefore assume that they are constant over time and across areas within a given sector.

Table 3 reports the decomposition of output per worker in the ten manufacturing sectors considered here. The upper part of the table shows that the level of TFP (calculated in 1991) accounts for more than one half of labour productivity, a result that is roughly comparable to those obtained by Bernard and Jones (1996) in a sample of OECD countries. The bottom part of the table presents a standard growth accounting exercise. As the second column shows, between 1986 and 1998 TFP grew on average at a rate ranging between 1.2 per cent and 4 per cent and was generally lower in

Table 3

Labour productivity decomposition

Sector		Y/l	PTF	α*k/l	γ*l
		Levels, 1991 (log)			
1	F	3.80	1.97	1.71	0.12
2	T&C	3.41	2.33	1.34	-0.26
3	L&F	3.33	1.78	1.32	0.23
4	W&C	3.40	1.86	1.32	0.22
5	T&Gl	3.72	1.86	1.64	0.22
6	BM	3.64	2.64	1.49	-0.49
7	Mach	3.58	2.56	1.02	0.00
8	Chem	3.90	2.76	1.27	-0.12
9	P&P	3.81	2.24	1.34	0.23
10	TEq	3.54	2.85	1.06	-0.37
		Growth rates, 1986-1998 (% per year)			
1	F	3.24	2.29	0.97	-0.02
2	T&C	3.32	2.22	1.07	0.03
3	L&F	3.20	1.64	1.51	0.04
4	W&C	3.40	3.18	0.13	0.09
5	T&Gl	3.61	3.34	0.29	-0.02
6	BM	4.60	4.03	-0.15	0.72
7	Mach	4.18	4.00	0.18	-0.00
8	Chem	3.53	3.18	0.29	0.06
9	P&P	3.15	2.70	0.49	-0.03
10	TEq	1.94	1.15	0.61	0.18

Note: The first column is overall labour productivity, the second is the TFP contribution, the third capital accumulation and the last returns to scale. See Table 1 for the sectoral labels.

traditional industry (textiles, footwear etc.) and the food sector than in basic metals and machinery.[10] The accumulation of capital per worker, on the contrary, accounted for large parts of the growth in productivity per worker in the traditional sectors. Further interesting differences emerge, driven by returns to scale. The last column indicates the amount of the productivity increase/decrease due to the change in the productive structure of the firms in the sample. In line with the previous discussion about the coefficients, contributions are generally small. The most noticeable exception is basic metals and, to a lesser extent, transportation equipment, where a substantial contribution to labour productivity growth came from the *decrease* (recall that both sectors are characterized by DRS) in the average scale of production of firms in the sample. This effect is not captured by the Solow procedure, which therefore overestimates TFP growth.

3. Sources of local growth

The existence and extent of dynamic externalities is generally inferred from the analysis of the relationship between employment growth at the local level and indexes of the local productive structure. Our data allows us to closely parallel the existing literature and compare the results obtained when testing for the impact of alternative sources of dynamic spillovers on TFP, as opposed to employment growth rates. Though the theory lacks clear indications of what the relevant variables should be, since the first empirical works by Glaeser *et al.* (1992) and Henderson *et al.* (1995) the focus has been on specific employment-based indicators of dynamic spillovers. First, specialized locations should benefit from intra-industry knowledge spillovers, as is argued by a strand of literature that goes from Marshall (1890) to Arrow (1962) to Romer (1986). These are called MAR-externalities. Empirically, the degree of sectoral specialization of a given location (city) c in a certain sector s is captured by the share of sectoral city employment:

$$Spec_{c,s} = L_{c,s} / L_c$$

[10] The relatively high level of TFP growth is attributable to the fact that the sample is biased towards higher than average quality firms.

On the other hand, positive externalities could be induced by the scale or diversity of local economic activities outside sector *s* as a result of cross-fertilization. The effects of production variety in the city, commonly called Jacobs externalities (1969), is captured here by a Hirschman-Herfindahl index (Henderson *et al.*, 1995):

$$Variety_{c,s} = \sum_{j \neq s}^{J_c} \left(\frac{L_{c,j}}{L_c - L_{c,s}} \right)^2$$

The index is defined as the sum of the (squared) shares of other sectors' employment in overall net manufacturing employment in city *c*. Clearly it will be close to 1 if sector *s* is surrounded by few, concentrated industries in the city, while it tends to $1/(J_c - 1)$ if city-employment (at two-digit level) is evenly distributed across different industries.

The variables mentioned have traditionally been the most important, according to the empirical literature, for discriminating between specialization and urbanization economies. Other characteristics of the production structure have been considered, though, as potentially relevant determinants of local productivity. First, some theories predict that fierce product-market competition at the local level could be a source of positive externalities by, for instance, fostering the adoption of innovations by firms (these are known as 'Porter externalities' after Porter (1990). Following Combes (2000) we measure local competition as a local Herfindahl index of concentration computed at the firm level

$$\mathrm{Comp}_{c,s} = \sum_{i \in c,s} \left(L_{c,s,j} / L_{c,s} \right)^2$$

where with $L_{c,s,i}$ is the employment level of firm *i* belonging to city-industry *c,s*. The index measures the distribution of the employment shares calculated at plant level within each city-industry: low competition should result in a less uniform distribution of employment across the existing firms.

Finally, we include the average size of plants in the city-industry, to allow for the possible effects of firm size structure on growth.[11] To facilitate comparison with the existing literature, we use the inverse of average firm size

$$Size_{c,s} = \frac{1}{\bar{L}_{c,s}}$$

that is, the number of firms over employment in the city industry, the same index used by Glaeser *et al.* (1992).

Indexes have been calculated using the 1986 INPS archive on the universe of firms. Summary statistics of the main variables used in the empirical analysis are in Table 4.

Table 4

Descriptive statistics

City-industry variables	Descriptive statistics		
	Mean	Median	Std
TFP average yearly growth	0.027	0.027	0.033
Specialization index	0.172	0.100	0.179
City size	16,533	6,082	49,193
Average firm size	25.20	12.12	89.51
Variety index	0.126	0.089	0.107
Competition index	0.199	0.122	0.216
Average years of schooling	7.516	7.495	0.757
Number of firms*	8.456	3.000	24.54

Note: Statistics based on the sample of 1602 city-industry observations used in the regressions shown in the paper. * = number of firm-observations available by city-industry to calculate aggregate TFP.

11 Pagano and Schivardi (2001), using cross-country data at the sectoral level, find that productivity growth is positively correlated with average firm size and offer evidence that the direction of causality goes from size to growth.

3.1 *Productivity growth and the local industrial structure*

This section illustrates the regression specification and results obtained regressing the average TFP growth rate calculated over the 1986-98 period on the above-mentioned employment indicators. The Company accounts sample is an open one, with entry and exit of firms over time. Hence, in principle, we can compute city-sector TFP growth rates applying a variety of sample selection rules. The results shown in this section are obtained using the most restrictive selection rule, i.e. considering only those city-industries that are represented by at least one firm over the entire time-span. However, our results are robust to alternative selection rules (see below).

The adopted specification follows closely that proposed by Combes (2000):

$$\hat{A}_{cs} = \beta_1 SPEC_{c,s} + \beta_2 VAR_{c,s} + \beta_3 COMP_{c,s} + \beta_4 SIZE_{c,s} + \beta_5 X_{c,s} + u_{c,s} \tag{1}$$

where capital letters indicate log transformation of the corresponding regressors, and the vector $X_{c,s}$ contains additional controls included on the right-hand side. In particular we controlled for the logarithm of city employment in 1986 l_c, so that the coefficient β_1 can be correctly interpreted as the effect of local relative concentration (sectoral employment share), holding total employment in the city constant (see Combes, 1999). We also accounted for the variability in human capital endowment across cities, measured by the average number of years of schooling of the city working-age population in 1981, for the initial level of city-sector TFP and for two sets of dummy variables accounting for the sector of activity and geographical location of the city (macro-area).

The TFP estimates in the L-S are obtained by averaging firm-level TFP, so the precision of the estimates increases with the number of firms. To reduce the noise from imprecise estimates of the L-Ss for which only a few firms are included, we use WLS, where each data point has been weighted by the number of firm-level observations by L-S. This implies that L-Ss with a higher number of firms will have more weight in determining the estimated coefficients.[12]

[12] The weighting scheme is the same as would be obtained if we used firm-level TFP growth directly as the dependent variable rather than its average in the C-S.

Table 5

City-industry productivity growth

	[1]	[2]	[3]	4]	[5]	[6]
Specialization	0.230**	0.199*	0.206*	0.346***	0.206*	0.394***
	(0.111)	(0.113)	(0.117)	(0.125)	(0.110)	(0.162)
City size	0.401***	0.390***	0.447***	0.492***	0.395***	0.619***
	(0.095)	(0.093)	(0.094)	(0.115)	(0.094)	(0.118)
Firm size	0.357*	0.296	0.321	0.563*	0.343*	0.596**
	(0.209)	(0.206)	(0.211)	(0.294)	(0.206)	(0.264)
Variety	-0.012	-0.013	0.091	0.015	-0.009	-0.142
	(0.119)	(0.106)	(0.159)	(0.119)	(0.116)	(0.161)
Competition	0.085	0.074	0.097	0.214	0.088	0.022
	(0.102)	(0.098)	(0.094)	(0.132)	(0.101)	(0.124)
Spatial controls	5	20	95	5	5	5
Weights	YES	YES	YES	NO	YES	YES
No. of obs.	1,602	1,602	1,602	1,602	1,810	2,876
R^2	0.43	0.45	0.49	0.41	0.44	0.19

Note: Dependent variable: annual TFP growth rate at the L-S level. All regressions include sector dummies. Spatial controls are macro areas, regions and provinces. The first four columns are based on the sample of C-S continuously in the database, the fifth in the database in 1986 and 1998, the last in the database in any year.
*** indicates significance at 1 per cent, ** at 5 per cent and * at 10 per cent.

Table 5 summarizes the results obtained estimating different versions of equation (1) by WLS. Column [1] reports the basic specification. First, we find that the elasticity of TFP growth to sectoral specialization, holding total city size constant, is positive and significant at the 5 per cent level. Our point estimate (β_1 = 0.23) implies that an increase in sectoral employment that shifted the median city-industry concentration index to the third quartile (raising the share of sectoral employment 3 times) would be associated with an average yearly increase in TFP of nearly 0.5 percentage points over the subsequent period. This result is in contrast with previous evidence for other countries, where industries are found to grow more slowly in relatively more concentrated locations (Glaeser *et al.,* 1992; Combes, 2000). Second, we find that TFP growth is positively affected by city size. Since we hold the sectoral composition of production in the city constant, this can be interpreted as the effect of the size of the local market, consistent with a broad literature on urban growth. The elasticity of productivity to total manufacturing employment in the city is 0.4 per cent, implying that moving from the median to the 75 per cent larger city increases yearly productivity growth by 0.8 per cent on average. This indicates that scale effects are important determinants of productivity growth at the local level.

We do not find that other possible sources of externalities at the local level are relevant to our measure of TFP growth. Both the initial range of productive variety and the degree of competition at the beginning of period, capturing Porter and Jacobs externalities respectively, seem to affect TFP growth positively, but their elasticities are not significantly different from zero (at the 10 per cent level). The same is true for our measure of average human capital in the LLS. We find weak indication that productivity in city-industries characterized by a smaller firm size tends to grow faster. On the other hand, the coefficient of initial TFP level in the city-industry is negative and highly significant, capturing convergence in growth rates across city-industries.

One problem with the results shown in Table 5, col. [1], is that the original specification might be missing important determinants of productivity growth at the local level. To control for this possibility we checked the robustness of our estimates to spatially correlated omitted variables. In practice, this amounts to increasing the number of spatial controls included in the baseline regression: as long as some part of the variation in omitted determinants of TFP growth across city industries is picked up by these spatial control variables, and if omitted variables do indeed affect the estimation of the key parameters, then adding such variables would change the effect of the included regressors. The results are shown in column [2], where 20 spatial control variables (corresponding to administrative regions) are included, and in column [3], where we control for the 95 Italian provinces in 1986. The estimates are affected only very slightly: in column [2] the specialization coefficient falls marginally and in column [3] the corresponding standard error increases marginally, but the estimate remains significant at the 10 per cent confidence level. The other results are unaffected.

The last three columns perform additional robustness checks. Our weighting scheme gives greater weight to L-S with a high population of firms. To make sure that this does not influence the results, we run the basic specification as in column [1] without weighting. The specialization coefficient increases to .346 and is estimated more precisely, while that of city size increases marginally. All other results are unaffected. We also control for different selection criteria. As noted, the baseline specification uses only L-Ss that were continuously present in the sample. In column [5] we use those that are in the sample in the first and the last year, and in column [6] all possible information, calculating average TFP growth using all available years, i.e. including all L-S that were ever in the sample for some years. The number of observations increases from 1,602 to 1,810 and 2,876 respectively. Again, the basic results are unchanged, the major difference being that in the case with the most observations (column [6]) the effects of sectoral specialization and average firm size are stronger.

Having established the existence of non-negligible MAR externalities at the city-industry level, we also examined how localized these forces are by adding two variables measuring scale and own-industry

Table 6

City-industry productivity growth: neighbourhood externalities

	[1]	[2]	[3]
Specialization	0.230**	0.262**	0.219*
	(0.111)	(0.124)	(0.126)
Neighbourhood's specialization	-	0.066	-0.043
	-	(0.111)	(0.112)
City size	0.401***	0.383***	0.373***
	(0.095)	(0.094)	(0.094)
Firm size	0.357*	0.352*	0.292
	(0.209)	(0.206)	(0.204)
Variety	-0.012	-0.004	-0.008
	(0.119)	(0.118)	(0.106)
Competition	0.085	0.077	0.068
	(0.102)	(0.101)	(0.097)
Spatial controls	5	5	20
Weights	YES	YES	YES
No. of observations	1,602	1,602	1,602
R^2	0.43	0.43	0.45

Note: Dependent variable: annual TFP growth rate at the L-S level. All regressions include sector dummies. Spatial controls are macro areas, regions and provinces.
*** indicates significance at 1 per cent, ** at 5 per cent and * at 10 per cent.

specialization in the neighbouring area, obtained by aggregating our city-sector data to the province level.[13] While the estimated localization effects are not affected in this specification, we do not find that own-industry specialization in neighbouring areas matters for TFP growth (Table 6, cols. [2] and [3]; the first column replicates column [1] in Table 5). This result is in line with previous work based on patents (Jaffe *et al.*, 1993) and employment levels in new establishments computed at the postal code level (Rosenthal and Strange, 2000), which found that localization economies attenuate rapidly with distance.

4. Local conditions, productivity growth and employment growth

Our findings regarding the determinants of local productivity growth are at odds with those of most of the urban-growth literature, which has found strong evidence of negative MAR-specialization externalities and positive Jacobs urbanization economies. Here we argue that this difference probably derives from a problem with the specification used in earlier studies rather than from the peculiarity of the Italian productive system.[14]

Given the lack of data on productivity at the local sectoral level, the literature has mostly used employment growth, on the hypothesis that changes in productivity result in proportional employment changes. To examine the assumptions that underlie this approach, we construct a very simple model of employment determination at the local level. Consider an economy organized in many different cities C_i, each representing a local labour market.[15] We take a partial equilibrium approach, since in our empirical specification each C-S is small with respect to the economy as a

[13] Each of the 95 Italian provinces in 1986 contained on average more than 7 local labour systems.

[14] The Italian productive system, characterized by areas with many small and medium sized enterprises (called ‘industrial districts’), could in principle be particularly conducive to interaction-induced externalities. Guiso and Schivardi (2000) study information spillovers among Italian district firms, finding that they significantly influence firms’ behaviour and performance. Interestingly, the motivating example of Porter’s competition effect (1990) was the tile industry in Sassuolo, an area around Bologna where there is a heavy concentration of successful tile firms.

[15] For simplicity, we exclude the sectoral dimension, but the analysis can easily be extended to include sectoral differences. In practice, due to human capital specificity, segmentation across labour markets can have not only a geographical dimension but also a sectoral one, so that the city in the model can be thought of as a city-sector. The hypothesis of sectoral segmentation and of costs of moving from one sector to another finds empirical support (see for example Shin, 1997).

whole, which sustains the assumption that the overall wage rate is given and not influenced by that prevailing in each individual location.

Within C_i there is a representative firm producing output with labour as the only input using the production function $F(A_i, l_i)$, where l is labour and A is the level of TFP. Following Glaeser *et al.* (1992) it is maintained that the growth of TFP depends on the local industrial structure, here captured by the vector of (log) variables X_i:

$$\Delta \log A_i = \theta \log X_i \qquad (2)$$

where the parameter vector θ captures the dynamic external effects of different local characteristics. The technology is Cobb-Douglas with decreasing returns to scale[16] owing to some scarce factor such as land: $F(A,l) = Al^{\alpha}$, $\alpha < 1$. The representative firm takes wages as given, so that profit maximization yields a standard labour demand schedule:

$$l^D(w_i, A_i) = \left(\frac{\alpha A_i}{w_i}\right)^{\frac{1}{1-\alpha}} \qquad (3)$$

Equation (3) is the basis for the use of employment changes as proxy for productivity changes in previous empirical work. In fact, given w_i, productivity changes result in proportional employment changes, with $\frac{1}{1-\alpha}$ as the factor of proportionality. One can then substitute $\Delta logA_i$ for $\Delta logl_i$ in equation (2) and perform the analysis with employment growth.

The problem with this approach is that it neglects the role of labour supply, an assumption that can have important consequences for the identification of θ. In fact, workers' mobility choices are influenced not only by wage differentials, but also by other aspects of the location and the job, such as amenities, housing prices, pollution, congestion, individual preferences regarding jobs, and so on. To see how this might affect

16 This assumption regarding the aggregate production function in no way restricts the degree of returns to scale of the accumulable factors at the firm level, and is not in contrast with the estimation procedure, which does not require any assumption as to the degree of returns to scale for capital and labour at the firm level.

identification, we model labour supply by assuming that each worker's utility function is defined over income and city-specific characteristics, with the constant-elasticity form:

$$U(w,Z) = \omega^{\delta} Z^{\eta} \tag{4}$$

where income is equal to the wage rate and Z is a vector of relevant city-specific characteristics. Outside C_i there is continuum of workers of mass 1 with reservation utility normalized to zero. A worker can decide to take the reservation utility or pay a moving cost m measured in utility units, and move to take up employment in C_i. The problem of the worker is then:

$$\max\{0, w_i^{\delta} Z_i^{\eta} - m\} \tag{5}$$

The worker will move if and only if $w_i^{\delta} Z_i^{\eta} \geq m$.

Workers are distributed around C_i at increasing cost according to the distribution function $g(m)$ which, to get an analytical solution, we assume to be uniform on the interval $[0,M]$. The parameter M measures the average distance of workers from: $E(m) = M/2$. At any given level of wage w_i and local characteristics Z_i, all workers with $m \leq w_i^{\delta} Z_i^{\eta}$ will move to take jobs in C_i so that the local labour supply is:

$$l^S(w_i, Z_i) = \int_0^{U(w,Z)} g(m)dm = \frac{1}{M} w^{\delta}{}_i Z^{\eta}{}_i \tag{6}$$

Given the vector Z_i, labour supply depends positively on the wage but is decreasing in M: the higher the average mobility cost, the smaller the labour force moving to C_i.

By equating labour demand (3) and supply (6), taking log, first differences[17] and substituting (2), the relationship between equilibrium labour growth and X_i can be written as:

[17] We are implicitly assuming that the previous period employment level plays no direct role in determining current-period labour supply, so that all persistence in employment comes from

(continues)

$$\Delta \log l_i = \frac{\delta\theta}{\gamma} \log X_i + \frac{\eta}{\gamma} \Delta \log Z_i \qquad (7)$$

where $\gamma \equiv 1+(1-\alpha)\delta$. Equation (7) makes two points. First, $\frac{\delta\theta}{\gamma}$ constitutes a transformation of the original parameter of interest θ, so that the employment coefficient cannot be interpreted quantitatively in terms of productivity growth. Second, and more importantly, its unbiased estimation requires that changes in all omitted variables that affect labour supply through Z_i should be independent of the set of variables that generate technological spillovers X_i. Stated differently, the identification of dynamic externalities from employment growth requires that the sources of such externalities should not influence labour supply: failing this, local conditions will shift both labour demand and labour supply, giving rise to a classical identification problem.

To check whether employment-based regressions are affected by such identification problems, we run the regression specification (1) using measures of labour growth recovered from the INPS dataset[18] as the dependent variable:

$$\hat{l}_{cs} = \beta_1 SPEC_{c,s} + \beta_2 VAR_{c,s} + \beta_3 COMP_{c,s} + \beta_4 SIZE_{c,s} + \beta_5 X_{c,s} + u_{c,s} \qquad (8)$$

To maximize comparability with the TFP regressions, we estimated this equation using the same WLS scheme and the same sub-sample of city-industry observations we used in the TFP regression. The results, reported in Table 7, indicate that productive concentration is associated with slower employment growth at the city-sector level, the opposite of what was found for TFP. In particular we find that doubling the share of sectoral employment in a given location will reduce average employment growth in the same sector by 0.75 percentage points per year. The partial

persistence in city characteristics Z. This assumption could be removed by directly modeling the mobility choices of workers in the city in the previous period, by assuming symmetrically that they can leave the city to get the reservation utility by paying a moving cost. This modification would complicate the analysis somewhat without adding any important insight.

18 The Social security data cover the universe of workers and so are preferable to Company accounts. We defer the discussion of the results using Company accounts data to the robustness analysis.

elasticity of employment growth to city-size is also estimated to be negative and substantial, whereas it had positive impact on TFP growth. Holding sectoral composition and other determinants constant, doubling employment in a given city would reduce the growth rate by more than 1 percentage point per year over the subsequent period. We also find that the variety, size and competition indicators, which apparently have no direct effect on TFP growth, do significantly affect local employment. In particular, and similarly to what was found by Combes (2000), we estimate that the average impact of productive diversity in the city on subsequent employment growth of the manufacturing sectors is negative. Similarly to findings for France and US cities, we also estimated a positive partial elasticity of employment to average firm size.

Table 7

City-industry employment growth

	[1]	[2]	[3]	[4]	[5]
Specialization	-0.750***	-0.586***	-0.698***	-1.380***	-1.080***
	(0.210)	(0.209)	(0.201)	(0.267)	(0.525)
City size	-1.105***	-1.047***	-1.080***	-1.380***	-1.440***
	(0.127)	(0.128)	(0.140)	(0.192)	(0.354)
Firm size	0.648*	0.914**	0.803**	0.129***	-0.134
	(0.360)	(0.369)	(0.384)	(0.480)	(0.837)
Variety	0.828***	0.813***	0.319	0.472**	0.249
	(0.164)	(0.177)	(0.226)	(0.217)	(0.390)
Competition	-0.839***	-0.796***	-0.895***	-0.795***	-1.250**
	(0.155)	(0.160)	(0.159)	(0.235)	(0.485)
Spatial controls	5	20	95	5	5
Weights	YES	YES	YES	NO	YES
Data source	INPS	INPS	INPS	INPS	Nat. Acc.
No. of observations	1,602	1,602	1,602	1,602	1,602
R^2	0.43	0.47	0.53	0.32	0.16

Note: Dependent variable: annual employment growth rate at the L-S level. All regressions include sector dummies. Spatial controls are macro areas, regions and provinces. All regressions based on the sample of C-S continuously in the database.
*** indicates significance at 1 per cent, ** at 5 per cent and * at 10 per cent.

We ran several robustness checks, as we did for the TFP specification. We controlled for the existence of spatially correlated omitted variables (cols. [2] and [3]) and found that most of the previous results are unaffected, but the variety coefficient now becomes insignificant. We ran the unweighted regressions (col. [4]) and found no significant differences with respect to the initial regression. Finally, we also considered a version of our regression where the dependent variable had been obtained from the Company accounts sample, as opposed to the population (INPS social security data). This is a particularly interesting check of the representativeness of the Company accounts data and therefore of the generality of the results of the TFP regressions. The results are reported in the last column of Table 7. We find that the three coefficients that are significant (specialization, city size and competition) are very similar to those in column [1], a result that we interpret as evidence in favour of the representativeness of the Company accounts data. Instead, unlike the population based regressions (INPS data), the coefficients of firm size and of variety are not significantly different from zero. This, together with the fact that the R^2 is substantially lower in the Company accounts regression (0.16 against 0.43), suggests that resorting to a sub-sample introduces noise in the estimates and reduces their precision; however, there is no evidence of any systematic bias.

4.1 *Discussion*

Although they are in line with most of the urban-growth literature, our employment growth regression results are not in accordance with those of the TFP regressions. Given that these do not require any identifying assumption, our findings cast serious doubt on the use of employment changes as alternative indicators of dynamic externalities and suggest that the identification problem might go beyond the Italian case. In fact, following a different approach and using US data, Henderson (2003) also recognizes that employment growth regressions might be problematic. Studying only capital goods and high-tech industries, he estimates plant level production functions with fixed effects that include variables capturing both specialization and urbanization economies at the local level. He finds that for high-tech industry the level of current output at the plant level is positively affected by lagged indicators of specialization, whereas employment growth is not affected.[19] Henderson does not explicitly recognize the identification issue but calls into question allocative shocks, although by definition they should not be systematically correlated with local conditions. Similar conclusions are reached by Dekle (2002) for the Japanese prefectures, although he finds no evidence of dynamic externalities in manufacturing.[20] Taken together, this evidence calls into question the previous interpretations of employment growth regressions in terms of dynamic externalities.

Indeed, there are evident reasons to believe that the assumption that labour supply changes are independent of local conditions is generally not verified. In terms of the city-size effects, it seems reasonable that indicators of the quality of life, such as pollution, congestion, park lands and so on, should deteriorate more rapidly in highly urbanized areas. Moreover, these are superior goods, so that demand for them increases more than proportionally with income, again reducing labour supply in densely populated areas.[21] While it is not so obvious why sectoral indicators could impact on labour supply, Glaeser and Kahn (2001) showed that, across US cities, workers' preferences are important determinants of industry-level equilibrium employment.[22] The link between sectoral indicators and labour supply could be explained by a 'fishing off the pond' problem, which would dynamically reduce the number of workers in the local market who

---

19 Unlike ours, his approach, based on within estimation, disregards the cross sectional variability in the data and therefore only uses changes in the industrial structure variables, an approach that, while eliminating the possibility that the results are driven by some unobserved fixed factor, is vulnerable to unobserved innovations that drive changes in both the industrial structure and productivity. Our estimating approach also more carefully controls for endogeneity of inputs and for firm selection. Finally, his specification is not directly comparable with those of the urban growth literature, making comparisons with employment growth regressions less straightforward.

20 This might be due to the fact that his analysis, based on national accounts data, is constrained to fairly aggregated levels geographically (49 prefectures) and sectorally (manufacturing as a whole). In fact, spatial decay of localization externalities has been shown to be very fast (Jaffe *et al*., 1993; Audretsch and Feldman, 1994; Rosenthal and Strange, 2000); moreover the relative importance of intra-industry as opposed to cross-industry externalities might be difficult to disentangle using a one-digit breakdown of the entire economy (nine groups).

21 Chatterjee and Carlino (2001) construct a model in which agglomeration economies have a linear effect on productivity, while congestion diseconomies increase more than proportionally with city size. They show that their model matches the evolution of US cities in the post-war period, characterized by a decrease in the dispersion of employment density across cities.

22 In particular they show that workers' residential preferences are crucial in explaining productive decentralization at the industry level. After calculating the 'average' worker-type for each 3-digit SIC industry at the national level, they find that firms belonging to a specific industry are more likely to suburbanize in a given city, the more suburbanized are the types of workers the industry is likely to hire in that city.

are willing or have the appropriate skills to be employed in overrepresented sectors. Also, sectoral concentration might increase the bargaining power of workers, particularly increasing the level of unionization and thus curbing labour supply growth. Finally, the regulation of the labour market, especially in terms of legislation that limits the extent of wage differentials across locations, could impact on the way productivity changes at the local level are reflected in employment and wage changes.[23]

This last point suggests a final check, based on a wage growth equation. Our model yields the following estimating equation for the rate of change of average wages in the city-industry

$$\Delta \log w_i = \frac{\theta}{\gamma} \log X_i + \frac{\eta(1-\alpha)}{\gamma} \Delta \log Z_i \qquad (9)$$

which is also subject to the identification critique. Based on firm-level average annual compensation of employees available in INPS archives, we construct a measure of average per capita wage growth at the city-industry level and, in line with the previous analysis, run the following regression:

$$\hat{w}_{cs} = \beta_1 SPEC_{c,s} + \beta_2 VAR_{c,s} + \beta_3 COMP_{c,s} + \beta_4 SIZE_{c,s} + \beta_5 X_{c,s} + u_{c,s} \qquad (10)$$

Results are reported in Table 8, organized like Table 7. The first point is that the estimates are less precise than those for TFP and employment: the R^2 is lower (mostly due to the effect of the unreported initial wage level, negative and strongly significant) and the coefficients tend to be not very precisely estimated; moreover, when compared with the other tables the point estimates tend to be relatively smaller. With respect to the previous estimates we find that, as in the TFP regressions, city size has a positive and significant effect on wage growth and specialization a positive but generally insignificant one.[24] Instead, almost all coefficients have opposite signs with respect to the employment growth regressions. When interpreted in a labour demand and supply framework, the opposite

23 From a statistical point of view, the negative specialization coefficient might also signal mean reversion induced by random measurement error in the local employment data.

24 The coefficient of firm size indicates that larger firm size is associated with higher wage growth, an effect that is well established in the labour literature.

response of employment and wages would suggest that equilibrium outcomes are dominated by labour supply movements.

The joint findings of the three sets of regressions are compatible with an interpretation in which wage compression induced by centralized bargaining generates convergence of wages across space and reduces their responsiveness to local conditions. In this case, labour supply becomes an increasingly crucial determinant of equilibrium employment. So if, for example, a highly populated area becomes increasingly congested, the failure of wages to compensate for the loss in utility might induce people to move away, with little role for labour demand and therefore for productivity in determining employment levels. Stated more generally, the results suggest strongly that labour market outcomes depend on a variety of factors, not just productivity changes.

Table 8

City-industry wage growth

	[1]	[2]	[3]	[4]	[5]
Specialization	0.050	0.035	0.015	0.009	0.162**
	(0.037)	(0.034)	(0.033)	(0.039)	(0.070)
City size	0.098***	0.106***	0.097***	0.083***	0.288***
	(0.028)	(0.024)	(0.030)	(0.031)	(0.056)
Firm size	-0.230***	-0.292***	-0.333***	-0.151**	-0.060
	(0.068)	(0.062)	(0.062)	(0.063)	(0.108)
Variety	0.033	0.074**	0.024	0.015	-0.013
	(0.037)	(0.035)	(0.034)	(0.033)	(0.070)
Competition	-0.060*	-0.073**	-0.064*	0.023	-0.025
	(0.035)	(0.030)	(0.030)	(0.035)	(0.066)
Spatial controls	5	20	95	5	5
Weights	YES	YES	YES	NO	YES
Data source	INPS	INPS	INPS	INPS	Nat. Acc.
No. of observations	1,602	1,602	1,602	1,602	1,602
R^2	0.29	0.35	0.41	0.67	0.38

Note: Dependent variable; annual wage growth rate at the L-S level. All regressions include sector dummies. Spatial controls are macro areas, regions and provinces. All regressions based on the sample of C-S continuously in the database.
*** indicates significance at 1 per cent, ** at 5 per cent and * at 10 per cent.

Taken together, our results indicate that identification problems could be very serious in both employment and wage growth regressions and that, while interesting information on the reduced-form relation between local conditions and employment and wage changes can be obtained, not much can be said about productivity growth. All in all, we think that there are clear indications that only the direct measurement of TFP can identify dynamic local externalities within this framework.

5. Conclusions

Important empirical works estimating the strength of dynamic externalities at the local level have produced conflicting results. The evidence based on employment growth regressions requires that equilibrium employment determination should be demand-driven and changes in *labour supply* independent of local conditions. In this paper we show that, while useful for investigating the determinants of *industry growth* at the local level, employment growth regressions could be misleading if used to discriminate among different sources of *productivity growth* and dynamic externalities, because of serious identification problems. In fact, we find that TFP and employment growth regressions yield almost opposite results. In particular, TFP growth is enhanced by specialization and city size but not by urban diversity. We conclude that employment-based equations might not be able to disentangle the determinants of local industry growth from the sources of productivity growth.

For future work, it will be important to extend the TFP analysis from Italy to other countries to check whether, as seems likely from our discussion, our insights apply to other economies. Also it will be important to develop models that, by explicitly considering labour supply and mobility choices, allow for differential effects of the local structure on productivity on the one hand and employment on the other. This will help to give a structural interpretation of results of employment growth regressions in the previous literature and further guidelines for future empirical work.

APPENDIX

Production function estimation: data and procedure

This Appendix briefly summarizes the data and procedure adopted in the production function estimation. Firm-level variables are drawn from the Company Accounts Data Service (*Centrale dei Bilanci*), which has accounting data on a sample of between 30,000 and 40,000 Italian firms for the period 1982-98. Both value added and investments have been adjusted using the appropriate two-digit deflators, derived from Istat's National Accounts. The capital stock at firm level was obtained from the book value of investment using the permanent inventory method, accounting for sector specific depreciation rates from Istat's National Accounts data. The initial capital stock was estimated using the deflated book value, adjusted for the average age of capital estimated from the depreciation fund. We take care of outliers by excluding firms with value added per worker or value added per unit of capital in the first or last percentile of the distribution. This procedure improves the stability of the results without introducing systematic biases.

We use the estimation approach proposed by Olley and Pakes (1996). Production takes place through a Cobb-Douglas technology using capital and labour, with parameters α and β, subject to an unobserved (to the econometrician) productivity shock ω. In logs, the production function is

$$y_t = \alpha k_t + \beta l_t + \omega_{t+} + \eta_t \tag{11}$$

where η is a random shock uncorrelated with the other variables. For simplicity, the theoretical model assumes that capital is irreversible (the estimation method works independently of this assumption); moreover, capital is a predetermined variable at t so that it is independent from ω_t, while labour can adjust to the productivity shock. The firm also decides whether to continue production or shut down, in which case it collects a salvage value Φ. The dynamic programming problem of the firm is represented by the Bellman equation:

$$V(k_t, \omega_t = \max\left\{\Phi, \max_{i_t, l_t}\left[\pi(k_t, l_t, \omega_t) - c(i_i) + E(V(k_{t+1}, \omega_{t+1}))\right]\right\} \tag{12}$$

$$\text{s.t. } k_{t+1} = (1-\delta)k_t + i_t, F_\omega(w_t + 1|\omega_t) \tag{13}$$

where π is current profit, $c(.)$ is the cost of investment and $F\omega(\omega_{t+1} / \omega_t)$ is the probability distribution of ω_{t+1} given ω_t , assumed to be stochastically increasing. The dynamic programming problem delivers three policy functions: a continuation function $\chi(k_t, \omega_t) = \{0,1\}$, an investment function $i(k_t, \omega_t) \geq 0$ and an employment function $l(k_t, \omega_t)$. The continuation decision takes the form of a threshold value $\underline{\omega}(k)$ for the productivity shock below which it is optimal to exit.

The continuation decision and the input choices depend on the capital stock and the unobservable productivity shock. This implies that OLS estimation of (11) has two sources of bias. First, the labour input is correlated with ω; second, it can be shown that $\underline{\omega}(k)$ is decreasing in k, which induces a selection issue: the higher the capital stock the more likely it is that firms will remain in the market even with low realizations of ω. This implies that if selection is not accounted for the capital coefficient will be downward-biased, because of the negative correlation between ω and k.

Olley and Pakes propose a procedure to correct for both biases. For the simultaneity bias they approximate the unobservable ω with a non-parametric function of investment and current capital stock. In fact, the investment function is invertible, so that there exists a function relating the productivity shock to the stock of capital and investment:

$$\omega_t = h(i_t, k_t) \tag{14}$$

As the shape of $h(.)$ depends on the functional forms of the primitives and in general has no analytical representation, it is approximated by a polynomial series in i and k. The coefficient of the labour input is therefore consistently estimated by OLS on:

$$y_t = \beta l_t + \phi(i_t, k_t) + \eta_t \tag{15}$$

where

$$\phi(i,k) = \alpha k + h(i,k) \tag{16}$$

define the estimated value $\widehat{\phi} = y - \widehat{\beta} l - \widehat{\eta}$.

To estimate the capital coefficient we need to account for selection. In a first step we estimate a probability of survival as a function of (i_t, k_t) via a probit estimation of the continuation decision in a power series of i and k. Define the estimated probability as $\widehat{P}$. We can now introduce a Heckman-type correction in the estimation of the capital coefficient. In fact,

$$E(y_{t+1} - \beta l_{t+1} | k_{t+1}, \chi_{t+1} = 1) = \beta \kappa_{t+1} + E(\omega_{t+1} | \chi_{t+1} = 1, \omega_t) \tag{17}$$

Using the definition of conditional expectation and (14), it can be shown that the conditional expectation of ω_{t+1} can be expressed as a function of P and h, say $g(P,h)$. Using (16), the estimating equation therefore becomes

$$y_{t+1} - \beta l_{t+1} = \beta k + g(\hat{P}, \hat{\phi} - \alpha k) + \xi_{t+1} + \eta_{t+1} \tag{18}$$

where ξ is the innovation in ω. The last step therefore requires the non-linear estimation of equation (18), where the unknown function g is replaced by a power series in $\widehat{P}$ and $\widehat{\phi} - ak$.

We implement the procedure using polynomial approximations of the fourth degree in all stages to approximate h, P and g. Results are stable when going from a third to a fourth degree, an indication that the polynomial approximations are sufficiently accurate. The simultaneity bias does not greatly affect the estimation of the labour coefficient, while selection is very important for the capital coefficient. This is the same pattern observed by Olley and Pakes with data from the telecommunications equipment industry in the US.

REFERENCES

Arrow, K. (1962), "The Economic Implications of Learning by Doing", *Review of Economic Studies*, Vol. 29, pp. 155-173.

Audretsch, D. and Feldman, R. (1994), "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review*, Vol. 86, pp. 630-640.

Bernard, A.B. and Jones, C. (1996), "Comparing Apples to Oranges: Productivity Convergence and Measurement across Industries and Countries", *American Economic Review*, Vol. 86, pp. 1216-1238.

Bradley, R. and Gans, J. (1996), "Growth in Australian Cities", *The University of South Wales Working Papers*, Vol. 13, pp. 329-355.

Cainelli, G. and Leoncini, R. (1999), "Externalities and Long-Term Local Industrial Development. Some Empirical Evidence from Italy", *Revue d'Economie Industrielle,* Vol. 90*,* pp. 25-40.

Centrale dei Bilanci (1992), *Economia e finanza delle imprese italiane 1982-1990: quinto rapporto*, Centrale dei Bilanci.

Chatterjee, S. and Carlino, G.A. (2001), "Aggregate Metropolitan Employment Growth and the Deconcentration of Metropolitan Employment", *Journal of Monetary Economics*, Vol. 48, pp. 549-583.

Ciccone, A. and Hall, R.E. (1996), "Productivity and the Density of Economic Activity", *American Economic Review*, Vol. 86, pp. 54-70.

Combes, P. (1999), "Marshall-Arrow-Romer Externalities and City Growth: A Methodological Note", CERAS, Working Paper No. 6.

Combes, P. (2000), "Economic Structure and Local Growth: France, 1984-1993", *Journal of Urban Economies* , Vol. 47, pp. 329-355.

Dekle, R. (2002), "Industrial Concentration and Regional Growth: Evidence from the Prefectures", *Review of Economics Studies*, Vol. 84, pp. 310-315.

Duranton, G. and Puga, D. (2001), "Nursery Cities", *American Economic Review*, Vol. 91, pp. 1454-1475.

Eaton, J. and Eckstein, Z. (1997), "Cities and Growth, Theory and Evidence from France and Japan", *Regional Science and Urban Economics,* Vol. 27*,* pp. 443-474.

Glaeser, E.L. and Kahn, M.E. (2001), "Decentralized Employment and the Transformation of the American City", NBER, Working Paper, No. 8117.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A. and Shleifer, A. (1992), "Growth in Cities", *Journal of Political Economy*, Vol. 100, pp. 1125-1152.

Glaeser, E.L., Scheinkman, J.A. and Shleifer, A. (1995), "Economic Growth in a Cross Section of Cities", *Journal of Monetary Economics*, Vol. 36, pp. 117-143.

Guiso, L. and Parigi, G. (1999), "Investment and Demand Uncertainty", *Quarterly Journal of Economics,* Vol. 114*,* pp. 185-227.

Guiso, L. and Schivardi, F. (2000), "Information Spillovers and Factor Adjustment", Banca d'Italia, Temi di discussione, No. 368.

Henderson, V. (1974), "The Sizes and Types of Cities", *American Economic Review*, Vol. 64, pp. 640-656.

Henderson, V. (2003), "Marshall's Scale Economies, *Journal of Urban Economics*, Vol. 53, pp. 1-28.

Henderson, V., Kuncoro, A. and Turner, M. (1995), "Industrial Development in Cities", *Journal of Political Economy*, Vol. 103, pp. 1067-1090.

Jacobs, J. (1969), *The Economy of Cities*, Vintage, New York.

Jaffe, A., Trajtenberg, M. and Henderson, R., (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, Vol. 108, pp. 577-598.

Marshall, A. (1890), *Principles of Economics*, London, Macmillan.

Olley, S. and Pakes, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica*, Vol. 64, pp. 1263-1297.

Pagano, M., Panetta, F. and Zingales, L. (1998), "Why Do Companies Go Public? An Empirical Analysis", *Journal of Finance*, Vol. 53, pp. 27-64.

Pagano, P. and Schivardi, F. (2001), "Firm Size Distribution and Growth", Banca d'Italia, Temi di discussione, No. 394.

Porter, M. (1990), *The Competitive Advantage of Nations*, Free Press, New York.

Romer, P. (1986), "Increasing Returns and Long Run Growth", *Journal of Political Economy*, Vol. 98, pp. S71-S102.

Rosenthal, S. and Strange, W. (2000), "Geography, Industrial Organization and Agglomeration" UBC, mimeo.

Sapienza, P. (2002), "The Effects of Banking Mergers on Loan Contracts", *Journal of Finance*, Vol. 57, pp. 329-367.

Segal, D. (1976), "Are There Returns to Scale in City Size?", *Review of Economics and Statistics,* Vol. 58, pp. 217-235.

Shin, K. (1997), "Inter- and Intra-sectoral Shocks: Effects on the Unemployment Rate", *Journal of Labour Economics*, Vol. 13, pp. 93-109.

Sveikauskas, K. (1975), "Productivity of Cities", *Quarterly Journal of Economics*, Vol. 89, pp. 393-413.