

Manufacturing Growth and Agglomeration Effects*

Marcel Fafchamps

Oxford University

June 2004

Abstract

This paper investigates the effect of location-specific competition and diversity on manufacturing growth. We find strong and robust evidence of agglomeration effects: competition is good for growth but diversity is not. However, none that the effects are due to productivity or wages. First, agglomeration variables have on growth the opposite effect than on individual firm productivity. Second, controlling for productivity directly does not reduce the significance or magnitude of agglomeration variables. Agglomeration variables measure something that is relevant for growth, but it is not productivity. We also find that a rise in productivity raises subsequent employment and investment, but has no effect on firm entry and exit.

*I am very grateful to the World Bank for financial assistance. I thank Peter Lanjouw for his constant support and Said El Hamine for assisting in preparing the data for analysis.

1. Introduction

Since Marshall, agglomeration externalities have long attracted the attention of economists (e.g. Henderson 1988, Fujita, Krugman and Venables 1999) and geographers alike (e.g. Isard 1956, Jacobs 1969, Dicken and Lloyd 1990). Various sources of externalities have been hypothesized in the literature. Some are thought to raise the productivity of individual firms directly, for instance through the sharing of technological or market-related information (e.g. Arrow 1962, Glaeser, Kallal, Sheinkman and Shleifer 1992). Others are believed to raise profits by reducing transport costs, for example because of closer proximity to consumers and input providers (e.g. Krugman 1991, Rodriguez-Clare 1996). The first case corresponds to Marshallian externalities, the second to pecuniary externalities.

Much of the empirical literature on agglomeration externalities focuses on employment growth (e.g. Glaeser et al. 1992, Ellison and Glaeser 1997, Henderson 1997, Combes 2000, Bun and El Makhoulfi 2004), with the exception of Combes, Magnac and Robin (2004) who also look at firm entry and exit. In order to disentangle pure locational advantages from agglomeration effects, the literature has relied on dynamic panel analysis whereby sectoral employment growth is regressed on proxy variables capturing agglomeration externalities that vary across locations. Panel analysis offers the advantage of controlling for time-invariant location effects, such as geographical advantage. Using this approach, agglomeration effects have been shown to be a strong determinant of employment growth (e.g. Glaeser et al. 1992, Henderson 1997, Combes 2000, Bun and El Makhoulfi 2004) and, more recently, of firm entry (Combes et al. 2004). All these studies assume that agglomeration factors affect manufacturing performance through their effect on productivity, prices, and costs. But, to our knowledge, this has never been tested formally.

This paper examines how the introduction of productivity and wage shock measures in dynamic firm growth and entry regressions affects the coefficients of agglomeration variables.

Productivity and wage shock measures are constructed from a large exhaustive panel dataset on individual manufacturing firms at a disaggregated geographical level. Fafchamps and El Hamine (2004) have shown that agglomeration effects affect the total factor productivity of individual firms and the wages they pay their employees. If variables supposed to proxy for agglomeration externalities influence firm growth and entry via productivity and wages, the inclusion of direct – albeit imperfect – measures of productivity and wages should reduce the magnitude of their coefficient. Results contradict this conjecture: although productivity and wage shocks are shown to have a strong effect on employment growth and firm entry, this effect is quite distinct from standard agglomeration variables. We also find that agglomeration variables have an effect on employment growth and firm entry that is quite different from the effect they have on firm productivity and wages. It therefore appears that agglomeration variables in dynamic employment growth regressions measure something else than productivity or wages.

This paper is organized as follows. Our testing strategy is described in Section 2, in relation to the existing literature. The data are presented in Section 3, together with some descriptive statistics. Econometric analysis is presented in Section 4.

2. Testing strategy

There is a large literature on agglomeration externalities, much of it focusing on manufacturing. The theoretical literature is particularly well developed and has identified many different types of agglomeration effects, some negative (e.g., congestion), some positive (e.g., shared infrastructure). Alfred Marshall, subsequently followed by Arrow (1962) and Romer (1990), identified knowledge spillover as an important source of externalities. To the extent that knowledge is transferred more easily through direct human contact, local information sharing is thought to give rise to agglomeration externalities of the 'Silicon Valley' type. The shared information need

not be on technology; it may also include business opportunities or market relevant knowledge (e.g. Rauch and Casella 2003, Fafchamps, El Hamine and Zeufack 2002).

Different views on what shared information is relevant and how it is exchanged have given rise to different theories regarding the nature of agglomeration effects. One view, attributed to Marshall, Arrow and Romer and hence referred to as the MAR hypothesis by Combes (2000), claims that monopoly and market power are associated with more innovation and hence with larger externalities. The opposite view is championed by Porter (1990) who argues that monopolies are stultifying and that it is competition that spurs innovation and growth. Both these hypotheses are seen as emphasizing externalities within a sector. In contrast, Jacobs (1984) argues that it is the diversity of industries within cities that is a source of externalities, as industries borrow ideas from each other. The empirical evidence is contradictory. Using US data, Glaeser et al. (1992) find in general that local competition and urban diversity, but not specialization, encourage employment growth. In contrast, Henderson (1997) finds that both specialization and diversity have positive effects on firm growth but that the former is larger. Using French data, Combes (2000) finds the opposite result that competition and specialization reduce employment growth while diversity is negative for most industries and positive for services.

Pecuniary externalities have also been proposed as possible explanation for spatial concentration (e.g. Henderson 1988, Fujita et al. 1999). For instance, in a large labor market, it is easier and faster for employers to find the specialized manpower they need. This phenomenon is called thick labor market externalities by Glaeser et al. (1992). Forward and backward linkages as initially proposed by Hirschman (1958) are another possibility. Rodriguez-Clare (1996), for instance, construct a model where a larger market triggers entry in intermediate input production, thereby generating gains from specialization (see also (e.g. Ciccone and Matsuyama 1996, Fafchamps and Helms 1996, Fafchamps 1997)). Market size also matters.

Krugman (1991), for instance, illustrates how proximity to larger market may attract industries if transport costs are neither too high nor too low. In this paper, we examine both types of externalities.

The empirical literature on externalities and industrial development remains unsettled (Tybout 2000). Glaeser et al. (1992), for instance, conclude that competition and diversity favor firm growth. In contrast, Henderson (1997) and Desmet and Fafchamps (2005) conclude that own-sector externalities are much stronger than those generated by other sectors. In his study of French manufacturing and services, Combes (2000) concludes that competition and total local employment have a negative effect on firm growth while Bun and El Makhoulfi (2004) concludes that diversity has a positive effect but competition a negative one. Using a different methodology, Ciccone and Hall (1996) find that employment density increases average labor productivity.

Combes et al. (2004) argue that contradictory results may be driven by slight differences in methodology. They insist that a consistent set of regressors needs to be used to obtain meaningful results. Using a set of agglomeration variables similar to that of Combes et al. (2004), Fafchamps and El Hamine (2004) use firm-level data to provide evidence of a significant effect on productivity and wages. They find that returns to specialization are strong and large in magnitude and that the net effect of competition on productivity and wages tends to be negative. Their analysis shows that competition tends to lower wages, probably because of thick labor market externalities. They also find some limited evidence in favor of the diversity argument put forth by Jacobs (1984). Thompson (2004) shows that input-output linkages matter for agglomeration externalities.

In this paper we seek to understand the growth of manufacturing over time. Probably due to data limitations, earlier papers have followed Glaeser et al. (1992) and focused on changes in employment levels over time as measure of manufacturing growth. Here we take advantage of

richer data to include not only employment but also total output, investment, and changes in the number of firms.

At the heart of most analyses of sectoral dynamics is the idea that a variable of interest y_{it} – typically the level of economic activity in a given country or location i – tends towards a steady state y_i^* . If we linearize the law of motion of y around its steady state or mean, we obtain a linear difference equation:

$$\Delta y_{it} = \rho(y_i^* - y_{it}) \quad (2.1)$$

Equation (2.1) implies that growth Δy_{it} is faster the further away y_{it} is from y^* . This is common sense: if y_{it} is converging towards y^* , it must eventually slow down as it reaches y^* . Parameter ρ expresses the speed with which y_{it} converges to its steady state: if $\rho = 1$, convergence is instantaneous; if ρ is small but positive, convergence is slow; if $\rho < 0$, y_{it} does not move towards y_i^* but in fact moves away from it.

Following Quah (1993) and Desmet and Fafchamps (2003), it is straightforward to extend the above model to allow for stochastic shocks v_{it} , in which case we have:

$$\Delta y_{it} = \rho(y_i^* - y_{it}) + v_{it} \quad (2.2)$$

In this context, y_i^* can be thought of as y_{it} 's conditional mean and λ as the speed at which y_{it} reverts to its mean.¹

In equation (2.2), it is customary to assume that y_i^* depends on specific conditions z_i prevailing in i , i.e. to posit that $y_i^* = f(z_i)$. We can thus write:

$$\Delta y_{it} = \rho(f(z_i) - y_{it}) + v_{it} \quad (2.3)$$

¹Of course, if all countries or locations begin well below their conditional mean, most will be seen to revert to their mean from below, that is, most will grow.

In general, economists are interested not so much in growth itself but in $f(z_i)$ because it is indicative of the long run behavior of y_{it} . By linking the two, equation (2.3) enables researchers to infer something about $f(z_i)$ from the speed of growth: conditioning on y_{it} , equation (2.3) indeed predicts that variables z_i that yield a higher steady state y_i^* also increase the growth rate Δy_{it} . This yields a testing strategy: regress growth Δy_{it} on initial condition y_{it} and a set of variables z_i thought to affect steady state y_i^* ; if they are seen to speed growth, they should also raise y_i^* , and vice versa. The same reasoning can be extended to time-varying factors z_{it} . In this case, y_{it} can be thought of as following a moving target. As the target moves further away, y_{it} must speed up in order to catch up with it.

The above ideas form the basis of our testing strategy: if variables measuring agglomeration effects speed up growth, this is seen as evidence that they generate positive feedbacks raising y_i^* .² To illustrate how this works, let Q_{ijt} be total manufacturing output in location i and sector j at time t . By definition we have $Q_{ijt} = \sum_{k \in I_{ijt}} Q_k$ where k is an individual firm index and I_{ijt} is the set of firms present in location i and sector j at time t . We wish to know whether Q_{ijt} converges to a steady state or conditional mean that is affected by agglomeration effects. Our core regression is of the form:

$$\log \frac{Q_{ijt+1}}{Q_{ijt}} = \alpha_{ij} + \beta A_{ijt} + \gamma P_{ijt} - \rho \log L_{ijt} + \tau_{jt} + u_{ijt} \quad (2.4)$$

where α_{ij} is a location-sector fixed effect, A_{ijt} is a vector of agglomeration variables, P_{ijt} is a set of direct productivity and cost measures, τ_{jt} is a sector-specific time dummy, and u_{ijt} is a residual. In our earlier notation, $y_{it} = \log Q_{ijt}$, $\rho f(z_{it}) = \alpha_{ij} + \beta A_{ijt} + \gamma P_{ijt}$, and $\tau_{jt} + u_{ijt} = v_{it}$. Location and sector specific fixed effects α_{ij} control for all time-invariant factors such as geographical

²This approach is but one way of studying agglomeration effects. Fafchamps and El Hamine (2004), for instance, examine the effect of agglomeration forces directly on firm productivity and wages.

location, proximity to borders, etc. Sector-specific time dummies control for all shocks that are common across sectors, such as changes in interest rate or exchange rate. They also control for technological change. Agglomeration effects are thus identified by variations in A_{ijt} and P_{ijt} over time, relative to their national average. This ensures that we do not erroneously attribute to agglomeration effects what is in fact due to unobserved heterogeneity across locations.

For estimation purposes, it is customary to rewrite equation (2.4) in the form of a dynamic panel regression:

$$\log Q_{ijt+1} = \alpha_{ij} + (1 - \rho) \log Q_{ijt} + \beta A_{ijt} + \gamma P_{ijt} + \tau_{jt} + u_{ijt} \quad (2.5)$$

In the transformed regression, the coefficient of L_{ijt} , which is supposed to be smaller than 1, represents the speed of adjustment: the smaller it is, the faster adjustment is. Given the presence of fixed effects α_{ij} , estimation of (2.5) by OLS is known to generate inconsistent estimates. To deal with the difficulty, a number of alternative instrumental variable and GMM estimators have been proposed in the literature.³ Differencing the data to eliminate α_{ij} , (2.5) can be written:

$$\Delta \log Q_{ijt+1} = (1 - \rho) \Delta \log Q_{ijt} + \beta \Delta A_{ijt} + \gamma \Delta P_{ijt} + \Delta \tau_{jt} + \Delta u_{ijt} \quad (2.6)$$

GMM estimators for (2.6) rely lagged values of $\log L_{ijt}$ to instrument $\Delta \log L_{ijt}$ (e.g. Anderson and Hsiao 1982, Arellano and Bond 1991). This is, for instance, the approach adopted by Combes (2000) and Combes et al. (2004).⁴

We are also interested in the channel through which agglomeration effects influence output.

We focus on three possible channels: capital investment, employment, and firm entry and exit.

³See for instance Arellano (2003) and Arellano and Honoré (2001) for recent summaries of the literature.

⁴Combes et al. (2004) estimate a version of (2.5) where the dependent variable is L_{ijt}/N_{ijt} , i.e., employment per firm. Given that the model is estimated in logs and that N_{ijt} appears as regressor (see *infra*), their models is basically equivalent to ours in that respect.

Manufacturing growth can occur through the expansion of existing firms or through an increase in the number of firms N_{ijt} . As Combes et al. (2004) have shown, which of these two avenues dominates depends on demand elasticity and on the nature of competition. Given that agglomeration externalities are at least in part due to competition, we suspect that agglomeration variables may have a different effect on firm expansion than on firm entry. Let L_{ijt} , K_{ijt} and N_{ijt} denote total employment, capital stock, and number of firms in location i and sector j at time t , respectively. We begin by estimating a model of the form:

$$\begin{aligned} \Delta \log Q_{ijt+1} = & (1 - \rho) \Delta \log Q_{ijt} + \beta \Delta A_{ijt} + \gamma \Delta P_{ijt} \\ & + \theta_l \Delta \log L_{ijt} + \theta_k \Delta \log K_{ijt} + \theta_n \Delta \log N_{ijt} + \Delta \tau_{jt} + \Delta u_{ijt} \end{aligned} \quad (2.7)$$

Comparing estimates of parameter vectors β and γ between equations (2.6) and (2.7) should yield the first hints on how agglomeration and productivity variables affect output growth, i.e., either directly through prices and total factor productivity, or indirectly through investment, hiring, and firm entry and exit.

We also examine how employment growth, investment, and firm entry and exit respond to A_{ijt} and P_{ijt} . Our regressions are of the form:

$$\Delta \log X_{ijt+1} = \varphi_1^x \Delta \log L_{ijt} + \varphi_2^x \Delta \log K_{ijt} + \varphi_3^x \Delta \log N_{ijt} + \varphi_4^x \Delta A_{ijt} + \varphi_5^x \Delta P_{ijt} + \Delta \tau_{jt}^x + \Delta u_{ijt}^x$$

where X stands for L, K , and N , respectively. Since L_{ijt} , K_{ijt} and N_{ijt} feed back into each other, they are to be regarded as pre-determined variables. Consequently, their first differences are all instrumented with lagged levels of L_{ijt} , K_{ijt} and N_{ijt} . We further refine this approach by decomposing net firm entry into sub-components. A net increase in N_{ijt} requires that gross entries N_{ijt}^e and firm in-migration N_{ijt}^n from other locations exceeds gross exit $N_{ijt-1}^x < 0$ and

firm out-migration to other locations $N_{ijt-1}^o < 0$:

$$\Delta N_{ijt} = N_{ijt}^e + N_{ijt}^n + N_{ijt-1}^x + N_{ijt-1}^o$$

To investigate whether agglomeration externalities and productivity shocks affect entry, exit, and firm relocation differently, we also estimate models of the form:

$$\begin{aligned} \log(N_{ijt+1}^z + N_{ijt+1}) - \log N_{ijt} = & \varphi_1^x \Delta \log L_{ijt} + \varphi_2^x \Delta \log K_{ijt} + \varphi_3^x \Delta \log N_{ijt} \\ & + \varphi_4^x \Delta A_{ijt} + \varphi_5^x \Delta P_{ijt} + \Delta \tau_{jt}^x + \Delta u_{ijt}^x \end{aligned} \quad (2.8)$$

for $z = \{e, n, x, o\}$. This formulation offers the advantage that all estimated coefficients are expressed in terms of their effect on the growth rate of the number of firms. Since, strictly speaking,, equation (2.8) does not include a lagged dependent variable the GMM estimator of Arellano and Bond does not apply. We nevertheless worry that L_{ijt} , K_{ijt} , and N_{ijt} may be correlated with the fixed effect because. Consequently, $\Delta \log L_{ijt}$, $\Delta \log K_{ijt}$ and $\Delta \log N_{ijt}$ are instrumented using lagged levels when estimating (2.8).

We now turn to a description of our regressors. We follow Fafchamps and El Hamine (2004) and identify four variables measuring agglomeration effects A_{ijt} :⁵ (1) total employment in the location $L_{it} = \sum_j L_{ijt}$; (2) the total number of sectors present in the location M_{it} ; (3) a diversity index D_{it}^* defined as

$$D_{it}^* = \frac{1}{\sum_{j \in \Gamma_{it}} \left(\frac{L_{ijt}}{L_{it}} \right)^2}$$

where Γ_{it} is the set of sectors present in location i at time t ; and (4) a competition index C_{ijt}^*

⁵Combes et al. (2004) also regard L_{ijt} and N_{ijt} as capturing agglomeration effects. We are not comfortable with this interpretation (see below).

defined as

$$C_{ijt}^* = \frac{1}{\sum_{k \in \Gamma_{ijt}} \left(\frac{L_k}{L_{ijt}} \right)^2}$$

Both D_{it}^* and C_{ijt}^* are Herfindahl indices. Complete concentration in a single sector (D_{it}^*) or firm (C_{ijt}^*) yields a value of 1. In contrast, if employment is equally shared among sectors, the diversity index becomes:

$$\begin{aligned} D_{it}^* &= \frac{1}{\sum_{j \in \Gamma_{it}} \left(\frac{L_{ijt}/M_{it}}{L_{it}} \right)^2} \\ &= \frac{1}{\sum_{j \in \Gamma_{it}} \left(\frac{1}{M_{it}} \right)^2} \\ &= M_{it} \end{aligned}$$

By the same token, when all firms are of equal size, $C_{ijt}^* = N_{ijt}$. To facilitate interpretation, we normalize D_{it}^* and C_{ijt}^* as follows:

$$\begin{aligned} D_{it} &= \frac{D_{it}^*}{M_{it}} \\ C_{ijt} &= \frac{C_{ijt}^*}{N_{ijt}} \end{aligned}$$

Normalized indices vary between 0 (most concentrated) and 1 (least concentrated).

Several of the above variables have been used in one form or another in the literature before, typically in log form. For instance, own sector employment $\log L_{ijt}$ is referred to by Henderson (2003) as a localization effect while $\log L_{it}$ is said to capture urbanization effects. Sometimes similar variables are given a different interpretation. Henderson (2003), for instance, uses N_{ijt} as the number of sources of local information spillover while Combes et al. (2004) regard N_{ijt} alone as a measure of competition. In the work of Glaeser et al. (1992), it is the (log of the) ratio N_{ijt}/L_{ijt} that is used as a measure of competition. The likely reason for these discrepancies is

differences in data availability: authors with different types of data end up using different sets of agglomeration variables.

We are not comfortable interpreting the coefficient of $\log L_{ijt}$ as measure of agglomeration effects. The reason is that firms would typically grow even in the absence of agglomeration externalities. Summing over all firms would generate a relationship between $\Delta \log L_{ijt}$ and $\log L_{ijt}$ even though agglomeration effects are absent. The same reasoning also applies to firm entry and exit: firms would enter and exit even in the absence of agglomeration effects. To identify the agglomeration effect of L_{ijt} and N_{ijt} , one needs firm-level data so that it is possible to distinguish between factors that are internal and external to individual firms.

The meaning of each variable depends on the presence or absence of the others: variables can only be interpreted in conjunction with each other. Assuming that all variables enter in logs, L_{it} captures the agglomeration effect due to the presence of a large manufacturing sector L_{it} . The effect of specialization L_{ijt}/L_{it} is captured in the coefficient of L_{ijt} , together with growth factors that are internal to firms. Competition is captured by C_{ijt} which, after normalization, can be interpreted independently from N_{ijt} . If competition generates positive agglomeration externalities, as suggested by Porter, then C_{ijt} should have a positive effect on productivity and hence on firm growth.

Diversity is captured by two variables, M_{it} and D_{it} . Since we are conditioning on sectoral specialization through L_{ijt}/L_{it} , the variables M_{it} and D_{it} measure the effect of diversity in sectors other than j . For a given level of specialization, firm performance may increase with the diversity of production in sectors other than the firm's own sector j . It is this effect that variables M_{it} and D_{it} seek to capture. If diversity in other sectors is good for manufacturing firms, then we expect both M_{it} and D_{it} to have a positive effect on firm performance. Comparison of their effects can tell us whether it is the mere presence of a sector that matters or whether it is the

equal distribution of employment across sectors.

As we pointed out earlier, agglomeration variables A_{ijt} are meant to proxy for productivity effects due to location externalities (e.g. Glaeser et al. 1992, Henderson 1997, Combes 2000, Combes et al. 2004, Bun and El Makhoulfi 2004). In the context of this literature, productivity should be understood in a broad sense: it encompasses the effect that Marshallian and pecuniary externalities can have on the joint productivity of all firms in a given location. Marshallian externalities raise total factor productivity directly while pecuniary externalities raise prices and/or lower wage and intermediate input costs. Fafchamps and El Hamine (2004) have shown that both total factor productivity and wages are strongly influenced by agglomeration effects. Consequently, let our P_{ijt} vector include measures of total factor productivity as well as factor costs.

We estimate all models with and without P_{ijt} variables. Intuitively, if agglomeration variables capture productivity effects, then the inclusion of direct productivity measures P_{ijt} should set the coefficients of agglomeration variables to 0. Of course, the validity of this test rests on the assumption that P_{ijt} is measured without error. If measurement error is present, we would expect some of the productivity effects to be captured by A_{ijt} variables, in which case their coefficient may remain significantly different from 0. But even in this case, we expect β to fall in absolute value.

Before turning to the empirical analysis, we need to recognize its limitations. Although the literature often describes agglomeration effects as Marshallian or pecuniary externalities, nothing in the analysis reported here enables us to distinguish between externalities and any general equilibrium effects that is location specific. While the concept of Marshallian or technological externality is well defined in our context – it raises or lowers total factor productivity – the concept of pecuniary externality is not clearly distinguished from other general equilibrium

effects.⁶ Ultimately, it is a matter of semantics whether we want to call agglomeration effects an externality or not. With this caveat in mind, we now turn to the data.

3. The data

To implement the above testing strategy, we use manufacturing census data from Morocco. The data were collected by the Moroccan Ministry of Commerce and Industry every year over the period 1985 to 2001. Coverage is universal and includes all manufacturing enterprises in all sectors and all parts of the country. Given that answering the annual census questionnaire is a legal obligation, the rate of non-answer is fairly small – 12% over the entire period.⁷ These data have already been studied by others. The first years of this data set have been used by Clerides, Lach and Tybout (1998) to examine export behavior. The relationship between exports and productivity is also studied by Fafchamps et al. (2002). Fafchamps and El Hamine (2004) test the effect of agglomeration externalities on total factor productivity and wages.

The sectoral decomposition identifies 17 different sectors corresponding roughly to the 2-digit ISIC classification. Because 3 of the sectors have very few firms, for the sake of the analysis we combine them with other similar sectors, bringing the total number of sectors to 14. Data is available for all years on output, employment, wage payments, investment, and disbursed capital (a balance sheet equity concept). Capital stock information is available for the year 2001. Employment figures are separated into permanent and casual workers, the latter figure being given in total number of days per year. We divide the number of man-days by

⁶In the theoretical literature, the term pecuniary externality is sometimes used to describe situations in which market interactions generate multiple Pareto ranked equilibria (e.g. Murphy, Shleifer and Vishny 1989, Ciccone and Matsuyama 1996). Elsewhere (e.g. Romer 1986, Rodriguez-Clare 1996, Fafchamps 1997), pecuniary externalities describe multiplier effects. Our empirical analysis can neither identify multiple equilibria nor distinguish multiplier effects that arise from normal general equilibrium feedbacks from those that arise from pecuniary externalities.

⁷For the purpose of generating national and regional statistics, the Ministry of Commerce and Industry imputed values individually for each non-responding firm. Imputation was typically done using previous year information. Imputed firms are ignored in the regression analysis but to minimize measurement error imputed employment figures for non-respondent firms are used in computing the agglomeration variables described in the previous section.

256 to transform man-days of casual labor into permanent employee equivalent.⁸ To facilitate comparison, we deflate all output figures using sector-specific GDP price deflator. Investment data is deflated using the price index for machinery.⁹

Location information varies over time. From 1985 until 1993, the manufacturing census only recorded the province in which the firm was located. This period correspond to a trade liberalization phase (Haddad and de Melo 1996). From 1994 until 1997, the data also recorded the city code and from 1998 on the precise commune location of each firm was recorded. Morocco is divided into 70 provinces, 67 of which count at least one manufacturing firm over the study period. Starting from 1993, the data distinguishes between 242 cities. From 1998, firm location data is available at the commune level. There are approximately 1300 communes in Morocco, 689 of which had at least one manufacturing firm over the study period.

These data are used to construct three sets of location and sector specific variables: at the commune, city, and province level. Commune and city aggregates can be computed from 1998 until 2001 and from 1994 until 2001, respectively. Province aggregates can be computed for the whole span of the data, that is, from 1985 until 2001. Summary statistics are presented in Table 2. The table is a way that mirrors the subsequent analysis: each observation corresponds to a sector, location, and year with at least one active firm. Put differently, locations with no manufacturing are omitted from the table. The three panels correspond to the varying level of geographical detail available in the census data.

Average sectoral employment at the commune level is around 360 workers. The median is much lower at 53. Total manufacturing employment in the commune is a little over 4000 on average. There are on average around 5 firms in each sector and commune, with a smaller median

⁸This corresponds to $365 - 52 \times 2$ (week-ends) - 6 (public holidays).

⁹Since we include sector-specific year dummies in all regressions, deflating is not really an issue. But it matters for back-predicted capital stock, as explained below.

of 2. The corresponding value of the (unnormalized) competition index C_{ijt}^* is 2.7, hence falling roughly between 1, which corresponds to complete concentration, and N_{ijt} which correspond to complete equality conditional on N_{ijt} . Around 8 of the 14 sectors are present in the average commune. The average (unnormalized) diversity index D_{it}^* is 3.3, which similarly falls between 1 and M_{it} . Entry, exit, and movement across locations are presented next. We see that the number of entering and exiting firms are roughly of the same order of magnitude. This is consistent with the relative stagnation of Moroccan manufacturing in the late 1990's. The exit rate is high: the data suggest that, in any given year, some 10% of all firms exit. This reflects the small nature of many of the firms in the firm population. Small firms are indeed known to have a higher churning rate (e.g. Barrett 1997, Daniels 1994). We also see that many firms move across locations. By the nature of the data, no information on firms exiting or moving out is available for the last year, hence the smaller number of observations. As suggested by the large difference between the mean and the median, investment I_{ijt} is highly skewed. This is normal given the predominance of small firms in the sample (Bigsten, Collier, Dercon, Fafchamps, Gauthier, Gunning, Oduro, Oostendorp, Patillo, Soderbom, Teal and Zeufack 2004).

The second and third panels show similar statistics when geographical location is defined at the level of the city and province, respectively. While there are on average 18.5 communes per province, the number of communes with manufacturing employment is only 2.9 times the number of provinces. This suggests that, within each province, manufacturing employment is concentrated geographically in a few communes only. Comparison between the three panels indicates that values are roughly multiplied by 2.5-3 between the commune and province data. In spite of the reduction in the number of locations, the much larger number of usable years when using province data means that there are more than twice as many usable observations than when using the more disaggregated communes. Furthermore, when estimating regressions

using the commune data, we lose many observations due to lagging.

4. Productivity and wages

Before we turn to the estimation of our model of interest, we need to generate our productivity variables. Ideally, we would want to have information on total factor productivity in volume, sale prices, wages, and input costs since theoretically they can all channel externalities. In practice, we do not have information on sale prices and input costs. Consequently, we focus our attention on wages and total factor productivity in value.

Wage w_{ijt} is obtained by dividing, for each firm, the total annual wage bill by the number of employees.¹⁰ We then take the median of the location and sector as our wage measure w_{ijt} . The median is preferred to the mean because it is less sensitive to measurement error.¹¹

To obtain an estimate of total factor productivity, we estimate, for each firm k in our sample, a Cobb-Douglas production function of the form:

$$\log Q_{kt} = \theta_0 + \theta_1 \log L_{kt} + \theta_2 \frac{L_{kt}^{casual}}{L_{kt}} + \theta_3 \log K_{kt} + \theta_4 Z_{kt} + e_{kt} \quad (4.1)$$

where Q_{kt} is the value of output, L_{kt}^{casual} is the number of casual workers (in permanent employment equivalent units), and Z_{kt} is a vector of control variables including the log of the firm's age, the squared log of age, the share of foreign and government ownership, dummies for limited liability and corporate status, as well as sector, region, and year dummies. Since capital stock information is only available for 2001, we estimate a predictive equation with the 2001 data and use the estimated coefficients to predict the capital in other years.¹²

¹⁰ Measured in permanent employee equivalent units.

¹¹ Since w_{ijt} is obtained by dividing two variables reported with error, its distribution has fat tails driven by very large and very small outliers. Using the median takes care of this problem. In practice, using the average wage instead of the median does not change subsequent results much.

¹² The predictive regression is presented in appendix and discussed in detail by Fafchamps and El Hamine

Predicted values of capital \hat{K}_{kt} are then used in lieu of capital in equation (4.1) as well as in subsequent analysis. This implies that capital is de facto instrumented. To avoid simultaneity bias, the two labor variables are also instrumented using the same variables used to predict capital, namely, lagged labor, firm equity, a dummy if the firm existed in the previous year, and variables measuring lagged investment. All values are deflated using sector-specific deflators. Results are presented in Table 2. They show that this simple, parsimonious model accounts for more than three fourth of the variation in firm output. Labor and capital share parameters take reasonable values.

The residuals \hat{e}_{kt} from equation (??) are then obtained. The median residual \hat{e}_{kt} for a given sector, year, and location is our measure of firm total factor productivity p_{ijt} . Together, w_{ijt} and p_{ijt} form our P_{ijt} vector. The reader should keep in mind that since we cannot construct an input price variable, we cannot control for possible productivity effects that take place via intermediate inputs or service costs.

5. Dynamic panel analysis

Equipped with our productivity measures, we are now ready to turn to the dynamic panel analysis. We begin with the output growth equation (2.6). All regressors are in logs, except for the productivity shocks \hat{e}_{kt} which, by construction, are already expressed in logs.

Results for equation (2.6) are reported in Table 3 using the GMM estimator proposed by Arellano and Bond (1991).¹³ Robust standard errors are used throughout. Since we do not

(2004). A small number of firm characteristics such as age, legal status, foreign or public ownership, as well as sectoral and location dummies are included as regressors. Time varying predictors include lagged labor and share of casual workers, investment and lagged investment, dummies for whether the firm invested in the current and previous period, and a dummy for whether is in its first year of existence, in which case all lagged values are set to zero. This parsimonious model explains two thirds of the variation in capital stock across firms in 2001.

¹³Following the recommendation of Arellano and Bond (1991), we report one-step GMM estimates throughout because two-step estimates are known to seriously underestimate standard errors in finite samples. This is confirmed in our case: two-step estimates are very similar to one-step estimates, but t values are unrealistically large. To check for robustness, we also estimate equation (2.6) using the slightly less efficient instrumental variable

know at what geographical level agglomeration effects are felt, three sets of two regressions are reported, each set corresponding to a different geographical unit of analysis. The first two columns refer to the commune data, which only runs from 1998 until 2001. The second set of two columns refers to the city data which runs from 1994 until 2001, and the last set refers to provinces, with data from 1985 until 2001. For each of these sets, two regressions are estimated: without and with productivity variables p_{ijt} and w_{ijt} . The number of observations increases as one moves from commune to city to province data, reflecting the increase in the number of usable years of data. At the bottom of this Table – and subsequent Tables – we report autocorrelation tests on the residuals.¹⁴ We also report a joint test of the agglomeration variables A_{ijt} and productivity variables P_{ijt} . Since we control for fixed effects, all reported coefficients are purged of time-invariant location and sector-specific effects, such as those that could be due to pure geographical advantages.

Results are broadly similar across the three sets of regressions: inference does not appear to depend on the geographical unit of analysis. This is probably due to the fact that communes with manufacturing activity tend to be located close to each other within each city and province. We find a large and positive coefficient on lagged output Q_{ijt} , suggesting a lot of persistence in economic activity. Total employment in the location L_{it} has a strong negative effect in all three regressions, suggesting that the presence of manufacturing employment has a negative effect on manufacturing growth. This is consistent with the existence of negative agglomeration externalities due to congestion. Our competition index C_{ijt} is positive and significant in all three

method suggested by Anderson and Hsiao (1982). All pre-determined differenced regressors are instrumented using lagged levels. Results are very similar to those obtained with GMM. Even with robust standard errors, inference is virtually identical.

¹⁴The Arellano and Bond estimator is known to be consistent under first order autocorrelation, but not under second order autocorrelation. A Sargan overidentification test was also conducted. When calculated under the assumption of homoskedastic errors, the Sargan test is known to over-reject the null in the presence of heteroskedasticity in this category of models. Sargan tests based on the two-step model, which corrects for heteroskedasticity, all fail to reject overidentification.

regressions: less concentration within a sector is beneficial to growth in this sector. The number of manufacturing sectors present in a location has a positive effect on manufacturing output growth, but the effect is only significant in the first regression. In contrast, the diversity index D_{it} is everywhere negative, significantly so in the city and province regressions. Taken together, these results appear to reject Jacobs' idea that manufacturing diversity is beneficial to growth.

Adding our two productivity variables to the regression does not, contrary to expectations, reduce the effect of agglomeration variables L_{it} , M_{it} , C_{ijt} and D_{it} ; in most cases it even magnifies their effect, as evidenced by higher individual t -values and a higher Wald test statistic for joint significance (see bottom of Table 3). This flies in the face of the idea that agglomeration effects on growth operate through productivity. Variable p_{ijt} also behave in an unexpected manner: contrary to expectations, it has a strong negative coefficient in all three sets of regressions. This means that, controlling for time-invariant sector and location effects, a rise in productivity at time t is associated with slower growth of output at time $t + 1$. This result also contradicts the idea that what fuels manufacturing growth at the local level is productivity gains. In contrast, the wage variable has the anticipated negative sign: a rise in wages at t leads to slower output growth at $t + 1$, possibly because firms leave the location for another one with lower labor costs. We revisit this hypothesis below.

One likely explanation for the negative sign on p_{ijt} is that productivity is subject to large temporary shocks: as productivity reverts towards its mean after a large positive shock, output tends to fall. This interpretation finds some support in observing that, once we control for past productivity, the coefficient on lagged output rises above one in all three regressions. This means that if output rises at t for reasons other than a productivity shock (e.g., because of investment or firm entry), this rise leads to an even faster increase in output in subsequent periods. In contrast, if output rises at t because of a productivity shock, it tends to fall subsequently,

probably because the productivity shock was short-lived.

To investigate these ideas further, we estimate equation (2.7) which include L_{ijt} , N_{ijt} and K_{ijt} as additional regressors. As indicated earlier, these variables are regarded as pre-determined in the estimation, and thus their differences are instrumented with lagged levels. Regression results are presented in Table 4. Inference regarding agglomeration and productivity variables is basically unchanged by the presence of L_{ijt} , N_{ijt} and K_{ijt} : whatever agglomeration and productivity variables are measuring, it is not past firm expansion. As anticipated, the introduction of these new regressors brings down the coefficient of lagged output below 1. But estimated coefficients for L_{ijt} , N_{ijt} and K_{ijt} are highly variable – switching sign and significance from one regression to the other. Results for employment and number of firms are by and large inconclusive, probably because multicollinearity with lagged levels of output and capital precludes reliable identification. Results are a bit more consistent for capital: once we control for past productivity shocks, investment at t is associated with a *fall* in output at $t + 1$. We revisit this puzzling result below.

To investigate issues more in detail, we turn to the effect that agglomeration and productivity variables may have directly on investment, employment growth, and firm entry and exit. We begin with employment growth, which has been the primary focus of much of the literature to date. Results are presented in Table 5, again using the Arellano and Bond GMM estimator. Other estimation details are the same as in Table 4.¹⁵

We again find evidence of a lot of persistence, with coefficients on lagged employment fluctuating between 0.68 and 1.05, depending on the regression. An increase in the number of firms at t is associated with employment growth at $t + 1$. This is consistent with the idea that new firms go through an initial period of rapid growth as they converge to their firm-specific steady

¹⁵In particular, L_{ijt} , K_{ijt} and N_{ijt} are regarded throughout as pre-determined variables and instrumented using lagged levels. Robust standard errors are reported throughout.

state. Surprisingly, passed changes in capital stock are not reflected in subsequent employment growth. To verify whether this result is due to the fact that we are using predicted capital in lieu of actual capital stock, we reestimate the model using the simpler Anderson and Hsiao approach. This enables us to use lagged investment instead of change in predicted capital stock.¹⁶ Results, not shown here to save space, are identical: it is not the reliance on predicted capital stock that accounts for the non-significant coefficient on capital.

The effect of agglomeration variables on employment is by and large identical to their effect on output: total employment L_{it} and diversity D_{it} have strong negative effects, while the competition index C_{ijt} is strongly positive. Within-sector competition thus appears beneficial to employment growth while total manufacturing employment and sectoral diversity have negative effects. The number of sectors M_{it} appears with a significantly positive coefficient in the commune data, but the effect disappears as we move to the province data. This suggests that the effect of M_{it} may have changed over time or depends strongly on the size of the geographical unit.

Productivity variables p_{ijt} and w_{ijt} , in contrast, behave in a completely different way compared to their effect on output: both variables have positive effects in the commune, city and province regressions. The level of significance of the productivity variable remains low, however. What these results imply is that a rise in productivity or wages at time t leads to a subsequent rise in employment at $t + 1$. For productivity, this effect is what theory predicts: as productivity increases, firms hire more workers. But the effect of w_{ijt} is contrary to theory: manufacturing employment is seen to increase after a rise in manufacturing wages.

Turning to investment, we again find evidence of persistence (see Table 6). But the coefficient of lagged capital is much lower than that of labor – between 0.37 and 0.58 compared to 0.68

¹⁶ Because capital is pre-determined, however, investment at t must be instrumented with lagged levels of predicted capital.

to 1.05. This is consistent with the observation that, in poor countries, investment is sporadic, perhaps due to convexity in adjustment costs (Bigsten et al. 2004). In further contrast with Table 5, we find that lagged employment growth has a strong effect on investment: locations and sectors that have expanded employment in the past tend to experience more investment more in the future. Put differently, employment growth tends to lead investment instead of the contrary, as is often assumed. This again is consistent with the existence of convex adjustment costs or option effects in investment (e.g. Dixit 1989, Dixit and Pindyck 1994).

Agglomeration variables have by and large the same effect on investment as they have on employment, so we need not discuss them again. But p_{ijt} and w_{ijt} behave in a different manner. Here we find a strong and robust association between past productivity increases and investment: locations and sectors that experienced a large increase in productivity at t are more likely to invest at $t + 1$. Since we have also seen that positive productivity shocks are associated with a subsequent fall in output, this suggests that firms may be behaving in a myopic manner, failing to see that the current productivity shock is short-lived. An alternative explanation is that firms are credit constrained. By raising current revenues, a favorable productivity shock enables firms to undertake investment that they could not previously undertake (e.g. Fazzari, Hubbard and Petersen 1988, Hubbard 1998, Bigsten, Collier, Dercon, Gauthier, Gunning, Isaksson, Oduro, Oostendorp, Patillo, Soderbom and Teal 1999, Fafchamps and Oostendorp 2002, Nkurunziza 2004).

Table 7 shows a similar regression analysis for the number of firms N_{ijt} . Results suggest a very high level of persistence for firm numbers: the coefficient on lagged N_{ijt} oscillates between 0.89 and 1.5. Lagged employment and capital are only significant in the province regression, labor with a positive coefficient and capital with a negative one. Put differently, employment growth at t is associated with net firm entry at $t + 1$ while investment at t seems to lead to firm

exit. The first effect suggests that when existing firms expand employment, new firms enter. The second effect may be due to the fact that investment by existing firms displaces smaller firms, hence leading to firm exit. While agglomeration variables again have the same effect on net firm entry, p_{ijt} and w_{ijt} are non-significant, except for wages which appears with a positive and significant coefficient in the province regression.

We further investigate firm entry by decomposing $\Delta \log N_{ijt}$ into gross entry, firm in-migration, firm out-migration, and gross exit and estimating equation (2.8). As explained in Section 2, the GMM estimator of Arellano and Bond does not strictly apply. The model is therefore estimate using the Anderson and Hsiao approach of instrumenting first differences with lagged levels. Results are summarized in Table 8. Robust standard errors and a 10% significance level are used for inference purposes. Since exit and out-migration appear as negative variables, the signs of all coefficient is immediately comparable.

We find that persistence affects all four dependent variables: a larger number of firms increases entry and in-migration and reduces exit and out-migration. The positive effect of employment on N_{ijt} that is significant in the province regression appears to take place through increased gross entry and reduced out-migration; exit and in-migration are not affected. The negative effect of investment on N_{ijt} in the province regression appears driven primarily by increased firm exit. Agglomeration variables affect our four dependent variables differently. Total employment L_{it} depresses N_{ijt} via its effect on entry, exit, and out-migration. In contrast, firm in-migration does not appear to be affected by changes in L_{it} over time. The number of sectors, which is non-significant in Table 7, tends to be non-significant here as well, except for a few regressions where the effect is significantly negative. Much of the negative effect of diversity on N_{ijt} appears to be due to reduced gross entry. Diversity also tends to reduce out-migration and firm exit, but the effect is only significant in the commune regression.

Competition has a positive and significant effect in most regressions: less concentration reduces firm exit and out-migration while encouraging more entry and in-migration. According to the literature (e.g. Nelson and Winter 1982, Clerides et al. 1998, Das, Roberts and Tybout 2001, de Melo, Haddad and Horton 2001), the effect of competition on productivity is thought to come from the elimination of inefficient firms and entry by newer, more productive firms. In this case, we would expect that increased competition increases both entry *and* exit. This is not what we find.

Finally, p_{ijt} is seen to have a negative effect on firm in-migration and a positive effect on exit in the city regressions. The wage variable w_{ijt} , in contrast, tends to raise entry and reduce exit in the province regression. Neither results are consistent with the agglomeration externality idea that locations with favorable productivity shocks attract new firms, or that higher wages drive firms away.

6. Conclusion

The literature has attempted to provide evidence of agglomeration externalities by regressing sectoral employment growth on variables meant to capture location-specific specialization, competition, and diversity (e.g. Glaeser et al. 1992, Henderson 1997, Combes 2000, Combes et al. 2004, Bun and El Makhoulfi 2004). Evidence of an influence of agglomeration variables on growth has generally been taken as evidence of productivity effects. In this paper, we have tested the validity of this approach in two ways: by controlling directly for productivity, and by comparing results obtained using employment growth with those using output, investment, and firm entry and exit. What makes these improvements possible is detailed and exhaustive firm-level data available over an extended period of time. In our econometric analysis, we control for location and sector specific fixed effects and correct for the fact that some regressors are

pre-determined.

We find strong and robust evidence of agglomeration effects. Moreover, none of our results depends on the level of geographical disaggregation: we obtain similar findings whether working with commune, city, or province data. The reason probably is that within provinces, manufacturing remains concentrated in a few nearby locations. Our results show that manufacturing employment is negatively associated with growth in output, employment, capital, and number of firms. We also find that sectoral concentration is inimical to growth: locations and sectors where firms are of equal size tend to grow faster in terms of output, employment, and capital. Firm entry is also higher. Finally, locations with equal distribution of employment across various manufacturing sectors grow significantly slower, again in terms of output, employment, capital, and number of firms. Interpreting these results as other authors have done, we would conclude that competition is good for growth – and diversity bad – because of their effect on productivity.

Our other findings, however, cast some doubt on this interpretation. First, agglomeration variables have on aggregate sectoral growth in a given location virtually the opposite effect than the one they have on individual firm productivity. Using the same data set, Fafchamps and El Hamine (2004) estimate the effect of agglomeration variables on wages and total factor productivity at the individual firm level. They find that competition reduces productivity while diversity raises it – the opposite result from what we find here.

Second, if agglomeration variables influence manufacturing growth through their effect on productivity, controlling for productivity directly should eliminate or, at the very least, reduce the significance of agglomeration variables. In our detailed analysis, we find instead that none of the agglomeration effects is seriously affected when we introduce measures of total factor productivity: in none of our regressions do productivity variables lower the significance of agglomeration variables. Agglomeration variables measure something that is relevant to firm growth, but it is

probably not total factor productivity.

To be fair, our productivity variables do not always behave as anticipated either. In particular, past productivity shocks tend to lower future output growth. We interpret this finding as consistent with the idea that productivity shocks are not very persistent; reversion to the mean implies slower future growth. We also find that a rise in productivity raises subsequent employment and investment, but has no effect on firm entry and exit. The effect is particularly strong on investment, suggesting that firms invest more in the wake of a positive productivity shock when revenues are high, possibly because they are liquidity constrained (e.g. Hubbard 1998, Bigsten et al. 1999, Bigsten, Collier, Dercon, Fafchamps, Gauthier, Gunning, Oduro, Oostendorp, Patillo, Soderbom, Teal and Zeufack 2003).

The analysis presented here raises many new questions. The literature has relied on certain variables thought to affect productivity in order to measure agglomeration externalities. Our findings suggest that this approach is unreliable: agglomeration variables do not have on firm-level productivity the same effect that they have on aggregate growth, and they do not influence manufacturing growth via their presumed effect on productivity. Yet agglomeration variables have strong and robust effect on growth. The question then is: what is this effect capturing? To answer this question adequately, we need to turn to individual firm-level data. This is the object of future work.

References

- Anderson, T.W. and C. Hsiao. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18 (1982): 47–82.
- Arellano, Manuel. *Panel Data Econometrics*, Oxford: Oxford University Press, 2003).

- and Bo Honoré. “Panel Data Models: Some Recent Developments.” in “Handbook of Econometrics,” Vol. 5, Chapter 53, Amsterdam: J.J. Heckman and E. Leamer (eds.), North-Holland, 2001).
- and Steven R. Bond. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *Review of Economic Studies* 58 (1991): 277–97.
- Arrow, Kenneth J. “The Economic Implications of Learning by Doing.” *Review of Economic Studies* 29 (1962): 155–173.
- Barrett, Christopher B. “Food Marketing Liberalization and Trader Entry: Evidence from Madagascar.” *World Development* 25(5) 1997): 763–777.
- Bigsten, Arne, Paul Collier, Stefan Dercon, Bernard Gauthier, Jan Willem Gunning, Anders Isaksson, Abena Oduro, Remco Oostendorp, Cathy Patillo, Mans Soderbom, and Francis Teal. “Investment in Africa’s Manufacturing Sector: A Four Country Panel Data Analysis.” *Oxford Bulletin of Economics and Statistics* 61(4) 1999): 489–512.
- , — , — , Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Abena Oduro, Remco Oostendorp, Cathy Patillo, Mans Soderbom, Francis Teal, and Albert Zeufack. “Credit Constraints in Manufacturing Enterprises in Africa.” *Journal of African Economies* 12(1) 2003): 104–24.
- , — , — , — , — , — , — , — , — , — , — , and — . “Adjustment Costs, Irreversibility, and Investment Patterns in Africa’s Manufacturing Sector.” 2004). (mimeograph).

- Bun, Maurice J.G. and Abderrahman El Makhoulfi. "Dynamic Externalities, Local Industrial Structure and Economic Development: Panel Data Evidence for Morocco." 2004). (mimeograph).
- Ciccone, Antonio and Kiminori Matsuyama. "Start-Up Costs and Pecuniary Externalities as Barriers to Economic Development." *J. Devel. Econ.* 49(1) 1996): 33–59.
- and Robert E. Hall. "Productivity and the Density of Economic Activity." *Amer. Econ. Rev.* 86(1) 1996): 54–70.
- Clerides, Sofronis, Saul Lach, and James Tybout. "Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics* 113(3) 1998): 903–947.
- Combes, Pierre-Philippe. "Economic Structure and Local Growth: France, 1984-1993." *Journal of Urban Economics* 47 2000): 329–55.
- , Thierry Magnac, and Jean-Marc Robin. "The Dynamics of Local Employment in France." 2004). (mimeograph).
- Daniels, Lisa. *Changes in the Small-Scale Enterprise Sector from 1991 to 1993: Results from a Second Nationwide Survey in Zimbabwe*, Bethesda, Maryland: Gemini Technical Report No. 71, Gemini, 1994).
- Das, Sanghamtra, Mark J. Roberts, and James R. Tybout. "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." Manuscript. Cambridge, Mass.: National Bureau of Economic Research, Working Paper 8269 2001.
- de Melo, Jaime, Mona Haddad, and Brian Horton. "Exports and Industrial Performance During the Moroccan Trade Liberalization." in "Producer Heterogeneity and Performance in the

- Semi-Industrialized Countries,” Oxford: James Tybout and Mark Roberts (eds.), Oxford University Press, 2001).
- Desmet, Klaus and Marcel Fafchamps. “Employment Concentration across US Counties.” 2003). (mimeograph).
- and — . “What Are Falling Transport Costs Doing to Spatial Concentration Across U.S. Counties?” *Journal of Economic Geography* 2005). (forthcoming).
- Dicken, Peter and Peter E. Lloyd. *Location in Space: Theoretical Perspectives in Economic Geography*, New York: Harper Collins, 1990).
- Dixit, Avinash. “Entry and Exit Decisions Under Uncertainty.” *J. Polit. Econ.* 97(3) 1989): 620–638.
- Dixit, Avinash K. and Robert S. Pindyck. *Investment Under Uncertainty*, Princeton, N.J.: Princeton University Press, 1994).
- Ellison, Glenn and Edward L. Glaeser. “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach.” *Journal of Political Economy* 105(5) 1997): 889–927.
- Fafchamps, Marcel. “Mobile Capital, Location Externalities, and Industrialization.” *J. Comparative Econ.* 25(3) 1997): 345–365.
- and Brigit Helms. “Local Demand, Investment Multipliers, and Industrialization: Theory and Application to the Guatemalan Highlands.” *J. Devel. Econ.* 49(1) 1996): 61–92.
- and Remco Oostendorp. “Investment.” in “Industrial Change in Africa: Zimbabwean Firms under Structural Adjustment,” New York: Jan Willem Gunning and Remco Oostendorp (eds.), Palgrave, 2002).

- and Said El Hamine. “Firm Productivity, Wages, and Agglomeration Externalities.” 2004). (mimeograph).
- , —— , and Albert Zeufack. “Learning to Export: Evidence from Moroccan Manufacturing.” 2002). (mimeograph).
- Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen. “Investment, Financing Decisions, and Tax Policy.” *American Economic Review* 78(2) 1988): 200–5.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge and London: MIT Press, 1999).
- Glaeser, Edward L., Hedi D. Kallal, José A. Sheinkman, and Andrei Shleifer. “Growth in Cities.” *J. Polit. Econ.* 100(6) 1992): 1126–1152.
- Haddad, Mona and Jaime de Melo. “Brendan Horton.” in “Morocco, 1984-89: Trade Liberalization, Exports, and Industrial Performance,” Oxford and New York: Mark J. Roberts and James R. Tybout (eds), Oxford University Press for the World Bank, 1996): pp. 285–313.
- Henderson, J. Vernon. *Urban Development: Theory, Fact, and Illusion*, New York: Oxford University Press, 1988).
- Henderson, Vernon. “Externalities and Industrial Development.” *Journal of Urban Economics* 42 1997): 449–70.
- . “Marshall’s Scale Economies.” *Journal of Urban Economics* 53(1) 2003): 1–28.
- Hirschman, Albert O. *The Strategy of Economic Development*, New Haven: Yale University Press, 1958).
- Hubbard, R. Glenn. “Capital Market Imperfections and Investment.” *Journal of Economic Literature* 36(1) 1998): 193–225.

- Isard, W. *Location and Space Economy*, Cambridge, Mass.: MIT Press, 1956).
- Jacobs, Jane. *The Economy of Cities*, New York: Random House, 1969).
- . *Cities and the Wealth of Nations*, New York: Random House, 1984).
- Krugman, Paul R. “Increasing Returns and Economic Geography.” *J. Polit. Econ.* 99(3) 1991): 483–99.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. “Industrialization and the Big Push.” *J. Polit. Econ.* 97(5) 1989): 1003–26.
- Nelson, Richard and Sidney Winter. *An Evolutionary Theory of Economic Change*, Cambridge MA: Harvard UP, 1982).
- Nkurunziza, Janvier. *Firm Growth and Survival in Kenyan Manufacturing*, Oxford: Department of Economics. Oxford University, 2004). Unpublished DPhil thesis.
- Porter, Michael. *The Competitive Advantage of Nations*, New York: Free Press, 1990).
- Quah, Danny. “Galton’s Fallacy and Tests of the Convergence Hypothesis.” *Scand. J. Econ.* 95(4) 1993): 427–443.
- Rauch, James E. and Alessandra Casella. “Overcoming Informational Barriers to International Resource Allocation: Prices and Ties.” *Economic Journal* 113(484) 2003): 21–42.
- Rodriguez-Clare, Andrés. “The Division of Labor and Economic Development.” *J. Devel. Econ.* 49 1996): 3–32.
- Romer, Paul M. “Increasing Returns and Long-Run Growth.” *Journal of Political Economy* 94 (5) 1986): 1002–1037.

———. “Endogenous Technological Change.” *Journal of Political Economy* 98 (5) pt.2 1990): S71–102.

Thompson, Fraser. “Industry Interaction, Knowledge Spillovers and Employment Growth in Moroccan Manufacturing.” 2004). (unpublished MPhil thesis).

Tybout, James R. “Manufacturing Firms in Developing Countries: How Well Do They Do and Why?” *Journal of Economic Literature* 38(1) 2000): 11–44.

Table 1. Variables used in the analysis

		Number of obs.	Mean	Median	St. dev.	Minimum	Maximum
A. By commune (1998-2001)							
Sectoral employment in the location	Lijt	5428	359.1	53.0	1050.7	1	15478
Total employment in the location	Lit	5428	4252.3	1134.0	7566.0	1	59219
Sectoral number of firms in the location	Nijt	5428	4.9	2.0	7.6	1	76
Competition index	Cijt	5428	2.7	1.6	3.2	1	34
Number of sectors in the location	Mit	5428	8.2	9.0	4.5	1	14
Diversity index	Dit	5428	3.3	2.9	1.9	1	10
Number of new firms	Ne	5428	0.6	0.0	1.4	0	35
Number of firms moving in	Nn	5428	0.9	0.0	2.6	0	44
Number of exiting firms	Nx	4049	0.5	0.0	1.2	0	14
Number of firms moving out	No	4049	0.5	0.0	1.3	0	22
Investment	lijt	5428	1651.5	51.0	6715.0	0	164627
Median productivity in sector-location	pijt	5355	0.02	0.01	0.78	-5.35	3.94
Median wage in sector-location	wijt	5427	6.87	5.51	11.32	0.12	485.79
B. By city (1994-2001)							
Sectoral employment in the location	Lijt	5570	683.9	96.0	1857.7	1	24378
Total employment in the location	Lit	5570	8947.5	2868.3	15856.2	1	103629
Sectoral number of firms in the location	Nijt	5570	9.3	3.0	16.5	1	163
Competition index	Cijt	5570	3.9	2.0	5.5	1	67
Number of sectors in the location	Mit	5570	9.5	11.0	4.5	1	14
Diversity index	Dit	5570	3.8	3.5	2.1	1	11
Number of new firms	Ne	5570	0.9	0.0	2.1	0	64
Number of firms moving in	Nn	5570	0.5	0.0	1.3	0	25
Number of exiting firms	Nx	4809	0.8	0.0	1.9	0	26
Number of firms moving out	No	4809	0.4	0.0	1.3	0	25
Investment	lijt	5570	2849.9	142.4	9647.1	0	186474
Median productivity in sector-location	pijt	5517	-0.00	0.01	0.76	-5.35	3.72
Median wage in sector-location	wijt	5570	5.97	5.00	11.69	0.03	485.79
C. By province (1985-2001)							
Sectoral employment in the location	Lijt	7945	901.7	163.6	2094.8	1	24378
Total employment in the location	Lit	7945	11954.3	5192.4	18497.2	1	116381
Sectoral number of firms in the location	Nijt	7945	12.4	5.0	20.5	1	200
Competition index	Cijt	7945	4.8	2.7	6.4	1	75
Number of sectors in the location	Mit	7945	11.0	12.0	3.6	1	14
Diversity index	Dit	7945	4.4	4.2	2.1	1	11
Number of new firms	Ne	7639	1.2	0.0	2.6	0	64
Number of firms moving in	Nn	7639	0.5	0.0	1.6	0	41
Number of exiting firms	Nx	7354	0.9	0.0	2.0	0	34
Number of firms moving out	No	7354	0.5	0.0	1.9	0	45
Investment	lijt	7945	3781.2	290.3	13134.0	0	350678
Median productivity in sector-location	pijt	7909	-0.01	0.00	0.68	-4.19	3.63
Median wage in sector-location	wijt	7945	5.51	4.82	3.64	0.01	47.73

Table 2. Production function

(the dependent variable is the log of output in constant terms; the estimator is instrumental variables)

	Coef.	t-stat.
Labor (log) **	0.627	93.78
Share of casual workers **	-0.275	-8.98
Predicted capital stock (log)	0.458	99.52
Firm age (log)	0.405	21.38
Squared log of firm age	-0.059	-15.28
Share of foreign capital ownership	0.000	0.79
Share of government capital ownership	-0.002	-5.90
Dummy if firm has limited liability status	0.147	14.75
Dummy if firm is a corporation	0.157	16.42
Sectoral dummies	included but not shown	
Region dummies	included but not shown	
Year dummies	included but not shown	
Intercept	1.483	19.15
Number of observations	92930	
R2	0.765	

** instrumented using lagged labor, share of casual workers, firm equity, and investment

Table 3. Reduced form on total output

(Estimator is Arellano and Bond GMM; see text for details. All regressors in log. Robust standard errors reported)

	Commune data				City data				Province data			
	1998-2001				1994-2001				1985-2001			
Number of observations	2110	2088	3582		3555	6536			6510			
Number of sector-location groups	1185	1169	801		792	589			585			
Average number of obs. per group	1.781	1.786	4.472		4.489	11.097			11.128			
Maximum number of obs. per group	2	2	6		6	15			15			
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
Total output in sector	0.956	5.74	1.339	6.64	1.053	7.27	0.793	7.95	1.115	14.39	0.061	13.56
Total employment	-0.276	-2.62	-0.517	-4.12	-0.467	-4.01	-0.273	-4.88	-0.514	-3.46	0.061	-5.73
Number of sectors	0.835	4.38	0.820	4.05	0.116	0.81	0.048	0.67	0.048	0.30	0.061	0.36
Competition index in sector	1.219	4.98	1.630	5.43	0.490	4.62	0.812	5.84	0.826	7.50	0.826	8.20
Diversity index	-0.024	-0.16	-0.241	-1.45	-0.262	-2.09	-0.390	-2.50	-0.431	-2.19	-0.431	-4.05
Median productivity in sector			-1.080	-6.48			-0.769	-6.82			-0.725	-11.74
Median wage in sector			-0.219	-2.48			-0.195	-2.84			-0.117	-2.29
Intercept	-0.145	-1.01	-0.205	-0.75	-0.130	-1.06	0.007	0.07	0.141	0.97	0.024	0.17
<i>Arellano-Bond test of autocorrelation</i>												
first-order	-5.580	0.00	-5.170	0.00	-7.510	0.00	-7.410	0.00	-9.850	0.00	-9.980	0.00
second order					1.620	0.10	0.410	0.68	2.890	0.00	1.520	0.13
<i>Joint test on 4 agglomeration variables</i>												
	chi-square	p-value	chi-square	p-value	chi-square	p-value	chi-square	p-value	chi-square	p-value	chi-square	p-value
	35.970	0.00	37.530	0.00	32.310	0.00	43.010	0.00	63.050	0.00	81.400	0.00
<i>Joint test on productivity and wages</i>			42.810	0.00			48.420	0.00			141.690	0.00

Table 4. Total output, investment, employment, and number of firms

(Estimator is Arellano and Bond GMM; see text for details. Qijt, Lijt and Nijt instrumented. All regressors in log. Robust standard errors reported.)

	Commune data				City data				Province data			
	1998-2001				1994-2001				1985-2001			
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
Number of observations	2088		2088		3555		3555		6510		6510	
Number of sector-location groups	1169		1169		792		792		585		585	
Average number of obs. per group	1.786		1.786		4.489		4.489		11.128		11.128	
Maximum number of obs. per group	2		2		6		6		15		15	
Total output in sector	Qijt	0.631	3.44	1.848	0.337	5.70	0.981	4.47	0.479	5.92	10.79	0.908
Employment in sector	Lijt	0.189	0.99	-0.700	0.089	-2.20	-0.329	0.87	0.335	-2.10	5.69	0.067
Number of firms in sector	Nijt	0.663	1.58	0.991	0.439	2.18	0.455	2.58	0.306	2.45	3.30	0.285
Capital in sector	Kijt	0.012	0.11	-0.435	0.101	-2.73	-0.167	1.78	-0.022	-1.96	-0.63	-0.227
Total employment	Lit	-0.365	-3.05	-0.418	-0.313	-3.23	-0.277	-3.90	-0.285	-3.25	-3.95	-0.286
Number of sectors	Mit	0.679	3.17	0.414	0.038	1.85	-0.022	0.27	-0.030	-0.16	-0.20	-0.048
Competition index in sector	Cijt	1.479	4.60	1.440	0.627	3.80	0.514	4.84	0.691	3.47	8.93	0.632
Diversity index	Dit	-0.122	-0.75	-0.229	-0.299	-1.29	-0.272	-2.47	-0.329	-2.06	-3.58	-0.384
Median productivity in sector	pijt			-1.508		-5.81		-0.755		-5.31		-0.594
Median wage in sector	wijt			-0.263		-2.33		-0.175		-2.49		-0.031
Intercept		-0.161	-1.21	-0.078	-0.177	-0.52	-0.152	-0.71	-0.026	-0.58	-0.17	-0.090
<i>Arellano-Bond test of autocorrelation</i>												
first-order		-5.330	0.00	-5.240	-7.310	0.00	-7.510	0.00	-9.950	0.00	-10.050	0.00
second order					1.150	0.25	0.310	0.75	2.490	0.01	2.070	0.04
<i>Joint test on 4 agglomeration variables</i>												
<i>Joint test on productivity and wages</i>		32.680	0.00	17.750	27.890	0.00	15.770	0.00	80.870	0.00	62.700	0.00
				33.800		0.00	28.180				84.130	

Table 5. Employment growth

(Estimator is Arellano and Bond GMM; see text for details. All regressors in log. Robust standard errors reported)

	Commune data				City data				Province data			
	1998-2001				1994-2001				1985-2001			
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
Number of observations	2088		2088		3555		3555		6510		6510	
Number of sector-location groups	1169		1169		792		792		585		585	
Average number of obs. per group	1.786		1.786		4.489		4.489		11.128		11.128	
Maximum number of obs. per group	2.000		2.000		6.000		6.000		15.000		15.000	
Employment in sector	Lijt	0.961	4.86	1.054	4.91	0.685	5.68	0.724	5.91	0.760	16.24	0.772
Number of firms in sector	Nijt	0.779	2.20	0.711	1.95	0.404	2.61	0.386	2.44	0.500	5.39	0.512
Capital in sector	Kijt	0.128	1.36	0.080	0.80	0.048	0.83	0.043	0.70	0.013	0.43	-0.005
Total employment	Lit	-0.595	-4.48	-0.588	-4.47	-0.499	-5.26	-0.510	-5.39	-0.393	-5.52	-0.378
Number of sectors	Mit	0.696	3.34	0.723	3.44	0.266	1.73	0.272	1.73	0.027	0.20	-0.007
Competition index in sector	Cijt	1.882	5.93	1.897	5.89	1.076	7.29	1.102	7.42	0.885	12.60	0.889
Diversity index	Dit	-0.389	-2.44	-0.371	-2.34	-0.419	-3.42	-0.425	-3.44	-0.468	-5.23	-0.449
Median productivity in sector	pijt			0.078	1.21			0.067	1.78		0.046	0.046
Median wage in sector	wijt			0.184	2.10			0.104	2.01		0.172	0.172
Intercept		-0.189	-0.96	-0.179	-0.93	0.035	0.54	-0.225	-1.70	0.270	1.99	0.297
<i>Arellano-Bond test of autocorrelation</i>												
first-order		-4.840	0.00	-4.770	0.00	-7.290	0.00	-7.260	0.00	-10.380	0.00	-10.200
second order						-0.020	0.99	-0.060	0.96	0.670	0.50	0.710
<i>Joint test on 4 agglomeration variables</i>												
<i>Joint test on productivity and wages</i>												
	chi-square	46.720	p-value	46.860	p-value	64.890	p-value	66.930	p-value	166.730	p-value	162.610
			0.00	6.940	0.03		0.00	9.270	0.01		0.00	27.760

Table 6. Investment

(Estimator is Arellano and Bond GMM; see text for details. All regressors in log. Robust standard errors reported)

	Commune data				City data				Province data			
	1998-2001				1994-2001				1985-2001			
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z
Number of observations	2084		2084		3550		3550		6506		6506	
Number of sector-location groups	1168		1168		792		792		585		585	
Average number of obs. per group	1.784		1.784		4.482		4.482		11.121		11.121	
Maximum number of obs. per group	2.000		2.000		6.000		6.000		15.000		15.000	
Capital in sector	Kijt	0.553	5.19	0.578	5.05	0.395	6.99	0.423	6.92	0.372	10.95	0.392
Employment in sector	Lijt	0.574	3.04	0.564	2.72	0.477	4.01	0.474	3.84	0.540	10.74	0.530
Number of firms in sector	Nijt	0.471	1.25	0.426	1.11	0.254	1.57	0.214	1.31	0.231	2.34	0.226
Total employment	Lit	-0.456	-3.46	-0.459	-3.47	-0.394	-4.49	-0.405	-4.57	-0.405	-5.79	-0.405
Number of sectors	Mit	0.612	2.70	0.642	2.78	0.184	1.14	0.197	1.21	0.058	0.39	0.046
Competition index in sector	Cijt	1.699	5.27	1.694	5.21	0.790	4.94	0.795	4.89	0.710	8.52	0.713
Diversity index	Dit	-0.215	-1.29	-0.212	-1.27	-0.292	-2.39	-0.305	-2.48	-0.396	-4.29	-0.384
Median productivity in sector	pijt			0.108	1.65			0.126	2.76			0.113
Median wage in sector	wijt			-0.122	-1.48			-0.075	-1.33			-0.027
Intercept		-0.424	-1.41	-0.410	-1.34	-0.171	-0.93	-0.196	-1.04	0.247	1.26	0.254
Arellano-Bond test of autocorrelation												
first-order		-5.070	0.00	-5.080	0.00	-7.490	0.00	-7.420	0.00	-9.500	0.00	-9.500
second order						0.860	0.39	0.820	0.41	-0.490	0.62	-0.400
Joint test on 4 agglomeration variables	chi-square	33.820	0.00	33.330	0.00	32.240	0.00	31.840	0.00	86.120	0.00	83.970
Joint test on productivity and wages			0.13	4.070	0.13			7.940	0.02			15.960

Table 7. Number of firms
(Estimator is Arellano and Bond GMM; see text for details. All regressors in log. Robust standard errors reported)

Commuene data										City data 1994-2001				Province data 1985-2001			
1998-2001										1994-2001				1985-2001			
	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z	Coef.	z			
Number of observations	2088		2088		3555		3555		6510		6510		6510				
Number of sector-location groups	1169		1169		792		792		585		585		585				
Average number of obs. per group	1.786		1.786		4.489		4.489		11.128		11.128		11.128				
Maximum number of obs. per group	2.000		2.000		6.000		6.000		15.000		15.000		15.000				
Number of firms in sector	1.502		1.520		7.45		0.893		10.62		0.894		19.51				
Employment in sector	-0.014		-0.022		-0.20		-0.037		-0.70		-0.032		10.46				
Capital in sector	0.011		0.010		0.18		0.015		0.58		0.013		-0.61				
Total employment	-0.173		-0.173		-2.79		-0.121		-3.16		-0.121		0.48				
Number of sectors	-0.071		-0.080		-0.76		-0.047		-0.66		-0.046		-3.12				
Competition index in sector	0.911		0.910		5.45		0.430		6.00		0.432		-0.65				
Diversity index	-0.222		-0.225		-2.56		-0.137		-2.30		-0.136		6.01				
Median productivity in sector			-0.029		-0.89						-0.020		-2.28				
Median wage in sector			0.003		0.07						0.006		-1.01				
Intercept	0.009		0.17		-0.34		0.035		1.09		-0.099		0.22				
													-1.00				
													0.259				
													2.95				
													0.265				
Arellano-Bond test of autocorrelation										z p-value				z p-value			
first-order	-5.580		0.00		-5.570		0.00		-8.780		0.00		-11.980				
second order									-0.520		0.60		-0.740				
													0.66				
													-0.440				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				
													p-value				
													chi-square				

Table 8. Summary of the decomposition of the change in number of firms

(Estimator is Anderson and Hsiao GMM; see text for details. All regressors in log. Robust standard errors reported.)

All dependent variables expressed in percentage change (positive for entry and moved in, negative for exit and moved out relative to initial number of firms).

	Nijt	Entry	Moved in	Moved out	Exit						
Number of firms in sector	Nijt	+	(all)	+	(mc)	+	(all)	+	(all)		
Employment in sector	Lijt	+	(p)	+	(p)	0	(all)	+	(mp)	0	(all)
Capital in sector	Kijt	-	(p)	0	(all)	+	(m)	+	⊖	-	(p)
Total employment	Lit	-	(all)	-	(cp)	0	(all)	-	(all)	-	(all)
Number of sectors	Mit	0	(all)	0	(all)	-	⊖	0	(all)	-	(mc)
Competition index in sector	Cijt	+	(all)	+	(cp)	+	⊖	+	(all)	+	(all)
Diversity index	Dit	-	(all)	-	(all)	0	(all)	-	(mp)	-	(m)
Median productivity in sector	pijt	0	(all)	0	(all)	-	⊖	0	(all)	-	⊖
Median wage in sector	wijt	+	(p)	+	(p)	-	(m)	0	(all)	+	(p)

(all means commune, city, and province regressions; p means province regression only; m means commune regression only; c means city only).

Appendix Table. Predictive equation for capital stock

Firm characteristics	Coef.	t-stat.
Age of firm (log)	0.408	4.10
Age of firm (log squared)	-0.069	-3.45
Share of capital held by foreigners	0.002	3.41
Share of capital held by government	0.001	0.16
Limited liability (sole proprietor omitted categ.)	0.367	8.11
Corporation	0.590	9.16
Instruments		
Lagged labor (log)	0.355	17.19
Lagged share of casual workers	-0.267	-2.17
Disbursed equity (log)	0.508	37.68
Firm existed in previous year (yes=1)	-0.886	-11.62
Investment (log)	0.201	15.16
No investment (no=1)	0.377	6.23
Lagged investment (log)	0.049	3.34
No lagged investment (no=1)	0.043	0.63
Sectoral dummies	yes	
Regional dummies	yes	
Intercept	1.544	4.34
Number of observations	6106	
R-squared	0.666	