Urban Attributes and Regional Differences in Productivity: Evidence from the External Economies of Brazilian Micro-Regions from 2000 to 2010

April, 2014

Rodrigo Simões

Elton Freitas

UFMG Cedeplar

Overview

Aim

Identify the extent to which the local productive structure of a city leverages external economies of scale.

Structure of paper

- INTRODUCTION
- THEORETICAL
- METHODOLOGY
  - EMPIRICAL MODEL
  - DATABASE
  - DESCRIPTION OF VARIABLES
- RESULTS
- FINAL REMARKS

INTRODUCTION (Motivation)

INTRODUCTION (Motivation)

This article aims to add **new methodological procedures** that broaden the interpretation spectrum for specialized productive agglomerations in order to identify the economies of scale present in local productive sectors. Furthermore, this paper proposes investigating the **relationship** between the **local economic structure and the local productivity level**. To do so, we analyze local industrial wage levels. The geographical dimensions considered are that of **Brazilian micro-regions**.

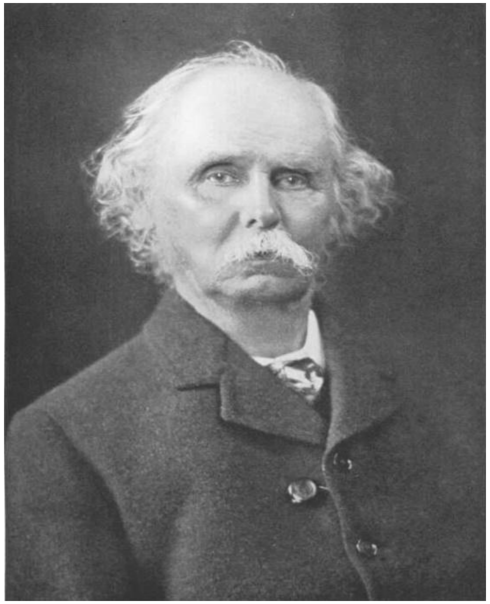
INTRODUCTION (Motivation)

Thus, **local productivity is influenced** not only by personal productive characteristics, such as elements related to human capital, regional attributes, features that impact upon industrial productivity differential or differences in regional productive structures. It is, on the other hand, also influenced by **urban attributes** – identified here as **centrality** and the **availability of complex services**.

THEORETICAL

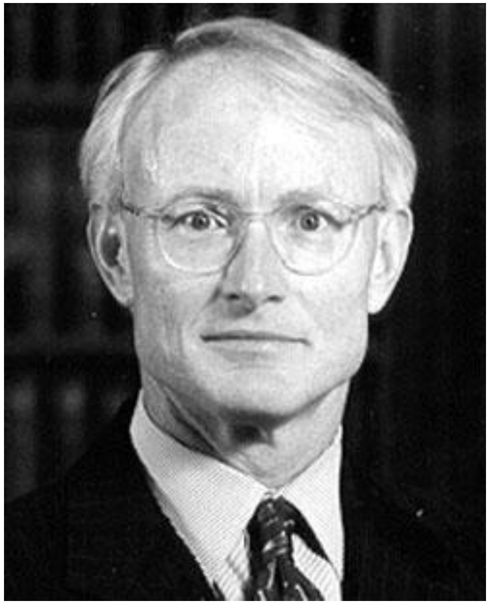
THEORETICAL

Agglomeration externalities: Marshall versus Jacobs versus Porter



X

X



Specialization **Diversification** **Competition**

METHODOLOGY

EMPIRICAL MODEL

EMPIRICAL MODEL

Identification Strategy

Capture the effects of agglomeration economies on productivity, indirectly investigating the variation in productivity through wages levels.

Model

- The **starting point** for our empirical research follows the work of **Combes et al (2008)**. The authors investigated the determinants of wage differentials in local labor markets in France, arguing that there are three major explanations for the spatial income gap.
- The first explanation assumes that spatial differences in wages are directly reflected by **spatial differences in the composition of the workforce and the skills**.
- The second explanation is based on **local endowments of attributes that are external to the employees**.
- The third explanation attributes the central role in the spatial differentiation of productivity gains – and hence wages – **to interactions in the labor market**.

EMPIRICAL MODEL

Model

- Combes et al. (2008) built their model based on the **profit equation of a representative firm** for a competitive area a, industry k and in year t:

$$\pi_{a,k,t} = p_{a,k,t} y_{a,k,t} - \sum_{i \in (a,k,t)} w_{i,t} l_{i,t} - r_{a,k,t} z_{a,k,t} \tag{1}$$

- where: $p_{a,k,t}$ is the price of the product $y_{a,k,t}$; $w_{i,l}$, and $l_{i,l}$ are the salary per day and the number of working days, respectively, for each employee at firm *i* in year *t*; $z_{a,k,t}$ represents other production factors and $r_{a,k,t}$ the prices.
- The **product follows a Cobb-Douglas function**:

$$y_{a,k,t} = A_{a,k,t} \left(\sum_{i \in (a,k,t)} s_{i,t} l_{i,t} \right)^{b} (z_{a,k,t})^{1-b} \tag{2}$$

- in which: the coefficient *b* is such that $0<b\leq1$; $s_{i,t}$ denotes the ability of the worker *i* in year *t*, and $A_{a,k,t}$ is the total factor productivity in *(a, k, t)*.

EMPIRICAL MODEL

Model

- If, in competitive equilibrium, the **worker receives wages equal to their marginal product**, then:

$$w_{i,t} = b p_{a(i,t),k(i,t),t} A_{a(i,t),k(i,t),t} \left(\frac{z_{a(i,t),k(i,t),t}}{\sum_{i \in (a,k,t)} s_{i,t} l_{i,t}} \right)^{1-b} s_{i,t} \qquad (3)$$

- Applying the first order condition for **profit maximization** with respect to other factors and inserting the result in (3), we have:

$$w_{i,t} = b(1-b)^{1-b/b} \left(p_{a(i,t),k(i,t),t} \frac{A_{a(i,t),k(i,t),t}}{(r_{a(i,t),k(i,t),t})^{1-b}} \right)^{1/b} s_{i,t}$$

$$w_{i,t} = B_{a(i,t),k(i,t),t} s_{i,t} \qquad (4)$$

EMPIRICAL MODEL

Model

- To make the model estimable from available data, Combes et al. (2008) made **two assumptions**. The first is that the **ability of the worker *i* is given by**:

$$\log s_{i,t} = X_{i,t}\varphi + \delta_i + \varepsilon_{i,t} \qquad (5)$$

- where: $X_{i,t}$ is a vector of the characteristics of workers; δ_i is a vector of the fixed effects for the worker; $\varepsilon_{i,t}$ is the *i.i.d*. error term.
- **The second one** considers $B_{a(i,t),k(i,t),t}$ as given by:

$$\log B_{a(i,t),k(i,t),t} = \beta_{a,t} + \mu_{k,t} + I_{a,k,t}\gamma_k \qquad (6)$$

- in which: $\beta_{a,t}$ is a vector of fixed effects indicating the area and year; $\mu_{k,t}$ is a vector of fixed effects indicating the industry and year; γ_k is a vector of associated coefficients and $I_{a,k,t}$ is a vector of variable interactions within the industry for each area/industry/year.

EMPIRICAL MODEL

Model

- Taking the log of equation (4) and combining it with equations (5) and (6) we have:

$$\log(w_{i,t}) = \beta_{a(i,t),t} + \mu_{k(i,t)} + I_{a(i,t),k(i,t),t}\gamma_{k(i,t)} + X_{i,t}\varphi + \delta_i + \varepsilon_{i,t} \qquad (7)$$

- Equation (7) is **the inverse labor demand equation**. This model takes the log of the wage rate of workers as a function of observable ($X_{i,t}$) and unobservable (δ_i) characteristics, the fixed effects of their geographical area ($\beta_{a(i,t),t}$) and sector ($\mu_{k(i,t)}$) and local characteristics of the sector in which they are employed: relative participation in the local economy, the number of establishments and the relative share of workers in professional occupations.

EMPIRICAL MODEL

Model

- This estimation allows for separately measuring the personal and area effects. One can thus assess the relative importance of **skills**, **local endowments** and **interactions (agglomeration economies)** for wage differentials in space. Therefore, Combes et al. (2008) adopt as their identification strategy a two-stage estimation. They first estimate equation (7), from which is obtained the vector of fixed effects by area, ($\beta_{a(i,t),t}$). They then regress the latter on variables representing the local endowments and intersectoral interactions. The specification takes the following form:

$$\beta_{a,t} = \varpi_0 + \theta_t + I_{a,t}\gamma + E_{a,t}\alpha + \upsilon_{a,t} \qquad (8)$$

- In this equation, θ_t are time fixed effects; γ is the vector of coefficients associated with the local intersectoral interactions $I_{a,t}$; α is a vector of coefficients associated with local capital endowments $E_{a,t}$; and $\upsilon_{a,t}$ is the i.i.d. error term that reflects local technological shocks.

EMPIRICAL MODEL

Model

- Combes et al (2008) also show that the model presented by equation (7) can be **aggregated and estimated for the geographical area**. Thus, equation (7) can be rewritten as:

$$\log w_{a,k,t} = \beta_{a,t} + \mu_{k,t} + I_{a,k,t}\gamma_k + \varsigma_{a,t}\phi_{a,t} + \varepsilon_{a,t} \qquad (9)$$

- where: $\log w_{a,k,t}$ is the average *log* of the wages of individuals in an industry k, in a given region *a*, in year *t*; $\phi_{a,t}$ is a vector of coefficients associated with $\varsigma_{a,t} = X_{a,t}\varphi + \delta_a$, which is a vector that captures the level of human capital in area *a* – or, as we would rather call it, the average skill level of workers in area *a*.

EMPIRICAL MODEL

Model

- In the first stage, information is available at the individual level; in the second, at the level of the area or region of study. Aggregating the equation in the first stage allows us to estimate a model with information from only the level of the area, which is appropriate for the database available for this work. Thus, substituting (8) into (9), we have:

$$\log w_{a,k,t} = \varpi_0 + \theta_t + I_{a,t}\gamma + E_{a,t}\alpha + \mu_{k,t} + I_{a,k,t}\gamma_k + \varsigma_{a,t}\phi_{a,t} + \varepsilon_{a,t} + \upsilon_{a,t} \quad (10)$$

- There is a problem in the aggregation of equation (9). The variable $\log w_{a,k,t}$ represents the average of the *log* of the wages of each individual *i* of an industry *k*, given region *a*. This becomes a problem because our database has no information on the individual level to measure the average. However, without loss of generality, the *log* of the mean wage is a good proxy for the average log of wages. We thus carry out the estimation with this proxy, since we have the average salary of an industry *k*, given a region *a*.

EMPIRICAL MODEL

Model

- Another important issue is that the **estimates are made separately for each segment**. One can then change the equation (10) once more. First, the subscript *k* may be removed. Then the $\mu_{k,t}$ that captures the fixed effects for each industry and time period can also be deleted. Moreover, $\varepsilon_{a,t}$ and $\upsilon_{a,t}$ are i.i.d. error terms, and we can therefore define $\xi_{a,t}$ as $\xi_{a,t} = \varepsilon_{a,t} + \upsilon_{a,t}$. The variable $E_{a,t}$ that captures the effects of local endowments may be included in the variable θ_t, which captures time fixed effects, thereby forming a component of fixed effects indicating the area and time. The equation can then be expressed as follows:

$$\log w_{a,t} = \varpi + \theta_{a,t} + I_{a,t}\gamma + \varsigma_{a,t}\phi + \xi_{a,t} \qquad (11)$$

- where: $\log w_{a,t}$ is the *log* of the mean wage in a given industry in the region *a* in year *t*; $\theta_{a,t}$ are fixed effects for area/year; $I_{a,t}$ captures the effects of the economic structure in an area in year *t*; $\varsigma_{a,t}$ captures the effects of the average skill of workers in the region *a* in year *t*; $\xi_{a,t}$ is the error term that reflects the local technological shocks and are assumed to be and i.i.d. for regions and periods.

EMPIRICAL MODEL

Model

- There are two features in **diversified urban centers**: one is **centrality** and the other one is the **availability of complex services**. However, variable $\theta_{a,t}$ captures the fixed effects for area/year. We will decompose it down as follows:

$$\theta_{a,t} = C_a + S_{a,t} + \tau_{a,t} \qquad (12)$$

- where: C_a is a variable to capture centrality; $S_{a,t}$ is a variable to capture the concentration of modern services; and $\tau_{a,t}$ is an i.i.d. error term for other unobserved regional influences.

EMPIRICAL MODEL

Model

- Thus, substituting (12) into (11) and making $e_{a,t} = \tau_{a,t} + \xi_{a,t}$, we have the equation estimated in this article:

$$\log w_{a,t} = \varpi + \underbrace{C_a + S_{a,t}}_{\text{Diversified Urban Centers}} + \overbrace{I_{a,t}\gamma}^{\text{Externalities}} + \underbrace{\varsigma_{a,t}\phi}_{\text{Human Capital}} + e_{a,t} \qquad (13)$$

- We must also consider, as a control, the different cost of living in the various regions, as they influence the pay gap in between them. We thus attempt to deal with this potential bias by using temporally constant monetary values that are, moreover, **regionally adjusted to account for spatial differences in the cost of living**.

DATABASE

DATABASE

DATABASE

- Annual Listing of Social Information (**RAIS**) for the period from **2000 to 2010**.

GEOGRAPHICAL AREA

- **Brazilian micro-regions**.

INDUSTRY CLASSIFICATION

- OECD classification of technological intensity
- Classify the sectors into four groups: high-technology, medium-high-technology, medium-low-technology and low-technology

Technological Intensity	CNAE 1.0 code
Low Technology	15; 16; 17; 18; 19; 20; 21; 22; 26; 27; 28; 36; 37
Medium-Low Technology	23; 24; 25
Medium-High Technology	29; 30; 33; 34
High Technology	31; 32; 35

DESCRIPTION OF VARIABLES

DESCRIPTION OF VARIABLES

Variable	Indicates	Expected Sign	Evidence
ql	Especialization	positive	Location Externalities (Marshall)
div	Diversification	positive	Urban Externalities (Jacobs)
comp	Competition	Positive or Negative	Porter Externalities or Location Externalities (Marshall)
den	Density	Positive	Urban Externalities (Jacobs)
educ	Human Capital	positive	Control for Human Capital
C	Centrality	positive	Urban centers leverages the sectoral productivity
S	Complex Services	positive	

RESULTS

RESULTS

Regarding the indicator of competition, **comp, the estimates are significant**, but the results showed **negative signs**. This indicates that it is **not a competitive structure that spurs productivity, but rather a monopolistic structure**, à la Marshall.

Table 1 - Estimation of the model (13) using as dependent variable regionalized real wages

Segments	const	ql	div	comp	den	educ	d_centralities	d_serv_complex	R2-adjust	F	N
Low Technology	5.968 (0.034)***	0.121 (0.005)***	0.072 (0.005)***	-0.241 (0.008)***	0.048 (0.003)***	0.032 (0.010)***	0.071 (0.008)***	0.011 (0.007)	0.62	497.66	6,084
Medium-Low	6.424 (0.063)***	0.146 (0.006)***	0.127 (0.017)***	-0.139 (0.012)***	0.084 (0.006)***	0.104 (0.018)***	0.057 (0.015)***	0.007 (0.012)	0.47	214.40	4,854
Medium-High	6.280 (0.052)***	0.191 (0.006)***	0.075 (0.021)***	-0.159 (0.011)***	0.083 (0.005)***	0.134 (0.016)***	0.020 (0.013)	0.017 (0.010)*	0.63	362.42	4,438
High Technology	6.409 (0.069)***	0.152 (0.005)***	0.122 (0.017)***	-0.070 (0.011)***	0.092 (0.005)***	0.287 (0.020)***	0.007 (0.012)	0.027 (0.013)**	0.57	176.26	3,345

Note: The standard error of each estimate is between brackets, * significant at 10%; ** significant at 5%, *** significant at 1%.

RESULTS

The dummy for **centrality** presented statistically **significant and positive values** in the segments of **Low- and Medium-Low-Technology**. With respect to the segments of Medium-High- and High-Technology, the results showed positive signs that were not, however, significant.

Table 1 - Estimation of the model (13) using as dependent variable regionalized real wages

Segments	const	ql	div	comp	den	educ	d_centralities	d_serv_complex	R2-adjust	F	N
Low Technology	5.968 (0.034)***	0.121 (0.005)***	0.072 (0.005)***	-0.241 (0.008)***	0.048 (0.003)***	0.032 (0.010)***	0.071 (0.008)***	0.011 (0.007)	0.62	497.66	6,084
Medium-Low	6.424 (0.063)***	0.146 (0.006)***	0.127 (0.017)***	-0.139 (0.012)***	0.084 (0.006)***	0.104 (0.018)***	0.057 (0.015)***	0.007 (0.012)	0.47	214.40	4,854
Medium-High	6.280 (0.052)***	0.191 (0.006)***	0.075 (0.021)***	-0.159 (0.011)***	0.083 (0.005)***	0.134 (0.016)***	0.020 (0.013)	0.017 (0.010)*	0.63	362.42	4,438
High Technology	6.409 (0.069)***	0.152 (0.005)***	0.122 (0.017)***	-0.070 (0.011)***	0.092 (0.005)***	0.287 (0.020)***	0.007 (0.012)	0.027 (0.013)**	0.57	176.26	3,345

Note: The standard error of each estimate is between brackets, * significant at 10%; ** significant at 5%, *** significant at 1%.

RESULTS

The dummy for **centrality** presented statistically **significant and positive values** in the segments of **Low- and Medium-Low-Technology**. With respect to the segments of Medium-High- and High-Technology, the results showed positive signs that were not, however, significant.

Table 1 - Estimation of the model (13) using as dependent variable regionalized real wages

Segments	const	ql	div	comp	den	educ	d_centralities	d_serv_complex	R2-adjust	F	N
Low Technology	5.968 (0.034)***	0.121 (0.005)***	0.072 (0.005)***	-0.241 (0.008)***	0.048 (0.003)***	0.032 (0.010)***	0.071 (0.008)***	0.011 (0.007)	0.62	497.66	6,084
Medium-Low	6.424 (0.063)***	0.146 (0.006)***	0.127 (0.017)***	-0.139 (0.012)***	0.084 (0.006)***	0.104 (0.018)***	0.057 (0.015)***	0.007 (0.012)	0.47	214.40	4,854
Medium-High	6.280 (0.052)***	0.191 (0.006)***	0.075 (0.021)***	-0.159 (0.011)***	0.083 (0.005)***	0.134 (0.016)***	0.020 (0.013)	0.017 (0.010)*	0.63	362.42	4,438
High Technology	6.409 (0.069)***	0.152 (0.005)***	0.122 (0.017)***	-0.070 (0.011)***	0.092 (0.005)***	0.287 (0.020)***	0.007 (0.012)	0.027 (0.013)**	0.57	176.26	3,345

Note: The standard error of each estimate is between brackets, * significant at 10%; ** significant at 5%, *** significant at 1%.

FINAL REMARKS

FINAL REMARKS

- We first observe **that there are evidences of the presence of location/MAR externalities**, mainly, **Medium-Low- and Low-Technology** segments.

- There is **still evidence of urbanization/Jacobs externalities**, but the results were most intense in the **segments of Medium-High- and High-Technology**.

- However, we found **no evidence of Porter externalities**.

FINAL REMARKS

- The results show that**, in the segments of Medium-High- and High-Technology, diverse urban centers have a positive impact on productivity**, which is not the case in the Medium-Low and Low segments.

- **Segments of Medium-Low- and Low-Technology, are subject advantages when located in a region of some centrality**. That is, traditional industries tend to gain advantages in smaller, highly specialized cities.

Obrigado!