

# MEASUREMENT OF THE AGGLOMERATION AND THE GEOGRAPHIC CONCENTRATION OF THE INNOVATION ACROSS MEXICAN STATES

VICENTE GERMÁN SOTO  
Universidad Autónoma de  
Coahuila.  
vicentegerman@uadec.edu.mx

LUIS GUTIÉRREZ FLORES  
Universidad Autónoma de  
Coahuila  
[luis.gutierrez@uadec.edu.mx](mailto:luis.gutierrez@uadec.edu.mx)

## ABSTRACT

*Nowadays, the extent of the innovation activities for the productivity and the economic growth is evident in regional economics. A large body of theoretical and empirical evidence suggests that to achieve higher well-being levels of the population it is essential to reinforce the innovation capacity of the economies. In this work we measure the extent of agglomeration and the geographic concentration across Mexican states using an endogenous innovation approach estimated through econometric techniques. The size of the regional economies to assess the importance of the scale effects is also a central concern. Using data from the Mexican states, evidence is found that innovation is geographically concentrated in the sense that biggest states (in terms of population density) tend to innovate more than smallest states. The latter implies that scale effects are present through agglomeration effects*

*Keywords: regional innovation, agglomeration, knowledge spillovers, scale effects, Mexican states.*

## INTRODUCTION

The last two decades of the 20<sup>th</sup> century witnessed the transformation in the model of accumulation of the Mexican economy. It passed from a regime of a relatively small, closed economy characterized by high government participation or direction into a mid-sized, modern open economy. In this last stage, we can distinguish two facts: firstly, the economy of the Mexican regions has maintained positive growth rates in its per capita income. Secondly, the dynamics of such growth has been dispersed, e.g. growth does not appear to be uniform across Mexican regions. Theoretically, this result can be explained by the accumulation of productive factors (mainly human and physical capital), but more interestingly, by the process of technical progress. In this matter, economists have attempted for some years to measure the effect of technological progress or innovation on productivity and/or growth. Several lessons can be learned from such attempts: for instance, human capital, research and development (R&D) expenditures carried out by firms and technical change can all be summed up to explain why advanced economies grow beyond than just considering

labor force and capital allocation efficiency. A superior ability to explode the new technology is required (Krugman, 1995) at the same time that at least two important trends in the world economy are reinforcing: first, technological innovations are becoming an ever more important contributor to economic well-being; second, the nations in the world economy are becoming increasingly open and increasingly interdependent.

So if we recognize that Mexico cannot escape to those trends, a relevant question becomes the following: How is the country doing in terms of its technological absorptive capacity, in order to enhance the growth perspectives of its regions?, does the process of knowledge transmission follows a geographically concentrated pattern, as does in the case of the more dynamic growth regions? As a consequence of trade liberalization, the transfer of technology has considerably grown due to the great increase in the number of patents. German-Soto and Gutierrez (2010) found the existence of a strong link among innovation and trade liberalization in Mexico using time series techniques. They found that until two structural breaks after 1994 suggesting an important increase of innovation activity with the trade openness. In this paper,

we seek to assess the importance of agglomeration and the geographic concentration of innovation across Mexican states.

In some sense, this area of research has not been fully analyzed in Mexico. Nevertheless, given the importance of the effects caused on growth by technology innovation diffusion, academics and policy makers have paid more attention to it. For instance, Capdevielle (2003) analyzed the technological composition of the Mexican manufacturing system by using indicators such as efficiency, production capacity and innovation capacity measured by patenting level for the period 1980-1996. He found a significant lack of integration among industrial sectors that limited the creation of synergies and the penetration of technological externalities, resulting in a very disperse national innovation system. He also found that the technological gap for technology-intensive sectors widened during the eighties. Aboites (2003) stated that NAFTA had exerted a decisive influence on the behavior of patent flows since 1994. An inverse tendency was observed: for non-residents, the flow of patents grew steadily for the whole period. In the case of residents, a negative rate of growth was detected. A possible explanation for this fact is that the need to innovate or transfer some technology caused an increase of the patenting demand while multinational corporations (MNC's) were increasing their economic activities in Mexico via their subsidiaries. More Recently, German-Soto *et al.* (2009) analyzed the factors that impact on the generation of innovation and showed the geographical relevance of such matter across the thirty-two Mexican states for the 1994-2006 periods. Their results showed that FDI and education levels were the innovation determinants in Mexico's states, with a concentration tendency towards the central part of the country.

In this work we find that as it occurs elsewhere, the innovation activities in Mexico are geographically concentrated around the center states, but also they are highly dependent of the population density. This suggests that the size and the scale of the economy are necessary to attempt to explain vast differences in rates of innovation across Mexican states.

Our results also demonstrate that rates of innovation vary greatly across states and are related to the geographic level of economic activity, because our parameter estimates suggest that bigger economies grow faster than smaller economies. It is difficult to assess the effects of political economy directed to

improve the rates of growth, but here we show that a part of the agenda of the federal government that aims to elevate the education levels and to support the researching programs, have the desired effects.

## **AGGLOMERATION AND INNOVATION: THEORETICAL ASPECTS**

### **The Extend of Knowledge Spillovers**

As a recently established fact, technological change is recognized as the primary engine for achieving economic growth and development. The fundamental element for achieving technological change is innovation, and the innovation process relies heavily on the accumulation and development of knowledge in a wide variety.

Nevertheless, the relationship between innovation and economic growth had not been carefully explored until Solow (1956) introduced it explicitly. Solow defined growth as an increase in GDP that arises from increases in capital (investments and machinery) input utilization. However, when Solow estimated his empirical model, capital accumulation accounted for roughly a quarter of the measured growth. Solow attributed the remainder of growth (the biggest share) to the "technical change" process. This result is known in the literature as the "Solow residual" and it is quite important because it placed innovation right in the core of the discussion about the determinants of growth.

Over the years following Solow's contribution, the relationship between innovation and growth has encountered many and more sophisticated ways to be modeled. Noticeably, the advances of Lucas (1988) and Romer (1986, 1990) appear to be of most significance. They introduced the concepts of human capital and knowledge spillovers, respectively, to understand economic growth. Human capital development is achieved through investments in education and training of the labor force. Particularly, Lucas (1988) included human capital in his model under the constant returns assumption, a feature that offered useful insights into the critical role of highly-skilled labor for long-term growth. Interestingly, Romer (1990) endogenized innovation in his growth model by introducing knowledge spillovers, providing a novel idea about the way in which economists would think of the economic growth process.

For the sake of simplicity, we can begin to describe how the Romer model works by recalling that firms engage in innovative activities since they expect to obtain positive profits from them. They allocate resources into

R&D expecting that the return on this investment is higher than that on any other potential allocation of the same resources. This kind of R&D investments generates two types of knowledge (Agrawal, 2002): one that is appropriable and other that is not. Appropriable knowledge refers to knowledge that a firm can use by itself excluding others from using it. Therefore, the firm obtains profits from it. On the contrary, knowledge that is not appropriable reflects the same characteristics of a public good: it is non-rivalrous (if one firm uses it, other firms can also use it) and non-excludable (no one can prevent others from using it). The more knowledge is created, the more productive R&D efforts using human capital will exist. In this way, while engaged in R&D activities, firms apply human capital to the stock of knowledge available, aiming at obtaining more profits. Of course, this same process contributes unintentionally to increase the stock of knowledge. This unintended contribution is referred to as knowledge spillovers. Nevertheless, according to Carlino (2001), knowledge spillovers can also be understood as an exchange of ideas among individuals. A knowledge spillover is an internal knowledge spillover if there is a positive impact of knowledge between individuals within an organization that produces goods and/or services. A knowledge spillover is an external knowledge spillover if there is a positive impact of knowledge between individuals without or outside of a production organization. Carlino (2001) classifies knowledge spillovers into three broad categories: firstly, the Marshall-Arrow-Romer (MAR) spillovers which denote that the proximity of firms to one another in the same, common industry affects how well knowledge travels among firms to facilitate innovation and growth. Therefore, the closer the firms are to one another, the greater the MAR spillover is. Secondly, Porter spillovers refer to local competition, in opposition to local monopolies, as the factor that fosters the pursuit and rapid adoption of innovation. Arguments of the Porter externality regarding geographical concentration and specialization are similar to those of the MAR spillovers. Thirdly, as regards the Jacobs spillover, the proximity of firms to one another in a different, diverse industry affects how well knowledge travels among firms to facilitate innovation and growth. In contrast to the MAR approach, the diverse proximity of a Jacobs's spillover brings ideas together from individuals with different perspectives to encourage an

exchange of ideas and foster innovation in an industrially diverse urban environment.

Now, the implications of Romer's model are increasing returns from growth on investments in human capital and R&D due to knowledge spillovers. Actually, this must be the case because the more human capital exists in an economy; the more value can be generated from the stock of public knowledge through R&D activities which will simultaneously increase the value of R&D. This process will originate a virtuous circle type of relationship. The discussion presented above explains why the concept of knowledge spillovers is so important to the way in which innovation relates to economic growth.

Although it seems clear that a positive outcome of growth is expected when knowledge spillovers are present in the economy, when facing the real world, one can wonder at least the following questions: Why is it that, if knowledge has public good properties, growth rates are still highly variable across countries? And, if knowledge spillovers are freely available, why has international prosperity not been uniformly achieved? The answers to these questions can have many interpretations, but one can generally argue, for instance, that not all countries have the same conditions to generate knowledge or adapt it to its production processes. In fact, if we think of innovation as the result of applied knowledge into the production process, aiming to create new production forms or new products, we must recognize that knowledge does not exist by itself (Corona, 1996). It remains codified in books, blueprints, drawings, production instruments; it is contained in human beings and institutions. In any society, knowledge grows, accumulates, enriches and preserves through individuals acting inside organizations or institutions such as firms. Additionally, knowledge can be present in a tacit form (Fischer and Varga, 2003) - the accumulated experience of workers (for example, researchers and engineers) at different levels of the innovation process, and it can also differ from person to person. Tacit knowledge is thought to be more geographically concentrated (Agrawal, 2002), a feature that has important implications for regional economic growth since, in its public good form, it can be easily and cheaply obtained by firms. Relevant studies that highlight the degree of geographically localized spillovers are, for instance, the work by Jaffe (1989) where the author examines the variance of new patents issued to federal funding at the state level, indicating that patents occur in those

states where public and private knowledge generators are the greatest. In a later work, Jaffe *et al.* (1993) infer the degree to which knowledge spillovers are geographically concentrated. The author's results suggest that patent citation is significantly localized in concordance with the localization of the organizational frame of knowledge generation such as universities. Audretsch and Feldman (1996) obtain results indicating the relative importance of new knowledge to the concentration of industrial production (local university funding vs. local industry value added). Zucker *et al.* (1998) report that local famous researchers and their collaborators can be thought of as a strong predictor of the geographical distribution of the biotechnology industry.

Agrawal (2002) mentions that there is another branch of analysis regarding knowledge spillovers where the characteristics of the firm to use the spillovers are examined. This concept is known as the "absorptive capacity" of the firm and is taken from the knowledge transfer literature that deals with firms' ability to assimilate and apply recent scientific information in their innovation development process. Cohen and Levinthal (1990) argue that firms' ability to use knowledge spillovers for their own gain relies heavily on their own R&D investment level. They conclude that R&D investments create a capacity to exploit new knowledge. Lim (2000) argues that the absorptive capacity of the firm is a function of its connectedness degree with other firms or knowledge generator organizations. He finds that the main factors which foster connectedness are: cultivating universities' relationship, hiring graduate students, participating in research associations, and, in general, any way of maintaining strong links with the scientific community.

The aspect of technology transfer is relevant in world ruled by globalization of economic activity. Particularly in Mexico, given the lack of a consolidated technology development sector, it may be considered as the most important factor to foster innovation. Viewed in this way, technology transfer is a helpful instrument to capture the innovation performance status between firms and, also importantly, between countries. There are several channels through which technology transfers are carried out. Romo and Hill (2006) mention two forms which characterize technology transfer: a direct form, implying the trade itself of production means (through the use of licenses, patents, entrepreneurial

practices and production processes), and an indirect form, related to the trade of goods, international factors of production mobility, and imports of capital goods that will adapt to local production processes. One of the most relevant forms that technology transfer takes is by means of the Foreign Direct Investment (FDI)<sup>1</sup>.

### **The Production of Knowledge**

According to Malecki (2010), almost any study about the geography of knowledge assumes away that knowledge is unevenly distributed across the economic scenario. There is a tendency to find that the innovative activities are spatially concentrated, a tendency reinforced over time. Of course, the production of knowledge is quite different than the production of goods or services. Knowledge can be shared freely but it is also very difficult to transmit to others.

In the literature, one of the main approaches to the analysis of the production of knowledge, which Malecki (2010) denominates as "standardized" or "conventional", views knowledge as the output of the combination of innovative activities of various economic agents, such as R&D laboratories or knowledge industries. Accordingly, the relationship is modeled through the knowledge production function (KPF) firstly introduced by Griliches (1979)<sup>2</sup>, whereas invention or innovation is represented by patents. Griliches (1990) provides the following definition of a patent: "A *patent* is a document, generally issued by a government authorized agency or institution, that grants the rights to exclude and to protect the production or use of a specific new device, apparatus or process for a determined number of years, which vary across countries". As it turns out, the patents system has been created with the main purpose of encouraging invention and technical progress by means of protecting such activities with the provision of a certain kind of monopoly power to the inventor. Several studies have used patent statistics to examine different aspects of knowledge and technological change. The relationship

---

<sup>1</sup> The importance of FDI as an innovation-generating factor is underlined in Alfaro *et al.* (2002). Of course, the capacity of each country to take advantage of the FDI-induced spillovers depends heavily on local conditions like environment policy regarding foreign activities, etcetera.

<sup>2</sup> A more detailed discussion of the KPF will appear later in the paper.

between patents and R&D expenditures, patents and knowledge spillovers, patents and the inventiveness rate of growth, patents and economic growth, and some others have been examined in some detail in the literature (Griliches, 1990 and 1992; Jaffe, 1989; Jaffe et al., 1993). There appears to be a consensus amongst economists that patent statistics somehow reflect the state-of-the-art technology level of a country, or its ability to embed technical change into the functioning of its economic system and its growth possibilities<sup>3</sup>. Tsur (1989) states that patent activities are a very rich and underestimated source of technology information.

Griliches (1990) offers a survey on the use of patent statistics as economic indicators, reflecting the extensive use of the US patent system to best understand how they can serve the purpose of economic growth. Mansfield (1986) shows that innovation tends to grow when the protection of the patent system and the propensity to patent have increased over time.

The knowledge production function appears to be one where the logical arguments stand: the more intensive use of innovation inputs will result in a more intensive generation of innovation indicators. Audretsch and Feldman (2003) mention that the knowledge production function works better when the unit of analysis is more aggregated (countries or industries), and it has also allowed inferring that less developed countries are associated with less innovative activities, in other words, LDC's are characterized by a paucity of new economic knowledge.

This approach has generated a vast amount of studies, for example, the recent work by Abdih and Joutz (2005), who attempt to investigate the knowledge production function in relation to intertemporal spillover effects by using cointegration techniques. Their results indicate the presence of strong intertemporal knowledge spillovers, which can be used to design better policies of R&D subsidies and achieve sustained growth. Also, the paper by Catozalla and Vivarelli (2007) tests the

possible catalyzing role of in-house R&D in a sample of Italian manufacturing firms. They categorize the sources of innovation into two broad groups: internal and external R&D, and embodied and disembodied technological acquisitions. These authors find that the innovation process is the result of a combination of diverse factors, but in-house R&D creates preconditions that will later enhance the synergy that amplifies the impacts of other innovative factors. The work by Ang and Madsen (2009) examines the extent to which growth has been driven by R&D and tests which second-generation endogenous growth model is consistent with the data from six Asian economies. The results indicate a Schumpeter type of growth, recognizing that R&D has effectively played a key role in Asian growth after World War II.

Although the KPF has been extensively used, it only focuses on aspects of quantifiable knowledge, neglecting other relevant aspects. For instance, it cannot provide quite exactly how knowledge is transmitted among individuals or firms. Audretsch and Feldman (2003) add that formal R&D data tend to ignore the complex process of technological accumulation in situations where tacit knowledge is built up.

## DATA AND METHODOLOGY

### Dataset

Our focus analysis is in the regional level of the Mexican economy, however, because the limited database only the federal entity level is considered, so data are collected for each of the thirty-one federal entity and the 1994-2006 period. Although Mexican regional system is integrated by thirty-two states, one of them, the Distrito Federal, was excluded from the database because some variables were not published for this state. The variables are collected from government institutions and they are disposing on the year-to-year basis. It means that one analysis in levels of the variables it can be carrying out, but such as we explain later, we have preferred use them like rates of growth. In this way, empirical evidence is based in the rates of growth of the periods 1994-1998, 1998-2002 and 2002-2006. One advantage of studying the innovation in terms of growing is that let us to infer the impacts by units of growth of such factors. In the econometric specification we use a measure of innovation as dependent variable; therefore this work is in line with the theoretical current of endogenous innovation,

---

<sup>3</sup> The selection of appropriate scientific and technological indicators bears several problems. Those problems include comparability, reliability and extensiveness. Some authors like Krugman (1991, p.53) argue that the empirical measurement of knowledge would be quite difficult. In fact, in his view, knowledge flows do not leave a paper trail which could make them traceable and measurable. Conversely, Jaffe *et al.* (1993) challenge Krugman's argument and mention that even knowledge sometimes leaves a paper trail behind in the form of patent citations.

such as was describe in the past section. An average patent rate per one-hundred thousand inhabitants within the state (PAT) is used as a proxy for the rate of innovation. Data on patents are collected from the Mexican Institute of the Industrial Property Rights (IMPI), while population data are collected from the population censuses published by the Mexico's National Statistic and Geography Institute (INEGI), although for the inter-census years we have lineally interpolated. Such measures of innovation are used by a number of authors and are typically found to be positively and significantly related to the rate of economic growth (see, for instance, Audretsch, 1998; Athreye and Keeble, 2002; Cheung and Lin, 2004; and Das and Finne, 2008). There are many other measures of innovation but patents has the advantage that is the most complete and proximate to innovation in the regional Mexican system on the year-to-year basis, also it has de advantage that constitute an output measure because patents are a direct measure of inventiveness and so of the innovation. Moreover, because majority of studies on innovation use patents as proxy, then it can serve with purposes of comparison. There are two, at least, potential problems with this patent variable. First, many authors of the innovation theory highlight that patents have a lag between the time an innovation is made and the time at which a patent is ultimately granted. For this reason and to reduce this limitation we have opted to use one innovation measure more extensive: the patent application. The patent application has de advantage that it has results in economic growth terms before of being granted. Also, the rates of growth of patents diminish the problem of the lag because in a lapse of time it there must be picked up something of the growth until it is granted. Maybe, one major problem is that not all innovations are equivalent in their economic significance. Some authors as Sedgley and Elmslie (2004) consider that "...the patent data have no way of distinguishing between very useful innovations and those of more dubious value", this is a more serious problem in the Mexican case because the absence of information about it. Despite this difficult we think that patent application is a preferred measure of innovative outcomes.

In our researching of the innovative outcomes we use two variables of agglomeration: population density (POPDEN) and one measure of the size of the region through the land area (AREA). According with the tradition if agglomeration effects exist then POPDEN must be important to the rate of

innovation, meanwhile if agglomeration economies do not exist then AREA must be appropriated to capture the effects in the rate of innovation. In the case of anyone results important to innovation then they must be excluded of the model. Data for these two variables are collected from INEGI and they are measured at the beginning of each period because the purpose is to assess its contribution to the innovation rates of growth. Heterogeneity between Mexican states is notorious almost of any point of view. To overcome this phenomenon we control for differences in the concentration of high technology industries through the calculus of the location quotient for those industries considered as "high technology" (LQ-HIGHTECH).<sup>4</sup> It is expected that a high concentration in these industries will lead to a higher rate of innovation. Also, we control for human capital differences including one education variable: the schooling average rate statistics (SCHOOL) obtained from the INEGI. Government subsidy to public universities as a measure of public expenditure is also considered to influence the patent rate (GSPU). These data were obtained from the ANUIES (National Association of Public Universities). Finally, due to high concentration of the innovation rate in the center states, we also have considered a dummy variable of that states belonging to the center region (Dum-Center).

### Methodology and Econometric Specification

Originally, Griliches (1979, 1990) constructs a model in which innovative output is related to knowledge founded on a knowledge production function. Being stimulated by this approach a great amount of studies have modified the baseline production function with the end to take into account additional channels of knowledge transmission and also to measure geographic spillovers of university research on regional innovative capacity. The knowledge production function for industry (region or firm)  $i$  at time  $t$  adopt the following theory specification:

---

<sup>4</sup> In this case, we use a classification carried out by Fischer and Varga (2003), who identify high tech sectors as a representation of the following industrial sector production: computers & office machinery, electronics & electrical engineering, scientific instruments, machinery & transportation vehicles, oil, refining, rubber & plastics, and chemistry & pharmaceuticals. Also Davies and Matraves (1996) suggest a similar classification to study the industrial organization in the European Union.

where  $P$  may be whichever index of innovative output, e.g. patent counts;  $R$  is one index of industry (region or firm) private spending and  $U$  is one variable or index of university research expenditures; in the Jaffe (1989) work  $X$  may be a term interacting  $U$  with a measure of proximity of university research to industry; and, finally,  $\varepsilon^*$  is an independent and identically distributed error term.

The theoretical framework suggested by equation (1) corresponds to endogenous growth models that resting under assumption that technological advancing (measured by  $P$ ) can be proxy by one adequate innovative output measure (patenting, for example). If we apply logarithms, then specification (1) is like follows:

$$P_{it} = \beta_0 R_{it}^{\beta_r} U_{it}^{\beta_u} X_{it}^{\beta_x} \varepsilon_{it}^* \quad (1)$$

$$\log(P_{it}) = \beta_0 + \beta_r \log R_{it} + \beta_u \log U_{it} + \beta_x \log X_{it} + \varepsilon_{it}^* \quad (2)$$

Other factors also are included in the specification (1) and (2) with the aim to capture regional heterogeneity, structural conditions such as education, geographical concentration of innovation and industrial output, and the possibility that scale effects are presents in the economies.

For example, Sedgley and Elmslie (2004) propose the next empirical specification to research if scale effects are significant in the U.S. states case:

where the dependent variable is in per capita terms and the equation (3) includes dummy variables of regions.

A similar specification like (3) can be essayed for the Mexican case. Recognition that the Mexican regional economy is also vastly heterogeneous, to better capture the geographical concentration of innovation this study considers some regional structural conditions such as education, industrial clustering, and the possibility that innovation capability be influenced by size of the economies (scale effects) in terms of population density and land extension.

Therefore, we use the following econometric specification to investigate if scale of economies explains the rate of innovation and also the presence of agglomeration within innovation across the states economies:

$$\begin{aligned} \log(PAT_i) = & \beta_0 + \mu_i + \gamma_i + \beta_1 \log(AREA_i) \cdot \\ & \beta_2 \log(POPDEN_i) + \sum_{j=3}^k \beta_j X_{j,k} \cdot \\ & + \alpha DUM_i + \varepsilon_i \end{aligned} \quad (4)$$

where  $X_{j,k}$  is one vector of explanatory variables (some control variables such as SCHOOL, LQ-HIGH-TECH and GSPU defined before),  $DUM$  is a dummy variable for the Center states and  $\mu$  and  $\gamma$  are the time periods effects and regional effects, respectively.

## EMPIRICAL RESULTS

We essay specification of equation (4) in three ways. The first regression will be a specification that includes only regional dummy variables and the scale variables AREA and POPDEN as independent variables. The second set of estimates allows controlling for differences in high technology industries, differences in human capital, and differences in university investment across economies. The last set of estimates excludes the scale variables with the end to test the consistency of the estimated parameters when scale effects are not considered in the model. This alternative also will allow us to assess the importance of scale effects in the innovation rate.

With the end to take into account the heteroskedasticity associated with the specification of equation (4) we have decided

$$\begin{aligned} \log(PAT_i) = & \beta_0 + \beta_1 \log(AREA_i) + \\ & \beta_2 \log(POPDEN_i) + \\ & \sum_{j=3}^k \beta_j X_{j,k} + \varepsilon_i \end{aligned} \quad (3)$$

to model the variance in a general form using the generalized method of moment (GMM) estimator. This method corrects the bias in the estimated standard errors of parameters estimates and it does not require information of the exact distribution of the disturbances.

The three essayed regressions with the equation (4) are also implemented from three perspectives: estimates for the global period

(Table 1), estimates by sub-periods 1994-1998, 1998-2002 and 2002-2006 (Table 2), and finally estimates in a panel data structure (Table 3). In this last case we use the panel data estimates with two methods: the GMM method and the two stage least squares (2SLS) method. The main reason to consider the 2SLS method is to test the possibility that the explanatory variables are correlated with the error term. It is important to highlight from the Table 3 that 2SLS estimates are not provided for the earliest period (the 1994-1998 period) because 2SLS use four years lagged values of the independent variables as instruments.

Before presenting the results it is worthwhile to comment about the expected signs of the estimated coefficients. First, the variables LQ-HIGHTECH, SCHOOL and GSPU are expected to be a positive impact on innovation because industry concentration, scholarship and federal government support at universities must surely to push the innovation activity within the states. A strong link between innovation and variables such as education, industry and university expenditure has been demonstrated in a number of studies (see, for the Mexican case, works of German-Soto, Gutierrez and Tovar, 2009; and German-Soto and Gutierrez, 2010).

Secondly, if we find that knowledge has a geographical effect then  $\beta_1$  and/or  $\beta_2$  is expected to be positive and significant.<sup>5</sup> Our first set of results is showed in Table 1, where the parameter estimates for the global period (1994-2006) are reported. The estimates for the three regressions are based in GMM method and considering a long-run vision the best adjustment of the model to the innovation rate seems to be the regression three, where scale effects are not included –the R-squared, calculated as the percentage deviation from the mean explained by the regression model, was 0.12, 0.26 and 0.47 for the regressions 1, 2 and 3, respectively.

In fact, when the innovation activity is not controlled by the differences in the economic structure none of the scale effects is significant; moreover, they result with the negative sign although not significant. However, once the structural variables are introduced scale effects become significant. This result implies two things. First, there are important structural differences between Mexican states and they are affecting the rate

of innovation. Second, the negative sign of the AREA and POPDEN variables seems to suggest that knowledge has not a geographic component, but results obtained from the Table 1 are not clear because once we exclude these two variables of the regression then the adjustment of the model is improved.

For the SCHOOL and LQ-HIGHTECH variables a positive and significant impact was estimated with bigger impact for the SCHOOL variable reflecting that the schooling average rate has elevated the innovation rate in 3.85% for each 1% of increases in the years of scholarly of the states. The index of productivity of the high-tech industry suggests a positive impact equal to 0.24% in the innovation rate for each one percent of increases in the productivity of the high-tech industries. It is observed that Center states are important to innovation rate during 1994-2006. This result confirms that we must to analyze the innovation rate controlling by these states because they experiment a path of growth notably dissimilar to other states.

Finally, we report in the bottom part of the Table 1 the J-statistic. The simplest use of the J-statistic is to test the null hypothesis that the over identifying restrictions are satisfied. The idea is that when GMM estimation is essayed then the parameters should satisfy a theoretical relation of orthogonality conditions between some function of the parameters  $f(\theta)$  and the set of instrumental variables  $z_t$ :

$$E(f(\theta)'Z) = 0 \quad (5)$$

where  $\theta$  are the parameters to be estimated. The used criterion by GMM estimation is to select that set of parameter estimates such a way the correlations between the instruments and the function  $f(\cdot)$  are as close to zero as possible. This criterion function is based in the following definition:

$$J(\theta) = (m(\theta))'Am(\theta) \quad (6)$$

where  $m(\theta) = f(\theta)'Z$  and  $A$  is a weighting matrix. Therefore, estimates are chosen to minimize the weighted distance between the theoretical and actual values.

In this exercise we employ the deviation from the mean of the explicative variables as instruments and it is possible to conclude that they are a good theoretical relation because the null hypothesis cannot be rejected for any of the regressions. Three aspects among innovation rate and its determinants are highlighted from the empirical results showed in the Table 1. First, scale effects are not

<sup>5</sup> Some studies have found a geographical component in the innovation, for instance Jaffe, Trajtenberg and Henderson (1993) and Sedgley and Elmslie (2004) in the US states case. For the Mexican case German-Soto, Gutierrez and Tovar (2009) constitutes a good and recent reference.

relevant and according with the theory the negative sign estimated in the regression two

could suggest that agglomeration effects does not exist in the states economies.

Table 1. Parameter estimates for the global period (1994-2006)

	1	2	3
Log(AREA)	-0.1503 (0.2274)	-0.7368 *** (0.2311)	
Log(POPDEN)	-0.0059 (0.2189)	-0.4484 ** (0.1731)	
Log(SCHOOL)		2.6398 *** (0.8381)	3.8579 *** (0.7309)
Log(LQ-HIGHTECH)		0.4059 *** (0.0977)	0.2462 *** (0.0812)
Log(GSPU)		-0.0846 (0.2719)	0.4325 (0.3215)
Dum-Center	0.4461 (0.4467)	0.5617 (0.4222)	1.2131 *** (0.3026)
R-SQR	0.12	0.26	0.47
J-Statistic	9.19E-30	0.1089	0.1906
Chi-SQR	0.0000	3.3777	5.9072
p-value	1.0000	0.7601	0.2062

Notes: Estimates based in GMM method. Standard errors are reported in parenthesis. Standard errors are based on White's (1980) estimator of the variance-covariance matrix. All regressions include the dummy variable for the center states. R-SQR = R squared, the percentage deviation from the mean explained by the regression model. \*, \*\*, and \*\*\*, indicate significant variable at 10%, 5%, and 1%, respectively.

Second, to judge from the explicative variables there are important structural differences between Mexican states that they are affecting the results of the innovation rate. And third, center states behavior differently of the rest of states.

It will be interesting to see if these conclusions are keeping in more shorts periods. This is the objective of the results displayed in the Table 2, where the analysis is conducted dividing the long period in three sub-periods: 1994-1998, 1998-2002 and 2002-2006 and apply the same methodology.

The empirical relation between innovation rate and some of its determinants as education, concentration of high-tech industries and university expenditure are similar to that obtained from a long run view exposing in the Table 1: the impacts are positive and highly significant almost in all the considered subperiods.

This result is according with the theoretical expectative. However, in the cases where the estimates of the AREA and POPDEN variables are significant we can see that prevails a negative relation with the innovation rate, which can be indicative that agglomeration forces are not presents in the regional innovation of the Mexican states. Also, in this set of regressions the J-statistic indicates that over identifying restrictions are satisfied except one of them: the first regression of the 1994-1998 period.

It is possible that obtained results until now (tables 1 and 2) are affected, besides other

factors, by the size of the sample, because these essays only comprise 31 observations. We know from the statistical methods that small sample can affect the robustness of the results.

With the aim to take into account this possibility we have experimented with a panel data structure. In this sense we have arranged a panel data set from a dimension of three temporal observations (the corresponding sub-periods previously defined) and 31 states. Now we apply GMM estimation and 2SLS method. Table 3 picks up the main results.

Parameter estimates by period

	1994-1998			1998-2002			2002-2006		
	1	2	3	1	2	3	1	2	3
Log(AREA)	-0.2366 (0.1957)	-0.7762 *** (0.2618)		-0.1589 (0.2248)	-0.6374 ** (0.2396)		-0.1381 (0.2323)	-0.0042 (0.2167)	
Log(POPDEN)	0.0004 (0.2632)	-0.4354 ** (0.2101)		0.0139 (0.2556)	-0.2637 (0.1865)		0.0177 (0.2275)	0.1423 (0.1945)	
Log(SCHOOL)		3.8132 *** (0.8426)	5.0539 *** (0.8054)		4.2849 *** (0.8703)	5.5012 *** (0.9068)		3.7837 *** (0.7804)	4.3713 *** (0.9098)
Log(LQ-HIGHTECH)		0.4159 *** (0.1088)	0.2838 *** (0.0854)		0.4296 *** (0.0990)	0.2556 *** (0.0701)		0.2390 *** (0.0762)	0.2471 *** (0.0691)
Log(GSPU)		-0.1332 (0.3425)	0.7482 ** (0.2985)		0.1075 (0.3182)	0.7186 ** (0.3447)		0.5348 * (0.2709)	0.3037 (0.3274)
Dum-Center	0.4572 (0.4353)	0.6531 (0.4750)	1.3031 *** (0.3275)	0.5152 (0.4870)	0.4394 (0.4978)	1.2025 *** (0.2824)	0.4217 (0.4518)	0.2052 (0.2561)	1.0412 *** (0.2953)
R-SQR	0.15	0.47	0.66	0.12	0.51	0.65	0.12	0.64	0.49
J-Statistic	0.2960	0.1481	0.2868	1.24E-31	0.1191	0.2095	4.50E-27	3.87E-25	0.1969
Chi-SQR	9.1762	4.5935	8.8907	3.8489	3.6928	6.4955	1.3953	1.1989	6.1043
p-value	0.0270	0.5969	0.1135	1.0000	0.8144	0.2609	1.0000	1.0000	0.2962

Notes: Estimates by GMM method. Standard errors are reported in parenthesis. Standard errors are based on White's (1980) estimator of the variance-covariance matrix.

All regressions include the dummy variable for the center states. R-SQR = R squared, the percentage deviation from the mean explained by the regression model.

\*, \*\*, and \*\*\*, indicate significant variable at 10%, 5%, and 1%, respectively.

We can see that 2SLS method has a total of 62 observations because values of the independent variables of the first period are used as instruments, while GMM estimation has three time periods and it is estimated with time period effects.

The estimates from the Table 3 are more consistent with the theory predictions. The adjustment measured through R-SQR is relatively high in almost all regressions. Now it is possible to see that POPDEN is positive and significant especially when other variables are included, suggesting that once we control by

structural differences the innovation rate has a geographical component in the sense that greater states in terms of population density tend to innovate and growth at major rates that smaller states. In this way, population density is established as scale effect through of which agglomeration is present among the Mexican states.

On the other hand, the AREA variable has been estimated with a negative sign in the regression 1 of both methods and it has not been significant in the other regressions. Interpreting the negative

finding for the AREA variable it is possible to say that greater economies (in terms of km<sup>2</sup>) trend to innovate to minor rates. However, for this variable it is possible that results in this direction can be affected by the fact that scale is not adequate, because it is constructed assuming that innovation is uniform in all geography

of the state, when truly, innovation and population are highly concentrated in some geographical regions as the great cities, for example. The other set of variables included in the regressions were estimated with positive sign and all significant, highlighting the relative importance of these variables for the innovative rate.

Table 3. Panel estimates						
	Method: GMM			Method: 2SLS		
	1	2	3	1	2	3
Log(AREA)	-0.1735 *** (0.0214)	0.0079 (0.0148)		-0.1485 *** (0.0079)	0.0197 (0.0162)	
Log(POPDEN)	-0.0017 (0.0144)	0.1635 *** (0.0167)		0.0158 *** (0.0014)	0.1751 *** (0.0215)	
Log(SCHOOL)		4.5037 *** (0.2240)	4.3108 *** (0.2331)		4.3957 *** (0.3871)	4.1490 *** (0.3764)
Log(LQ-HIGHTECH)		0.2443 *** (0.0078)	0.2717 *** (0.0108)		0.2455 *** (0.0079)	0.2757 *** (0.0120)
Log(GSPU)		0.5488 *** (0.0456)	0.5235 *** (0.0343)		0.5846 *** (0.0456)	0.5619 *** (0.0244)
Dum-Center	0.5083 *** (0.0401)	0.3342 *** (0.0834)	0.5579 *** (0.0803)	0.4686 *** (0.0345)	0.2409 *** (0.0274)	0.4663 *** (0.0291)
R-SQR	0.23	0.75	0.73	0.16	0.72	0.70
Observations	93	93	93	62	62	62
J-Statistic	4.70E-23	2.31E-19	1.20E-21			
Chi-SQR	4.3732	2.1442	1.1120			
p-value	1.0000	1.0000	1.0000			

Notes: Standard errors are reported in parenthesis. Panels estimated through generalized method of moments (GMM) and two-stage least squares (2SLS). GMM estimates pool has three time periods and includes time period effects. 2SLS estimates use four-years lagged values of the independent variables as instruments. For example, POPDEN in 1994 is used as an instrument for POPDEN in 1998 and so on. RS-QR = R squared, the percentage deviation from the mean explained by the regression model. \*, \*\*, and \*\*\*, indicate significant variable at 10%, 5%, and 1%, respectively.

The greater impact in the innovation rate was estimated for SCHOOL variable suggesting that rates of growth in average years of education have pushed the innovation activity in the states. The parameter estimates for this variable fluctuate about 4.1% and 4.5% of impact in the innovative rate; meanwhile the index of high-tech industries exerted an impact between 0.24% and 0.27%. It is worthwhile to highlight that university expenditure (mainly public universities) certainly has an important contribution in the rates of growth of the regional innovation. The impact of this variable fluctuates between 0.52% and 0.58%, an interesting finding because commonly these types of variables do not usually have a significant impact.

### CONCLUDING REMARKS

The innovation activity of the economies is a path through of which they can significantly improve the economic performance. The theorist of endogenous growth have emphasized that one key element to push the economic growth is the innovation level existent in their activities. Economic development generates innovation, therefore it can be considered as endogenous to the model, which suggests that rates of innovation should be pushed through channels such as education, technology and more support to investigation activities in universities and researches of the centers and institutes linked with firms and production activities.

Empirical results indicate that innovation activity in Mexico is geographically concentrated around the center states, but also it is highly dependent of the population density suggesting that size and scale of the economy are necessary to attempt to explain the vast differences in rates of innovation across Mexican states.

Our results demonstrate that rates of innovation vary greatly across states and are related to the geographic level of economic activity, because also parameter estimates suggest that larger economies grow faster than smaller economies. It is difficult to assess the effects of political economy directed to improve the rates of growth, but here we have demonstrated that part of the agenda of the federal government that seeks to elevate the education level and to support the researching programs have desired effects. With the purpose to assess the relative importance of agglomeration effects, it is necessary to include a view of scale in the model, because in the Mexican case some results are influenced by the vast extension of the greater states of the country. In this sense, it is necessary to design one

measurement of scale more refined, for example to use a measure of scale that represents the land area available in the economy.

### REFERENCES

1. Abdih Y., F. Joutz (2005) Relating the Knowledge Production Function to Total Factor Productivity: An Endogenous Growth Puzzle, *IMEF Working Paper No. 05/74*.
2. Aboites J. (2003) Innovación, Patentes y Globalización, In J. Aboites & G. Dutrénit (eds). *Innovación, Aprendizaje y Creación de Capacidades Tecnológicas*. Universidad Autónoma Metropolitana: Porrúa, pp. 163-206.
3. Agrawal A. (2002) Innovation, Growth Theory and the Role of Knowledge Spillovers, *Innovation Analysis Bulletin* 4(3): 3-6.
4. Alfaro L., A. Chanda, S. Kalemli-Ozcan, S. Sayek (2002) FDI Spillovers, Financial Markets and Growth, *mimeo*, Harvard Business School.
5. Ang J.B., J.B. Madsen (2009) Can Second-Generation Endogenous Growth Models Explain the Productivity Trends and Knowledge Production in the Asian Miracle Economies?, *MPRA Paper No. 17543*.
6. Athreye S. S., D. Keeble (2002) Sources of Increasing Returns and Regional Innovation in the UK, *Regional Studies*, 36(4): 345-357.
7. Audretsch D. B. (1998) Agglomeration and the Location of Innovative Activity, *Oxford Review of Economic Policy*, 14(2): 18-29.
8. Audretsch D.B., M.P. Feldman (2003) Knowledge Spillovers and the Geography of Innovation. In V. Henderson and J.F. Thisse (eds) *Handbook of Regional and Urban Economics, Volume 4: Cities and Geography*. North Holland, pp. 2713-2740.
9. Capdevielle M. (2003) Composición Tecnológica de la Industria Manufacturera Mexicana. In J. Aboites & G. Dutrénit (eds) *Innovación, Aprendizaje y Creación de Capacidades Tecnológicas*. Universidad Autónoma Metropolitana: Porrúa.
10. Cheung K., P. Lin (2004) Spillover Effects of FDI on Innovation in China:

- Evidence from the Provincial Data, *China Economic Review*, 15: 25-44.
11. Carlino G. (2001) Knowledge Spillovers: Cities' Role in the New Economy, *Business Review*, 4: 17-26.
  12. Catazolla A., M. Vivarelli (2007) Beyond the Knowledge Production Function: The Role of R&D in a Multi-Faceted Innovative Process, *Jena Economic Research Papers # 2007-87*.
  13. Cohen W.M., D.A. Levinthal (1990) Absorptive Capacity: A New Perspective on Learning and Innovation, *Administrative Science Quarterly*, 35: 128-152.
  14. Corona J. M. (1996) Organización, Aprendizaje e Innovación en la Empresa: Un Estudio de Caso, *Comercio Exterior*, 46 (10): 786-794.
  15. Davies S., C. Mataves (1996) The EU Market Share Matrix, pp. 34-46. In Davies and Lyons (eds.) *Industrial Organization in the European Union. Structure, Strategy, and the Competitive Mechanism*, New York: Oxford University Press.
  16. Das S., H. Finne (2008) Innovation and Co-location, *Spatial Economic Analysis*, 3(2): 159-194.
  17. Fischer M., A. Varga (2003) Spatial Knowledge Spillovers and University Research: Evidence from Austria, *The Annals of Regional Science*, 37(2): 303-322.
  18. German-Soto V., L. Gutierrez Flores, S. H. Tovar Montiel (2009) Factores y relevancia geográfica del proceso de innovación regional en México, 1994-2006, *Estudios Económicos*, 24(2): 225-248.
  19. German-Soto V., L. Gutierrez Flores (2010) Determinants of the Regional Patent Activity in Mexico: A Standardized Coefficients Model, Paper presented in the 57<sup>th</sup>. *North American Meetings of the Regional Science Association International (RSAI)*, Denver, U.S. November of 2010.
  20. Griliches Z. (1979) Issues in Assessing the Contributions of Research and Development to Productivity Growth, *Bell and Journal of Economics*, 10: 92-116.
  21. Griliches Z. (1990) Patent Statistics as Economic Indicators: A Survey, *Journal of Economic Literature*, 28(4): 1661-1707.
  22. Griliches Z. (1992) The Search for R&D Spillovers, *Scandinavian Journal of Economics* 94, Supplement: 29-47.
  23. Jaffe A. (1989) Real Effects of Academic Research, *The American Economic Review*, 79(5): 957-970.
  24. Jaffe A., M. Trajtenberg, R. Henderson (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *Quarterly Journal of Economics*, 108(3): 577-598.
  25. Krugman P. R. (1991) *Geography and Trade*, Cambridge: The MIT Press.
  26. Krugman P. R. (1995) *Development, Geography and Economic Theory*, Cambridge: The MIT Press.
  27. Lim K. (2000) The Many Faces of Absorptive Capacity: Spillovers of Copper Interconnect Technology for Semiconductor Chips, Manuscript unpublished, National University of Singapore.
  28. Lucas R. (1988) On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22(74): 3-42.
  29. Malecki E. (2010) Everywhere? The Geography of Knowledge, *Journal of Regional Science*, 50(1): 493-513.
  30. Romer P. (1986) Increasing Returns and Long Term Growth, *Journal of Political Economy*, 94(5): 1002-1037.
  31. Romer P. (1990) Endogenous Growth and Technical Change, *Journal of Political Economy*, 98: 71-102.
  32. Romo A., P. Hill (2006) Los Determinantes de las Actividades Tecnológicas en México, *Documentos de Trabajo en Ciencia y Tecnología 06-01*, Centro de Investigación y Docencia Económicas, pp. 1-63, marzo.
  33. Sedgley N., B. Elmslie (2004) The Geographic Concentration of Knowledge: Scale, Agglomeration, and Congestion in Innovation across U.S. States, *International Regional Science Review*, 27(2): 111-137.
  34. Zucker L., M. Darby, M. Brewer (1998) Intellectual Capital and the Birth of US Biotechnology Industries, *The American Economic Review*, 88: 290-306.

#### ACKNOWLEDGMENTS

Authors acknowledge financial support by Autonomous University of Coahuila.

