Regional Impacts of High Speed Rail in China

Working Paper 2

Spatial proximity and productivity in an emerging economy: econometric findings from Guangdong Province, People’s Republic of China

June 30, 2013

OFFICIAL USE ONLY
NOT FOR CIRCULATION

Final
V6.0

World Bank Office, Beijing 世界银行驻华代表处
Preface

This Report has been prepared for the World Bank by Dr. Ying Jin, Mr. Richard Bullock, and Dr. Wanli Fang. This Bank Team was first led by John Scales (Transport Sector Coordinator), then by Gerald Ollivier (Sr. Infrastructure Specialist).

This report has been drafted as part of evaluation of the NanGuang Railway Project (P112359) and completed as part of the Technical Assistance activity called “Impact of High Speed Rail on Regional Economic Development” (P143907). This activity aims at developing a standard approach to identify and quantify regional economic impact of High Speed Rail (HSR) projects, extending beyond traditional economic benefits associated with reduction of transportation costs.

The econometric work reported in this paper is carried out as part of a wider investigation of the concept of regional development benefits of transport improvements. The paper presents the econometric findings that are based on the most detailed spatial economic data that is available in Guangdong Province in Southern China, with an aim to analyze and monitor the regional economic effects of recent and on-going major transport improvements.

Acknowledgement

We acknowledge the help, cooperation and information provided to the Bank team by the municipal government of Yunfu during the field trip in April 2010. We are grateful to Mr. Paul Amos for his pertinent comments on this study.

Disclaimer

The findings, interpretations, and conclusions expressed in this volume do not necessarily reflect the views of the Executive Directors of the World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of the World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

All rights reserved/Rights and Permissions

The material in this publication is copyrighted. Copying and/or transmitting portions or all of this work without permission may be a violation of applicable law.

©2013 The International Bank for Reconstruction and Development / The World Bank
1818 H Street NW
Washington DC 20433
Telephone: (202) 473-1000
Internet: www.worldbank.org E-mail: feedback@worldbank.org/gollivier@worldbank.org
Contents
Summary ................................................................................................................................. 4
1 Introduction ........................................................................................................................ 6
   1.1 Skepticisms .................................................................................................................. 9
   1.2 New Economic Geography studies ............................................................................. 10
   1.3 Aim of this paper ........................................................................................................ 12
2 Theoretical framework ...................................................................................................... 13
   2.1 Basic assumptions ....................................................................................................... 13
   2.2 Basic model form ......................................................................................................... 14
   2.3 Definition of economic mass ...................................................................................... 15
   2.4 Definition of control variables ................................................................................... 19
3 Data .................................................................................................................................. 20
   3.1 Employment and earnings ......................................................................................... 20
   3.2 Zonal economic mass ................................................................................................. 21
   3.3 Zone specific control variables .................................................................................. 23
   3.4 Preliminary cross-section data analysis .................................................................... 24
4 Results from econometric models ..................................................................................... 28
   4.1 Overview of the models tested ................................................................................... 28
   4.2 Cross-section regressions for 2005-2008 ................................................................ 28
   4.3 Time-Series Regressions for 1999-2009 ................................................................ 30
5 Summary of findings and further work ............................................................................ 33
   5.1 Summary of findings ................................................................................................... 33
   5.2 Further work ............................................................................................................... 34
Bibliography .......................................................................................................................... 35
Summary

A growing body of recent literature in economics and regional science suggests that there is a statistically significant correlation between spatial proximity and productivity. The findings are corroborated by economic theories that account for urban agglomeration effects, particularly the conjecture that improved spatial proximity raises productivity. However, the empirical findings are highly context-specific. The magnitude of effects appears to vary greatly across geographic locations and stages of economic development.

In this paper we investigate the relationship between spatial proximity and productivity in Guangdong Province, China. Guangdong is a leading regional economy and its ways of doing business are being widely emulated in other provinces in China. We assemble economic data by county and urban district for 1999-2009 and develop correspondingly detailed business travel cost and time matrices, so that the effects of transport accessibility on productivity can be discerned at a level that is appropriate for appraising major regional infrastructure projects such as expressways and the high speed rail. So far as we are aware, this is the first study to provide such empirical underpinnings for major transport project appraisal in China. The methodology is theoretically rigorous, yet it is operational with modest data availability; it opens a new avenue for assessing the evidence of agglomeration effects across the emerging economies.

We start with existing ways to measure spatial proximity and existing New Economic Geography models that relate hourly earnings of workers to spatial proximity. We test new measures of spatial proximity that are more consistent with known trade and travel behaviour, define control variables that address other key influences on productivity including spill-over from neighbouring areas, and use dynamic panel-data models to control for endogeneity that is a natural component of agglomeration effects.

A comprehensive set of cross-section, pooled and time series regressions find that there is a stable and statistically significant relationship between spatial proximity and productivity in Guangdong. The results show that for Guangdong at its current stage of development the elasticity of productivity with respect to spatial proximity is around 0.14, which implies that doubling the size of the economic mass of an urban district or county is associated with 10% increase in productivity. This is considerably higher than the range of values for predominantly developed economies where ‘doubling city size seems to increase productivity by … roughly 3-8%’ (Rosenthal and Strange, 2004). The findings are in line with theoretical expectations, are corroborated by existing studies using other methods, and provide the first empirical evidence for productivity effects of major transport projects in China.

We should make it clear that the findings are as yet associated with considerable uncertainties. By the very nature of agglomeration, it is difficult to distinguish precisely spatial proximity effects from other influences. Furthermore, productivity elasticities may change over time. At
the end of this paper we consider further empirical modeling and micro-level surveys that can address these issues in China and other emerging economies.
1 Introduction

Railway construction in China has attracted worldwide attention especially the expansion of high-speed railways. More than 9,300 km of high-speed railways are in operation in China (December 31, 2012), and an additional 8,700 km is expected to be completed by 2015.

The World Bank’s China Transport team in Beijing has initiated research into regional economic impacts of improvements of high-speed rail. The first three years of the operation have seen the new HSR competing strongly on short to medium distances routes up to 1000km; generated traffic often accounts for more than half of the total, which is remarkably high among the world’s HSR networks (Bullock, Salzberg and Jin, 2012).

However, there is insufficient monitoring data to enable us to carry out specific regional impact analysis on China’s new HSR network yet. Meanwhile, the economic assessment of HSR investment proposals require urgently evidence-based investigation of their regional effects, particularly regarding the more controversial branch lines off the main HSR network. To meet this need, our investigation here adopts a more general approach through relating spatial proximity and productivity, which enables us to use economic and transport data from the recent past to measure productivity effects of transport investments.

Spatial proximity is a result of both concentrating human activities in one location and, more relevant to a contemporary economy, connecting locations with fast transport services such as expressways or HSR. Do improved transport infrastructure and services contribute in any significant way to productivity growth? Kopp (2007; 2012) shows that doubling road stock in a country will lead to about 10% growth in total factor productivity in Western Europe. For an emerging economy like China, empirical evidence has just started emerging. For instance, Roberts and Goh (2012) show that distance has a significant role in determining spatial productivity disparities in Chongqing municipality. Roberts, Deichmann, Fingleton and Shi (2012) show that China’s national expressway network has brought sizeable aggregate benefits to the Chinese economy, although its impact on regional disparities may be contingent upon factors such as migration.

Nevertheless, in contrast to the considerable volume of research on the relationship between spatial proximity and productivity in the OECD countries, there are few geographically detailed econometric investigations of this relationship in China and other emerging economies (see, for example, comprehensive reviews of urban agglomeration studies in Rosenthal and Strange, 2004 and Melo et al, 2009). This working paper aims to start filling this gap in that literature by developing a methodology that is theoretically rigorous but can be made operational with modest data availability typical in the emerging economies.
Before delving into the details of econometrics, it is helpful to review the big picture of regional developments in China. The NASA nightflight picture of East Asia (Figure 1) provides a glimpse of the main urban agglomerations: the apparent concentration of human activities as shown by the mass of light emitted in and around the mega-city regions of Beijing, Shanghai, and the Pearl River Delta in Guangdong Province corroborates the statistics and daily experience about these Chinese mega-city regions: high capital investments, better educated work forces, clusters of productive and innovative industries, relative ease of encountering industry and technology leaders, higher per capita earnings, and above all, higher business productivity when measured in per employee output and earnings.

Of the three mega-city regions in China, Guangdong province in the south seems to have most to offer in such an investigation of spatial proximity and productivity. It contributes to the highest provincial share of national GDP for more than two decades. Its ways of doing business are being widely emulated by other provinces in China, thus are likely to represent what is to come in the rest of China. Its land boundaries consist primarily of mountain chains which makes it straightforward to delineate a study area boundary for this investigation.

Furthermore, the patterns of Guangdong’s spatial development may be informative. Guangdong contains three Special Economic Zones (SEZs) out of a national total of four in the first wave that was announced by the national government in 1979. Those SEZs are expected to lead economic growth because of their business potential, overseas trade links and (equally applied) special policy incentives. The Guangdong SEZs however have had markedly different growth trajectories:

1. Shenzhen, which is next door to the largest city in the area at the time, Hong Kong, flourished: it grew from a sleepy border town to a metropolis of over 10m residents, and its annual average growth rate of GDP during 2000-2008 was 15%.
2. Zhuhai, which is further from Hong Kong but adjacent to Macau (a sizeable town dominated by international tourism) had a growth rate of 13% per year in the same period.
3. Shantou, the third SEZ in Guangdong, which was a well established historic town and had strong links to overseas Chinese communities but was more than 450km away from the centres of regional economic activity represented by Hong Kong and Guangzhou (the provincial capital), had the lowest GDP growth among all Guangdong municipalities (9% per year for 2000-2008). Transport links to the rest of the province had been a problem, but the relatively low growth rates did not help the city to gain project investments – Shantou did not get any expressway connection until 2003.

---

1 According to Wang and Zheng (2008, pp50-51), Guangdong’s contribution to national GDP was 10.7% over the period 1981-2005, which was the highest amongst China’s provinces and provincial-level municipalities.
2 The fourth one was Xiamen (Amoy), which is in the neighbouring Fujian Province.
3 All data quoted comes from Guangdong Statistics Yearbooks.
Meanwhile, those municipalities that are close to Guangzhou and Shenzhen, such as Foshan, Dongguan, Zhongshan and Qingyuan achieved the highest GDP growth rates (16-19%) over this period.

For those who are used to dealing with low GDP growth rates in developed countries, 9% per year in Shantou may look just as admirable as 15% per year in Shenzhen. This is not the case in Guangdong. As GDP growth in this period has generally a large property investment component, the differences between a single digit and double digit annual growth rate could imply significant productivity growth differentials.

It is clear that the growth patterns are not all down to spatial proximity. It would be of significant policy as well as academic interest to know whether spatial proximity engendered by transport improvements has played any role. To date, there is little quantified understanding of whether spatial proximity contributes to higher productivity within China. The existing studies tend to explain high productivity by endowments of natural advantages, capital investment, labour skills, industrial composition, and foreign trade links.
1.1 Skepticisms

Although recent empirical studies have found a statistically significant relationship between spatial proximity and productivity in many developed countries, China is not a developed country. The bulk of its economy in the more developed regions such as Guangdong consists of manufacturing and local commerce, which are a far cry from the knowledge-based industries in the developed world. Guangdong, albeit being one of the richest provinces in China, had a per capita GDP of US$6500 per capita in 2008, which in real terms is equivalent to the level of the US per capita output in the 1930s. The primary and manufacturing industries, mostly low-tech and labour intensive, account for over 70% of the provincial output, and the high-end R&D and business services are a small, unknown fraction of the tertiary sector output. Empirical evidence for the developed economies may not therefore be transferrable to Guangdong or elsewhere in China.

There is also some doubt with the impact of transport infrastructure improvements on the productivity of manufacturing industries in the developed countries. For instance, Hulten and Schwab (2000, pp.158-159) find that their work on the US manufacturing industries during 1970-1986 ‘leaves little room for convergence explanations of regional growth that rely on the technological diffusion or learning-by-doing, or for endogenous growth explanations that rely on increasing returns to scale or the differential growth of public capital’. They consider that their findings are

‘consistent with a model of regional growth in which the location and scale of economic activity are strongly influenced by historical evolution and geographical factors: i.e. the US developed from East to West, with the South initially specialized in agriculture, the North in commerce and manufacturing, and the Midwest, with its resource endowments, in manufacturing and agriculture. In this paradigm, the overall growth and structural changes in the economy (e.g. the huge increase in output per worker in the economy as a whole between 1880 and 1930, and the decline in the importance of the agricultural sector) unleashed forces that, at the level of regional economies, created significant factor market disequilibria: an excess supply of labour in the agricultural and resource regions of the South and West, but also opportunities for capital formation in those regions, which, in turn, raised the demand for manufacturing labour.’

Nevertheless, Hulten and Schwab (2000, p.159) think that there is an important caveat to the explanation above: they make it clear that their finding ‘does not mean that public capital formation4 is irrelevant. Indeed, it is likely to have played an essential role in facilitating the

---

4 e.g. infrastructure and other investment that engender spatial proximity and agglomeration.
movement of capital, labour, and intermediate inputs among the regions, and thus enabled the main sources of differential regional growth’.

Furthermore, the contribution of infrastructure and other investments that facilitate positive externalities (or public capital as mentioned above) such as spatial proximity may be more or less effective depending upon the stage of development. Hulten and Schwab (2000, p158) offer a possible explanation of their work on data from 1970-1986: their interpretation is that manufacturing technology and organizational practice had already diffused widely throughout the country before the start of this period in the US. However, they argue (p159) that excess returns due to spillovers are not an important component of the overall rate of return, at least for manufacturing industry, and therefore cannot be used to rationalize the very large rates of return to infrastructure found in the US, e.g. by Aschauer (1990).

1.2 New Economic Geography studies

Recent years have seen a growing body of research on the relationship between spatial proximity and productivity, and more generally, the effects of agglomeration economies. The arguments are primarily built upon the New Economic Geography literature, which gives due recognition to (1) consumers’ and producers’ love of variety in their use of products and services, (2) increasing returns to scale in production and (3) the importance of transport costs in shaping the economic landscape. This has led to theoretical models that can identify reasons why modern firms tend to concentrate production close to large markets, i.e. the ‘agglomeration’ effects. Empirical studies have so far built up a substantial body of evidence which suggests that production and income are correlated with spatial proximity in the way suggested by the theories. Ciccone and Hall (1996), Rosenthal and Strange (2004) and Melo et al (2009) provide systematic surveys of the empirical evidence, and Redding (2010) reports more recent progress in empirical research and remaining issues with identification and interpretation.

Given the New Economic Geography’s original preoccupation with international trade (see Krugman, 1991a; 1991b), it is not surprising that there has been extensive empirical investigation in that field. For example, Redding and Venables (2004) use bilateral trade data to estimate market capacity, supply capacity, and the determinants of transport costs, and construct measures of market and supplier access that are consistent with the theoretical models; these measures are found to explain a substantial proportion of the cross-country variation in per capita income and manufacturing wages, after controlling for a large number of other influences. More recently, Mayer (2008) shows that the correlation between per capita income and market access holds not only in the cross-section but also in the time-series analysis.

At the inter-regional and city scale, theoretical models emerged about a decade after the initial trade models (see Fujita, Krugman and Venables, 1999; Fujita and Thisse, 2002). Empirical studies followed. Rice, Venables and Patacchini (2006) outline an analytical framework within which interactions between the different aspects of regional inequality in per employee productivity can be investigated econometrically using aggregate data. Kopp (2007; 2012) uses
a panel data model to address the issue of endogeneity, and identifies contribution from transport investment to productivity, showing that that doubling road stock in a country will lead to about 10% growth in total factor productivity in western Europe. Combes, Duranton and Gobillon (2008) and Venables (2010) develop general frameworks to investigate respectively the sources and mechanism that lead to wage disparities across regional labour markets through sorting and self-selection. Graham and Kim (2008) investigate the relationship between spatial proximity and productivity using a large sample of financial accounting information from individual firms in the UK. Meré and Graham (2010) further investigate the effect of firm level heterogeneity and non-random sorting of firms across space using a dataset from New Zealand.

For the emerging economies, Deichmann, Kaiser, Lall, and Shalizi (2005) distinguish between natural advantage, including infrastructure endowments, wage rates, and natural resource endowments, and production externalities that arise from the co-location of firms in the same or complementary industries in their examination of the aggregate and sectoral geographic concentration of manufacturing industries for Indonesia. Lall, Gun Wang, Hyoung and Deichmann (2010) differentiate local and national infrastructure supply in India, and find that a city’s proximity to international ports and highways connecting large domestic markets has the largest effect on its attractiveness for private investment.

In China, there has been a growing volume of literature that associate productivity benefits with agglomeration in Chinese cities and city regions (e.g. The World Bank, 2006, p145, Lu et al, 2007, p163, The World Bank, 2009, pp126-129). Using two nation-wide Censuses of Establishments of 1996 and 2001, Lu (2010) outlines the spatial distribution of economic activities across China and finds through multivariate analysis that, during that period, the micro-economic explanations of agglomeration do not work well with the publically owned institutions, although they do work well with the non-publically owned institutions. Roberts, Deichmann, Fingleton and Shi (2011; 2012) use counterfactual analysis based on a general equilibrium model to show that China’s national expressway network has brought sizeable aggregate benefits to the Chinese economy, although its impact on regional disparities may be contingent upon factors such as migration.

These studies have shed an important light both on the statistical relationship between spatial proximity and productivity, and on a variety of complex issues of empirical modeling. Nevertheless, the studies have also shown that such statistical relationships may be highly context specific. Melo et al (2009) carry out a meta-analysis of 729 measurements of agglomeration effects from 34 studies, and conclude that ‘agglomeration estimates for any particular empirical context may have little relevance elsewhere’.

At the heart of the difficulties of empirical measurements is the very nature of agglomeration as a process of circular cumulative causation, which has become known since the work of Gunnar Myrdal: Agglomeration propels endogenous growth – higher productivity leads to higher wages, which attract employees of a higher caliber, which in turn draws in new investment, more
productive technologies and so on; these lead to a new round of productivity growth. Conventional, instrumental variables are used to overcome endogeneity issues in regressions; but by its very nature, agglomeration studies rarely have good instrumental variables for dealing with cumulative causation (Redding, 2010).

1.3 Aim of this paper

The aim of this paper is to start quantifying the extent to which improved spatial proximity as a positive externality contribute to productivity growth in China. It is clear from the discussions above that there are many challenges.

First, a rigorous theoretical framework is required. Within as well as across the economic regions in China, there are enormous variations in per capita productivity; The influences that give rise to such patterns of spatial disparity are extremely complex: they involve the interplay of a multitude of historic, economic and socio-political forces in a dynamic process within which there is a great deal of serendipity, uncertainty and chance. In order to gain an insight into the role of spatial or transport costs within the context of this process we need a theoretical framework that takes due account of the complexity and at the same time offers an opportunity to explore it analytically and systematically. The New Economic Geography models that are reviewed above would seem to be the most appropriate starting point given that their assumptions about product varieties, scale economies and transport costs match more closely than neo-classical location models with the outlook of the business communities in Guangdong (for further details see Working Paper 1 on Yunfu).

Secondly, the empirical work needs to be built on a thorough understanding of the data available. Even in the relatively developed Chinese regions such as Guangdong, only a limited amount of economic data is available, often with restricted spatial resolution. The rapid structural changes in China’s economy make it difficult to have a long and consistent time series. An emphasis must be placed on a thorough understanding the data. In order to gain insights into the data, it would seem necessary to carry out micro-level case studies of local firms and institutions to examine how firms and institutions are affected by spatial proximity in their day to day operations. These micro-level case studies, carried out in one of the peripheral municipalities in Guangdong (see Working Paper 1) may also shed light on the process and mechanisms of cumulative causation.

Thirdly, given that a modern, open economy has a multitude of trade linkages reaching far and wide in geographic expanse, administrative regions are more often than not inappropriate case study areas. It is however fortunate that, to date, the natural geography of Guangdong Province (delineated by mountain chains along the northwest, north and northeast borders, and the coast in the south) makes it a relatively self-contained economic region in terms of the resident labour market, transfer of technologies and know-how, political and legal administration, and cultural conventions/customs (Lang, 2006). This enables us to focus on the provincial data rather than
having to incorporate datasets from different provinces (as would be necessary for the metropolitan areas around Beijing and Shanghai).

Finally, since the question of spatial proximity and productivity is of general interest across the emerging economies in the world, a general analytical approach is called for, particularly if it enables comparisons across regions in the developing as well as the developed world.

The aim of this paper is to address these challenges through an econometric study of Guangdong. Section 2 below considers the theoretical framework, building on models that have already been established in developed countries. Section 3 discusses the data. The main econometric results are reported in Section 4. Section 5 concludes.

2 Theoretical framework

The basic framework we adopt in this paper follows the general approach of New Economic Geography to examining spatial costs and productivity (Fujita et al., 1999; Redding and Venables, 2004). In particular, we take the analytical frameworks that have been put forward by Rice, Venables and Patacchini (2006) and Combes, Duranton and Gobillon (2008) and relate employees’ earnings to spatial proximity and control variables. We then extend their frameworks through both alternative measurements of spatial proximity and the dynamic panel approach in the development of the empirical models for Guangdong.

2.1 Basic assumptions

Our first assumption is that the economy in Guangdong is an open, market-oriented one in terms of business operations. Although this hardly seems a novel claim for anyone who knows Guangdong, it is important to acknowledge the role central planning has played in shaping the contemporary economic landscape of Guangdong, especially in the decisions to integrate the economies of Guangdong and Hong Kong, and develop new cities such as Shenzhen and Dongguan in that process\(^5\). Nevertheless, we postulate that the businesses, once located in the province, are profit-seeking to an extent that is comparable with the capitalist economies in the developed world.

More specifically, following Rice, Venables and Patacchini (2006) we assume that the businesses in Guangdong operate under perfect competition and constant internal returns to scale and face the same price of capital everywhere, that good land and floorspace availability in the last 30 years have given firms the freedom to choose where to produce and at what scale, and at equilibrium price equals unit cost (including returns to capital) in all activities in all locations.

\(^5\) Shenzhen has grown from a small town into a metropolis of 10m people since its establishment as a Special Economic Zone in 1979, meanwhile Dongguan grew into a coonurbation of 7m people from a collection of small villages and towns over an area of 2500 sq km.
If we subdivide the study area into a number of different spatial units (‘zones’ below), then each zone contains workers of different skills and occupational types. We further assume that labour productivity is zone-specific, and these productivity variations apply equally to all skills or occupations. The productivity variations may be a physical productivity difference or a value effect, as would be the case if, for example, in one zone all output prices were higher or all non-labour input prices lower.

It follows from these assumptions that any spatial variations in labour productivity will be equal to spatial variations in nominal earnings (i.e. wages plus employees’ social costs). The mobility of production activities bids up earnings in high productivity zones. Furthermore, under the assumptions above, spatial earning differences are proportionately the same for all skill or occupation types. No production activities have an incentive to move, as all earn zero economic rent in all zones. The production structure of each spatial unit is either determined directly by the skill or occupation mix of the labour force (if there are as many skill or occupation types as production activities) or is indeterminate. Labour can move between zones. This will bid up land and property prices in high productivity and high earning zones until the real income of each and every skill or occupation type is the same everywhere. There is therefore an equilibrium in which firms and workers are fully mobile, and the ultimate beneficiaries of spatial productivity differences are the property owners.

Thus at the equilibrium the nominal earnings of each type of worker vary across zones, and these variations are equal to the productivity differences between zones. This makes it possible to measure productivity through the variations in per worker nominal earnings.

Rice, Venables and Patacchini (2006) point out that the standard assumptions above give a benchmark case. Relaxing them adds more detail but does not change the main conclusion. For example, spatial productivity differences may be greater for some types of workers or for some activities than others, in which case the model would provide a theory of regional specialization. In this paper we follow the standard assumptions as current data availability in Guangdong does not yet allow the differentiation of workers or activities by type among the zones.

2.2 Basic model form
The hypothesis regarding regional productivity differences is that increasing external returns cause labour productivity in any given firm to be high in regions that have better access to other firms, labour pool, and other inputs – or to put it more generally – regions that have a large, aggregate economic mass. The main mechanisms that underlie such effects are, following Fujita and Thisse (2002): (1) technological externalities – firms learn from co-presence with other firms in related activities, so they can innovate and implement new technologies efficiently; (2) thick markets for labour and other factor inputs – they work more efficiently by having lower search costs and generating improved labour market matching between employers and the labour force, and also improved matching for other factor inputs; (3) firms gain from having good access to
their customers, thus enhancing competition among the producers and, providing a spur to product differentiation and innovation.

In other words, these mechanisms have the potential of raising the productivity of a worker of a given type in a given job through accelerated technology-learning across firms, better match between jobs and personal skills and aptitudes, and innovation in technology including business management.

The underlying empirical model can thus be presented in a general form

\[ y_i = f(M_i, X_i) \]  \hspace{1cm} (Eq. 1)

Where \( y_i \) is a measure of per worker income or productivity in zone \( i \), and \( f(M_i, X_i) \) is a function of the economic mass of zone \( i \), \( M_i \) and a set of control variables \( X_i \) that reflect the zone specific characteristics that are also believed to affect per worker income and productivity. Below we define the specific functions of economic mass and the control variables in turn.

### 2.3 Definition of economic mass

As stated above, economic mass (‘EM’ below) measures the level of market access to economic activity in any given location. Since firms today interact not only with local firms in the home zone, but also to an ever increasing extent with other zones within a radius that is dependent upon among other things the ease of transport, the EM of a given zone (‘home zone’) is a sum of the measures of market access to each relevant zone modulated by the economic distance between that zone and the home zone. In other words, the intensity of interactions between firms, e.g. information sharing, labour pooling, competition etc, are weighted by a suitable measure of the cost of travel between all relevant zone pairs.

Obviously, there are different ways to measure both the level of economic activity and the economic cost of travel. We first review two specific measurements of the economic mass that have been used for empirical analysis elsewhere, because they underpin empirical models that are potentially comparable with those we aim to develop here for Guangdong.

**Economic Mass Type A**, which is based on zonal number of employees and generalized cost of car travel as defined by Graham and Kim (2008). This EM measure underpins the models that support the UK DfT’s assessment of agglomeration effects of major transport projects (UK DfT, 2006). The same EM measure has been adopted in the assessment of the World Bank loan projects of the Guiyang-Guangzhou and Nanning-Guangzhou High Speed Railways.

**Economic Mass Type B**, which is based on zonal levels of the working age population and car travel times as defined by Rice, Venables and Patacchini (2006). The extent of spatial aggregation in their analysis of Great Britain is comparable with the empirical analysis that can be permitted by the available data in Guangdong.
The two EM measures above are isotropic in the sense that trade linkages between any cities, towns and so on are considered in an identical way. In fact, this has been a common approach in the wider New Economic Geography literature. It is nevertheless inconsistent with the Central Place geography (as originally defined by Christaller, 1933) where the cities and towns are central places of different orders in a regional hierarchy, and the linkages between different orders often tend to be stronger than those among centers of the same order.

This is not a criticism upon the existing EM measures, because they have largely been defined for regions of developed countries where the inter-city and inter-regional transport networks today are so well connected that they enable nearby central places at the same level of hierarchy to specialize and cross-trade to an extent that was not seen in Christaller’s time. Extensive analyses of inter-city and inter-regional travel in Europe and Australia during the 1960s and 1970s indicate that the spatial patterns of travel in that era still exhibit features of the central place hierarchies (Bullock, 1980). Our field work in Guangdong have also shown that regional hierarchies are important when firms consider their suppliers, markets and linkages for technology transfer (see Working Paper 1).

For this reason, we will test a further alternative formulation of the economic mass, based on our empirical analysis of trip distribution patterns in China. We name this alternative EM measure Type C and will discuss it below after reviewing the precise formulae of Types A and B.

**Economic Mass - Type A**

Graham and Kim (2008) defines the economic mass as

\[ M_i = \sum_j \left( \frac{E_{ij}}{g_{ij}} \right), \text{ for all zones } j \text{ including } j = i \quad \text{(Eq. 2)} \]

where

- \( i \) Location of the ‘home’ zone, for which the economic mass is computed
- \( j \) zones in the study area, including \( j = i \)
- \( g_{ij} \) Cost of travel between \( i \) and \( j \), which may include time and monetary costs.
- \( E_j \) A measure of economic activity in zone \( j \).
- \( \alpha \) A parameter that controls the distance decay effect; it is set to 1 in Graham and Kim (2008)
We note that with this measure, the calculation of economic mass includes the contribution from the home zone (i.e. for $j = i$). However, for travel within the zone (i.e. $g_{ii}$), it is difficult to define a precise distance or cost of travel. Given that the contribution of $\frac{E_j}{g_{ii}}$ to $M_i$ can be large in the case of a dense employment zone, we split the EM measure into two components in order to test the stability of the model without $M_i = \sum_j \left( \frac{E_j}{g_{ij}} \right)$:

$$M_i^{\text{Ext}} = \sum_j \left( \frac{E_j}{g_{ij}} \right), \text{ for all zones } j \neq i \quad (\text{Eq. 3})$$

$$M_i^{\text{Int}} = \frac{E_i}{g_{ii}}, \text{ for home zone } i \quad (\text{Eq. 4})$$

It goes without saying that the economic mass of location $i$ increases if:

1) there is an increase in the level of economic activity in $i$,

2) there are decreases in the generalized costs of travel between $i$ and $j$ (e.g. through some transport intervention).

All being equal, increased level of traffic congestion or dispersion of economic activity around a zone will reduce its economic mass.

We will test this EM measure empirically below as it is originally defined. In addition, we will also test for $M_i^{\text{Ext}}$ to verify stability of the empirical models – in this case for zones $i$ that are not major metropolitan areas.

**Economic Mass - Type B**

Rice, Venables and Patacchini (2006) proposed a way to calculate the economic mass that take account of the distance decay effects that are observed in travel:

$$M_{ik} = \sum_k \left( P_a e^{-\theta(T_k - T)} \right) \quad \text{for all travel time bands } k \quad (\text{Eq. 5})$$

where

- $i$ Location of the ‘home’ zone, for which the economic mass is computed
- $k$ Travel time bands (i.e. ranges), and for intra-home zone travel, $k = 0$
\( \theta \) An empirically estimated parameter to approximate the rate of decay in the influence of economic activity as travel time increases

\( T_k \) The upper limit of travel time for band \( k \) in minutes; \( T_k = 30, 40, 50, 60, \ldots 120 \) minutes

\( T^\ast \) A constant, set to be 30 in the model

\( P_{ik} \) A measure of economic activity - the working age population is used as a proxy – within each travel time band that is reachable from a given home zone \( i \)

This EM measure makes it possible to investigate travel time decay impacts through the introduction of travel time bands. However, the use of the exponential function requires non-linear regressions. Rice, Venables and Patacchini (2006) have grouped the zonal EM measure into travel time band measures in order to derive reasonably robust regression results.

There are issues associated with breaking down continuous distribution of employment into discrete times bands: the accuracy of this measurement depends on the areas of administrative units, because the employment within each administrative units will be assigned to a travel time band based on the travel time from/to the center of each units. The larger the administrative units are, the less accurate are the banded EMs. Therefore, this method tends works better at the detailed geographic level.

**Economic Mass - Type C**

Analysis of a wide range of regional travel patterns in Europe and Australia during the 1960s and 1970s shows that they are dominated by trips between the different levels of the regional hierarchy (Bullock, 1980). An index of influence of zone \( j \) to home zone \( i \) can be defined as:

\[
I_{ij}^0 = E_j^\alpha f\left(t_{ij}\right) \quad \text{(Eq. 6)}
\]

where

\[
f\left(t_{ij}\right) = e^{-2t_{ij}} \quad \text{for} \quad t_{ij} < T^\ast
\]

\[
f\left(t_{ij}\right) = at_{ij}^{-\beta} \quad \text{for} \quad t_{ij} \geq T^\ast
\]

\( t_{ij} \) is the travel time, and \( \alpha, a, \lambda, \) and \( \beta \) are parameters to be estimated subject to the constraint

\[
e^{-\lambda t_{ij}} = at_{ij}^{-\beta} \quad \text{when} \quad t_{ij} = T^\ast
\]
An index to reflect the ease or otherwise of the spread of innovation through personal contact in Guangdong can then be constructed by using \( I_{ij}^0 \) to identify the hierarchy. This involved the following steps:

1. For each centre, identify all other centres with a greater number of functions, using aggregate GDP as an indicator. These centres are then candidates to be the next link in the hierarchy.

2. For each candidate centre, calculate \( I_{ij}^0 \).

3. Select the candidate which has the highest \( I_{ij}^0 \) value

4. Repeat steps (1) - (3) until it terminates at the largest regional centre (in our case Hong Kong).

An index to reflect the potential for innovation for each centre was then constructed from the centres included in its hierarchy:

1. The potential from the next link in the hierarchy (centre j at level 1 say) was constructed relative to the home zone potential (\( Q_i = E_i^\alpha f(t_i) \)) as

\[
I_{ij}^1 = \frac{(E_j / E_i)^\alpha}{f(t_j) / f(t_i)}
\]

The numerator gives the additional functions that are available at the next level of the hierarchy whilst the denominator discounts this influence for the difficulty of access.

2. The calculation is then repeated for the next centre in the hierarchy (centre k at level 2 say) with the potential for innovation at this level constructed as

\[
I_{jk}^2 = \frac{(E_k / E_j)^\alpha}{f(t_k) / f(t_j)}
\]

and so on until the last centre is reached.

The overall potential index OPI for innovation for each centre was then derived as the product of the base potential and the incremental improvement at each step of the hierarchy:

\[
OPI = Q I_1^1 I_2^2 I_3^3 \ldots I_N^N \quad (Eq. 7)
\]

2.4 Definition of control variables

Other than spatial proximity that is represented by economic mass, per employee earnings in a given zone are influenced by a range of factors such as working hours, skills, industry composition, and capital investment and natural endowment. It is intuitive that if workers in a
given zone work longer hours (e.g. through routine over-time working) they get higher nominal total pay. All being equal, higher-skilled workers are paid more and a high proportion of skilled workers in zonal employment would raise the level of average earnings. Similarly, employees working in some industries, such as finance, business services, IT and research & development are often seen to be paid more than in other industries. These influences on per worker earnings must be tested, and if significant, controlled for.

Here we follow Rice, Venable and Patacchini (2006) and control the effect of working hours by modeling the average hourly earnings per employee as the dependent variable, i.e. the annual average per employee earnings are divided by the average number of working weeks and the average working hour per week. We also follow Rice Venable and Patachini (2006) to control for employee skills using as a proxy the proportions of those who achieved college, university and post graduate qualifications among the employees. In addition, we follow Combes, Duranton and Gobillon (2008) and include control variables to represent industry composition and capital investment.

3 Data
Data from Guangdong is available at two different spatial scales: the province is first divided into 21 municipalities, and the municipalities are in turn subdivided into 67 counties/county-level cities and 21 urban districts of the municipalities (therefore, 88 county-level units in total). The municipal level data is readily available and our first empirical models were tested at this level. Interesting econometric results at the municipality level encouraged us to assemble over many months the economic and transport data at the county and urban district level. The county/urban district level data is the most detailed currently releasable by the provincial statistics authorities. Once the county-level data was assembled, the municipal level database was refined and made consistent (e.g. for average travel costs).

3.1 Employment and earnings
Currently available employment statistics report the Employed Person\(^6\) in Urban Establishments and Fully Employed Staff and Workers\(^7\) in Urban Establishments at the municipality level, but only the latter at the county/urban district level. The earnings data is available in the same way, i.e. only the earnings of Fully Employed Staff and Workers in Urban Establishments are

---

\(^6\) **Employed Persons** include: 1) fully employed staff and workers; 2) employers of private enterprises; 3) self-employed workers; 4) employed persons in private enterprises and individual economy; 5) employed persons in township enterprises; 6) employed persons in rural areas; 7) other employed persons (including re-employed retirees, teachers in schools run by the local people, foreigners and Chinese compatriots from Hong Kong, Macao and Taiwan working in various units, and people engaged in religious profession, etc).

\(^7\) **Fully Employed Staff and Workers** refer to persons who have work posts, work in and receive payment from units of state ownership, collective ownership, joint ownership, share holding ownership, foreign ownership, and ownership by entrepreneurs from Hong Kong, Macao and Taiwan, and other types of ownership and their affiliated units during the data collection period.
available at the county level. The employment and earnings data excludes farmers and other workers in rural areas. Although we do not have a choice with what employment and earnings data to use at the county/urban district level, the Fully Employed Staff and Workers dataset is remarkably fit for our purpose, i.e. because of the relative long term commitments the employers need to make in recruiting and retaining these workers, we would expect that their earnings would reflect reasonably well their productivity. Also, given that the forces of agglomeration in rural areas are expected to be weak, the data for workers in urban establishments would seem appropriate for our investigation.

The earnings data is reported in the National Statistical Yearbook for Regional Economies for each year 1999-2009. The 2005 1% sample mini Population Census contains the weekly average working hours for each county, and we adopt the working hours of employees in urban areas only. This is the only working hour data available and for time series regression we have used the 2005 working hour data to calculate the hourly earnings of every year from 1999-2009. In addition, per worker GDP has also been calculated as an alternative measurement to earnings.

3.2 Zonal economic mass
The additional data set for calculating the economic mass is the costs and times of business travel, because these trips are most directly related to business linkages, technology transfer, commercial transactions and negotiations. The generalized travel cost is calculated using values of time for employer’s business trips to convert travel times into money units.

We first developed a road transport network for 2009. The road transport network was built within a GIS spatial analysis tool, which made it possible to estimate the travel distances, costs and times of travel between all pairs of municipalities, and between all pairs of counties/urban districts. Up to 2009 the use of rail for business travel was minimal within the province, and thus it is not necessary to include rail costs and times in the network data.

The GIS-tool-estimated travel times and distances are verified through data collected on the web-based travel directions service on the Google Map interface (www.ditu.google.com) during 2009 which provides distances and prevailing driving times between centers of the counties and urban districts. All pairs of counties/urban districts that have at least one end being a municipality capital (thus potentially have a large contribution to the economic mass) have been verified individually against the web data, and any discrepancies corrected. A sample of all other pairs of counties which represent minor economic links are verified against web data. All estimated travel distances and times that have been checked are within ±20% of the web data where they are comparable, and the resulting calibrated transport network is used to estimate

8 We note here that in addition to testing with fully employed staff and workers of urban establishments at the municipal and county/urban district level, we have also run the same empirical models on the earnings of both fully employed staff and workers of urban establishments, and all employees in urban establishments at the municipality level where data is available. We have found no change in the magnitude between the two sets of model results at the municipal level that would alter the conclusions of the study.
travel distances and times on all remaining pairs. The estimation of travel distances and times conform to the standard that is applied to transport cost and benefit assessment.

Road construction data has then been assembled over the period of 1999-2009 from a variety of provincial sources. Road links from the 2009 road network are then modified backwards in time. For cross-section regressions, we developed a road network for 2006, the year by which the majority of Guangdong’s existing inter-city expressways were completed. For time series analysis a road network has been produced for each year of 1999-2009 within the GIS tool. The resulting travel distance, cost and time matrices at the county/urban district level for 1999-2009 are checked using our transport modeling experience. The calculations of the economic mass varies by type, as detailed below.

**Economic mass - Type A**

To calculate the Economic Mass as defined by (Eq. 2), the first step is to use the number of fully employed workers and staff as \( E_j \), and the generalized travel cost in monetary terms as \( g_{ij} \). In the second step, we replace the generalized travel cost with travel time to develop a new set of EMs. After that, we repeat the first and second steps but replace the number of employment with gross regional product. By doing so we could test whether or not the estimation results are sensitive to different measurement of the indicators that compose economic mass.

Following formula (Eq. 3) we calculate a separate series of external EMs by removing the part generated from home municipalities from the total economic mass. The reason for doing so is that the intrazonal travel distances and times are not estimated by the GIS models; instead we manually assigned intra-zonal travel times according to daily observation and experience. Our concern is that, for large cities where the EMs of themselves take up a large proportion of the total, the absence of properly estimated intra-zonal travel time and cost may lead to the inaccuracy of regression results. A separate set of economic mass data has been produced that excludes intrazonal travel, and this set will be used below to test whether the inclusion of intrazonal travel has a material effect on the econometric results.

**Economic mass - Type B**

In Rice, Venables and Patacchini (2006), working age population has been used as measurement of economic activity \( P_{ia} \). To make our study comparable to previous studies, we also use employment data. Since the data they have used is for Britain, where less than 1% of the working population is now engaged in agricultural work, it is appropriate for us to use the number of workers and staff in urban establishments for Guangdong.

In the empirical tests below for Guangdong, we have also modified this EM measure in order to adapt it to the Guangdong conditions, where travel speeds are generally lower and travel times are longer: \( T_k \) is set at 60-minute intervals thus \( T_k = 60, 120, 180, \ldots, 360 \) plus a further band
that includes all travel that is $>360$ minutes. In line with this, $T^+$ is set to 60 and $\theta$ estimated simultaneously with the coefficient of banded EMs in the non-linear regression.

**Economic mass - Type C**

Analysis undertaken on a wide range of data in Australia and Europe in the 1970s (Bullock, 1980) showed that trip rates to and from a central place were proportional to $P^{0.36}$, where $P$ is the population as a proxy for the size of the central place. This reflects the number of functions\(^9\) (e.g. banking, department stores, specialist physicians etc) that are available in towns with different populations. The number of trips between the central place and any particular location, after allowing for the location’s population\(^10\), also varies with distance. Trips within the immediate catchment decline with time $T$ according to $e^{-0.05T}$. Beyond this daily commuting catchment, trips decline with $T$ according to $T^{-1.7}$.

The parameter values adopted in the analysis are: $\alpha = 0.4$, $\alpha = 344$, $\lambda = -0.02$, and $\beta = -1.70$. The catchment boundary is set at 90 minutes.

### 3.3 Zone specific control variables

There are two detailed datasets that have contributed to many of the calculations of the control variables: First, the published Economic Census 2004 contains a detailed survey of productive activities on municipality level, including labour composition statistics. Secondly, a more detailed survey of labour market is found in the 1% sample mini Population Census of 2005. These contribute directly to the cross-sectional regressions and also help to define the time series dataset.

**Educational background**: The 2005 1% sample population census provides the number of workers in all enterprises in terms of their education level (i.e. postgraduate, university, college, senior middle school, junior middle school and below) by county and urban district. Education level indices are calculated as a ratio of the proportions of workers at each education level in each zone to the average provincial proportions for each respective level. Specifically, we use the percentage of workers with college degree and above as a proxy for labour skills from the 2005 Census at the county/urban district level. At the municipality level, we use the data provided by the Economic Census for 2004. This is used for cross-section regressions. For time series analysis, a slightly different indicator of the percentage of workers with university education or above as a proxy of labour skills, which comes from the statistical yearbooks.

---

\(^9\) In central place literature, functions are binary variables which reflect the existence or not of a particular activity. The aggregate number of such distinct activities is termed ‘functional units’. Thus a centre with five dentists would classify these as one function (dentistry) but five functional units.

\(^10\) Larger towns generate fewer external trips per capita to a regional centre, presumably because of the greater availability of goods and services within the town. Studies in Australia showed an external trip rate varying at about $Q^{0.7}$, where $Q$ is the population of the town. Similar estimates have been obtained from studies in the US.
**Capital stock per worker**: The statistical yearbooks report the levels of fixed asset investment per year. The Economic Census of 2004 also reports the total capital stock for production purposes per municipality. We estimate the county/urban district level capital stock through these sources, and build up the yearly capital stock for the entire time series that incorporates a standard capital stock depreciation rate of 5% per year. Investment in residential properties is excluded. We divide the zonal total capital stock by the total full time workers and staff in that zone.

**Industry composition**: According to the National Labour Statistics Yearbook 2009, finance, information technology and R&D industries are ranked as the top three high-earning sectors in Guangdong Province. We use the number of employees by region in these three sectors to control for the effects that can potentially arise from such differences in industrial composition. Specifically, we construct the index of sectoral composition following the definition of location quotient (LQ):

\[
LQ_{ik} = \frac{EMP_{ik}}{EMP_k} \div \frac{EMP_i}{EMP}
\]

Where \(LQ_{ik}\) is the location quotient of sector \(I\) in county/district \(k\); \(EMP_{ik}\) is the employment of sector \(i\) in region \(k\); \(EMP_k\) is the total employment of county/district \(k\); \(EMP_i\) is the employment of sector \(i\) in the whole province; \(EMP\) is the total employment of the whole province.

This data is available for the cross-section regressions only.

### 3.4 Preliminary cross-section data analysis

Figure 2 through Figure 7 show the geographic variation of the earnings, economic mass indices Type A and C, and the control variables for the cross-section regressions. Broadly speaking, they follow similar patterns, although an inspection of the correlation between the dependent variable hourly earnings and the independent variables such as economic mass Type A, capital endowment, proportion of employees with college and above qualifications, proportion of workers working in R&D, IT and financial service industries are not excessively high (ranging from 0.28 to 0.60).

High correlation between hourly earnings and proportion of employees with college and above qualifications indicates employee spatial self-selection and sorting across the counties and urban

---

11 See Guangdong Statistics Yearbook 2007, Table 5-10; 2005 1% sample population census, Table 4-1.
12 Economic mass Type B cannot be analyzed prior to model estimation, as its quantification depends on the model parameters to be estimated.
districts (i.e. higher skilled workers gravitating towards high pay jobs). Very high correlations are found between the proportion of employees with college and above qualifications and workers in IT and financial services (0.75 and 0.83), and between IT and financial service workers (0.81), suggesting that there are possible colinearity problems to be considered when choosing the regression models.

**Figure 2** Hourly Earnings of Full-time Workers and Staff by County/District (2008)

![Map of Hourly Earnings](image1)

Data Source: Guangdong Statistic Yearbook 2009

**Figure 3** Per Employee Capital Stock by County/District (2008)

![Map of Capital Stock](image2)

Data Source: Guangdong Statistic Yearbook 2000 – 2009
Figure 4  Economic Mass Type A: with Generalized Travel Cost by County/District (2006)

Source: Authors’ estimation.

Figure 5  Economic Mass Type C: with Generalized Travel Cost by County/District (2006)

Source: Authors’ estimation.
Figure 6  Index of Labour Force with College Degree and Above (2005)

Source: 1% sample population census of Guangdong, 2005.

Figure 7  Index of R&D Sector by County/District (2005)

Source: 1% sample population census of Guangdong, 2005.
4 Results from econometric models

4.1 Overview of the models tested
A large number of regression models have been carried out on the earnings data, initially using cross-sectional Ordinary Least Squares (OLS) regressions (of each year of 2005-2008 and the four-year average) to explore the data. The main parameter estimations are then carried out using the time series data for 1999-2009 using pooled OLS, fixed effects (FE) models and dynamic panel data (DPD) models. Models based on alternative measurements of the earnings, economic mass and control variables are tested. Econometric models have been estimated mainly at the county/urban district level, with some tests carried out at the municipal level for verification.

We first outline the exploratory cross-section regression results in Section 4.2, and then the main model estimation using the time series data in Section 4.3.

4.2 Cross-section regressions for 2005-2008
We first exploit the relative simplicity and efficiency of the OLS models and in particular use them to test the alternative economic mass (EM) measurements. The main cross-section model uses the average earnings of 2005-2008 as the dependent variable, and the travel costs of 2006 and the control variables of 2005 as explanatory variables. We have also tested the regression models on the earnings data for individual years of 2005 through 2008, using the EM measures off the 2006 transport network and control variables of 2005. There are N=88 county-level zones which are used in the regressions.

The OLS models are estimated first on EM Type A. They are calculated using the number of fully employed workers and staff in urban establishments and generalized travel cost, which includes time value, fuel and toll cost. Further, an alternative EM Type A calculated from travel time only is tested, in order to be comparable with those models in the literature that are estimated on drive time only (e.g. adopted by the preliminary models in Rice, Venables and Patacchini, 2006). In all the models, the EM Type A variable is shown to be highly statistically significant, whilst the coefficients of the control variables remain stable. The EM Type A coefficients vary within a range of 0.12-0.27.

In models on EM Type B, we use the number of fully employed workers and staff in urban establishments within each travel time band at an interval of 60 minutes to represent the economic mass, and assign a factor \( \theta \) to account for the spatial decay effects. The coefficient of EMs and the value of \( \theta \) are estimated simultaneously through non-linear regression. Non-linear regression using the distance decay model of Rice et al (2006) give also reasonably stable results for all zones and for inter-zonal EMs excluding Guangzhou and Shenzhen. The coefficient values for EM Type B range between 0.05-0.07. The models, however, have low R-Squared values with or without the control variables, reflecting the fact that the small number of travel time bands may be too coarse in this case to capture the spatial patterns of earnings.
In models on EM Type C, we incorporate the influence of the regional hierarchies on the economic mass. From a theoretical point of view, we would expect that the hierarchical EM measure would be a more realistic in representing the interactions among the counties and cities in Guangdong. This is corroborated by the higher R-Squares of the models. The EM Type C models are statistically significant and have stable coefficients. The same range of alternative models have been run as above, and the coefficient values for EM Type C range between 0.12-0.22.

As discussed above, to test whether the estimation of intra-urban travel costs and times fundamentally affect the estimation results, we test a separate set of models that are based on an alternative formulation of the EM that excludes all intra-zonal components: this is done by separating the economic mass that arise from influences within own zone (intra-EM) and from those outside own zone (inter-EM). This controls for the influence of the considerable economic mass that arise from economic activities within own municipalities: this ‘intra’ component of the economic mass is extremely sensitive to the assumption of intra-zonal average distance of travel, which cannot be as precisely measured as the interzonal travel costs. We therefore exclude the EM values generated by the interaction within each zone and only consider the influence of interzonal EM to see how sensitive the estimation results are to the intrazonal travel time. However, this alternative set of models produce similar, statistically significant answers.

In Guangdong Province, the economic activities show a dual-centered pattern, where Guangzhou and Shenzhen play a dominant role. In order to ensure these two large centers did not distort the results, a further set of estimations were also done in which these two large cities are excluded from the estimation (i.e. Sample N=86 instead of N=88). In doing so we exclude Guangzhou and Shenzhen in the econometric estimations, but the contribution of these two cities to the economic mass of the other 86 county/urban district areas is included. The results remain statistically significant, although the numerical values of the coefficients are generally lower as expected.

In addition, we have also replaced average earnings with per employee GDPs and this has not changed the nature and statistical significance of the results. The majority of the models are also tested at the municipal level and the results are similar – generally the EM coefficients tend to be higher. The details of the estimation results above are available upon request.

Among the control variables, per employee capital stock is shown to be statistically highly significant, the education level modestly so, and the industry composition shown to be generally insignificant in the models. The current classification system for the industries may be too crude to discern the industry sectors in terms of productivity. We also take note of the higher correlation between the EM and the education level variables which could suggest a particularly high degree of endogeneity, i.e. zones of higher pays tend to attract higher skilled workers and vice versa. Such endogeneity issues will be addressed in the time series models in the next section.
4.3 Time-Series Regressions for 1999-2009

The complete time series for economic data is assembled for 1999-2009 whilst the EM data is estimated by the study team through a road network model, as mentioned above. Based on the results of the explorative OLS runs above, we have focused on using the average hourly earnings rather than the per employee GDP as the dependent variable. We choose EM Type C as the main economic mass variable because of our field survey findings and the better OLS model fit. For the control variables, we focus on per employee capital stock and the education level (in the panel data the latter is represented by the percentage of college-and-above graduates among the fully employed employees). In line with our field survey findings, we assume a lag of 2-3 years for the EM, capital stock and education level to have an effect – for any year t, we produce an average of t, t-1 and t-2 as those explanatory variables.

The spatial econometrics literature suggests that there can be significant spill-over effects between neighbouring counties/urban districts. A formal way to deal with such spill-over effects is to construct a weight matrix such that the lagged dependent and independent variables of all the near and distant neighbours are tested as explanatory variables, in addition to own independent variables of each county/urban district. Given that the EM variable has by definition already accounted for spatial proximity to employment centres, and a weight matrix containing influences of both near and distant neighbours would make the regression model over-complicated if used simultaneously with the dynamic panel data models, we adopt here a simplified approach of only include only the control variables from the nearest neighbour of each county/urban district as additional explanatory variables. As a rule, including the nearest neighbour in the spatial spill-over analysis should take account of 70-80% of the effects (LaSage, 2012).

We also exploit what is known in statistical theory about the nature of the OLS, fixed effects (FE) panel data models, and dynamic panel data models in terms of parameter estimation bias when used with a dataset such as ours which is autoregressive in nature. It is well known that the pooled OLS estimation will bias the parameter upwards, whilst the corresponding FE model will bias them downwards (Brülhart and Mathys, 2008).

Given that our aim is to identify the causal effects that run directly from the economic mass to hourly earnings where all explanatory variables are considered as potentially endogenous, and possibly correlated with the error term of the regression model. An increasing popular regression technique in dealing with such requirements is the linearized generalized method of moments (GMM) technique (Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998). The idea of the dynamic panel data model is to use the past realisations of the model variables as instrument variables. The time invariant zonal effects are differenced out. Past changes of a variable can influence its current year level and rate of change, but not the reverse. The method suits well our requirements because proper exogenous instrumental variables are hard in investigations of agglomeration effects.
In large samples and given some weak assumptions, GMM models can be free of the estimation bias inherent in the OLS and FE models. However, the two variants of the GMM methods, namely DIFF-GMM and SYS-GMM, have different properties when used with small samples. Whilst the DIFF-GMM technique may be unreliable under small samples (Bond et al, 2001), the SYS-GMM technique yields considerable improvements in such situations (Blundell and Bond, 1998). Our dataset for 88 zones and 8 periods (after taking 3-year averages), although having required substantial effort to assemble, is not a big sample. It is therefore necessary to test all the above models in order to clarify the robustness of the models. In turn, a comparison with the theoretical, prior expectations may also serve as a reassurance test.

Table 1 presents the series of statistical models that test the theoretical expectations above. Model (1) is a pooled OLS model that generate a similar pattern of coefficients to that found in the cross-section analyses above. As expected, the EM (Type C) coefficient, at 0.24 is at a high end of the values we have seen so far. Although both the EM and the control variables (for capital stock and education level) are statistically significant and the R-Squared = 0.69, we have grounds to suspect that the coefficients are biased upwards.

This is confirmed by the FE model in column (2): the EM coefficients drops in this model to 0.115 when the period dummies (representing the period specific effects) are included. However, our prior expectations suggest that this may form a lower bound to the EM coefficient.

This is reflected in the DIFF-GMM model in column (3). The EM coefficient output from this model is at 0.151, between the upper and lower bounds as we expect, although the coefficients are not statistically significant. The SYS-GMM model in column (4) gives a similar EM coefficient at 0.141, although both the EM and the capital stock coefficients are now significant. Note that this model includes additional explanatory variables that represent the spill-over effects from the nearest neighbours in terms of capital stock and education level of the employees.

Model (5) is a standard test to assess the robustness of the model by reducing the number of instrument variables (from 115 to 69), which has raised somewhat the significance of the education level variable but has not altered the nature of the model results. The standard tests of the GMM models suggest that there are no misspecification problems. The Hansen test for overidentification restrictions, and the Difference Hansen tests for the validity of the GMM and IV instruments fail to reject the null hypothesis that the instruments are valid. The Arellano-Bond AR2 test fails to reject the null hypothesis that no second order residual auto-correlation is present.

The results show that spatial proximity as represented by the EM is statistically significant after controlling for endogeneity and spatial spill-over effects. In particular, the SYS-GMM model that has included both own county and nearest neighbour county control variables show that the productivity elasticity with respect to the EM is 0.14, with a robust standard error of 0.067. We
consider this to be an appropriate central estimate for Guangdong on the basis of the time series dataset for 1999-2009.

Table 1  Time-series model results

<table>
<thead>
<tr>
<th>Dependent variable = ln {hourly earnings [t]}</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooled OLS</td>
<td>0.240***</td>
<td>0.151</td>
<td>0.141***</td>
<td>0.162***</td>
<td></td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td>0.115***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIFF-GMM</td>
<td></td>
<td></td>
<td>0.067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS-GMM</td>
<td></td>
<td></td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYS-GMM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln (average economic mass of t-2, t-1 and t)</td>
<td>0.240***</td>
<td>0.151</td>
<td>0.141***</td>
<td>0.162***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.016]</td>
<td>[.152]</td>
<td>[.067]</td>
<td>[.077]</td>
<td></td>
</tr>
<tr>
<td>ln (average per worker capital stock of t-2, t-1 and t)</td>
<td>0.155***</td>
<td>0.007</td>
<td>0.026</td>
<td>0.133***</td>
<td>0.151***</td>
</tr>
<tr>
<td>ln (average percentage of college graduates of t-2, t-1 &amp; t)</td>
<td>0.193***</td>
<td>-0.021</td>
<td>-0.046</td>
<td>0.031</td>
<td>0.062</td>
</tr>
<tr>
<td>Period dummies</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ln (average per worker capital stock of t-2, t-1 and t) - nearest neighbour county</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[.033]</td>
</tr>
<tr>
<td>ln (average percentage of college graduates of t-2, t-1 and t) - nearest neighbour county</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[.041]</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.569***</td>
<td>1.975</td>
<td>.170</td>
<td>0.104</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>704</td>
<td>704</td>
<td>616</td>
<td>704</td>
<td>704</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.69</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of instrument variables</td>
<td>64</td>
<td>115</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(2) test (p-value)</td>
<td>0.430</td>
<td>0.170</td>
<td>0.790</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hansen test of overidentification restrictions (p-value)</td>
<td>0.274</td>
<td>.980</td>
<td>0.352</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference-in-Hansen tests of exogeneity of instrument subsets for GMM instruments (p-value)</td>
<td></td>
<td></td>
<td>0.713</td>
<td>0.367</td>
<td></td>
</tr>
<tr>
<td>for iv instruments (p-value)</td>
<td>0.115</td>
<td>0.909</td>
<td>0.365</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: For the GMM models, the variables for the economic mass and average percentage of college graduates are treated as GMM instrument variables whilst the year dummies and average per worker capital stock are treated as IV variables; For GMM models they are based on two step Windmeijer finite sample adjustment. For the FE and GMM models the fixed effects are modeled but not reported here.

Robust standard errors in square parentheses

*** p<0.01, ** p<0.05, * p<0.1
5 Summary of findings and further work

5.1 Summary of findings
The models based on alternative measurements of the economic mass and different sets of control variables have been estimated using ordinary least squares and non-linear least squares for cross-section tests of 2005-2008, and time series tests for 1999-2009.

The results show a consistent pattern regarding the alternative measurements of the economic mass (EM). In particular the main econometric estimates using the dynamic panel data (DPD) models and time series data for 1999-2009 show that:

a) Its central estimate for the economic mass (EM) coefficient, which represents the productivity elasticity, is 0.14 (with a corrected standard error of 0.07 and Student’s t = 2.09) when controlled for the level of education of the employees, capital investment and the spatial spillover effects of the level of education of the employees, capital investment in the nearest neighbour county. This GMM model estimate sits in the middle of the coefficient estimate of 0.24 from the pooled OLS model (which is expected to overestimate the coefficient) and that of 0.11 from a Fixed Effects panel data model (which is expected to underestimate the coefficient).

b) The coefficients of the control variables for the level of education of the employees and capital investment in both the own county and the nearest neighbour county are all positive as expected, although their statistical significance vary.

c) All associated statistical tests show that the central estimate is robust. Reducing the GMM instrumented variables, which is a standard test for the GMM models, does not alter the magnitude of the coefficient estimate.

Almost all models show that proximity to economic mass is generally associated with higher level of average hourly earnings and this positive relationship remains robust after controlling for a range of control variables, endogeneity, and nearest neighbour spill-over effects. Analyses on different geographical scales, municipality and county/district, return consistent results.

These central productivity elasticity estimate of 0.14 implies that a doubling of the economic mass would give rise to an increase of per worker productivity of 10% (i.e., $2^{0.14} - 1 = 0.10$). This is at the higher end of the consensus range of productivity elasticities from a comprehensive review of such evidence in predominantly developed economies that ‘doubling city size seems to increase productivity by an amount that ranges from… roughly 5-8%’ (Rosenthal and Strange, 2004).

In assessing the estimates we may also compare them with our prior expectations: Spatial proximity and agglomeration is thought to play an important role in knowledge spillover and technological improvements in China. The empirical findings in this paper are to an extent supported by those of Au and Henderson (2006). Their analysis using data of 1990 and 1997
from 205 Chinese cities suggests that there are significant urban agglomeration benefits: e.g. moving from a city of 635,000 to one of 1.27m increases the real output per worker by 14%, after controlling a wide range of other influences. Recent findings by Ren and Lin (2007) and Jorgenson et al (2007) may also be relevant – they find that the contribution of technological improvements to productivity to be much higher in China than in developed countries – although this is circumstantial rather than direct evidence.

However, although the econometric results fit the expectations, it may nevertheless be that the current econometric models have not fully controlled for other differences between zones, for example the spatial self-selection and sorting of employees within and amongst the counties and urban districts. Recent research has highlighted the importance of such mechanisms upon productivity, through e.g. improving the quality of matches among the workers in cities and reinforcing the development of informational networks (Combes et al, 2005; Venables, 2010). Clearly, spatial proximity resulting from transport improvements plays an enabling role in spatial self-selection and sorting. Nevertheless, it is yet difficult to discern the precise contribution of transport improvements to such mechanisms within the available data sources.

Also, it is not for econometric studies alone to establish causality between spatial proximity and productivity where there is a process of significant cumulative causation – that task should be supported by an in-depth understanding of the actual mechanisms at work, e.g. through specific case studies.

On balance, we recommend retaining the current productivity elasticity value of 0.075 in project assessment, since this is within one standard deviation of the central estimate, and it is preferable to be conservative when quantifying the agglomeration effects.

5.2 Further work
The following four pieces of additional work may improve the robustness of the findings presented here:

First, it may be possible to expand the time series under consideration both in years covered and the range of explanatory variables, which is likely to make the model more robust and improve the precision of the coefficient estimates.

Secondly, similar econometric models can be estimated for the economically less developed regions in China (e.g. inland regions such as Sichuan), as well as other affluent regions along the Eastern Coast (e.g. the Yangtze River Delta centered upon Shanghai and the Bohai Bay Metropolitan Area centered upon Beijing). This would make the findings more robust as well as providing evidence whether there are significant differences among regions of different levels and contexts of development.

Thirdly, if and when the disaggregate Economic Census data becomes available from the Chinese statistics bureaux, enterprise-level production functions (e.g. the Translog type) can be
estimated, which would provide more precise estimates of the agglomeration effects. The Economic Census data was collected by enterprise, although so far has not been released for use in research in China.

In addition, micro-level case studies of firms and institutions will help us understand how firms actually respond to transport improvements, and through what mechanisms they gain from agglomeration effects or otherwise. For further considerations on this, see Working Paper 1.

Bibliography


