## Dissertation

## **Essays on China's Economic Growth and Regional Economic Development**

Ausgeführt zum Zwecke der Erlangung eines akademischen Grades des Doktors der Volkswirtschaften unter der Leitung von

Prof. Dr. Michael Funke

Institut für Wirtschaftssysteme, Wirtschafts- und Theoriegeschichte Arbeitsbereich Makroökonomik und Quantitative Wirtschaftspolitik

> eingereicht an der Universität Hamburg Fachbereich Wirtschaftswissenschaften

> > von

Hao, Yu MA Economics

Hamburg 2012

Vorsitzender:	Prof. Dr. Thomas Siedler
Erstgutachter:	Prof. Dr. Micheal Funke
Zweitgutachter:	Prof. Dr. Erich Gundlach

Datum der Disputation: 17.09.2012

### Contents

Introduction	3
1. The emergence and spatial distribution of Chinese segment	
cities	
Michael Funke and Hao Yu	9
2. The Role of Inter-Provincial Transfers in the	
Convergence Process - Evidence for China	
Hao Yu	23
3. Economic growth across Chinese provinces: In search of	
innovation-driven gains	
Michael Funke and Hao Yu	53
4. Uncertainty and risk analysis of the Langrun Chinese GDP	
Forecast: Fan Charts revisited	
Hao Yu	73

#### Introduction

The extraordinary emergence of China in the world economy since the late 1970s has been a hallmark of the global economic trends, heralding a 21st-century world that looks fundamentally different from that of the previous century. As both a transition economy and a developing country, China still represents a complex case with a mix of characteristics even after over three decades of economic reform. According to the WDI database, by 2010 China's GDP (\$5.93 billion USD) had surpassed that of Japan and became the second largest economy after the U.S., but the per capita GDP in USD (\$4428) ranked only 99th out of 190 countries and regions in the world in 2010.<sup>1</sup> The achievements of China's economy over the last three decades have provided economists with abundant research topics as well as with an increasing number of puzzles.

The four papers in this thesis all concern China's economy, each focusing on a specific aspect thereof. The first two papers are theoretical studies, in which new economic geography and endogenous growth theory are utilized to analyze the emergence of China's port cities and the evolution of China's regional disparity in the post-reform era since the late 1970s, respectively. The last two papers are empirical studies, in which, first, the importance of innovation to China's economy and secondly, the evaluation of the uncertainty and risks embodied in Chinese GDP growth forecasts are investigated. Although seemingly unrelated, the four topics have fundamentally strong interconnections and reflect different features of China's economy, as will later be interpreted in detail. The internal logic of these topics follows a loose chronological order, accordingly, this introduction will offer a brief historical overview.

In order to better understand the contemporary Chinese "economic miracle", we have to take a backward glance at its history. In fact, China's role as a leader in the global economy is nothing new. According to Maddison's (2007) estimation, from the Song dynasty (960-1279) until the late eighteenth century, the level of Chinese technology had led the rest of the world, and its productivity was approximately one third higher than that in Europe. However, due to the lack of an industrial revolution, which began in Britain in the mid-eighteenth century and then spread rapidly to other Western European countries and to the U.S., China's economy stagnated in terms of per capita GDP after the Song dynasty. For over a century after the first Opium War (1839-1842), due to constant social and economic instability, China's economy suffered from a downturn with its GDP ratio in the world decreasing from an estimated 32.4% in 1820 to 5.2% in 1952 (Maddison, 2007).

Despite the unfavorable social and political environment and nationwide stagnation of productivity, several Chinese port cities still emerged during this period, quickly changing China's landscape dramatically. Given the importance of these port cities, the first paper, entitled "The emergence and spatial distribution of Chinese seaport cities,"<sup>2</sup> analyzes the mechanism and process of Chinese port cities' emergence since 1840. The analysis utilizes a new economic geography analytic framework, which provides an integrated and micro-founded approach to spatial economics. It emphasizes the role of clustering forces in generating an uneven distribution of economic activity and income across space (Fujita and Krugman, 2004). Since the new economic geography was developed in the early 1990s, the approach has been applied to the economics of cities, the emergence of regional disparities, and the origins of international inequalities (see Fujita et. al, 1999). The purpose of this paper is to investigate the formation of the leading Chinese port cities and the dynamics that drives it. Concretely, the paper provides an answer to the following question: How are the number and the geographical locations of the emerging Chinese ports endogenously determined? The key feature in the emergence of Chinese port cities is the agglomeration of population and industry.<sup>3</sup> Regional economic imparity arises because agglomeration creates growth, and certain regions experience forces encouraging agglomeration and others experience the opposite. In order to delineate this agglomeration process, Fujita and Mori's (1996) model is adapted as a formal analytic framework, highlighting monopolistic competition and transfer cost as the forces promoting agglomeration. Taking the geographical characteristics of China's coastline into account, the emergence of several most representative and influential port cities is successfully calibrated, including Shanghai, Qingdao and Fuzhou along the coastline, as well as Wuhan and Chongqing along Yangtze River.

The rise of China's port cities has had long-lasting and far-reaching impact on China's regional disparity. The port cities experienced modernization processes earlier than other regions of China because they had the earliest assess to western scientific and administrative technologies. The port cities' prosperity propelled economic growth in the coastal regions and established the economic advantages of coastal area. This regional economic pattern even persisted into the post-reform era. In 1978, an economic reform was launched as a response to preferential economic and fiscal policies given to eastern provinces under the developing strategy "allowing some to get rich first." In the first

<sup>&</sup>lt;sup>1</sup> In PPP terms, China's total GDP was \$10.17 billion in 2010 and also ranked second after the U.S., while its GDP per capita was \$7599 and ranked 96th globally.

<sup>&</sup>lt;sup>2</sup> See <u>http://www.sciencedirect.com/science/article/pii/S1043951X11000034</u>.

<sup>&</sup>lt;sup>3</sup> As an illustration, the population of Shanghai ballooned from 400,000in 1840 to 3.35 million in 1934, when Shanghai became world's sixth largest city in terms of population. In the meantime, industrial agglomeration also took place in Shanghai. By the 1930s Shanghai had become the nation's commercial center and the most important financial center in the Far East. Because of highly developed modern industry and commercial activities and broad international relationships, Shanghai earned the nickname "the Great Athens of China" in the 1930s.

two decades after this reform, the eastern provinces rapidly developed export-oriented industries and soon became the powerhouse of China's economy. At the same time, the western provinces experienced slower economic growth and regional disparity increased. In 2000, after realizing that the regional disparity would threaten the political stability if it were allowed to spira out of control, the central government launched the "Western Development" programme. This programme, featuring large-scale infrastructure development in western provinces and dramatic increase of interregional transfer payment from eastern to western provinces, has proved successful in promoting the productivity of western provinces and narrowing the regional income gap.

With this background, the second paper, "*The Role of Inter-Provincial Transfers in the Convergence Process - Evidence for China*," offers an interpretation of China's regional disparity by employing Funke and Strulik's (2000) endogenous growth modeling framework which highlights the function of interregional transfer payment system. These fiscal policies are highly significant for China's regional economic development because Chinese government tends to play an active role in economic affairs (Huang, 1999). This study outlines the characteristics of China's transfer payment systems in different stages of economic reform by emphasizing the importance of government public spending to the growth of productivity. The calibration results fit well with China's actual key economic indicators between 1992 and 2010 and therefore offer a reasonable prediction for future development of regional disparity: if the current "Western Development" programme continues, the west-east relative per capita income would reach approximately 80% in 2030, up from nearly 55% in 2000.

There are different explanations for China's economic success since the economic reform took place in the late 1970s. In an influential paper, Krugman (1994) asserts that the rapid growth in the newly industrializing countries of Asia, including China, was achieved simply through input accumulation, while innovation progress played little role in the growth process. The increasing investment rate since the mid-1980s is considered by some scholars as evidence of China's Solow-style growth pattern which relies on input (especially capital) accumulation (Garnaut and Huang, 2006). However, some empirical studies estimating Chinese total factor productivity (TFP) have verified that the China's TFP growth accelerated and became quantitatively nontrivial since the 1990s (Chow and Lin, 2002; Guo, 2006). As China has attached more importance to the sustainability of economic growth in recent years, innovation-driven growth could be promoted to become the main dynamic of China's economy in the future.

In order to better understand the effect of innovation on China's economic growth during the postreform era, the third paper, "Economic growth across Chinese provinces: In search of innovationdriven gains"<sup>4</sup>, the contribution of innovation to China's economic performance

is empirically estimated. One serious problem of most previous empirical studies is the ignorance of the endogeneity between investment and TFP. This endogeneity stems from one simple fact: the regions with high productivity generally tend to attract more investment. Obviously, neglecting this endogeneity may lead to biased estimation of capital coefficient in production function (Levinsohn and Petrin, 2003). In order to remedy this endogeneity problem, the semiparametric (control function) estimator suggested by Olley and Pakes (1996) is utilized. The idea of this estimator is to invert demand for capital to infer unobserved productivity shocks and then use the estimated productivity shock as a regressor in the production function. The additional attrition problem is addressed by using attrition probabilities. With Chinese provincial panel data from 1993 to 2006, and taking into account possible spatial spillovers from technology progress, China's production function is estimated by employing a semiparametric estimator. The estimation results not only verify the role of innovation as a source of growth in China but also reveal the important effect of technology spillovers: the R&D activities that took place in neighboring provinces may to some extent compensate for weak contributions of the R&D activities pursued locally. Moreover, the technology spillover effect is stronger for coastal provinces than the non-coastal inland provinces.

As China's economy rapidly expanded, economic and social problems also accumulated, which may potentially affect China's economic growth.<sup>5</sup> As a result, it is necessary to incorporate the uncertainty analysis into the GDP growth rate forecast. In the final paper, "*Uncertainty and risk analysis of the Langrun Chinese GDP Forecast: Fan Charts revisited*"<sup>6</sup>, a fan chart approach developed by Elekdag and Prakash (2009) is adapted to estimate the density distribution for China's GDP growth forecasts and show it intuitively. The fan chart is a graphic representation of the density forecast, which uses different forms of lines and shades to demonstrate the bands of different percentiles (for example, the 10th, 20th, 30th . . . and 90th percentiles) of the estimated probability distribution. Because the bands for the same percentiles become wider and spread out over time, such representation is called a 'fan chart'. The prominent advantages of the fan chart include the ability to highlight the baseline forecast, indicate the level of uncertainty surrounding the baseline forecast, and show the balance of risks (Elekdag and Kannan, 2009). Compared with the traditional point forecast method, which contains no uncertainty information, the fan chart method effectively highlights the overall uncertainty and balance of risks embodied in the GDP growth rate forecasts, especially as China's economy slowed down during the past international financial crisis between 2007 and 2009.

<sup>&</sup>lt;sup>4</sup> See <u>http://www.suomenpankki.fi/bofit\_en/tutkimus/tutkimusjulkaisut/dp/Pages/dp1009.aspx.</u>

<sup>&</sup>lt;sup>5</sup> For a brief review of problems faced by China's economy at the moment, see

http://www.economicshelp.org/essays/problems-chinese-economic-growth.html.

<sup>&</sup>lt;sup>6</sup> See <u>http://www.tandfonline.com/doi/abs/10.1080/17538963.2011.666056</u>.

The four papers introduced previously highlight different aspects of China's economy. As an old Chinese saying goes, to look at several spots on a leopard helps you to visualize the whole animal. Hopefully, through the discussion of these specific aspects of China's economy, the views described in these papers could serve as a catalyst and promote new and intriguing thoughts about China's economy.

#### **References:**

Chow, Gregory C., and Lin, An-Ioh. 2002. "Accounting for Economic Growth in Taiwan and Mainland China: A Comparative Analysis", *Journal of Comparative Economics*, Vol. 30(3), pp. 507–530.

Elekdag, Selim, and Prakash Kannan. 2009. "Incorporating market information into the construction of the fan chart." *IMF Working Paper*, August.

Fujita, M., and Mori, T. 1996. "The role of ports in the making of major cities: self-agglomeration and hub-effect." *Journal of Development Economics*, Vol 49(1), pp. 93-120.

Fujita, M., and Krugman, P. 2004. The new economic geography: past, present and the future. *Papers in Regional Science*, Vol. 83(1), pp. 139–164.

Fujita, M., Krugman, P., and Venables, A. J. 1999. *The spatial economy: cities, regions, and international trade*. Cambridge: MIT Press.

Funke, M., and Strulik, H. 2000. "Growth and Convergence in a Two-Region Model of Unified Germany", *German Economic Review*, Vol. 1(3), pp. 363-384.

Garnaut, R., and Huang, Y. P. 2006, "The Turning Point in China's Economic Development", in Garnau and Song (eds), *The Turning Point in China's Economic Development*, Canberra: Asia Pacific Press at the Australian National University.

Guo, Y. Q. 2006. "Capital Accumulation, Technical Change and Aggregate Production Function: A Study Based on Empirical Data from China between 1980 and 2005" (in Chinese), *Nankai Economics Studies*, 2006(3), pp. 79-89.

Huang, Y. 1999. Inflation and investment controls in China: the political economy of central-local relations during the reform era. Cambridge: Cambridge University Press.

Krugman, P. 1994 "The Myth of Asia's Miracle", *Foreign Affairs* 73, Number 6 (November/December), pp. 62-78.

Levinsohn, J. and A. Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies*, Vol. 70(2), pp. 317-341.

Maddison, Angus. 2007. Chinese Economic Performance in the Long Run (2nd edition). Paris: OECD.

Needham, Joseph. 1981. *Science in Traditional China: A Comparative Perspective*. Cambridge, MA: Harvard University Press.

Olley, S.G. and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica*, Vol. 64(6), pp. 1263-1297.

ELSEVIER

Contents lists available at ScienceDirect

China Economic Review

## The emergence and spatial distribution of Chinese seaport cities

#### Michael FUNKE\*, Hao YU

Department of Economics, Hamburg University, Von-Melle-Park 5, 20146 Hamburg, Germany

#### ARTICLE INFO

Article history: Received 16 September 2010 Received in revised form 22 December 2010 Accepted 5 January 2011 Available online 12 January 2011

JEL classification: R11 R12 R41

Keywords: Seaports Cities Economic geography model China

#### 1. Introduction

# In a rapidly globalising world, the maritime sector has been a significant factor in facilitating the integration of markets for manufactured goods, with seaports operating a natural gateway to countries across the globe. Since world trade grows significantly faster than world output, foreign trade plays an increasingly important role in the development of national economies. Particularly in East and South East Asia, the enormous industrial success has been closely linked to the development of intermodal transport and has led to Asian dominance in container traffic. In fact, seaports constitute the backbone of the transport network without which today's global economy could not exist in its present form. It should be noted that, of the top 25 seaports in the world port traffic league ranked by containers handled 1999–2003, 15 form what can be likened to a string of pearls, stretching from Singapore to Tokyo. In all, 7 of these seaports are located in China.<sup>1</sup>

The new economic geography, new trade theory, and endogenous growth models have highlighted the nexus between geographic location and economic growth. Conclusions emanating from this line of inquiry are: (a) landlocked regions and countries trade less with coastal regions or countries, and (b) coastal regions and maritime countries on average post higher growth than landlocked regions and countries. For example, Démurger (2001) and Démurger et al. (2002) have demonstrated that transport facilities are a key differentiating factor in explaining regional growth disparities across China. Bruinsma, Gorter, and Nijkamp (2000) find that transport infrastructure is a significant determinant of the location decisions of footloose multinational firms, and that these

\* Corresponding author.

#### ABSTRACT

Seaports have historically played a key role in facilitating trade and growth. This paper is the first attempt in the literature to analyse the formation of Chinese seaport cities and the dynamics that drives it. First, we aim to identify theoretically the emergence of urbanized seaports with the help of a formal economic geography model. Second, employing an empirically plausible parameterisation of the model, we calibrate the evolutionary process and spatial distribution of seaports along the Chinese coastline.

© 2011 Elsevier Inc. All rights reserved.

E-mail addresses: funke@econ.uni-hamburg.de (M. Funke), hao.yu@wiso.uni-hamburg.de (H. Yu).

<sup>&</sup>lt;sup>1</sup> Arguably, this development of waterborne container traffic illustrates the shift in the gravity centre of global economic activity. The emergence, and now dominance, of Asia in container traffic stands out.

<sup>1043-951</sup>X/\$ – see front matter 0 2011 Elsevier Inc. All rights reserved. doi:10.1016/j.chieco.2011.01.002

firms tend to locate in particular high-value-added sectors in close proximity to a seaport. Furthermore, Kawakami and Doi (2004) have presented empirical evidence that seaport capital has Granger-caused economic development in Japan.

The tremendous explosion of foreign trade has also created a situation wherein ports today are not only redistribution centres but also generate a wide range of commercial activities and become entrepôts, i.e. commercial centres with related industries. Given these developments and their immense economic importance, it is a surprising omission in the literature that the emergence and growth of seaport cities in China have not been studied in greater detail. In order to fill this gap and explore this unchartered territory, the present paper will employ a theoretical modelling framework that facilitates rigorous investigation of the hypothesis that seaports are a fertile ground for industrial concentrations. The geographical economics modelling approach laid out below incorporates location, increasing returns to scale, imperfect competition, and transport costs, in conjunction with mobility of factors of production. These modelling ingredients enable us to discover what allocation of seaport cities in China has been the most efficient when China opened its economy and began to industrialise.

The remainder of the paper is divided into three parts. Section 2 provides the historical context for the subsequent analysis. Section 3 is devoted to the modelling, which serves to map out the formation of Chinese seaports with a developed industrial urban character. In Section 4 we present some model simulation results. Finally, some conclusions are given in Section 5.

#### 2. Historical background

China before the First Anglo-Chinese War (1839–42), popularly known as the "Opium War" was closed to the West, and foreign trade was strictly controlled by the Chinese government. Ultimately, this was much more than a war about opium trade. Rather, the fundamental change which brought confrontation was the Industrial Revolution. Technological innovations, advances in communication and improvements in organizational capacities in Europe had enabled Britain, as the leading European nation, to capture markets and project power overseas.

Brought about by the insulate attitude of the Chinese Empire, late imperial China exercised strict control over its foreign trade. External trade was organized by the "Guangzhou Trade System", since only the seaport of Guangzhou in southern China was open to foreign trade.<sup>2</sup> Having reached Guangzhou, the Western merchants could only deal with a group of government appointed merchants that had a monopoly on the trade with the West. The volume of the trade and the prices were also regulated.

After its defeat in the Opium War, China was forced to open up. The unequal Treaty of Nanjing (August 1842) and subsequent treaties signed after the war were the primary means of opening China along with its markets and resources. They radically increased China's trade openness and opened additional treaty seaports, including Shanghai. Under the Treaty of Nanjing, China also ceded the island of Hong Kong. In the 1850s, the western countries grew increasingly dissatisfied with both the terms of their treaties with China, and the Qing Government's failure to adhere to them. The British forced the issue by attacking the Chinese port cities of Guangzhou and Tianjin in the second Opium War (1857–1860). In 1860 a combined British French army attacked Beijing and burned down the Old Summer Palace. The Treaty of Tianjin signed after the war granted the western countries further rights and privileges. The number of treaty seaports increased further, with new ports opened to Western trade along the Chinese coast, and along the Yangtze River in the interior. The era of the Treaty Port System lasted until 1948.

On the economic side, the opening-up of the country meant that China had lost its own economic protection against the influx of cheap foreign goods. The domestic handicraft industry was hard hit, and this engendered social and economic dislocations in China. But new technologies also arrived in China – the railroad, the telegraph – along with new administrative technologies and new ways of organizing financial institutions. The Chinese were quick to take advantage of these opportunities, and a tide of modernization and integration was on the rise.<sup>3</sup> In other words, the Industrial Revolution changed the landscape.<sup>4</sup> Small towns grew into huge cities, and urbanized seaports began to develop. Urbanization and economic development went hand in hand. A schematic spatial mapping of Chinese seaports at that time is given in Fig. 1.

The emergence of Chinese seaports took place in various stages and was shaped by the Opium Wars. Originally, the corner pillars Guangzhou and Tianjin served as the main ports. Guangzhou's natural advantages – its location and its local topography – gave it the preferred position on China's southeast coast for foreign trade. By the start of the Qing dynasty, Tianjin had become the leading economic centre of North China because of its nearby capital city, Beijing, and its location at the northern terminus of the Grand Canal. Historically, the Grand Canal was the designated channel for the transport of tribute-grain from the south and the east to the imperial capital. Tianjin Port grew rapidly as a port and commercial centre, and it became the chief storage, transfer, and distribution point for grain and other foodstuffs from central and southern China. After the two Opium Wars, further treaty

linked to the price of rice all along the Yangtze River Valley and to the international price (Shiue and Keller (2007)).

<sup>&</sup>lt;sup>2</sup> Additionally, Portugal had established a foothold in Macau in 1557 and other nations were allowed to trade with China through the seaport of Macao.
<sup>3</sup> One indication of China's increasing economic integration was price convergence. For example, by the end of the 1880s, the price of rice in Shanghai was

<sup>&</sup>lt;sup>4</sup> An interesting question is why the Industrial Revolution did not take place in China although China was clearly the world leader in scientific developments, production technologies and administrative organization until around the 16th century. And once there was an Industrial Revolution, why did it take so long to get to China? Many economists favour institutional explanations of growth and stagnation around the time of the Industrial Revolution. The account has at its core the hypothesis that China was unable to break through the institutional and structural barriers to the industrial-scientific revolution. See, for example, Lin (2008). As an alternative thereto, cultural norms may also help to explain divergent development. Nunn and Wantchekon (2009) have disentangled various channels through which cultural norms influence development. For a textbook treatment of institutions, culture, and development, see Aghion and Howitt (2009), pp. 237–264 and pp. 419–442. Williamson (2010) has also explained when, where, and why industrialisation in various countries took place.



Fig. 1. Formation and location of urbanized chinese seaports.

seaports were gradually opened. Among them was the conveniently located seaport of Shanghai, with its easy access to the Yangtze waterway and other main trading routes. Further coastal cities opened up to foreign trade included Fuzhou and Qingdao and the ports of Wuhan and Chongqing on the Yangtze River. The mighty Yangtze River was the most important waterway for trade and communications in the richest part of China. We will argue below that this historical context offers an unprecedented and so far unexplored event to analyse the explanatory power and guidance of regional economics modelling frameworks.

#### 3. Seaport cities' genesis - tools and applications

The focus of this section is to understand the interplay of seaport development and city growth.<sup>5</sup> The basic hypothesis is that the role of seaports goes far beyond the nautical dimension. Seaports were trade centres that generated a wide range of economic activities. First, we will develop a modelling framework to understand the formation and expansion of Chinese seaports and their connectivity with the hinterland.<sup>6</sup> Let us describe the conceptual framework informally. The new economic geography research agenda fills the gap left by traditional trade theories, as it describes the formation of economic agglomerations in geographic space [Fujita, Krugman, and Venables (1999) and Fujita and Krugman (2004)]. The rationale behind regional economic imparity is that agglomeration creates growth, and certain regions experience forces that encourage agglomeration and others experience the reverse forces. Fujita and Mori (1996) analyse the role of seaports in the development of cities with a developed industrial urban character. Their evolutionary model of spatial economic development explains how agglomerations and seaports arise from the interaction of random historical events, increasing returns, transport costs, and seaborne trade. Below we analyse in detail the mechanisms underlying the formation and growth of urbanized Chinese seaports during the Industrial Revolution.

#### 3.1. Basic framework

We begin with a sketch of the seaport model with non-neutral space which is adapted from Fujita and Mori (1996). For simplicity, the quality of land is the same everywhere and all non-land factors are mobile. Labour is assumed to be the only mobile factor of production and each worker is endowed with one unit of labour. Workers can change jobs and locations. Consumers derive utility from consumption; and there are two types of consumption goods, a homogeneous agricultural good (*A*) and differentiated manufactured goods (*M*).<sup>7</sup> Preferences are represented by the utility function

$$U = A^{1-\mu}M^{\mu},$$

(1)

<sup>&</sup>lt;sup>5</sup> Despite widespread agreement that industrial co-location can generate positive economic externalities, academics and policymakers still strive for a thorough understanding of the mechanisms through which clusters can be expected to deliver economic growth. Nevertheless, clustering has become a policy panacea for many governments, and international agencies that see clusters as drivers of regional and national competitiveness and growth. We do not deal explicitly here with the subtle policy implications of the formal model.

<sup>&</sup>lt;sup>6</sup> Rawski (1969) has shown that even though treaty ports were opened to foreign merchants, western merchants relied heavily on Chinese middlemen (the so-called *compradors*) to conduct business in China. Therefore, there was no sharp distinction between the treaty ports and the hinterland in which business practices remained largely unaffected by western influence. In a similar manner, Murphy (1970) has argued that western merchants did not make any institutional change to centralised Imperial China. Putting it bluntly, Murphy (1970) has compared the treaty port system with a fly on the elephant.

<sup>&</sup>lt;sup>7</sup> Defining two sectors as being agriculture and manufacturing is arbitrary. The main point is that one sector employs an immobile factor of production producing a homogeneous good that is freely tradable under perfect competition, and that the other sector employs a mobile factor of production, producing heterogeneous varieties that are costly to trade between regions under monopolistic competition.

where the expenditure share of manufactured goods ( $\mu$ ) is strictly between 0 and 1. The CES aggregator of differentiated manufacturing goods is given by

$$M = \left[\int_{0}^{n} m_{i}^{\rho} di\right]^{1/\rho},\tag{2}$$

where *n* is the range of manufacturing varieties and  $0 < \rho < 1$ .<sup>8</sup> The utility function exhibits the feature that, ceteris paribus, the larger the variety of differentiated manufacturing goods, the greater the utility. The budget constraint of consumers is

$$p^{A}A + \int_{0}^{n} p_{i}m_{i}di = Y, \tag{3}$$

where  $p^A$  is the price of food and Y is income. The solution to this program leads to the demand functions for food and for manufacturing good *i*:

$$A = \frac{(1-\mu)Y}{p^A} \tag{4}$$

$$m_i = \frac{\mu Y p(i)^{-\sigma}}{G^{-(\sigma-1)}},\tag{5}$$

where  $\sigma \equiv 1/(1-\rho)$  is the elasticity of substitution between any two differentiated manufacturing goods and

$$G = \left[\int_{0}^{n} p_i^{1-\sigma} di\right]^{1/(1-\sigma)} \tag{6}$$

is the price index for manufacturing goods. We now turn to the supply side. The agricultural-good production is subject to constant returns to scale, requiring one unit of land and one unit of labour. The production of differentiated manufactured goods exhibits increasing returns such that labour input  $l^M$  for producing  $q^M$  manufactured goods is given by

$$l^M = F + c^M q^M,\tag{7}$$

where *F* represents a fixed input and  $c^{M}$  is the variable labour input.<sup>9</sup> We impose symmetry across varieties by assuming the same inputs across varieties. Producers set a price that maximises instantaneous profits:

$$\pi = p^M q^M - w^M \left( F + c^M q^M \right),\tag{8}$$

where  $w^M$  is the wage in the manufacturing industry. Free entry implies zero profits. In turn, this implies that the price for manufactured goods is

$$p^{M} = \frac{c^{M} w^{M}}{(1 - 1/\sigma)}.$$
(9)

The resolution of this program leads to the following equilibrium output and input of manufacturing goods:

$$q^* = \frac{F(\sigma - 1)}{c^M} \tag{10}$$

$$l^* = F\sigma \tag{11}$$

The equilibrium number of varieties is then determined by equating the profits of the marginal firm to zero. Denoting the endogenously determined variety by  $n^*$ , we have

$$n^* = \frac{L^M}{l} = \frac{L^M}{F\sigma}.$$
(12)

<sup>&</sup>lt;sup>8</sup> Using information from the Chinese Maritime Customs, Keller, Li, and Shiue (2010) have shown that there was a very notable expansion in the diversity of product categories imported into China during the Treaty Port era. Overall, the number of differentiated imports rose from 80 to 483, or 504% for the period 1868 to 1947.

<sup>&</sup>lt;sup>9</sup> This issue will become important below. The Marshallian externalities arise from three sources: labour market pooling, the creation of specialised suppliers, and the emergence of knowledge spillovers.

We assume that the transport cost of each good takes Samuelson's iceberg form. The "deglaciation" of goods shipped over distance *d* follows the exponential function  $e^{-\tau^{id}}$  (*i*=*A*, *M*), where  $\tau^{i}$  is a positive constant.<sup>10</sup> Thus the delivered price of manufacturing goods produced at location *s* and consumed at location *r* is  $p^{M}(r)e^{-\tau^{Mi}|r-s|}$ . Substituting this price in (6) yields the manufactured good price index at location *r*:

$$G(r) = \left\{ \int_{s \in \mathbb{R}} n(s) \left[ p^{M}(s) e^{\tau^{M} | r - s|} \right]^{1 - \sigma} ds \right\}^{\frac{1}{1 - \sigma}},\tag{13}$$

where R is the geographic range of the economy. Following (5), the consumption demand at location for a certain manufacturing variety produced at r is

$$m_i(s) = \mu Y(s) \left[ p^M(r) e^{\tau^M |r-s|} \right]^{-\sigma} G(s)^{\sigma-1},$$
(14)

where  $m_i(s)$  and Y(s) denote demand for manufacturing variety *i* and total income at location *s*, respectively. In order to supply amount  $m_i(s)$  of product to location *s*, the amount  $m_i(r)e^{-\tau^{Mi}|r-s|}$  should be produced at *r*. The total sales of manufacturing variety *i* at location *r*, denoted  $q^M(r)$ , thus amounts to

$$q^{M}(r) = \mu \int_{s \in \mathbb{R}} Y(s) \left[ p^{M}(r) e^{\tau^{M} |r-s|} \right]^{-\sigma} G(s)^{\sigma-1} e^{\tau^{M} |r-s|} ds.$$
(15)

Next, nominal wages can be determined. By the zero profit condition equilibrium output is equal to

$$q^* = \mu \int_{s \in R} Y(s) \left[ p^M(r) \right]^{-\sigma} e^{-(\sigma - 1)\tau^M |r - s|} G(s)^{\sigma - 1} ds.$$
(16)

Reverse engineering of (16) yields

$$\left[p^{M}(r)\right]^{\sigma} = \frac{\mu}{q^{*}} \int_{s \in \mathbb{R}} Y(s) e^{-(\sigma-1)\tau^{M}|r-s|} G(s)^{\sigma-1} ds.$$
(17)

Substituting this equation for  $p^{M}$  into Eq. (9), yields the nominal wage of a manufacturing worker at location r:

$$w^{M}(r) = \left(\frac{\sigma-1}{\sigma \cdot c^{M}}\right) \left[\frac{\mu}{q^{*}} \int_{s \in \mathbb{R}} Y(s) e^{-(\sigma-1)\tau^{M}|r-s|} G(s)^{\sigma-1} ds\right]^{\frac{1}{\sigma}}.$$
(18)

Without loss of generality, we assume

$$c^{M} = \frac{\sigma - 1}{\sigma}.$$
(19)

This definition simplifies the notation in the equations below. It follows that  $p^M = w^M$  and  $q^* = l^*$ . For further simplification we assume

$$F = \mu / \sigma. \tag{20}$$

It can be easily verified that the equilibrium number of firms in each location (Eq. (12)) is constant:

$$n^* = \frac{L^M}{\mu}.$$
(21)

Accordingly, the equilibrium output level at which firms make zero profit becomes

$$q^* = l^* = \mu. \tag{22}$$

<sup>&</sup>lt;sup>10</sup> The cost of transport is assumed to be a constant. Alternatively, one could assume economies of transport density. The average cost of processing freight may fall with the quantity processed at a particular port, creating economies of transport at seaports and river junctions with access to the sea. This leads to circular causation: processing industries prefer agglomerations and this leads to some reinforcing force due to an endogenous improvement in the efficiency of transport. For a geographic economy model with such economies of transport density, see Mori and Nishikimi (2002).

Finally, the price index and the wage can be written as

$$G(r) = \left\{ \int_{s \in \mathbb{R}} n(s) \left[ p^{M}(s) e^{-\tau^{M} |r-s|} \right]^{1-\sigma} ds \right\}^{\frac{1}{1-\sigma}} = \left\{ \frac{1}{\mu} \int_{s \in \mathbb{R}} L^{M}(s) \left[ w^{M}(s) e^{-\tau^{M} |r-s|} \right]^{1-\sigma} ds \right\}^{\frac{1}{1-\sigma}}$$
(23)

201

$$w^{M}(r) = \left(\frac{\sigma-1}{\sigma \cdot c^{M}}\right) \left[\frac{\mu}{q^{*}} \int_{s \in \mathbb{R}} Y(s) e^{-(\sigma-1)\tau^{M}|r-s|} G(s)^{\sigma-1} ds\right]^{\frac{1}{\sigma}} = \left[\int_{s \in \mathbb{R}} Y(s) e^{-(\sigma-1)\tau^{M}|r-s|} G(s)^{\sigma-1} ds\right]^{\frac{1}{\sigma}}$$
(24)

Thus, the model is fully determined and we can now obtain the spatial equilibrium and the emergence of Shanghai as a further seaport city.

#### 3.2. Monocentric spatial equilibrium

Armed with this framework, we can now turn to our specified abstract Chinese economy. Space is one-dimensional and stretches between the two peripheral seaport cities Beijing/Tianjin and Guangzhou (see Fig. 1).<sup>11</sup> The entire production of manufacturing goods is assumed to take place in both cities, and the surrounding agricultural area extends from each of the peripheral cities towards the central region (hub).<sup>12</sup> One question is how to draw the catchment areas and hence the borders of both peripheral cities. Another question is whether a new seaport city will emerge over and above one of the existing cities.<sup>13</sup> Suppose that initially the population size is small and therefore only one city (either Beijing/Tianjin or Guangzhou) has already emerged at location 0. This city is specialised in manufacturing goods and exports manufactures for agricultural goods.

Next, assume that the economy grows. This leads to a larger city, and more manufactured varieties will be produced, leading to increasing returns at the city level. As the city population grows, further farmland has to be developed to support the growing city.<sup>14</sup> Eventually, as long as the population keeps growing, manufactured goods and agricultural products will be transported over increasing distances, leading to a dissemination of growth benefits across the country. Finally, beyond some adjacent catchment area, a new seaport city will emerge from the seaport-hinterland dynamics. In the course of this process, the trigger point for the emergence of a new agglomeration is where the cost of setting up production in a new city is less than that of transporting goods.

To clarify, we have depicted the geometry of our stylised Chinese economy in Fig. 2.<sup>15</sup> In our abstract schematic graph, Shanghai is at the location of the hub, and the branch stretching in direction *bz* from the hub is the Yangtze River Valley, the longest inland river in China. To make the presentation comprehensible, we denote the line extended from the segment of the centre city to the hub, or *Obx*, as baseline; the branch on this line, *bx*, therefore is the baseline branch. All the other branches are nonbaseline branches. The number of nonbaseline branches at the hub is *k*. Without loss of generality, we assume k = 1.

The above intuitive explanation is inaccurate because it is not clear when the new city will emerge. To answer this question precisely, we introduce a new variable  $\delta$  to indicate whether the farm hinterland reaches the existing hub location.

$$\delta = \begin{cases} 0 & f \le b \\ 1 & f > b \end{cases}$$
(25)

The economic distance between the marginal farmland and the peripheral city is denoted by f and there is a hub at location b. For  $f \le b$ , the hub is beyond the existing city's sphere of influence and therefore no new city will emerge. On the contrary, for f > b the necessary condition for the emergence of a new city at the Yangtze River Valley hub is fulfilled. The apparent next step is to determine the critical value of f. Suppose that the agricultural good has to be transported from the hinterland to the city. At each location r in the hinterland, it must hold that

$$p^A(r) = p^A e^{-\tau^A r}.$$
(26)

<sup>&</sup>lt;sup>11</sup> See Fujita, Krugman and Venables (1999), pp. 136–140. One must bear in mind that the model makes a number of simplifications. One limitation is the implicit assumption that the coastline is uniform, i.e. we don't model coastlines with different water depths. This may lead to limited coastline resources with deep water, which is critical to constructing port facilities.

<sup>&</sup>lt;sup>12</sup> Due to the geographic restriction, the farm hinterland of Guangzhou and Beijing/Tianjin can only stretch in one direction. This is different from Fujita and Krugman's (1995) monocentric equilibrium model, where farm hinterland develops symmetrically at both sides of the centre city.

<sup>&</sup>lt;sup>13</sup> New cities need not be seaport cities. Self-organizing forces may also lead to new non-port cities in the municipal area of the existing city. But a seaport city has the natural advantage of being a transport node for trade. Shiue and Keller (2007) estimated the current relative costs ratio of sea transport vs. inland water transport vs. overland transport in China at 1:2.7:9.5. Thus, seaports had an eminent comparative advantage in transport. Needless to say, the comparative advantage of navigable waterways is not just the result of geography but also of investment in seaport facilities and port expansion programmes.

<sup>&</sup>lt;sup>14</sup> Traditionally, seaports have offered connectivity towards land- and ocean-side. In the current globalised world, container ports are part of a larger logistic chain, i.e. a global distribution channel.

<sup>&</sup>lt;sup>15</sup> The implicit assumption in Fig. 2 is that the farm hinterlands of the existing two cities do not overlap and so there is no economic integration between the two peripheral cities. Numerical simulations of the model in Section 4 support this assumption.



Fig. 2. Genesis of Shanghai.

By the zero-profit condition, the land rent and agricultural wage at each location r in the hinterland are given by

$$R(r) = p^{A}(r)e^{-\tau^{A}r} - c^{A}w^{A}(r)$$
(27)

and

$$w^{A}(f) = \frac{p^{A}e^{-\tau^{A}f}}{c^{A}},$$
(28)

respectively. Turning to manufacturing, let  $p^{M}(0) = w^{M}(0) = 1$  be the price of manufactured goods at the central location. From Eq. (23) we obtain the price index for location *r*:

$$G(r) = \left(\frac{L^{M}}{\mu}\right)^{1/(1-\sigma)} e^{\tau^{M}r}$$
<sup>(29)</sup>

where  $L^M = N - c^A(f + \delta k(f - b))$  is the labour force working in the manufacturing sector, which is equal to the total workforce less the number of farmers.<sup>16</sup>

Let us now define the supply-demand relationship determining *f*. The urban worker's income share spent on food is  $1 - \mu$ . Total food demand in the city is therefore  $D^A = (1 - \mu)w^M L^M / p^A$ . Likewise, farmers consume the fraction  $1 - \mu$  of their harvest and transport the remaining part  $\mu$  to the city. Therefore, food supply in the city is  $S^A = \mu \left( \int_0^f e^{-\tau^A s} ds + \delta k \int_b^f e^{-\tau^A s} ds \right)$ . Market clearing  $D^A = S^A$  implies

$$p^{A} = \frac{(1-\mu) \left[ N - c^{A} (f + \delta k (f - b)) \right]}{\mu A(f)},$$
(30)

where  $A(f) = \int_0^f e^{-\tau^A s} ds + \delta k \int_b^f e^{-\tau^A s} ds$ . Finally we must ensure that the real wages of farmers in the hinterland and manufacture workers in the city are identical. This requirement leads to another equilibrium relationship of  $p^A$  with f. The real wage at each location r is proportional to the nominal wage deflated by the cost-of-living index,  $G(r)^{\mu}[p^A(r)]^{1-\mu}$ . Therefore we can solve for the real wage of farmers at the fringe of farm hinterland:

$$\omega^{A}(f) = w^{A}(f)[G(f)]^{-\mu} \Big[ p^{A}(f) \Big]^{-(1-\mu)} = \frac{1}{c^{A}} \Big( p^{A} \Big)^{\mu} G^{-\mu} e^{-\mu \left(\tau^{M} + \tau^{A}\right) f}.$$
(31)

Since manufacturing is geographically concentrated in the city, the manufacturing real wage is

$$\omega^{M} = G^{-\mu} \left( p^{A} \right)^{-(1-\mu)}.$$
(32)

The requirement of real wages of farmers and manufacturing workers being equal implies the no-arbitrage condition:

$$p^{A} = c^{A} e^{\mu (\tau^{A} + \tau^{M}) f}.$$
(33)

The above information is enough to determine the equilibrium. We can readily see that (30) and (33) together uniquely determine the equilibrium farm hinterland range f, as well as the agricultural good price in the city,  $p^A$ .

<sup>&</sup>lt;sup>16</sup> The cultivated area on the baseline is always equal to f, while in every nonbaseline branche it is  $\delta k(f-b)$ . The total cultivated area is therefore  $f + \delta k(f-b)$ .

#### 3.3. Market potential function and possible emergence of new urbanized seaports

In this subsection we derive a condition for alternative equilibrium configurations. To this end, let us define the market potential function for manufacturing firms, which was introduced by Fujita and Krugman (1995). The market potential function  $\Omega(r)$  measures the relative real wage of manufacturing workers and farmers at a certain location *r*:

$$\Omega(r) = \frac{\left[\omega^{M}(r)\right]^{\sigma}}{\left[\omega^{A}(r)\right]^{\sigma}},\tag{34}$$

where  $\omega^{M}(r)$  and  $\omega^{A}(r)$  denote the real wage of manufacturing workers and farmers at location *r*, respectively, and  $\sigma \equiv 1/(1-\rho)$  is the elasticity of substitution between any two differentiated manufacturing goods defined above. The market potential function measures the relative profitability at each location *r* for manufacturing firms. Hence, the location monocentric (single-city) equilibrium requires that market potential function not exceed 1 anywhere in the economy, i.e.

$$\Omega(r) \le 1 \tag{35}$$

for all *r*. In other words, for  $\Omega(r) < 1$  the centripetal force created by the existing city is so strong that no new city can emerge. Put differently, for  $\Omega(r) > 1$  manufacturing workers would be attracted to location *r*, triggering a self-enhancing feedback effect of spatial agglomeration there. It is straightforward to show that

$$\Omega(r) = \frac{\left[\omega^{M}(r)\right]^{\sigma}}{\left[\omega^{A}(r)\right]^{\sigma}} = \frac{\left\{w^{M}(r)G(r)^{-\mu}\left[p^{A}(r)\right]^{-(1-\mu)}\right\}^{\sigma}}{\left\{w^{A}(r)G(r)^{-\mu}\left[p^{A}(r)\right]^{-(1-\mu)}\right\}^{\sigma}} = \frac{\left[w^{M}(r)\right]^{\sigma}}{\left[w^{A}(r)\right]^{\sigma}} = \left[w^{M}(r)\right]^{\sigma}e^{\sigma\left[(1-\mu)\tau^{A}-\mu\tau^{M}\right]r}.$$
(36)

Recalling the definition of  $w^{M}(r)$  in Eq. (24), we can decompose the wage into three parts as follows:

$$\begin{bmatrix} w^{M}(r) \end{bmatrix}^{\sigma} = L^{M} w(0) e^{-(\sigma-1)\tau^{M}r} [G(0)]^{\sigma-1} + \int_{0}^{f} p^{A}(s) e^{-(\sigma-1)\tau^{M}|r-s|} [G(s)]^{\sigma-1} ds + \delta k \int_{b}^{f} p^{A}(s) e^{-(\sigma-1)\tau^{M}d(r,s)} [G(s)]^{\sigma-1} ds,$$
(37)

where d(r,s) is the distance of farmers from the baseline, given by

$$d(r,s) = \begin{cases} s-r & f \le b\\ (r-b) + (s-b) & f > b \end{cases}$$
(38)

In Eq. (38) we divide the whole economy visualised in Fig. 2 into three segments: the original city, the segment from the centre city to the edge of the hinterland on the baseline *obx*, and the farm hinterland on the nonbaseline branch starting from junction point *b*. Substituting (26) and (30) into (37) yields

$$\left[w^{M}(r)\right]^{\sigma} = \mu e^{-(\sigma-1)\tau^{M}r} + \left(\frac{\mu p^{A}}{L^{M}}\right) \left[\int_{0}^{f} e^{-\tau^{A}s} e^{(\sigma-1)\tau^{M}(s-|r-s|)} ds + \delta k \int_{b}^{f} e^{-\tau^{A}s} e^{(\sigma-1)\tau^{M}(s-d(r,s))} ds\right].$$
(39)

It follows straightforwardly from (31) that  $\mu p^A/L^M = (1 - \mu)/A(f)$ . Let us repeat the dynamic system we have arrived at. The potential function  $\Omega(r)$  is a piecewise function to the left-hand side and to the right-hand side of the hub as follows:

$$\Omega(r) = \begin{cases} \Omega_1(r) & \text{for } r \le b \\ \Omega_2(r) & \text{for } r > b \end{cases},$$
(40)

where the potential functions for the two regions can be solved as

$$\Omega_{1}(r) = e^{\sigma[(1-\mu)\tau^{A} - \mu\tau^{M}]r} \Big[ \mu e^{-(\sigma-1)\tau^{M}r} + \psi(r,f)(1-\mu)e^{(\sigma-1)\tau^{M}r} \Big] \quad \text{for } r \le b$$
(41)

and

$$\Omega_{2}(r) = e^{\sigma[(1-\mu)\tau^{A}-\mu\tau^{M}]r} \Big\{ \mu e^{-(\sigma-1)\tau^{M}r} + \psi(r,f)(1-\mu)e^{(\sigma-1)\tau^{M}r} - (1-\mu)\phi(f)e^{(\sigma-1)\tau^{M}r} \Big[ 1 - e^{-2(\sigma-1)\tau^{M}(r-b)} \Big] \Big\} \quad \text{for } r > b, \quad (42)$$

where  $\phi(f) = \delta k \int_{b}^{f} e^{-\tau^{A}s} ds / A(f)$  and  $\psi(r, f) = 1 - \int_{0}^{r} e^{-\tau^{A}s} \left[ 1 - e^{-2(\sigma-1)\tau^{M}(r-s)} \right] ds / A(f)$ , and A(f) is defined above in Eq. (30).

The model considered in this section has an obvious merit. Once the  $\Omega(r)$  bifurcation criterion is established for the emergence of further seaports, we can calibrate possible equilibrium configurations. Thus, we can examine whether new hubs will be formed and therefore whether the framework is important for explaining the observed geographic dispersion of further Chinese seaports. Or to put it somewhat differently: The calibrations will shed light on whether the initial seaport hierarchy in the Beijing/Tianjin– Guangzhou range is likely to be challenged.

#### 4. Spatial dynamics in the Beijing/Tianjin-Guangzhou range

Section 3 developed and discussed the main features of the model and paved the way for the numerical calibrations exercise. The next task is to calibrate the emergence of urbanized Chinese seaports in the setup and to explore the sensitivity of the results to changes in the parameters. This allows to get a feel for the model and the space–time dynamics.

Before initiating the model calibrations, the slope of the potential function at the centre city (r=0) is calculated from (41) as

$$\frac{d\Omega_1(0)}{dr} = \sigma \Big[ (1-\mu)\tau^A - (2\rho\mu + \mu - \rho)\tau^M \Big].$$
(43)

The stability of the monocentric equilibrium requires that the slope of the potential function is negative, i.e.  $d\Omega_1(r)/dr < 0 \Leftrightarrow (1-\mu)\tau^A < (2\rho\mu + \mu - \rho)\tau^M$ . Otherwise, the relocation of an arbitrarily small number of manufacturing firms would lead to the formation of a new city. On the other hand, a new city will emerge only when the farm hinterland of the city exceeds a critical threshold *f*. From Eq. (41), the limit of the potential function for  $f \rightarrow \infty$  can be derived as

$$\overline{\Omega}_{1}(r) = \Omega_{1}(r,f) | f \to \infty = K e^{\sigma(\rho - \mu) \left(\tau^{A} + \tau^{M}\right)r} + (1 - K) e^{-\left[(1 - \mu)(\sigma - 1)\tau^{M} - d\Omega_{1}(0)/dr\right] \cdot r},$$
(44)

where *K* is a positive constant, and  $\overline{\Omega}_1(r)$  is thus the upper limit of potential curve  $\Omega_1(r)$ . It can be verified that the condition  $\overline{\Omega}_1(r) > 1$  for  $f \to \infty$  is that  $\rho > \mu$ .

Given that these conditions are met, the historical evolution of Chinese port cities can be simulated. How useful is the theoretical modelling framework for China's economic history? Can the framework be (loosely) fitted to a variety of different circumstances? These questions are addressed below. Rather than presupposing the existence of seaports, we simulate the spatial arrangement endogenously. Our mode of conduct is straightforward: First, we calculate the potential function  $\Omega(r)$  for given parameter values. Once the maximum value of the potential function exceeds 1, a bifurcation occurs and a new city emerges. Because methodological issues related to calibration are not the focus of this paper, a pragmatic stance is taken. The coastline between Guangzhou and Shanghai is approximately 1200 km long, and the coastline distance between Beijing/Tianjin and Shanghai is about 1800 km long. Therefore we assume the corresponding lengths of these two coastlines are 0.4 and 0.6, respectively. Hence, in Fig. 4 Guangzhou locates at the origin, Shanghai at b = 0.4 and Beijing/Tianjin at r = 1. Next we focus on the population data. To obtain benchmark population data, we first determine the former prefectures comprising the catchment areas of Guangzhou (denoted by  $N_G$ ) and Beijing/Tianjin ( $N_B$ ).<sup>17</sup> Fig. 3 illustrates the resulting boundaries of the two catchment areas.<sup>18</sup> Using the prefecture-level population data of Cao (2001, Table 16-1), the resulting population numbers for 1820 were  $N_G = 47.18$  million and  $N_B = 27.44$  million, respectively. The ratio of the two is approximately 0.78. Therefore we assume that the population parameter for Guangzhou city and its periphery is 1.0 while that for Beijing/Tianjin and its periphery is 0.78.

The remaining parameters, which were chosen for realism, are  $\sigma$ =3,  $\mu$ =0.45,  $\tau^{A}$ =1.2,  $\tau^{M}$ =1.4, and  $C^{A}$ =0.78. The substitution elasticity between differentiated products is notoriously difficult to estimate. Following Mitchener and Yan (2010) in their general equilibrium analysis for China in the 1920s, we set the value of  $\sigma$  to equal 3. During the Qing Dynasty, agricultural production was believed to be a fundamental prerequisite. On the contrary, higher taxes and tariffs were levied on manufacturing products and trade. Therefore we assume the transporting cost parameter of manufacturing goods ( $\tau^{M}$ =2.4) to be higher than for

<sup>&</sup>lt;sup>17</sup> The catchment area of Guangzhou comprises the former prefectures Guangzhou Fu, Huizhou Fu, Chaozhou Fu, Zhangzhou Fu, Quanzhou Fu, Xinghua Fu, Fuzhou Fu, Funing Fu, Wenzhou Fu, Taizhou Fu, Ningbo Fu, Shaoxing Fu, Hangzhou Fu, Jiaxing Fu, Songjiang Fu, and Suzhou Fu marked in blue. The catchment area of Beijing/Tianjin comprises the former prefectures Zunhua Zhou, Shuntian Fu, Tianjin Fu, Wuding Fu, Jinan Fu, Qingzhou Fu, Laizhou Fu, Dengzhou Fu, Yizhou Fu, Haizhou, Huaian Fu, and Yangzhou Fu marked in red.

<sup>&</sup>lt;sup>18</sup> The chosen delimitation of both catchment areas is also consistent with the fact that the Grand Canal ended in Hangzhou. The Chinese Grand Canal, also known as the Beijing–Hangzhou Grand Canal, interconnected the Yangtze, Yellow, Huaihe, Haihe, and Qiantang Rivers and flowed through Beijing, Tianjin, Hebei, Shandong, Jiangsu and Zhejiang with Hangzhou at its southernmost end. The Grand Canal offered much facility to transport foods and goods from south to north in past times.



Fig. 3. Prefectures comprising the catchment areas of Guangzhou and Beijing/Tianjin.



Fig. 4. Potential curves for Guangzhou and Beijing/Tianjin.

food ( $\tau^A = 1.4$ ). The spending ratio on manufactured goods is assumed to be  $\mu = 0.45$ . Given these parameters, the parameter constraints in Eqs. (43) and (44) are satisfied.<sup>19</sup>

Armed with an empirically plausible parameterisation of the model and given the two initial peripheral cities, the key question is where are new seaport cities likely to emerge? There are several noteworthy features represented in Fig. 4. To begin with, the solid line on the left indicates the baseline potential curve for Guangzhou, while the dashed curve starting from the right side represents the potential curve for Beijing/Tianjin. For robustness checks, we have also drawn both potential curves for smaller population sizes (solid and dashed thin lines for  $N_G = 0.8$  and  $N_B = 0.63$ ; all other parameters as in the baseline case). As expected, both curves shift downward as population size decreases. As shown in Fig. 4, irrespective of the assumed population size, the potential curve for Guangzhou exceeded the  $\Omega(r) = 1$  threshold.<sup>20</sup> Thus, the calibrations indicate the emergence of another urbanized seaport along the southern coastline. However, due to the Qing government's closed-door policy and attempts to limit contacts with the outside, no further seaport cities emerged at first.

<sup>&</sup>lt;sup>19</sup> It may be noteworthy that the calibrations primarily serve as a communication device. The modelling approach introduced in this paper goes some way towards achieving the purpose of understanding the dynamics of the theoretical framework. Yet we do not claim empirical accuracy for the model, which we use rather for qualitative features and predictions.

<sup>&</sup>lt;sup>20</sup> We haven't considered regional wage differentials and consequentially thresholds  $\Omega(r) \neq 1$ . The scarcity of data makes it virtually impossible to construct Chinese city-hinterland wage differentials for the mid 19th century. Furthermore, Yan (2007, p. 14, footnote 19) shows that regional wage differentials were small until 1890 when the ban on foreign direct investment was lifted. Thus,  $\Omega(r) = 1$  is a defendable assumption in our view.



Fig. 5. Potential curves for Guangzhou, Shanghai and Beijing/Tianjin.

The area  $\Omega(r) > 1$  comprehended the location of Shanghai (b = 0.4). Accordingly, the emergence of Shanghai as a seaport city was within the realm of possibility and finally occurred once the Qing government ended its policy of seaport closings. Its location at the mouth of the Yangtze River Delta initially led to its development as a coastal trade port. After Shanghai became an international treaty trade port in 1842, foreign ships, shipyards and related business increased rapidly and finally Shanghai developed into an international transportation hub. Industry was another of the impetuses for urban development and interactive, but the development of the seaport has been and still is one of the important dynamics of urban expansion in Shanghai.<sup>21</sup> On the other hand,  $\Omega(r)$  for Beijing/Tianjin is always smaller than 1. This implies that the lock-in effect of Beijing/Tianjin prevented the formation of a further seaport city along the northern coastline at that time.<sup>22</sup>

Given this evolutionary process, the additional potential function for Shanghai, denoted as  $\Omega_S(r)$ , can be derived in line with (41) and (42) as follows<sup>23</sup>:

$$\Omega_{S}(r) = e^{\sigma \left[(1-\mu)\tau^{A} - \mu\tau^{M}\right]r} \left\{ \left[ \frac{1+\mu}{2} - \frac{(1-\mu)\phi_{S}(f)}{2} \right] e^{-(\sigma-1)\tau^{M}r} + \psi_{S}(r,f) \left( \frac{1-\mu}{2} \right) e^{(\sigma-1)\tau^{M}r} - (1-\mu)\phi_{S}(f) e^{(\sigma-1)\tau^{M}r} \left[ 1 - e^{-2(\sigma-1)\tau^{M}(r-b)} \right] \right\}$$

$$\tag{45}$$

where  $\phi_S(f) = \int_0^f e^{-\tau^A s} ds / A_S(f)$ ,  $\psi_S(r, f) = 1 + \phi_S(f) - 2\int_0^r e^{-\tau^A s} \left[1 - e^{-2(\sigma-1)\tau^M(r-s)}\right] ds / A_S(f)$ , and  $A_S(f) = 3\int_0^f e^{-\tau^A s} ds$ . Given the city population data of Cao (2001) for 1910, the population size of Guangzhou, Shanghai (denoted by NS) and Beijing/Tianjian at this historical stage is assumed to be NG = 0.9, NS = 0.3 and NB = 1.05. All remaining parameters are the same as in Fig. 4. Again we present the outcomes graphically instead of with unreadable, large tables. The new potential curves after the development of Shanghai are provided in Fig. 5.

As before, the solid and dashed lines represent the potential curves for Guangzhou and Beijing/Tianjin, respectively. The tapered dotted line represents the potential curve for Shanghai. The significance of the diagram is that along the northern coastline a new seaport was likely to emerge in a range of 0.6 < r < 0.9. It is particularly noteworthy that this calibration result is consistent with the formation of Qingdao as a further seaport city. Before the 17th century, the Port of Qingdao was little more than a small fishing village. In 1891, the Qing Dynasty government began to extend and fortify the seaport of Qingdao. In 1898, the German navy overcame these defenses, and the city was ceded to Germany. They made the Port of Qingdao a free port in 1899. As a result, the economy swelled and the city was undergoing a period of rapid growth based upon light manufacturing industries.

Along the southern coastline, Guangzhou's potential curve touched  $\Omega(r) = 1$  near to r = 0.25. Thus it was advantageous to establish a new seaport city there. This new seaport was Fuzhou, which was opened to foreign trade after the Opium War, in 1842, as one of the five unequal treaty ports. Fuzhou immediately benefited from this and became the chief port for tea trade. The opening also enhanced the development of Fuzhou's urban market economy.<sup>24</sup>

<sup>&</sup>lt;sup>21</sup> A comprehensive review of Shanghai's port origin and city formation is provided in Zhongmin and Jianzhong (1990).

<sup>&</sup>lt;sup>22</sup> It is worth mentioning that due to  $f_G + f_B = 0.9 < 1$ , the farm hinterlands of the two peripheral cities indeed do not overlap. This fact validates our previous assumption that the two monocentric cities co-exist.

<sup>&</sup>lt;sup>23</sup> Contrary to both peripheral cities, Shanghai had three directions to extend its hinterland: along the northern and southern coastline and along the Yangtze River Valley. To get the potential curves for this special case, set b=0 and k=2. Note that the potential curves in Fig. 5 only typify the stretch-out along the northern and southern coastline. In other words, the potential curve along Yangtze River is not depicted. For this point, see Figs. 6 and 7.

<sup>&</sup>lt;sup>24</sup> Fuzhou's port history goes back to the northern Song Dynasty (960–1127). Shipping routes from Fuzhou to Japan and to Arabic countries were established. During the Yuan Dynasty Marco Polo passed through Fuzhou. During the Ming Dynasty, the famous Chinese navigator, Zhenghe, called at the port of Fuzhou. However, as recently as the Opium War, the traditional connection between seaports and hinterlands began to change. Openness and trade led to an increase in the division of labour and major changes in the organization of production.





Fig. 7. Potential curve for the emergence of Chongqing.

A final issue concerns the emergence of inland ports. From the mid 19th century to 1900, the population of Shanghai more than doubled to over 1 million, with an expanding international community. After 1900 Shanghai was characterised by further remarkable growth. In 1920 the total population was already estimated at 2 million, a census then showing 950,000 in the International Settlement and French Concession. Among other things, Shanghai's growth as a shipping hub was linked to the needs of the Yangtze River Valley. As Shanghai developed rapidly, the city's farm hinterland extended further and further along the Yangtze River. As a result, the emergence of new inland port cities along the Yangtze River Valley became a possibility.

To shed some light on this issue, the same modelling framework as above can be used to calibrate the emergence of such inland port cities. Given Shanghai's rapid population growth,  $N_S = 1.2$  is assumed. Fig. 6 shows the in-depth numerical exercise results for the branch along Yangtze River Valley, i.e. we limit our analysis to *bz* in Fig. 2.

As explained before, where the potential curve hits the threshold 1 a new city will eventually emerge in the evolutionary process, due to the bifurcation of the spatial system. One can readily see that the likely emergence of a navigable inland port city and hub occurs in the range 0.1 < r < 0.5. This coincides with the development of Wuhan as an industrial agglomeration inland port city. Lying where the Han and Yangtze Rivers meet, it was formed in 1949 from the consolidation of three cities: Wuchang, Hanyang, and Hankou. Located centrally between Beijing/Tianjin and Guangzhou and between Shanghai and Chongqing, it is sometimes called the "thoroughfare of nine provinces".<sup>25</sup> The port is accessible to oceangoing ships. The city of Wuhan is China's traditional manufacturing industry base, and one of the origins for China's modern industry.

Finally, the emergence of Chongqing is calibrated. Set in the middle reaches of the Yangtze, Chongqing has long been the economic hub of southwestern China. In 1891, Chongqing was opened to foreign trade and a customs house was established there. The open port marked the beginning of the history of steamboat navigation from Yichang through the treacherous gorges to Chongqing. Shipping and trade and the processing industries in Chongqing grew steadily as the city came to link southwestern

<sup>&</sup>lt;sup>25</sup> Prior to that, Hankou was already a designated treaty trading port for western powers after the 2nd Opium War.

China and the upper reaches of the Yangtze River with the rest of the world. In 1929, Chongqing was formally declared a city. Fig. 7 indicates that the likely emergence of another inland port in Yangtze River Valley occurs in the viable new port range 0.15 < r < 0.35 which comprehends Chongqing. This echoes our Wuhan-finding above.

All in all, one may say that while the calibration evidence shown in this section is merely suggestive, it is consistent with the predictions of the modelling framework and the above-mentioned historical facts.

#### 5. Conclusions

The miracle of China's growth based on exporting of manufactures is above all a maritime one and it would have been inconceivable without the ship-borne container. Urbanization also took place differently along the coast, with cities growing more rapidly than in the interior.<sup>26</sup> In this paper we have therefore tried to motivate a focus of attention on the genesis of urbanized seaports in China using the analytical modelling tools of economic geography. The paper investigates the extent to which the new economic geography model can answer the where-do-seaport-cities-form question in a particular historical episode. A fascinating feature of the underlying economic geography modelling framework is that seaport city growth is path-dependent, but the path does not seem to be entirely determined by sheer luck but is rather constrained by the geographic economic conditions, as mentioned above. The model makes no presumption on which location might become an urbanized seaport, but once a location gets a headstart via the initial emergence of a seaport, the process of cumulative causation begins to unfold. What were initially small GDP per capita differences across locations can evolve over time into large income differences. In other words, the interaction of agglomeration and spreading forces implies that history is decisive.<sup>27</sup>

Let us conclude our journey into the economics of Chinese seaport cities by pointing out what we have learned. In a nutshell, we have demonstrated that the spatial distribution of Chinese seaport cities in modern times can be explained in the context of a new economic geography model framework, even though the emergence of Chinese seaports took place under very special circumstances. Thus we have contributed to efforts to map the contour of China's development process. However, one has to acknowledge that the established economic geography modelling toolbox reflects a compromise in representing the real economy. Behrens and Robert-Nicoud (in press) are quite right pointing to at least two notable shortcomings. First, the fact the model can be calibrated to illustrate the real world does not prove that the effects emphasised were at work. Second, the calibration exercises in the economic geography literature fall short of the standards in the state of the art macroeconomic literature. For example, the models are never asked to compare the moments implied by the calibrated model with those measured in the data.<sup>28</sup> Therefore the economic geography toolbox is no more than an incomplete summation of the full range of issues related to the emergence and growth of Chinese seaport cities. Nevertheless, the modelling approach forms a useful point of departure for future work on model formulation, calibration, and interpretation.

#### Acknowledgements

We would like to thank an anonymous referee and the editor for helpful comments on an earlier draft. We would also like to thank Ruixue Jia (Stockholm) for her discussion of the historical background. Finally, we thank seminar participants at Academia Sinica (Taipei) for their comments. The usual disclaimer applies.

#### References

Aghion, P., & Howitt, P. (2009). The economics of growth. Cambridge: MIT Press.

Behrens, K., & F. Robert-Nicoud, Tempora Mutantur: In search of a new testament for NEG, Journal of Economic Geography (in press). doi:10.1093/jeg/lbq039. Bruinsma, F., Gorter, C., & Nijkamp, P. (2000). Multimodal infrastructure, transport networks and the international relocation of firms. *Transport Planning and Technology*, 23, 259–281.

Cao, S. (2001). History of Chinese population: Qing Dynasty (Zhongguo Renkou Shi: Qing, in Chinese). Shanghai: Fudan University Press.

Démurger, S. (2001). Infrastucture development and economic growth: an explanation for regional disparities in China? *Journal of Comparative Economics*, 29, 95–117.

Démurger, S., Sachs, J., Woo, T. W., Bao, S., Chang, G., & Mellinger, A. (2002). Geography, economic policy and regional development in China. Asian Economic Papers, 1, 146–197.

Fujita, M., & Krugman, P. (1995). When is the economy monocentric?: von Thünen and Chamberlin unified. *Regional Science and Urban Economics*, 25, 505–528. Fujita, M., & Krugman, P. (2004). The new economic geography: past, present and the future. *Papers in Regional Science*, 83, 139–164.

Fujita, M., Krugman, P., & Venables, A. J. (1999). The spatial economy: cities, regions, and international trade. Cambrige: MIT Press.

Fujita, M., & Mori, T. (1996). The role of ports in the making of major cities: self-agglomeration and hub-effect. Journal of Development Economics, 49, 93-120.

<sup>&</sup>lt;sup>26</sup> Not all ports are of equal stature and their success has been variable over time. As trade shifts, so does the fortune of ports. Changes in the seaport sector are the result both of new technologies and of structural shifts in the world's trading patterns. The increasing concentration of traffic in a few giant ports has been a part of this development. Furthermore, institutional, regulatory, and government policies can help promote sea transport services. For example, the decentralisation reforms adopted in China after 2001 provide strong incentives for local authorities to commit more emphasis upon seaport development and the necessary institutional framework.

<sup>&</sup>lt;sup>27</sup> The model of Fujita and Mori (1996) also entails the lock-in effect of agglomerating forces. In other words, seaport cities continue to prosper even after their initial geographic advantage has ceased to play an important role.

<sup>&</sup>lt;sup>28</sup> It is obvious that Chinese time series data for the mid 19th century aren't available. Unfortunately, it is therefore impossible to run diagnostic tests as in the current macro literature.

Kawakami, T., & Doi, M. (2004). Port capital formation and economic development in Japan: a vector autoregression approach. Papers in Regional Science, 83, 723 - 732

Keller, W., Li, B., & Shiue, C. H. (2010). China's foreign trade: perspectives from the past 150 years. *NBER Working Paper No. 16550, Cambridge (Mass.)*. Lin, J. Y. (2008). The Needham puzzle, the Weber question, and China's miracle: long-term performance since the Sung dynasty. *China Economic Journal*, 1, 63–95. Mitchener, K. J., & Yan, S. (2010). Globalization, trade and wages: what does history tell us about China? NBER Working Paper No. 15679, Cambridge (Mass.).

Mori, T., & Nishikimi, K. (2002). Economies of transport density and industrial agglomeration. Regional Science and Urban Economics, 32, 167–200.

Murphy, R. (1970). The treaty ports and China's modernization: what went wrong? Ann Arbor: University of Michigan.

Nunn, N., & Wantchekon, L. (2009). The slave trade and the origins of mistrust in Africa. NBER Working Paper No. 14783, Cambridge (Mass.).

Rawski, T. G. (1969). Chinese dominance of treaty port commerce and its implications, 1860–1875. Explorations in Economic History, 7, 451–473.

Shiue, C., & Keller, W. (2007). Markets in China and Europe on the eve of the Industrial Revolution. American Economic Review, 97, 1189-1216.

Williamson, J. G. (2010). When, where, and why? Early industrialization in the poor periphery 1870-1940. NBER Working Paper No. 16344, Cambridge (Mass.). Yan, S., (2007). Real wages and skill premia in China, 1858-1936, UCLA, Los Angeles, mimeo.

Zhongmin, Y., & Jianzhong, T. (1990). Areal expansion of urban Shanghai. GeoJournal, 21, 57-64.

## The Role of Inter-Provincial Transfers in the Convergence Process - Evidence for China

Hao Yu

Department of Economics Hamburg University Von-Melle-Park 5 20146 Hamburg GERMANY hao.yu@wiso.uni-hamburg.de

Hamburg, June 2012

#### Abstract

In this paper we adapt Funke and Strulik's (2000) two-region endogenous growth framework to analyse Western China's development and regional divergence and convergence processes. Using data reflecting the observed features of China's eastern and western regions, we also made calibrations of our model after 1992. The calibration results provide evidence of the speed of Western China's relative development and shed light on future regional convergence processes under the "Western Development" programme launched in 2000.

#### 1. Introduction

As China's economy as a whole has expanded at a remarkable average annual rate of 10.1% in the last thirty years, the regional disparity inside China has increased significantly. For over two decades since the reform and opening-up initiated in 1978, China has promoted economic development with particular priority given to its coastal regions, following Deng Xiaoping's famous development philosophy of "allowing some to get rich first." The unbalanced development policy has led to huge regional disparity. According to data from the National Bureau of Statistics (NBS), in 2009 per capita GDP of the eastern megacity Shanghai had reached \$11,000 (in U.S dollars, USD), a level comparable with moderately developed countries such as Hungary (\$12,980 USD) and Lithuania (\$11,410 USD). In the same year, the per capita GDP of Gansu, the poorest remote western province, was only \$1,881 USD, which is approximately one sixth of Shanghai's level.

However, creating large interregional disparity actually runs contrary to the original intention of Deng's economic reform. There is in fact a second half of his development philosophy policy, although not often mentioned, that "the regions which get rich first should carry along the backward regions and gradually achieve common prosperity". Unfortunately, he did not achieve this goal when he was in power. In 1999, two years after the death of Deng Xiaoping, then Premier Zhu Rongji proposed his own solution to this problem. His proposal, the Western Development programme, which featured high transfer payments to western regions to boost infrastructure development and foster capital accumulation, has proved to be successful in narrowing China's interregional relative income gap.<sup>1</sup> More details of this programme will be analysed in the next section.

Given the fact that China's market transition has been under the guidance of the Chinese government, the role of government in the development of China's regional disparity needs to be carefully investigated. First, China's government (at all levels) is keen on participating in economic affairs directly. For instance, China's local governments are usually enthusiastic about investing in large-scale projects which may boost local GDP growth rate quickly (Huang, 1999). Secondly, China's government tends to exercise its authorities to promote regional economic growth or achieve particular political goals by designing preferential policies for certain areas or certain interest groups. Essentially, the widened regional income gap before 2000 and the regional convergence afterwards are both closely related with Chinese government policies. We will elaborate on this point in the next section.

<sup>&</sup>lt;sup>1</sup> Due to the huge original gap, although the relative income disparity has been reduced (see Figure 1 in section 2), the absolute difference of per capita GDP in the two regions has nevertheless constantly increased from 4179 yuan in 2000 to 10532 yuan in 2009.

Barro (1990) was the pioneer researcher to investigate the effect of government spending with the framework of an endogenous growth model. Following his paradigm, many subsequent studies have treated transfer payment as one kind of government fiscal policy employed to maximize economic growth (Cashin, 1995; Gong and Zou, 2002). However, studies on the influence of transfer payment to regional convergence are still insufficient. Among these include the influential studies by Funke and Strulik (2000, 2005), in which the authors highlight the role of government in public capital accumulation and the effect of transfer payments on regional convergence progresses in unified Germany and an imaginary reunified Korea, respectively.

In the case of China, the existing empirical studies on the influence of transfer payment on regional disparity have led to rather controversial results, depending on the concrete data set used and the indicators designed to reflect disparity. For instance, using data sets whose sample periods are mainly before 2000, Ma and Yu (2003), Tsui (2003), Tochkov (2005) and Yin (2008) find either a negative or no statistically significant effect of inter-governmental transfer payment on regional disparity. In a recent study, however, Zhang and Qin (2011) utilize provincial panel data between 1994 and 2009 and verify a positive effect of inter-governmental transfer payment on long-term economic growth. They also report that the interregional transfers could narrow disparity across regions overall. These empirical studies reflect the actual development of China's regional disparity in terms of real GDP per capita (see Figure 2 in the next section): Before 2000 the income gap between Eastern and Western China gradually widened, but this trend was effectively reversed after Western Development was initiated in 2000.

In this paper, we adapted the analytic framework of Funke and Strulik (2000, 2005) to analyse the regional disparity problem in China before and after the implementation of Western Development in 2000. The effect of the transfer payment system on the development of underdeveloped Western China is carefully examined. The calibration results fit China's actual data well and therefore offer a prediction for future regional development in China: Under the current Western Development programme, the catch-up process of western provinces will likely continue but at a decreasing speed. In 2010, the relative income of Western China had reached approximately 65% of that of Eastern China, while it should take another 20 years before this ratio reaches 80%.

The remainder of the paper is arranged as follows: Section 2 provides some historical background and relative information about the development of China's regional disparity. Section 3 is devoted to modelling the development of this disparity since 1992. And then we present some model simulation results in section 4. Finally, some conclusions are provided in Section 5.

#### 2. Historical Background

As shown in Figure 1 below, Western China covers 6 provinces (Gansu, Guizhou, Qinghai, Shaanxi, Sichuan, and Yunnan), 5 autonomous regions (Guangxi, Inner Mongolia, Ningxia, Tibet, and Xinjiang), and 1 municipality (Chongqing).<sup>2</sup> As of the end of 2009, this region contains 71.4% of mainland China's area, but only 28.6% of the population and 19.9% of the total economic output. In this paper we refer to all the provinces that do not belong to the Western China as Eastern China.<sup>3</sup>





2.1 Regional Divergence before 2000

<sup>&</sup>lt;sup>2</sup> In this paper, we use the term "Western China" to refer to the western regions that benefit from China's Western Development policy adopted in 2000. In the context of geography, however, Inner Mongolia and Guangxi province are sometimes classified as part of Central China.

<sup>&</sup>lt;sup>3</sup> The Eastern China as defined in this paper could be further divided into coastal area and central area. The coastal area is most economically prosperous and has the highest per capita GDP. However, because both coastal and central areas lead the western region in terms of per capita income, and also because Western China is the area which benefits most from interregional transfer payment under the Western Development programme, we incorporate coastal and central areas as a whole bloc and refer it as "Eastern China" in this paper.

Throughout Chinese history, Western China has been the least developed area of China. Before the foundation of People's Republic of China in 1949, Western China was mainly home to minorities and had almost no industry.<sup>4</sup> During the Mao era (1949-1976), some efforts were paid to promote the economic and social development in western regions, including the abolition of serfdom in Tibet and other ethnic communities as well as the Third Front construction to boost western industrial development.<sup>5</sup> However, these efforts did little more than hold inequalities in check.



Figure 2. The relative per capita GDP of Western China to Eastern China, 1978-2010, %

Data Sources: Author's calculation from China Statistical Yearbook (various issues).

As shown in Figure 2 above, the regional productivity difference broadly increased in the 1980s and 1990s, despite two fluctuations in the 1980s. The widened gap was largely a consequence of both historical and geographic factors interacting with the liberalization of the economy and the decentralization of the fiscal system, which allowed the eastern regions, well favoured by these factors, to develop economically. The eastern regions were naturally favoured by their geographical location

<sup>&</sup>lt;sup>4</sup> In the beginning of second Sino-Japanese war (1937-1945), Chinese government managed to shift some industrial equipment and machinery from coastal regions that were vulnerable to Japanese attack to the remote western regions, which established the preliminary industrial foundation in Western China.

<sup>&</sup>lt;sup>5</sup> Since the mid-1960s, the central authorities launched a campaign of *Third Front*, which refers to a large-scale programme -- in response to the then volatile international situation (the deteriorating ties with the Soviet Union and the escalating Vietnam War) -- to build a range of industrial bases in its remote yet strategically secured south-western interior. Although the *Third Front* campaign has objectively boosted the industrial development in Western China, its effect on economic growth was rather limited, because most projects were military-industrial complexes and located in remote and inaccessible areas, not easily reachable by air attack. In economic terms, the *Third Front* projects were basically inefficient and usually relied on subsidies to operate. From the 1980s,

closer to the fast-growing economies along the Pacific Rim. Soon after the reform and opening-up was launched in 1978, eastern provinces were given preferential policies to attract foreign direct investment (FDI) and develop external trade. They took great advantage of their geographical position to develop export-oriented industries and gradually became the powerhouse of China's economy.<sup>6</sup> Since the late 1980s, these preferential policies were gradually extended to the western provinces and all interior western regions were eventually opened up in 1994. The time lag in benefitting from preferential policies has already caused great regional disparity and left western provinces in an unfavourable position in attracting investments and generating growth (Yao and Zhang, 2001; Demurger et al., 2002). Except for a few years in the 1980s, the growth rate of regional real GDP per capita and the growth rate of capital stock in Eastern China were constantly higher than those in Western China before 2000.

The widening East-West gap before 2000 was also partly attributed to fiscal policies designed to favour eastern provinces. Chinese fiscal reform history could be divided into two stages: first, the Fiscal Responsibility System (FRS) between 1980 and 1993 and second, the ensuing Tax Sharing System (TSS). Generally speaking, the main objective of FRS was to transfer some fiscal autonomy from central to provincial governments: Provinces were given much greater control over how much revenue they could retain and in which way they could allocate it. The key fiscal rules under the 1980s fiscal reform were that all China's provinces agreed to a fiscal contract specifying either a certain amount of fiscal revenue to remit to the centre (wealthy provinces) or a subsidy to be received from the centre to maintain fiscal balance (poor provinces). Despite these agreements, in practice the wealthier eastern provinces were always able to find ways to reduce their remittance to central government (such as channelling larger amounts of their revenue into extra-budgetary accounts that were beyond the control of the centre). Moreover, they even received higher earmarked grants in the form of price subsides from central government under the dual price system (1984-1992).<sup>7</sup> The implementation of FRS caused a trend of reduction in fiscal resources available to central government, which objectively weakened the government's ability to make transfers to western provinces.

with the post-Mao economic reform, a number of moribund factories were shut down while others have struggled to shift to non-military production and move close to urban areas.

<sup>&</sup>lt;sup>6</sup> The cities and provinces that benefitted from preferential policies are called special economic zones (SEZs). These preferential policies include special tax incentives (such as various tax deductions and exemptions) to attract FDI and greater independence on foreign trade activities. In the SEZs, the economic activities are primarily driven by market forces, and export-oriented industries are especially encouraged.

<sup>&</sup>lt;sup>7</sup> After a few years' experiments, the dual price system was officially established in 1984, in which state-owned industries were allowed to sell any production above the plan quota, and commodities were sold at both plan and market prices, allowing citizens to avoid the shortages of the Maoist era. Because the state-set price was usually lower than the market-set price, central government had to provide fiscal subsidies to maintain the lower official diktat price. The price subsidy increased rapidly in the 1980s and it amounted to nearly 60% of total earmarked grants in 1990. In 1992 the price reform was completed and the dual price system was replaced by the single price system under which commodity prices are essentially determined by the market mechanism.

Province	1980	1985	1990	1995	2000	2005	2010
East Total	-10.3	-3.5	-0.2	-0.5	-3.5	-3.9	-5.1
Beijing	-26.2	-7.6	-1.5	-2.7	-15.3	-21.2	-27.4
Tianjin	-25.4	-12.1	-1.5	-2.8	-10.6	-14.4	-19.7
Hebei	-3.0	-0.9	0.7	-0.1	-0.7	-1.2	-0.6
Shanxi	-1.2	4.8	0.7	-0.6	0.3	-3.8	-1.4
Liaoning	-18.8	-5.5	-0.7	-0.1	0.7	-0.1	0.2
Jilin	3.0	6.4	4.9	0.0	2.0	3.7	3.9
Heilongjiang	3.9	2.0	2.3	-0.4	-1.8	-0.5	5.3
Shanghai	-49.9	-29.6	-12.1	-4.8	-17.3	-18.3	-30.4
Jiangsu	-10.5	-5.9	-2.5	-0.4	-4.1	-5.5	-9.2
Zhejiang	-7.7	-4.9	-2.4	-0.1	-3.0	-2.4	-5.6
Anhui	-2.6	1.1	1.3	0.2	0.7	1.6	3.0
Fujian	-0.3	2.8	2.2	0.6	-1.8	-1.5	-2.2
Jiangxi	3.2	4.1	2.4	1.1	1.8	3.7	4.6
Shandong	-6.2	-2.4	1.0	-0.7	-2.4	-2.3	-2.9
He'nan	-2.2	0.1	0.6	0.0	-0.6	1.1	3.6
Hubei	-3.8	-1.7	0.8	0.6	0.0	0.1	1.7
Hu'nan	-3.2	0.3	1.3	-0.1	0.4	1.7	3.9
Guangdong	-4.5	-0.8	1.3	-0.3	-5.5	-5.2	-7.8
Hainan	4.7	6.3	9.8	3.0	1.8	3.1	2.3
West Total	6.2	6.5	4.6	0.6	2.3	3.7	6.8
Inner Mongolia	20.8	12.8	8.7	3.3	5.3	3.0	1.8
Guangxi	5.3	5.3	4.0	1.1	0.8	3.2	5.4
Chongqing	N.A.	-4.6	-1.6	6.3	1.8	2.1	2.7
Sichuan	N.A.	2.4	2.8	-2.7	0.9	2.8	8.8
Guizhou	9.9	7.6	4.8	1.5	4.4	6.5	12.9
Yun'nan	6.7	5.6	3.0	-6.8	-4.4	-1.9	2.8
Tibet	46.8	54.6	46.0	52.9	47.2	70.7	94.8
Shannxi	2.6	4.0	3.1	0.8	3.1	2.0	0.9
Gansu	-3.6	6.2	4.8	2.0	6.1	8.3	15.0
Qinghai	23.8	23.3	14.1	6.5	15.1	20.7	34.7
Ningxia	23.2	22.9	13.5	4.3	10.4	12.3	11.9

Table 1. The provincial net transfers received from central government (as a percentage of provincial GDP), 1980-2010

Notes: 1. Before 1994, because there was no institutionalized transfer payment under FRS, the provinces with fiscal surplus simply handed in their surplus as transfers to the central government, while the provinces running fiscal deficits received transfers from the centre to maintain a balanced

deficit. 2. After 1994, under TSS, each province received transfers from the centre in accounting terms. The net transfers from the centre were equal to the difference of nominal transfers from the centre and the tax revenues handed to the central government.

Data Sources: Author's calculation from *China Statistical Yearbook* and *China Financial Yearbook* (various issues).

The trend of decreasing financial support to western provinces during this period can be seen clearly in Table 1 above. The net transfers to the centre as a percentage of GDP decreased from 10.3% in 1980 to 0.2% in 1990 for eastern provinces, meanwhile the net transfers received by the western provinces (as the percentage of the provincial GDP) from the centre also decreased from 6.2% to 4.6%.

Because fiscal decentralization in the 1980s eroded the fiscal position of the central government, in 1994 a comprehensive fiscal reform put an end to the FRS. The brand-new "Tax Sharing System" (TSS) for the first time explicitly defined central, shared, and local taxes between the central government and the provinces. The tax collection was also split into central and provincial administrations, with the former collecting the central and shared taxes and the latter collecting the provincial taxes. The implementation of TSS successfully reversed the trend of declining central government revenue share, which stood at around 50% after TSS was carried out, up from approximately 24% in 1993. However, the transfer payment system in the first few years under TSS still favoured more developed eastern regions.

There are three kinds of transfer payments under TSS: tax rebate, general transfer payment and special transfer payment. The tax rebate had been the main form of central-province transfer payment before 2000, which counted for more than 90 percent of total transfer payment amount. Because the amount of tax rebate received by a province depended on its tax revenues, the more prosperous eastern provinces with a higher tax base could receive more tax rebate from the centre. As a result, the transfer payment system during the period 1994-1999 still had an eastern bias.

#### 2.2 Regional Convergence after 2000

Observing the continuously widening East-West gap since the reform and opening-up initiated in 1978, the Chinese Communist Party (CCP) leaders became increasingly worried about further expansion of regional disparity and its threat to political and social stability. As early as 1992, some tentative policies were formulated to try to address the regional problem.<sup>8</sup> However, these attempts were proven unsuccessful and the regional gap continued to grow in the 1990s.

<sup>&</sup>lt;sup>8</sup> In 1992, during his southern tour, Deng Xiaoping proposed a solution to regional divergence by "twinning" eastern cities and provinces with western ones, thereby promoting a transfer of skills and finance to narrow the

In 2000, a systematic, integrated programme of "Western Development" was drawn up to solve the problem of regional disparity. The main components of the programme included the development of infrastructure (transport, hydropower plants, energy, and telecommunications), enticement of foreign investment, increased efforts on ecological protection (such as reforestation), promotion of education, and retention of talent flowing to richer provinces. Among these, the infrastructure development was of primary importance. The inferior and insufficient infrastructure has long been one of the greatest obstacles for Western China to attract investment at home and abroad. The improvement of infrastructure would in the long run help western provinces to attract investment and talents, thereby boosting economic growth. The direct and indirect productivity promotion caused by rapid infrastructure development has been crucial to the catch-up process of western provinces after 2000. The following Table 2 summarizes the main infrastructure projects under the "Western Development" plan.

Table 2. The main infrastructure constructions under "Western Development" programme

Project	Construction Period	Total Investment	Beneficial Western Provinces
Three Gorges Project	1994-2009	180 billion yuan	Provinces along Yangtze River
			(Sichuan and Chongqing)
West-East Gas Pipeline	2002-2004	300 billion yuan	Xinjiang, Gansu and Ningxia
Returning Grazing Land	2003-	14.1 billion yuan	All western provinces
to Grassland		(2003-2006)	
Qinghai-Tibet Railway	2001-2006	33 billion yuan	Qinghai and Tibet

Sources: Western Development department of National Development and Reform Commission

An important feature of the transfer payment system under the Western Development plan is the significant increase in interregional transfers from east to west: both general transfer payments and specific transfer payments to western provinces dramatically increased after 2000. At the same time, the central government also increased the share of financial construction funds used to support investment projects in western provinces. In terms of purpose, the specific transfer payments and financial construction funds from the centre were intended to finance western government investment, and the general transfer payments were designed to improve the livelihood of western households. The ratios of these two types of transfers to eastern provincial government revenues are shown in Figure 3 below.

regional gap. The Ninth Five Year Plan (1996-2000) carried out Deng's proposal under the name of *hengxiang jingji lianxi* (horizontal economic cooperation), under which each eastern province had a specified western province to support and help. For example, Beijing was twinned with Inner Mongolia, Shanghai with Yunnan, and the special economic zones with Guizhou.

Figure 3. The ratio of interregional transfer payment used for western government investment to eastern government revenues (TRINV) and the ratio of interregional transfer payment used for improvement of western households' livelihood to eastern government revenues (TRLLH), 2000-2010





Data Source: Author's calculation from Financial Yearbook of China (various issues).

During the period 2000-2010, the total interregional transfer payment from east to west amounts to approximately 17% of Eastern China's government revenues. The high transfers supported the implementation of the Western Development programme and proved to be central to the regional convergence process.

#### 3. Western Development and Regional Convergence - Tools and Applications

The focus of this section is to understand the interplay of Chinese fiscal policy (especially transfer payment system) and regional disparity. The basic hypothesis is that government spending may play a key role in promoting regional productivity, which contributes to higher regional growth rate and leads to regional convergence. In terms of of methodology, we adapt Funke and Strulik's (2000) analytical framework which highlights the specific government spending that could foster regional productivity

Note: Total government revenue of local government consists of budgetary and extra-budgetary fiscal revenues and land granting income. Before 1994 the institutional transfer system was not established, therefore the data before 1994 are not shown in this figure.

growth.<sup>9</sup> In the following subsection we first investigate the catch-up process of Western China since the Western Development programme was carried out in 2000. In the second subsection we modify the basic analysis framework based on the features of Chinese transfer payment system during the period 1992-1999 to analyse the regional divergence process.

## **3.1** The catch-up process of Western China after the implementation of Western Development in 2000

#### **3.1.1 Basic Analytical Framework**

In this subsection we investigate the catch-up process of western regions since 2000. Following Funke and Strulik (2000), we assume that there is a huge number of identical firms in each region. Each firm operates under perfect competition and employs private capital  $k_i$  and labour  $l_i$  to produce an output  $y_i$ . For simplicity, the production function is based on the Cobb-Douglas form:

(1) 
$$y_i = A_i k_i^{\alpha} l_i^{1-\alpha}$$
,  $i = W, E$   $\alpha \in (0,1)$ 

where the subscript i represents the region where the firm is located.  $A_i$  is a productivity parameter which is determined in equation (12) below. Without loss of generality, the price of goods is normalized to unit. Labour input and capital investment are determined by firms. The private capital accumulates according to the following equation:

(2) 
$$k_i = i_i - \delta_K k_i ,$$

where  $\delta_{\kappa}$  is a positive capital depreciation ratio. To avoid double taxation, we assume that the government levies a corporate tax on cash flow with exogenous constant rate  $\tau$ . Then the optimal problem of the firm is to maximize the present value of their intertemporal net profit flow

(3) 
$$V_i(0) = \int_{0}^{+\infty} \exp[-\bar{r}_i(t)t] \{(1-\tau)[y_i - w_i l_i] - i_i\} dt$$
,

<sup>&</sup>lt;sup>9</sup> There is one important difference in terms of government expenditure between Funke and Strulik's (2000) original setting and the model setting in this paper. Funke and Strulik (2000) count government expenditure on infrastructure development as the sole factor that determines regional productivity. However, in China's context, not only infrastructure expenditure but also expenditure on technological progress and education-related public spending may help promote regional productivity.

where  $w_i$  denotes the wage rate. The average capital return  $\bar{r}_i(t)$  is defined as following:

(4) 
$$\overline{r_i}(t) = \frac{1}{t} \int_0^t r_i(s) ds$$

The factor prices (wage rate and capital return) are easily derived from the first-order conditions:

(5) 
$$w_i = (1 - \alpha) A_i \left(\frac{K_i}{L_i}\right)^{\alpha}$$
,

(6) 
$$r_i = (1 - \tau) \alpha A_i \left(\frac{K_i}{L_i}\right)^{1 - \alpha} - \delta_K$$

Eqs. (5) and (6) reveal a fact that all firms in the same region have an identical capital-labour ratio, which could be replaced by the aggregate capital-labour ratio,  $k_i/l_i = K_i/L_i$ . The capital letters represent the variables for the whole region.

Next we consider government behaviour. After the Western Development plan was carried out in 2000, the fractions of government fiscal expenditure on productive spending for the both regions were nearly the same and fluctuated around 50 per cent.<sup>10</sup> The main driving force of regional convergence stems from the transfer payment from Eastern China to Western China used for government productive spending. The fraction of transfer payment used to improve the livelihood of western citizens has a function similar to government subsidies to households as in the case of reunited Germany. Given these basic observations, the behaviour of government in two regions could be summarized in following equations:

(7) 
$$\dot{G}_E = q \tau Y_E - f(\theta) \tau Y_E - \delta_G G_E$$

(8) 
$$\dot{G}_W = q \tau Y_W - \delta_G G_W + f(\theta) \tau Y_E$$

(9) 
$$Z_E = (1 - q - x)\tau Y_E$$

(10) 
$$Z_W = (1-q)\tau Y_W + x\tau Y_E$$

Eqs. (7) and (8) describe the investment behaviour of regional governments, in which q represents the ratio of government fiscal expenditure that is used for productivity spending;  $\delta_G$  is the depreciation

<sup>&</sup>lt;sup>10</sup> In this paper, government productive spending is defined as the government expenditure that helps to promote the growth of productivity. In the case of China's economy, concretely, government productive spending includes expenditure on infrastructure development, expenditure for science and technology promotion, expenditure for education, etc.

rate of productive capital accumulated by local government.  $f(\theta)$  is the ratio of eastern provinces' government revenue delivered to western provinces to finance government productive spending. Eqs. (9) and (10) interpret the transfer payment system.  $Z_i$  represents the transfer payment within region i, and x represents interregional transfer payment used to improve western households' livelihood. Adding up Eqs. (7) – (10) yields the balanced budget run by the government:

(11) 
$$\tau (Y_E + Y_W) = \dot{G}_E + \delta_G G_E + \dot{G}_W + \delta_G G_W + Z_E + Z_W$$

Following Funke and Strulik (2000), we assume that the productive parameter  $A_i$  depends on per person productive capital stock accumulated by the government of region i

(12) 
$$A_i = A(0) \left(\frac{G_i}{L_i}\right)^{1-\alpha},$$

where A(0) is the initial productivity parameter when government's productive capital stock is zero. Substituting Eq. (12) into the production function (1) yields

(13) 
$$y_i = A(0)k_i^{\alpha}g_i^{1-\alpha}$$

where  $g_i=G_i/L_i$  denotes the government productive capital stock per person. Eq. (13) essentially replicates Barro's (1990) original assumption that production exhibits constant returns to scale in k and g together but diminishing returns in each input separately. In other words, the economy follows a general form of "AK Model" growth pattern.

Finally, in order to close the model, we turn to household behaviour. Following Funke and Strulik's (2000) setting, the household's utility function is completely determined by household consumption

(14) 
$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}$$

where  $\sigma >1$  is the reciprocal of the intertemporal elasticity of substitution; *c* represent the household's private consumption. Each household in region i supplies one unit labour and earned wage w<sub>i</sub>. The optimization problem for a representative household in area i is to maximize the present value of its lifetime utility by choosing consumption *c* at each time spot
(15) 
$$U_i^j = \int_0^{+\infty} u(c_i^j) e^{-\rho t} dt$$
,

Where  $\rho$  is the time preference rate and the utility function u is defined in Eq. (14). At each time spot t, each household faces the budget constraint

(16) 
$$\dot{a}_i^j = r_i a_i^j + (1 - \tau) w_i - c_i^j + z_i^j$$
.

where  $a_i^j$  denotes the assets of the representative household of region i, and  $z_i^j$  represents the transfer payment he receives. The growth rate of private consumption determined by households follows the familiar Ramsey Rule

(17) 
$$\frac{\dot{c}_i}{c_i} = \frac{r_i - \rho}{\sigma}$$

which applies to all households in both regions.

#### **3.1.2 Economic Convergence**

Armed with the basic analytic framework developed above, we can now turn to the analysis of economic convergence. Under neoclassic assumptions of efficient production factor market and free flow of production factors, once factor prices differ across regions, the equilibrium would be rebuilt instantly through instantaneous sufficient factor flow from the region where the factor price is lower to the region where it is more expensive. Unfortunately, despite the rapid urbanization in the last two decades, and in part due to the strict household registration (*hukou*) system, most labour flow took place within a region instead of across regions (Xu, 2002; Cai et al, 2002; Yang, 2002). Comparatively, the mobility of capital is much higher in China than labor force. As a result, we make the assumption that cross-regional labour flow does not exist but cross-regional capital flow is sufficient, which implies that the capital returns in the two regions are always identical while the wage rates may differ across regions.

Because capital returns are always identical in the two regions, from Eqs. (6) and (12) we can derive the following condition

(18) 
$$\frac{G_W}{K_W} = \frac{G_E}{K_E}$$

Assuming the West-East labour ratio is  $\lambda$ , the West-East relative GDP per capita  $\theta$  could be calculated based on the production function (13) and condition (18):

(19) 
$$\theta = \frac{y_W}{y_E} = \frac{k_W}{k_E} = \frac{K_W}{K_E} \cdot \frac{1}{\lambda} = \frac{G_W}{G_E} \cdot \frac{1}{\lambda}$$

As western government receives transfer payments to invest,  $G_W$  grows faster than  $G_E$ , which implies that the productivity gap between the two regions narrows according to Eq. (12). This in turn leads to lower income inequality. Concretely, the growth rate of  $\theta$  is derived as

(20) 
$$\gamma_{\theta} = \frac{\dot{G}_{W}}{G_{W}} - \frac{\dot{G}_{E}}{G_{E}} = \left(1 + \frac{1}{\theta\lambda}\right) A(0) f(\theta) \tau \left(1 + \theta\lambda\right)^{-\alpha} g_{E}^{-\alpha}$$

where  $g_E = G_E/K$  is defined as the ratio of productive capital accumulated by eastern government to total private capital stock. It is clear that the ratio of transfer payment that is spent on western government investment,  $f(\theta)$ , is crucial to the convergence process: Only when  $f(\theta)$  is positive does regional convergence exist. For the convergence progress, we further assume that  $0 < f(\theta) < 1$  and  $f'(\theta) < 0$ , which implies that the transfer payment diminishes as regional disparity decreases.

From the national perspective, any household income net of tax and transfers that is not spent on consumption is accumulated as private capital stock; therefore the nationwide private capital stock K accumulates according to

(21) 
$$\dot{K} = (1 - q\tau)Y - C - \delta_K K = (1 - q\tau)A(0)\left[\left(\frac{G_E}{K_E}\right)^{1 - \alpha}K_E + \left(\frac{G_W}{K_W}\right)^{1 - \alpha}K_W\right] - C - \delta_K K$$

Therefore the growth rate of private capital stock is

(22) 
$$\gamma_{K} = \frac{\dot{K}}{K} = A(0)(1-q\tau)(1+\theta\lambda)^{1-\alpha}g_{E}^{1-\alpha} - \chi - \delta_{K}$$

where  $\chi \equiv \frac{C}{K}$  is nationwide consumption-capital ratio. Combining Eqs. (7), (17) and (22), the growth rate of  $\chi$  and  $g_E$  are given by

(23) 
$$\gamma_{\chi} = \frac{\dot{C}}{C} - \frac{\dot{K}}{K} = \frac{1}{\sigma} \Big[ A(0) \alpha (1 - \tau) (1 + \theta \lambda)^{1 - \alpha} g_E^{1 - \alpha} - (\delta_K + \rho) \Big] - \gamma_K$$

(24) 
$$\gamma_{g_E} = \frac{\dot{G}_E}{G_E} - \frac{\dot{K}}{K} = \tau A(0)[q - f(\theta)](1 + \theta\lambda)^{-\alpha} g_E^{-\alpha} - \delta_G - \gamma_K$$

With Eqs. (20), (22), (23) and (24), the dynamic of the economy during the regional convergence process is fully determined by a three-dimensional differential equation system in  $\theta$ ,  $\chi$  and  $g_E$ .

#### 3.1.3 Equilibrium Analysis When Regional Convergence Is Achieved

In equilibrium the gap between East and West has been completely bridged, i.e.  $\theta^* = 1$  (the asterisk denotes the steady-state level, hereinafter). At steady state the growth rates of the three key variables are all zero, i.e.  $\gamma^*_{\theta} = \gamma^*_{\chi} = \gamma^*_{g_E} = 0$ . Equalizing Eqs. (23) and (24) yields

(25) 
$$F(g_E) = \frac{1}{\sigma} \Big[ A(0)\alpha (1-\tau)(1+\lambda)^{1-\alpha} g_E^{1-\alpha} - (\delta_K + \rho) \Big] - q\tau A(0)(1+\lambda)^{-\alpha} g_E^{-\alpha} + \delta_G = 0$$

where  $F(g_E)$  is an implicit function to determine  $g_E$  at steady state. Because

$$F'(g_E) = \frac{\alpha(1-\alpha)}{\sigma} A(0)(1-\tau)(1+\lambda)^{1-\alpha} g_E^{-\alpha} + \alpha q \tau A(0)(1+\lambda)^{-\alpha} g_E^{-\alpha-1} > 0, \ F(g_E) \text{ is strictly}$$

monotonic increasing. Since  $\lim_{g_E \to 0} F(g_E) = -\infty$  and  $\lim_{g_E \to +\infty} F(g_E) = +\infty$ , there must exist a unique equilibrium value  $g_E^*$  satisfying Eq. (25). The nationwide consumption-capital ratio  $\chi$  could be determined by inserting Eq. (22) into Eq. (23)

(26) 
$$\chi^* = A(0) \left[ 1 - q\tau - \frac{\alpha(1-\tau)}{\sigma} \right] (1+\lambda)^{1-\alpha} \left(g_E^*\right)^{1-\alpha} + \frac{\delta_K + \rho}{\sigma} - \delta_K$$

Generally, the equilibrium values  $g_E^*$  and  $\chi^*$  could only be solved numerically. If we consider a special parameterization when  $\frac{\delta_K + \rho}{\sigma} = \delta_G$ , the steady state level of  $g_E$  could be calculated analytically from (25) as

(27) 
$$g_E^* = \frac{\sigma q}{\alpha (1+\lambda)} \cdot \frac{\tau}{1-\tau}$$

Substituting Eq. (27) into Eq. (23) yields steady state growth rate

(28) 
$$\gamma_C^* = \frac{1}{\sigma} \left[ A(0)\alpha(1-\tau)(1+\lambda)^{1-\alpha} \left(g_E^*\right)^{1-\alpha} - \left(\delta_K + \rho\right) \right] = A(0) \left(\frac{\alpha}{\sigma}\right)^{\alpha} q^{1-\alpha} \tau^{1-\alpha} (1-\tau)^{\alpha} - \delta_G$$

Under this special parameterization the maximum steady state growth rate is achieved when  $\tau = 1 - \alpha$ . This result practically replicates Barro's (1990) basic finding that the optimal tax rate maximizing the long-term growth rate equals the production elasticity of government spending, which is  $1-\alpha$  in our model.

With the equilibrium values  $\theta^* = 1$ ,  $g_E^*$  and  $\chi^*$ , the steady state of the economy is uniquely determined. Next we will prove that this equilibrium is a stable saddle point. First we calculate the Jacobian matrix

(29) 
$$J = \begin{bmatrix} J_{11} & 0 & 0 \\ \bullet & 1 & J_{23} \\ \bullet & 1 & J_{33} \end{bmatrix}$$

where the dots denote the elements that are irrelevant for the analysis, and

(30) 
$$J_{11} = \frac{\partial \gamma_{\theta}}{\partial \theta} \bigg|_{\theta^* = 1, \chi^*, g_E^*} = \left(1 + \frac{1}{\lambda}\right) A(0) q \, \mathcal{T}'(1) (1 + \lambda)^{-\alpha} \left(g_E^*\right)^{-\alpha} < 0$$

$$(31) \qquad J_{23} = \frac{\partial \gamma_{\chi}}{\partial g_E}\Big|_{\theta^* = 1, \chi^*, g_E^*} = \frac{\alpha(1-\alpha)}{\sigma} (1-\tau) A(0) (1+\lambda)^{1-\alpha} (g_E^*)^{-\alpha} - \frac{\partial \gamma_K}{\partial g_E}\Big|_{\theta^* = 1, \chi^*, g_E^*}$$

$$(32) \qquad J_{33} = \frac{\partial \gamma_{g_E}}{\partial g_E}\Big|_{\theta^* = \mathbf{1}, \chi^*, g_E^*} = (-\alpha)q \, \tau A(0)(1+\lambda)^{-\alpha} \Big(g_E^*\Big)^{-\alpha-1} - \frac{\partial \gamma_K}{\partial g_E}\Big|_{\theta^* = \mathbf{1}, \chi^*, g_E^*}$$

The Jacobian determinant at the steady state is given by

(33) 
$$|J| = J_{11} (J_{33} - J_{23}) = J_{11} \cdot A(0) (1 + \lambda)^{-\alpha} (g_E^*)^{-\alpha} \left[ (-\alpha) q \tau g_E^* - \frac{\alpha (1 - \alpha)}{\sigma} (1 + \lambda) (1 - \tau) \right] > 0$$

Therefore the steady state equilibrium is a saddle point. The eigenvalues of the Jacobian matrix (29) are determined by the determinant

(34) 
$$|J - \xi I| = 0 \implies (J_{11} - \xi)[(1 - \xi)(J_{33} - \xi) - J_{23}] = 0$$

Therefore the three eigenvalues are

(35) 
$$\xi_1 = J_{11} < 0$$
 and  $\xi_{2,3} = \frac{1}{2} \left[ (1 + J_{33}) \pm \sqrt{(1 + J_{33})^2 - 4(J_{33} - J_{23})} \right]$ 

Because  $J_{33} - J_{23} < 0$ , the last two eigenvalues  $\xi_{2,3}$  are both real but carry different signs. As a result, the dynamic system has two negative real eigenvalues and one positive eigenvalue at steady state, implying that the equilibrium is monotonous on a two-dimensional stable manifold.

#### 3.1.4 Regional Income Redistribution

The problem of regional productivity convergence has been solved independently from interregional income redistribution issue. Despite the identical consumption growth rates in both regions according to Ramsey rule (17), eastern households enjoy higher absolute consumption level than their western counterparts. The reason is twofold: First, a representative western household has less initial wealth than its eastern counterpart, i.e.  $a_w(0) < a_E(0)$  when the convergence progress initiated in 2000; Second, the western wage rate remains persistently below the eastern wage rate due to its lower productivity.

The transfer payment delivered to western households could reduce or even eliminate the income disparity which emerged during the catch-up process. If the relatively worse-off western households' wage income were completely compensated by the interregional transfer payment, there would be no economic incentive for western labour to immigrate to eastern regions, which would greatly reduce the

economic and social pressure of the relatively prosperous eastern provinces. If we use  $\phi$  to denote the West-East relative wage income net of taxes and transfers, then  $\phi$  could be derived as

(35) 
$$\phi = \frac{z_W + w_W}{z_E + w_E} = \frac{[1 - \alpha(1 - \tau) - q\tau]\theta + (1/\lambda)x\tau}{1 - \alpha(1 - \tau) - q\tau - x\tau}$$

The necessary amount of transfer payment as western households' income subsidy is determined by the level of  $\phi$  endogenously from Eq. (35)

(36) 
$$x = \frac{1}{\tau} \cdot \frac{\lambda}{1 + \phi \lambda} \left[ 1 - \alpha (1 - \tau) - q \tau \right] (\phi - \theta)$$

Because the transfer payment to households within Eastern China could not be negative, the upper limit of ratio x is 1-q (refer to Eq. (9)). Additionally, because transfer payment should be non-negative, the reasonable range of x is between 0 and 1-q. As a result, we define x as

$$(37) \qquad x = \begin{cases} 0, & \text{if } \frac{1}{\tau} \cdot \frac{\lambda}{1+\phi\lambda} [1-\alpha(1-\tau)-q\,\tau](\phi-\theta) \le 0; \\ \frac{1}{\tau} \cdot \frac{\lambda}{1+\phi\lambda} [1-\alpha(1-\tau)-q\,\tau](\phi-\theta), & \text{if } 0 < \frac{1}{\tau} \cdot \frac{\lambda}{1+\phi\lambda} [1-\alpha(1-\tau)-q\,\tau](\phi-\theta) < 1-q; \\ 1-q, & \text{if } \frac{1}{\tau} \cdot \frac{\lambda}{1+\phi\lambda} [1-\alpha(1-\tau)-q\,\tau](\phi-\theta) \ge 1-q \end{cases}$$

#### 3.2 Regional divergence before 2000

China's reform and opening-up history could be roughly divided into two phases. The period between 1978 and 1992 was the initial stage for the development of China's market economic system, during which there were still apparent ideological conflicts within CCP about the direction of China's economic reform. The preferential policies to attract investment and foreign trade were to a great extent tentative and only tested in the special economic zones and eastern coastal regions, while the traditional planned economic system continued to operate in Western China. The difference in economic policies and institutions has caused a higher capital return in Eastern China, causing net capital to flow from West to East (Bai et al. 2006; Zhao and Lv, 2007). In this case, the key condition (5) of equal capital return in the two regions does not hold; our basic analytic framework is therefore not suitable to explain the widening interregional gap across regions before 1992.

In 1992, after Deng Xiaoping's tour in southern China calling for acceleration of reform and opening up, the market-oriented economic system was officially set as the ultimate goal of economic reform during the 14th Chinese Communist Party National Congress held that autumn. Since then, the preferential policies which were previously only available to the eastern regions were gradually applied to the interior of Western China, and the institutional obstacles to capital flow across regions have been gradually removed. Given the fact that capital mobility was much higher after 1992, the key assumption (5) is basically satisfied since 1992. As a result, the interregional divergence between 1992 and 2000 could be examined by our basic analytic framework. According to CBS statistic data, during this period the share of government expenditure used in productive government spending was higher in Eastern China than in Western China, and the interregional transfer payment was considerably small.

Given these fiscal policy features between 1992 and 2000, the government behaviour of Eastern and Western China could be depicted in following equations

$$(38) \qquad \dot{G}_E = q_E \tau Y_E - \delta_G G_E$$

$$(39) \qquad \dot{G}_W = q_W \tau Y_W - \delta_G G_W$$

÷.

where  $q_E$  and  $q_W$  represent the ratio of government revenue used in the productive capital accumulation by local government. Here we have  $q_E > q_W$ . Because interregional transfer payment was rather small before 2000, the accumulation of government productive capital stock was almost financed entirely by government revenues. The heavier spending of eastern provincial governments on productive capital led to faster productivity growth in Eastern China thereby widening the regional gap.

The dynamics of other variables are the same as regional convergence occurs after 2000. Therefore the whole economic dynamic system during 1992-2000 is given by the following equations.

(40) 
$$\dot{K} = (1 - q_E \tau)\tau Y_E + (1 - q_W \tau)\tau Y_W - C - \delta_K K$$

(41) 
$$\gamma_{K} = \frac{K}{K} = A(0)(1+\theta\lambda)^{-\alpha} g_{E}^{1-\alpha} [1-q_{E\tau}+\theta\lambda(1-q_{W}\tau)] - \chi - \delta_{K}$$

(42) 
$$\gamma_{\theta} = \frac{\dot{G}_{W}}{G_{W}} - \frac{\dot{G}_{E}}{G_{E}} = (q_{W} - q_{E})A(0)\tau(1 + \theta\lambda)^{-\alpha}g_{E}^{-\alpha}$$

(43) 
$$\gamma_{\chi} = \frac{\dot{C}}{C} - \frac{\dot{K}}{K} = \frac{1}{\sigma} \Big[ A(0) \alpha (1 - \tau) (1 + \theta \lambda)^{1 - \alpha} g_E^{1 - \alpha} - (\delta_K + \rho) \Big] - \gamma_K$$

(44) 
$$\gamma_{g_E} = \frac{\dot{G}_E}{G_E} - \frac{\dot{K}}{K} = A(0)q_E\tau(1+\theta\lambda)^{-\alpha}g_E^{-\alpha} - \delta_G - \gamma_K$$

The mechanism of regional disparity during this period is straightforward: According to Eq. (42), because  $q_E > q_W$ ,  $\gamma_{\theta} < 0$ , which means that the relatively worse-off of western households would deteriorate over time. Without intervening by central government, the regional disparity would constantly increase and therefore no equilibrium would be reached.

#### 4. Model Calibration

#### 4.1. Parameterization Setting

In section 3 we have developed and discussed the main features of the model and paved the way for the numerical calibrations exercise. In this section we will calibrate the evolvement of West/East relative income since 1992 - i.e., the regional divergence period before 2000 and the catch-up process of Western China afterwards under the Western Development programme.

From the Western and Eastern China's population sizes we set population ratio  $\lambda = 0.4$ . As Funke and Strulik (2000) point out, tax rate  $\tau$  could also be calculated as the government share in GDP. According to the estimates of Chinese researcher Li (2007) and various issues of *China's Fiscal Policy Report* released by Chinese Academy of Social Sciences, the ratio of Chinese government total revenue to GDP is around 30% between 1992 and 2009.<sup>11</sup> As a result, we set  $\tau = 0.3$ .

Parameter  $\alpha$  plays two roles in our model simultaneously: On the one hand, it is the output elasticity of private capital stock in production function (1); on the other hand, it also determines the production elasticity of productive capital stock accumulated by regional government through Eq. (12). Therefore the appropriate value of  $\alpha$  should be consistent with the empirical estimates in both aspects. Given the importance of capital accumulation to China's economic growth, the estimated capital output elasticity is considerably high. Chow (1993) and Chow and Lin (2002) have used the Cobb-Douglas production function to estimate China's capital output elasticity between 0.6 and 0.75, a range higher than that in most developed countries. Sturm et al. (1998) summaries that most researchers estimate an output elasticity of government public spending (equal to  $1-\alpha$ ) between 0.1 and 0.3, implying the value range of  $\alpha$  should be between 0.7 and 0.9. As a result, the appropriate range of  $\alpha$  should be between 0.7 and 0.75. We choose  $\alpha = 0.75$  in the basic scenario.

The estimation of China's capital depreciation rate  $\delta_{\kappa}$  remains a controversial task. We do not intend to discuss all existing estimates extensively here. Instead, we only focus on the relatively new estimates and determine an appropriate value of  $\delta_{\kappa}$  accordingly. Using the perpetual inventory method, recent studies (He et al., 2007; Shan, 2009) have tended to calculate a relatively high capital depreciation rate compared with traditional estimation results.<sup>12</sup> According to Shan's (2009) summary, the capital depreciation estimated through perpetual inventory method is between 0.07 and 0.11. Therefore we adopt  $\delta_{\kappa} = 0.09$  in the calibration. According to Wang and Wu's (2003) estimation, the depreciation of Chinese infrastructure and municipal construction is 3.6 (see Table 2 of their paper), therefore we take  $\delta_{G} = 0.04$ . Because the share of government investment in total government revenue of both regions fluctuated around 50% after the Western Development project launched in 2000, we set q=0.5. Because the share for Eastern China was about 60% and that for Western China was nearly 30%,  $q_{E}=0.6$  and  $q_{W}=0.3$ . Following numerous other calibration studies we adopt  $\rho = 0.02$  and  $\sigma = 2.3$ .

The remaining parameter, A(0), determines the growth rate and capital-output ratio at steady state. Despite continuous rapid growth during last three decades, China's economy is generally believed to be still far from its steady state given its relatively low per capita GDP level. Taking industrial countries' development experience as a reference, China's steady state economic growth rate should be between 2% to 4%. Considering the current robust growth of the Chinese economy, we take the upper limit of the range (i.e., 0.04) as China's steady state growth rate in the basic scenario, which corresponds to a steady state capital-output ratio of 2.58. Such settings imply  $\alpha = 0.75$  and A(0)=0.42 in the basic scenario. So far, we can summarize the basic parameterization setting in Table 3 below

Table 3 Model Parameterizations of Basic Scenario

α	A(0)	τ	q	$q_{\rm E}$	$q_{W}$	ρ	σ	$\delta_{K}$	$\delta_{G}$	λ	$(\gamma_Y)^*$	(K/Y)*
0.75	0.42	0.3	0.5	0.6	0.3	0.02	2.3	0.09	0.04	0.40	4%	2.58

<sup>11</sup> In China, budgetary revenue is only one part of total government revenue because of high off-budget revenues, which include extra-budgetary revenue, social security contributions, land granting revenue and the profits of state-owned enterprises.
<sup>12</sup> Earlier estimates of China's capital depreciation rate were relatively lower. Chow (1993) calculates capital

<sup>&</sup>lt;sup>12</sup> Earlier estimates of China's capital depreciation rate were relatively lower. Chow (1993) calculates capital depreciation according to national income accounting system (depreciation=GDP-national income+subsidies-indirect taxes) and reports an estimate of 4%. Another influential early estimation method was to utilize the capital accumulation data under MPS system of planned economy. Perkins (1988) applies this approach and estimates China's capital depreciation rate to be 5%,

We choose the year 2000 as the base year when Western Development was officially implemented. As shown in Figure 2, the per capita GDP in Western China was about 55 % of the corresponding Eastern level in 2000, i.e. the catch-up process of western regions starts at  $\theta$ =0.55. According to various issues of *Financial Yearbooks of China*, the initial ratio of transfer payment to western government to total eastern financial revenue,  $f(\theta)$ , equals to 0.12.<sup>13</sup> Furthermore, we specify the transfer payment policy rule as

(45) 
$$f(\theta) = a(1 - \theta^b)$$

where parameter *a* controls the absolute weight given to the interregional fiscal transfer payment, while b>0 is a sensitive parameter to control for the policy reaction on the relative income level. A larger b means the transfer payment policy is more sensitive when the relative income is high enough because  $f(\theta)$  reduces more rapidly to zero as  $\theta$  approaches to 1. We set parameters *a* and *b* to make the ratio  $f(\theta)$  close enough to the actual data between 2000 and 2010 as shown in Figure 3. Therefore we adopt b=3/5 and a=0.395.

Following Funke and Strulik (2000, 2005), we utilize the method of backward integration introduced by Brunner and Strulik (2002) to simulate the evolvement of the economic dynamic to the steady state. Compared with other numerical calibration approaches, the backward integration method is easy to calculate and has satisfactory high accuracy. Because the whole dynamic process contains two stages, with the year 2000 being the turning point from regional divergence to regional convergence, we apply the backward integration method twice. First we simulate the adjustment path of the catch-up process starting in 2000, then we utilize the simulated values of key variables  $\theta$ ,  $\chi$  and  $g_E$  of year 2000 as ending values of the regional divergence process to simulate the dynamic of the pre-2000 period. According to Eq. (19) and production function (1) the growth rates of several interesting variables during the whole dynamic process are derived as follows

(46) 
$$\gamma_{K_E} = \gamma_K - \gamma_\theta \frac{\theta \lambda}{1 + \theta \lambda}$$

$$(47) \qquad \gamma_{K_W} = \gamma_{K_E} + \gamma_{\theta}$$

<sup>&</sup>lt;sup>13</sup> Because of the fluctuation of transfer payments shown in Figure 4, we use a three-year average to calculate the initial value of ratio  $f(\theta)$ . In the first three years of Western Development (2000-2002), the total transfer payment used to finance the government investment of western provinces amounted to 372 billion yuan (including special transfer payment of 272 billion yuan and central financial construction funds of 100 billion yuan), while the total government revenue of eastern provinces during the same period was 3220 billion yuan. As a result, the initial ratio of  $f(\theta)$ as of the start of regional convergence was 372/3220=0.12.

(48) 
$$\gamma_{Y_E} = (1 - \alpha) \left( \gamma_{g_E} + \frac{\theta \lambda}{1 + \theta \lambda} \gamma_{\theta} \right) + \gamma_{K_E}$$

(49) 
$$\gamma_{Y_W} = (1 - \alpha) \left( \gamma_{g_E} + \frac{\theta \lambda}{1 + \theta \lambda} \gamma_{\theta} \right) + \gamma_{K_W}$$

#### 4.2. Results





Notes: 1. In all panels, the solid lines represent the stimulated values while the dashed lines represent the actual outcomes. In panel E and F, the green (grey) lines represent the variables of Western China while the red (dark) lines represent the variables of the Eastern China. 2. In panel C, the solid and dotted black lines correspond to the situations according to  $\Phi$ =0.95 and  $\Phi$ =0.9, respectively. 3. Because the interregional transfer payment was very low before the Western Development plan was implemented in 2000, the curves in panel B and panel C begin from the year 2000. 4. The graph is generated by modifying corresponding MATLAB codes for Funke and Strulik (2000) and Brunner and Strulik (2002).

The main calibration results are shown in Figure 4 above. Panel A demonstrates the transformation of the relative backwardness of the western area in terms of per capita GDP. In the basic scenario, the simulated ratio  $\theta$  fits the evolution of the actual outcome well before 2010. After 2000, the western

area catches up with the eastern area at a decreasing rate: In 2010, ten years after the launch of the Western Development programme, the ratio of per capita GDP of Western China to Eastern China has already been approximately 65%, but it will take another 20 years before this ratio reaches 80%.

Panels B and C depict the interregional transfer policy. Between 1992 and 2010, the simulated ratio  $f(\theta)$  in panel B is broadly in line with the actual data, while the actual data for ratio x fluctuates between corresponding transfer payment requirement for  $\Phi$ =0.95 and  $\Phi$ =0.9. Note that the income redistribution problem is solved independently from the interregional productivity convergence issue. The simulated results for ratio x imply that the redistribution policy of the Chinese government is not to fully compensate western households to achieve absolute income equalization across regions. Because of tight control in labour's interregional flow, such income inequality is not likely to trigger mass interregional migration.

Panel D describes the development of capital return rate. In the basic scenario, the simulated return to capital decreases steadily from 34% in 1992 to about 28% in 2030 and then will further decrease gradually to its steady state level 11%. This simulated capital return rate, although considerably higher than most industrial countries, is still broadly consistent with corresponding empirical studies such as those of Bai et al. (2006) and Sun et al. (2010), which estimate China's capital return between 20% and 40% since 1990. Note that China's interest rate has been under tight governmental control and therefore does not reflect the actual capital return rate in China.

As shown in panels E and F of Figure 4, there is a jump for capital and GDP growth rates at 2000, which is caused by applying backward integration approach in the calibration. The apparent change in transfer payment policies in 2000 has caused the discontinuity of the growth rates. Before 2000, the capital accumulation and per capita GDP growth rate in Eastern China were both higher than those in Western China, while after 2000 the opposite is the case. Under the Western Development programme, the Western China grew faster in terms of capital accumulation and GDP per capita than Eastern China.

#### 4.3. Further Discussion: The Factors Promoting Regional Convergence beyond the Model

So far with our basic model we have explained the evolvement of relative development of Western China since 1992. The productivity difference caused by different levels of government public spending in the west and east could basically explain the regional divergence and convergence processes before and after the implementation of the Western Development programme in 2000. However, taking the features of Western China and China's economic development after 2000 into consideration, some other factors affecting convergence progresses beyond mere capital accumulation (both by private firms and government) should be investigated in more detail.

The first potential factor is the abundant natural resources in Western China. Western China has been the most important energy production base for a long time. According to statistic data from NBS, as of 2009, Western China contained over 80% of China's hydropower resources, more than 85% of China's natural gas, nearly 40% of national coal reserves and also supplies approximately 30% of crude oil production. Since 2000, international energy prices have increased sharply. The price of oil has grown from less than 25 dollars per barrel in 2000 to more than 100 dollars per barrel in 2010; and the price of natural gas in the international market had also doubled in the same period. Climbing energy prices and abundance in energy reserves provide a good opportunity for western provinces to increase income and fiscal revenues. However, because the energy price in China has not yet integrated with the world market and also because the energy tax system has not yet been established, the western provinces have so far not fully benefitted from their resource advantages.<sup>14</sup> Fortunately, China's government has recently recognized these problems and has begun to handle with them by advancing energy price reform and pilot implementation of a resource tax reform.<sup>15</sup> It is reasonable to expect that the abundant natural resources might play a more important role in the future development of Western China.

Another potential factor is the migrant workers from Western China. An important hypothesis we made in the model setting is that the west-east labor flow is tightly controlled and therefore negligible. According to Zheng's (2008) estimation, the number of migrant workers moving from the western provinces to eastern provinces was about 2.5 million in 2005 (Table 7 on p. 9).<sup>16</sup> Because the migrant workers tend to send their income back to their ancestral western homes, the labour movement helps to increase western households' income. Moreover, according to the current household registration

http://www.jpmorgan.com/cm/BlobServer?blobkey=id&blobwhere=1158595093036&blobheader=application% 2Fpdf&blobcol=urldata&blobtable=MungoBlobs&blobheadername1=Contentdisposition&blobheadervalue1=attachment;filename=Resource\_taxes\_to\_help\_local\_governments\_boost\_Weste

<sup>&</sup>lt;sup>14</sup> Recently there has been debate among Chinese scholars about whether there exists a so-called "curse of resources" in China, a phenomenon also known as the "paradox of plenty," which assumes that the abundant deposits of natural resources are a real curse instead of a blessing posed to a country or a region's economic growth. As pioneers in this field, Xu and Wang (2006) used panel-data methods and found evidence supporting the hypothesis of a resource curse at the provincial level. However, using a data set at city level from 1997 to 2005, Fang et al. (2009) and Fan et al. (2010) found that there is no evidence to support the statement of a resource curse in China. On the contrary, the level of natural resources in a city extends a significant positive diffusion effect on the economic growth of neighbor cities within the same province.

<sup>&</sup>lt;sup>15</sup> For a more detailed interpretation of resource taxes reform, see J. P. Morgan's report "Resource Taxes to Help Local Governments, Boost Western Development", which is available at

rn development.pdf. <sup>16</sup> Compared with the population base of 360 million in Western China, the scale of interregional population movement (0.7% of total western population) was rather limited. This estimation is in line with the finding of Cai et al (2002) and Yang (2002) that most labour flow took place within a region instead of across regions.

system (*Hu Kou*), migrant workers are not counted in the eastern population, although the GDP they produce is still counted in the statistics of eastern provinces. These statistical methods essentially cause an upward bias of per capita GDP in eastern provinces and a corresponding downward bias in western provinces. Given the small scale of interregional labour flow, these biases are not too large.

#### 5. Conclusion

The miracle of China's rapid economic growth in the last decades is accompanied by the fact of widening regional disparity across the country. After two decades of expanding west-east regional disparity, the Western Development programme launched in 2000 has efficiently reversed this trend and led to catch-up progress of western regions. From a historical perspective, the development of China's regional economies has been closely related to government policies as well as regional historical and geographical features. Among these, the regional transfer payment system proves to be of key importance in determining whether historically backward western regions could grow faster than more prosperous eastern counterparts to realize regional convergence. Given these observations, we adapted Funke and Strulik's (2000) analytical framework to investigate the regional divergence and convergence courses since 1992. The model is essentially an "AK model"-style endogenous growth model emphasizing the importance of capital accumulation – massive capital accumulation is one of the most prominent phenomena in China's "economic miracle" and Asian Tigers' economic takeoffs. Unsurprisingly, the calibration results fit well with actual data of China's economy up to 2010.

The success of using our model to explain the economic development of underdeveloped regions in China also sheds some light on an important question in development economics: How to promote regional convergence in a country with unbalanced regional development? From China's experience after the implementation of the Western Development programme in 2000, we can conclude that one important measure, if not the most important measure, is to arrange large-scale interregional transfer payments to less developed regions. It is also especially important for the less developed regions to devote the transfers to the areas that may help to boost regional productivity, such as infrastructure development, research and development projects and education.

#### **References:**

Bai, Chong-En., Hsieh, Chang-Tai., and Qian, Yingyi. 2006. "The Return to Capital in China", *NBER Working Paper*, No. 12755.

Barro, R. J. 1990. "Government spending in a simple model of endogenous growth." *Journal of Political Economy*, Vol. 98, No. 5, pp. 103-125.

Brunner, M., and Strulik, H. 2002. "Solution of perfect foresight saddle point problems: a simple method and applications", *Journal of Economic Dynamics and Control*, Vol. 26(5), pp. 737–753.

Cai F., Wang D., and Du Y. 2002. "Regional Disparity and Economic Growth in China: the Impact of Labour Market Distortions", *China Economic Review*, Vol. 13, pp. 197-212.

Cashin, P. 1995. "Government spending, taxes, and economic growth", *IMF Staff Papers*, Vol. 42(2), pp. 237-269.

Chow, Gregory C. 1993. "Capital Formation and Economic Growth in China," *Quarterly Journal of Economics*, CVIII (3), pp. 809-842.

Chow, Gregory C., and Lin, An-loh. 2002. "Accounting for Economic Growth in Taiwan and Mainland China: A Comparative Analysis", *Journal of Comparative Economics*, Vol. 30(3), pp. 507–530.

Demurger, S., Sachs, J. D., Woo, W. T., Bao, S., Chang, G., and Mellinger, A. 2002. Geography, economic policy, and regional development in China. *Asian Economic Papers*, vol. 1(1), pp. 146-197.

Di, Z., Hou, M., and Qin, M. 2008. "Correlation between Tax Burden and Economic Growth", Journal of Northeast Normal University (Philosophy and Social Science), 2008(6), pp. 155-159.

Fan, R., Fang, Y., and Park, S. Y. 2010. "Resource Abundance and Economic Growth in China", Discussion Paper, available at <u>http://www.wise.xmu.edu.cn/paperInfor.asp?id=181</u>.

Fang, Y., Qi, L., and Zhao, Y. 2009. "The 'Curse of Resources' Revisited: A Different Story from China", Discussion Paper, available at <u>http://www.wise.xmu.edu.cn/paperInfor.asp?id=131</u>.

Funke, M., and Strulik, H. 2000. "Growth and Convergence in a Two-Region Model of Unified Germany", *German Economic Review*, Vol. 1(3), pp. 363-384.

Funke, M., and Strulik, H. 2005. "Growth and convergence in a two-region model: The hypothetical case of Korean unification", *Journal of Asian Economics*, Vol. 16(2), pp. 255-279.

Gong L. T., and Zou H, F. 2000. "Public Expenditures, Taxes, Federal Transfer, and Endogenous Growth", *Chinese Science Funding*, vol. 1, pp. 15-18.

Gong L. T., and Zou H, F. 2002. "Fiscal Federalism, Public Capital Formation, and Endogenous Growth", *Annals of economics and finance*, Vol. 4, pp. 471-490.

He, D. Zhang, W. and J. Shek. 2007. "How Efficient Has Been China's Investment? Empirical Evidence from National and Provincial Data", *Pacific Economic Review* 12, pp. 597-617.

Huang, Y. 1999. Inflation and investment controls in China: the political economy of central-local relations during the reform era. Cambridge: Cambridge University Press.

Jia, L.1998. "Regional catching up and productivity growth in Chinese reform period", *International Journal of Social Economics*, Vol. 25 Iss: 6/7/8, pp.1160-1177.

Li, B. 2007. "On Measurement and Evaluation of China's Macro-tax-burden Level" (in Chinese), *Journal of Finance and Economics*, 2007(12), pp. 18-22.

Litwack, J. M., & Qian, Y. 1998. "Balanced or unbalanced development: special economic zones as catalyst for transition." *Journal of Comparative Economics*, Vol. 26, 117–141.

Ma, S. Y., and Yu, H. X. 2003. "Inter-Governmental Transfer payment and Regional Economic Convergence in China" (in Chinese), *Economic Research Journal*, 2003(3), No. 26-33.

Perkins, D. H. 1988. "Reforming China's Economic System", *Journal of Economic Literature*, Vol. 26(2), pp. 601-645.

Shan, H. J. 2008. "Reestimating the Capital Stock of China: 1952-2006" (in Chinese), *Journal of Quantitative and Technical Economics*, 2008(10), pp. 17-31.

Sturm, J.-E., G.H. Kuper and J. de Haan. 1998. Modelling Government Investment and Economic Growth on a Macro Level: A Review, in: Brakman, S., H. van Ees, and S.K. Kuipers, eds., Market Behaviour and Macroeconomic Modelling, London (MacMillan/St. Martins Press).

Sun, W. K., Xiao, G., and Yang, X. K. 2010. "The influence of capital return to investment rate" (in Chinese), *World Economy*, 2010(6).

Tochkov, K. "Interregional transfers and the smoothing of provincial expenditure in China", *China economic review*, Vol. 18(1), pp. 54-65.

Tsui Kai-yuen. 2005. "Local tax system, intergovernmental transfers and China's local fiscal disparities", *Journal of comparative economics*, Vol. 33(1), pp. 173-196.

Wang, Y. X., and Wu, Y. 2003 "The Preliminary Measurement of Fixed Capital Stocks of Stateowned Economy of China" (in Chinese), *Statistical Research*, 2003(5).

Yang D. T. 2002. "What Has Caused Regional Inequality in China?" *China Economic Review*, Vol. 13, pp. 331- 334.

Yao, S., and Zhang, Z. 2001. "On regional inequality and diverging clubs: a case study of contemporary China." *Journal of Comparative Economics*, vol. 29, pp. 466–484.

Xu, K. and Wang, J. "An empirical study of a linkage between natural resource abundance and economic development." (in Chinese). *Economic Research Journal*, pp. 78–89, 2006 February.

Xu X. 2002. "Have the Chinese provinces become integrated under reform?" *China Economic Review*. Vol. 13, pp.116-133.

Yang, D. T. 2002. "What has caused regional inequality in China?" *China Economic Review*. Vol. 13, pp.331-334.

Yin, H. "Fiscal disparities and the equalization effects of fiscal transfers at the county level in China", *Annals of economics and finance*, Vol. 9, pp. 117-151.

Zhang, H. L., and Qin, P. L. 2011. "The Statistical Appraisal on Effects of Inter-governmental Transfers to Interprovincial Economic Convergence in China" (in Chinese), Shanghai Journal of Economics, 2011(8), pp. 90-98.

Zhang, W. G., Ren, Y. Y., and Hou, Y. Y. 2010. "The Local Government Investment Effects on Longterm Economic Growth -- Evidence from China's Economic Transformation" (in Chinese), China Industrial Economics, 2010(8), pp. 23-33.

Zhao, Z. Y., and Lv B. Y. 2007. "Capital Flow, Capital Supply and the Unbalanced Regional Economic Development" (in Chinese), *China Soft Science*, 2007(12), pp. 152-160.

Zheng. B. W. 2008. "30 Years of Reform and Openness in China: Development and Challenges of Social Security System for Floating Population" (in Chinese), Chinese Journal of Population Science, No. 5, 2008(1), pp. 2-17.

### Michael Funke and Hao Yu

Economic growth across Chinese provinces: In search of innovationdriven gains

# Abstract

In this paper we analyse the impact of R&D on total factor productivity across Chinese provinces. We introduce innovations explicitly into a production function and evaluate their contribution to economic growth in 1993 - 2006. The empirical results highlight the importance and the interaction between local and external research. The evidence indicates that growth in China is not explained simply by factor input accumulation.

Keywords: China, R&D, R&D Spillovers, patents, regional economic growth, semiparametric estimators JEL-Classification: C14, O47, R11, R12

Department of Economics, Hamburg University, Von-Melle-Park 5, 20146 Hamburg, GERMANY Email: <u>funke@econ.uni-hamburg.de</u>, Email: <u>hao.yu@wiso.uni-hamburg.de</u>

## 1 Introduction

Despite China's remarkable achievements and miraculous growth, a great deal of debate and attention has focussed on China's uneven regional developments. Urban and rural standards of living continue to be poles apart. Therefore, the country's leadership has recently been contemplating a smoother ride on its development path by setting forth a guideline prioritising "harmony". The sustained reforms and opening-up over the past two and half decades have resulted in prosperity for many Chinese citizens, but the cross-country income gaps are among the top concerns of the Chinese government. The government's uneasiness stems from the fact that China's history is littered with rebellions, uprisings, and revolutions sparked by economic inequalities. Against this historical experience, Chinese leaders have placed the concept of a "harmonious socialist society" for renewed political legitimacy and political cohesion of the country at the top of their list of things to do. It is envisaged that this harmonious society should enable all the people to share the social wealth brought by reform and development.

Productivity growth is probably the single most important indicator of an economy's health and driver of its real GDP in the long run. The more productive an economy, i.e. the more effectively it uses its capital and labour, the greater its prosperity and standard of living. Productivity can be measured in different ways. Labour productivity is a widely used and transparent measure, but it provides only a partial view of the relationship between inputs and outputs. Total factor productivity (TFP), which can be traced back to the seminal paper by Solow (1957), takes into consideration all inputs and is thus a better measure of technological change.<sup>1</sup> There is growing theoretical and empirical evidence that innovation is among the main sources of TFP growth. Product-related R&D activity creates new markets and process-related R&D activity reduces production costs. When an innovation is commercially successful, its effects spill over to other firms and across regional and national boundaries.<sup>2</sup>

The emergence of the R&D-based endogenous growth literature has re-emphasized the strategic role of technological advance in economic dynamics over time and space. The technological gap is frequently cited as an important factor in explaining income disparities between countries and regions. The recent rise of new growth theory has also led to an overlap between the macroeco-

<sup>&</sup>lt;sup>1</sup> Labour productivity growth may reflect "extensive" growth - doing more with more inputs – while TFP tries to measure "intensive" growth - doing more with less inputs. If China's fast labour productivity growth is entirely the result of capital deepening, then questions about its sustainability would arise.

 $<sup>^{2}</sup>$  A large body of literature has been devoted to these social benefits of R&D investment. See, for example, Coe and Helpman (1995) and Coe et al. (1997).

nomic growth literature and the empirical literature on R&D. R&D-based growth models emphasize the idea that economic growth results from the increasing returns associated with new knowledge. These are models with two sectors: producers of final output and an R&D sector. The R&D sector develops ideas that lead to a monopoly situations. R&D firms are assumed to be able to make monopoly profits by selling ideas to production firms, but the free entry condition means that these profits are dissipated on R&D spending. The implications of this approach are that the higher investment in R&D, the higher the innovative capacity and the faster the economic growth.<sup>3</sup>

In the paper below we try to disentangle the importance of R&D for growth across Chinese provinces. We also investigate different capacities to innovate and to assimilate innovation. These differences may be important for explaining persistent differences in economic performance.

The remainder of the paper is divided into three parts. Section 2 provides a brief overview of the methodological issues of estimating production functions and the building blocks of existing estimators for resolving them. In section 3, we detail our estimation of the structural parameters of the production function and discuss the estimation results in the light of previous studies. Finally, conclusions and implications for Chinese policy making are presented in Section 4.

## 2 Growth accounting: Methodological issues

As stated above, the objective of this paper is to shed light on the determinants of regional economic growth in China. Production functions are used to examine the role of human capital, physical capital, R&D and R&D spillovers in China's innovations. For the purpose of exposition, we introduce our empirical approach by means of the simplest conceivable two-factor Cobb-Douglas production function

(1) 
$$Y_{it} = A_{it} K_{it}^{\beta} L_{it}^{\gamma},$$

where  $Y_{it}$  is *GDP* of province *i* in period *t*,  $K_{it}$  and  $L_{it}$  and inputs of capital and labour, and  $A_{it}$  is the efficiency level of province i in period *t*.<sup>4</sup> The subscripts *i* and *t* refer respectively to the province

<sup>&</sup>lt;sup>3</sup> Izushi (2008) examined the role of R&D underlying the Romer (1990) model and its subsequent modifications, and compared the models against productivity growth of European regions in the 1990s.

<sup>&</sup>lt;sup>4</sup> The estimation methods discussed extend immediately to more factors of production and/or other functional forms, provided that variable inputs have positive cross-partials with respect to productivity.

and the year. While  $Y_{it}$ ,  $K_{it}$  and  $L_{it}$  are observable,  $A_{it}$  is not observable to the researcher. Taking the logarithm of both sides and appending an iid error term yields

(2) 
$$y_{it} = \alpha + \beta k_{it} + \gamma l_{it} + \varepsilon_{it}$$
,

where lower-case letters refer to natural logs and  $\ln(A_{it}) = \alpha + \varepsilon_{it}$ .  $\alpha$  measures the mean efficiency across China and  $\varepsilon_{it}$  is the province- and time-specific deviation from that mean.  $\ln(Ait)$  can further be decomposed into an observable (or at least predictable) and an unobservable component according to

(3) 
$$y_{it} = \alpha + \beta k_{it} + \gamma l_{it} + \kappa_{it} + u_{it},$$

where  $\kappa_{it}$  represents province-level productivity and  $u_{it}$  is an iid component representing measurement errors and/or omitted variables. In other words, the difference between  $\kappa_{it}$  and  $u_{it}$  is that the former is a state variable and hence impacts regional economic performance. Estimated productivity can then be calculated as

(4) 
$$\kappa_{it} = y_{it} - \hat{\beta}k_{it} - \hat{\gamma}l_{it}$$

and TFP in levels can be obtained as the exponential of  $\kappa_{it}$ , i.e. TFP<sub>it</sub> = exp( $\kappa_{it}$ ). Three prominent econometric difficulties arise when TFP is estimated applying ordinary least squares (OLS) to equation (4). First, when using a balanced panel of Chinese provinces no allowance is made for endogenous location decisions, resulting in a selection bias. Second, since productivity and input choices of provinces are likely to be correlated, OLS introduces an endogeneity problem. Third, the product mix is likely to be related to TFP.

The location bias results from the fact that investment decisions of Chinese firms are related to productivity.<sup>5</sup> To make a long story short, high productivity may trigger investment. If

<sup>&</sup>lt;sup>5</sup> Because of their geography and preferential policies, the more developed coastal regions have also been able to attract more foreign direct investment. It is very likely that this exposure to foreign direct investment had a positive impact on TFP. Aziz and Duenwald (2001), Zou and Zhou (2007) and Maasoumi and Wang (2008) have identified club convergence across Chinese provinces. While per capita GDP of poor provinces is catching up with those of the rich ones, the relative income distribution appears to be stratifying into several modes or clubs.

firms have prior knowledge about regional productivities levels kit prior to the investment decision, this will generate correlation between uit and fixed capital. In sum, the problem of endogeneity of attrition or location will generate a negative correlation between uit and kit, causing the capital coefficient in the production function to be biased downwards. The endogeneity problem arises because the inputs in the production function (2) are determined by the characteristics of the province, including its efficiency level. This endogeneity of inputs is defined as the correlation between the level of inputs chosen and unobserved productivity shocks.<sup>6</sup> Finally, Bernard et al. (2005) noted that TFP estimates encounter the product mix problem: the provinces' output composition is likely to be correlated to their productivity.<sup>7</sup>

To remedy these problems, instrumental variables (IV) or fixed effects may be used. IV estimation tries to achieve consistency of coefficient estimates by instrumenting the independent variables that cause the endogeneity problem with regressors that are correlated with these inputs, but uncorrelated with unobserved TFP. In practice, IV estimators have not been particularly useful. One of the obvious shortcomings is the lack of instruments. Furthermore, IV estimation techniques do not provide any solution to the endogeneity-of-location problem .

By assuming that  $\kappa_{it}$  is province-specific but time-invariant, it is possible to estimate (2) using a fixed effects estimator. Moreover, when the endogeneity of location is determined by the time-invariant fixed effects, the fixed effects estimator should also sort out the attrition problem. In spite of the attractive properties of the fixed effects estimator, however, it often leads to unreasonably low estimates of the capital coefficient  $\beta$ . To sum up, traditional estimation methods are vitiated by endogenity and attrition problems and estimation biases.<sup>8</sup>

Below, we briefly introduce the more suitable semiparametric (control function) estimator suggested by Olley and Pakes (1996), with special attention to its advantages and drawbacks. Olley and Pakes (1996) were the first to introduce an estimation algorithm that controls for both the endogeneity and the attrition bias and yields reliable production function estimates. The idea of the

<sup>&</sup>lt;sup>6</sup> In the presence of many issues of inputs and simultaneity it is usually impossible to determine the direction of bias in the estimated  $\beta$  coefficient in equation (2). Levinsohn and Petrin (2003) illustrated, for a two-input production function where labour is the only freely variable input and capital is quasi-fixed, that the estimated  $\beta$  coefficient will be biased downward if a positive correlation exists between labour and capital. The coefficients of the variable inputs will be biased upwards.

<sup>&</sup>lt;sup>7</sup> In a complementary strand of literature, Rodrik (1996) recently emphasized the importance of the product mix. China's exports are as sophisticated as those of a country three times richer. The goods it sells to America overlap to a surprising extent with the merchandise America buys from members of the OECD, argues Schott (2006).

<sup>&</sup>lt;sup>8</sup> Olley and Pakes (1996) applied fixed effects to various samples and found that the time-invariant nature of  $\kappa_{it}$  underlying the model is invalid. This also obtains for Chinese provinces experiencing rapid growth and structural change.

estimator is to invert demand for capital to infer unobserved productivity shocks and then use the estimated productivity shock as a regressor in the production function.<sup>9</sup> The additional attrition problem is addressed by using attrition probabilities.<sup>10</sup>

To be specific, the amount of domestic and foreign investment in region i is assumed to depend on capital and productivity according to

(5) 
$$i_{it} = f_t(k_{it}, \kappa_{it}).$$

Provided investment is strictly increasing in productivity and conditional on capital, it is straightforward to invert factor demands:

(6) 
$$\kappa_{it} = g_t(k_{it}, i_{it}),$$

where  $g_t(\cdot) = f_t^{-1}(\cdot)$ . Inserting (6) into (2), yields

(7) 
$$y_{it} = \alpha + \beta k_{it} + \gamma l_{it} + g_t(k_{it}, i_{it}) + u_{it}$$

Define the function

(8) 
$$\phi(k_{it},i_{it}) = \alpha + \beta k_{it} + g_t(k_{it},i_{it}).$$

Estimation then proceeds in three steps. In the first stage of the semiparametric algorithm, the following equation is estimated using OLS

(9) 
$$y_{it} = \alpha + \gamma l_{it} + \phi(k_{it}, i_{it}) + u_{it},$$

<sup>&</sup>lt;sup>9</sup> Other inversion (control function) estimators have been suggested by Levinsohn and Petrin (2003), Pavcnik (2002) and Doraszelski and Jaumandreu (2008). To address the endogeneity bias, the GMM system estimator is also a suitable estimation method provided lagged values and lagged differences are good instruments [see Blundell and Bond (2000)]. The degree to which these instruments are a good choice is subject to some discussion in the literature. Below we focus upon techniques which are more structural in nature.

<sup>&</sup>lt;sup>10</sup> Below we give a brief overview with emphasis on the mechanics of the estimator. More elaborated expositions can be found in Olley and Pakes (1996) and Ackerberg et al. (2007).

where  $\phi(\cdot)$  is approximated by a higher-order polynomial in  $i_{it}$  and  $k_{it}$ . Before moving further, let us recall that estimation of (9) leads to a consistent estimate of  $\gamma$ . However, estimation of (9) does not identify  $\beta$ , so some further effort is required to disentangle the effects of capital on investment from the effect on output.

In order to recover the coefficient of the capital variable, information on regional dynamics is exploited. We assume that productivity evolves over time as an exogenous first-order Markov process:

(10) 
$$\kappa_{it+1} = E\left(\kappa_{it+1} \mid \kappa_{it}\right) + \xi_{it+1},$$

where the stochastic nature of productivity improvement is captured by  $\xi_{it+1}$  which is treated as an *iid* shock with zero mean and variance  $\sigma_{\xi}^2$ .<sup>11</sup> Regions will remain as attractive investment locations  $(\chi_{it+1} = 1)$  as long as productivity exceeds a lower bound, i.e.  $\chi_{it+1} = 1$  for  $\kappa_{it+1} \ge \tilde{\kappa}_{it+1}$ , where  $\chi_{it+1}$  is a survival indicator. Considering the expectation of  $E[\gamma_{it+1} - \gamma_{it+1}]$  yields

$$E[\gamma_{it+1} - \gamma_{it+1} | k_{it+1}, \chi_{it+1} = 1] = \alpha + \beta k_{it+1} + E[\kappa_{it+1} | k_{it}, \chi_{it+1} = 1]$$

(11) 
$$y_{it+1} - \gamma_{it+1} = \alpha + \beta k_{it+1} + E[\kappa_{it+1} | k_{it}, \chi_{it+1}] + \xi_{it+1} + u_{it+1}$$

$$y_{it+1} - \gamma_{it+1} = \alpha + \beta k_{it+1} + g(\mathbf{p}_{it}, \phi_{it} - \beta k_{it}) + \xi_{it+1} + u_{it+1}$$

where  $E[_{\kappa_{it+1}|k_{it}, \chi_{it+1}}] = g(p_{it}, \phi_{it} - \beta k_{it})$  follows from the law of motion for productivity shocks and  $p_{it}$  is the probability of attrition of region *i* in period *t*+1, i.e.  $p_{it} = \text{prob}\{\chi_{it+1} = 1\}$ . In our implementation, we estimate the probability of attrition in the second stage by fitting a probit model of  $\chi_{it}$  on  $i_{it-1}$  and  $k_{it-1}$  and their squares and cross products. Denote the predicted probabilities from the probit model as  $\hat{p}_{it}$ .

In the third stage, this finally leads to the following equation by nonlinear least squares, which enables identification of the coefficient of capital,  $\beta$ :

<sup>&</sup>lt;sup>11</sup> In equation (10), productivity is modelled as an exogenous Markov-process. Recently, Doraszelski and Jaumandreu (2008) endogenised productivity by allowing it to depend on R&D. The Markov transition matrix methodology was adopted by Curran et al. (2007) and Sakamoto and Islam (2008) to study real GDP per capita convergence across Chinese counties and provinces and to capture the dynamics embodied in the Chinese data.

(12) 
$$y_{it+1} - \hat{\gamma}_{it+1} = \alpha + \beta k_{it+1} + g(\hat{p}_{it}, \hat{\phi}_{it} - \beta k_{it}) + \xi_{it+1} + u_{it+1},$$

where the unknown function  $g(\cdot)$  is approximated by a second-order polynomial.<sup>12</sup> Because the estimation routine involves three steps, deriving analytical standard errors is nontrivial. Therefore bootstrapped standard errors are used and the variation in the bootstrapped samples provides an estimate of the standard errors of the original point estimates.

### 3 Empirics: Economic growth across Chinese provinces

Empirical research on economic growth is faced with considerable uncertainty given a set of multiple, overlapping theories emphasizing different growth channels. Brock and Durlauf (2001) referred to this as "openendedness" of economic theories, in the sense that the truth of one theory does not imply the falsity of another. Furthermore, within each channel there may be alternative measures representing the same theory. Against this background, we introduce our specification of the production function and our dataset.

The data are a balanced panel for 30 Chinese provinces for the period 1993 to 2006.<sup>13</sup> GDP data were obtained from the *Statistical Yearbook of China*.

Measuring provincial capital stocks is a challenging task. Of particular relevance are the assumptions concerning depreciation rates, initial capital stock and the appropriate deflators. We compiled the capital stock data using the perpetual inventory method recently suggested by He et al. (2007).

Needless to say, human capital stock data for China are hard to come by. Therefore, various previous papers have not paid attention to the differences between unadjusted labour input and human capital. Neither have they addressed the differences between the accumulation and the stock of human capital. In our approach below we use a proxy for the stock of human capital, since we believe that the stock of human capital generates technological innovations and facilitates learning

<sup>&</sup>lt;sup>12</sup> The estimates are based on Newton's method. We have also used grid search for confirming that the procedure finds the global minimum of the objective function.

<sup>&</sup>lt;sup>13</sup> China currently has 23 provinces, four Centrally Administered Municipalities, and five autonomous regions. Since these entities are administratively equal, we will use the term "province" throughout the paper. Chongqing became a new municipality only in 1997. In our database we therefore combined Chongqing and Sichuan.

and technology diffusion.<sup>14</sup> We first calculate the number of workers at different education levels using microcensus data. The Chinese microcensus provides the official representative statistics for the population and the labour market (*National Sample Survey of Population Changes*), covering each year 1‰ of all households in China. The survey allows one to calculate for each province i (i = 1, ..., 30) the shares of workers having finished primary school ( $S_1$ ), junior secondary school ( $S_2$ ), senior secondary school ( $S_3$ ), special school ( $S_4$ ) and higher education ( $S_5$ ). Primary schooling is assumed to last 6 years, junior secondary school 9 years, senior secondary school 12 years, special school 11-12 years and higher education 16 years. The average length of schooling of the provincial workforce in period t can then be calculated according to

(13) 
$$l_{it} = \frac{\left(6S_{1,it} + 9S_{2,it} + 12S_{3,it} + 11.5_{4,it} + 16_{5,it}\right)}{\text{Population}_{it}}.$$

Multiplying  $l_{it}$  with the provincial labour force then enables calculation of the stock of human capital as

(14) 
$$H_{it} = l_{it} \text{ Labour Force}_{it}$$

It is well known that it is difficult to disentangle technological improvement from production inputs. Therefore, it is hardly surprising that contributions within this strand of research differ widely as to how R&D activities are inferred. In our study, patent applications are used as a proxy for technological innovations. Patent applications are easily accessible data and have been widely used as indicators of innovation and diffusion.<sup>15</sup> Another reason is that most studies find a very strong relationship between R&D and patent applications at the cross-sectional level: the median R<sup>2</sup> is around 0.9.<sup>16</sup> Because patent applications are disclosed only 18 months after filing, 2006 is the latest year for which data are available.

<sup>&</sup>lt;sup>14</sup> Similar methodologies have been used by Wang and Yao (2003) and Islam et al. (2006). Although on-the-job training and firm-specific human capital investments contribute to the improvement of the human capital stock, they were excluded due to missing data and measurement problems.

<sup>&</sup>lt;sup>15</sup> One might be inclined to think that the number of patent filings is only a sketchy indicator of a country's ability to generate ideas. This is because the usefulness of patents can vary widely, nor are all invention patents commercially successful. A good reference on the caveats of using patent data as a measure of innovation is Griliches (1990).

<sup>&</sup>lt;sup>16</sup> See e.g. Griliches (1990) and Hausman et al. (1986) for the international evidence. Sun (2000) analysed the innovation landscape across Chinese provinces. The analysis supports the view that state-supported R&D activities are the major source of invention patent applications. Agglomeration does not seem to be a significant factor for patents in China, in contrast to findings of other studies.

China's patent system is evolving fast, and enforcement, though lagging, is improving. China's Patent Law went into effect on 1 April 1985. Generally, the technology to be patented must pass four tests: that it is novel, useful, non-obvious and man-made. The law grants three types of patents: invention, utility model and design patents. Applications for invention patents are more rigorously scrutinized for novelty and non-obviousness before the patents are granted. Invention patents receive 20 years of protection, up to the global standard. On the contrary, the utility and design patents generally cover more incremental innovations and are not subject to examination for novelty and inventive step. In the empirical work below we use invention patent applications because we consider them to represent high quality ideas.

The first amendment of the Chinese Patent Law entered into effect on 1 January 1993. The duration of invention patent protection was extended from 15 to 20 years. The Law was then revised again in August 2000. The amended law simplified the procedures of patent application and examination.<sup>17</sup>

Chinese patent applications increased at an annual rate of 3.5 per cent in 1993 – 1999; from 2000 to 2006, the annual increase was 29.7 per cent.<sup>18</sup>

Using the invention patent application data, we compiled the provincial patent stock data using the perpetual inventory method. Before turning to the estimation results, it is interesting to look at the R&D intensity across provinces.<sup>19</sup> Figures 1 and 2 show the distribution of invention patent stocks by province. The evidence is quite striking and reveals that R&D intensity varies substantially between subsets of provinces. The figures provide a landscape and league table of China's most intellectually creative regions. In 2006, the geographical distribution of patents exhibits a clear pattern: the coastal and central provinces, such as Guangdong, Beijing, Shanghai, Jiangsu, Zhejiang, Shandong, Tianjin, Liaoning, Hunan, and Sichuan, are all among the top 10 innovative regions. These top 10 regions account for 80 per cent of the total invention patent stocks in 2006.

<sup>&</sup>lt;sup>17</sup> One caveat is worth mentioning. Focussing on patents as an indicator of innovativeness overlooks an important fact. What matters for economic innovation is turning scientific discoveries into new products and smart processes. In other words, a venturesome array of products and venturesome consumption patterns matter. One may argue that Chinese firms have a large pool of domestic customers that do not have the same high expectations as Western customers typically have. Chinese firms can therefore practise on their domestic customers while they improve quality to the point where they can begin to export.

<sup>&</sup>lt;sup>18</sup> This trend is important because countries that create intellectual property eventually enforce it as well.

<sup>&</sup>lt;sup>19</sup> Chinese patents include filings from domestic and foreign firms. Unfortunately, it is impossible to winnow domestic from foreign patents neatly because filings from joint ventures are always classified as having a Chinese origin. This implies that some international R&D spillovers arising from China's "open door policy" are reclassified as intranational spillovers and hence included in the analysis. Cheung and Lin (2004) found supporting evidence that spillovers from FDI have sparked patent applications. Likewise, Hu and Jefferson (2009) found that China's surge in patent activity is due to the amendments to property right laws and the expansion of FDI.



Figure 1 Inter-regional differences in invention patent stocks, 1993

Notes: The Chinese patent data are available online at http://www.sipo.gov.cn/sipo2008/



Figure 2 Inter-regional differences in invention patent stocks, 2006

Notes: The Chinese patent data are available online at http://www.sipo.gov.cn/sipo2008/

One element that is missing is the possibility that regions benefit from spatial spillovers. The idea is that provinces can benefit from external knowledge. This requires a proxy for the ability of regions to learn or assimilate knowledge from others. The knowledge spillovers approach has been adopted by economists using different quantitative methods. To guide our thinking, we follow Funke and Niebuhr (2005) and Kuo and Yang (2008) and define a patent externality variable as

(15) 
$$PS_{i} = \ln \left[ \sum_{\substack{j=1\\j\neq i}}^{30} P_{j} w_{ij} \right],$$

where  $P_j$  represents the stock of patents in province *j*,  $w_{ij}$  represents the spatial weight and  $PS_j$  is the patents spillovers variable. Knowledge external to a province is obtained as a combination of the stock of patents obtained by other provinces and weighted by a measure of proximity within the geographic space.<sup>20</sup> We assume that the spatial weights  $w_{ij}$  decline with geographic distance and follow a negative exponential function with distance decay parameter  $\beta_E$  defined as

(16) 
$$w_{ij} = e^{-\beta_E d_{ij}},$$

where  $d_{ij}$  is the rail travel time between the provincial capital cities *i* and *j*.<sup>21</sup> As  $\beta_E$  increases, the frictional effect of distance rises and interaction declines.<sup>22</sup> Lastly, the parameter  $\beta_E$  is defined as

<sup>&</sup>lt;sup>20</sup> The decisive factor in the knowledge transmission process is the difference between codifiable and tacit knowledge. This distinction helps one to understand why modern information technologies do not erase the importance of proximity. It is indeed difficult to visualise any barriers to the diffusion of codifiable information. But tacit knowledge is clearly different. Tacit knowledge is embedded in the minds of people and the routines of firms and so does not move easily from place to place. Even codified innovations like those covered in patents do not flow freely from place to place. Frequently, in order to make full advantage of the insights provided in a patented (codified) innovation, one needs to have the complementary tacit knowledge to apply it. Below we investigate this idea in more detail.

<sup>&</sup>lt;sup>21</sup> We choose rail travel time, rather than straight line distance, as it gives a more realistic representation of the cost of interaction and contacts across space. The use of kilometres would not allow us to take into account different types of train connections which significantly affect real world interactions.

<sup>&</sup>lt;sup>22</sup> Adams and Jaffe (2002) have emphasized that the flows of interregional knowledge spillovers are likely to wane with distance, as the potential for face-to-face and other forms of interaction decay. At the EU level, Greunz (2003) found a bounded effect on local patenting activity of innovative efforts pursued in the neighbourhood of up to 360 kilometres. In the same vein, Bottazzi and Peri (2003) found regional spillover effects with a positive impact of neighbouring regions' R&D efforts within a 200-300 kilometre limit. This implies that knowledge spillovers are mainly an intra-national phenomenon.

(17) 
$$\beta_E = \frac{-\ln(1-\gamma_E)}{D_{Min}},$$

where  $D_{Min}$  is the average distance between the capital cities of adjacent provinces and  $0 \le \gamma_E \le 1$  is a transformed distance decay parameter. As is customary in the literature, we chose  $\gamma_E = 0.8$ . The amount of knowledge flowing from outside the province is thus proxied by the magnitude of all other provinces' R&D activity weighted by the inverse of the bilateral travel time.

In order to provide a sense of the data, the calculated neighbouring provinces' patent stocks are given in Figures 3 and 4. The data shed light on the role of technological diffusion, geographic distance and external accessible patents in the Chinese innovation process.

Figure 3 Neighbouring regions' patent stocks, 1993





Figure 4 Neighbouring regions' patent stocks, 2006

The above ideas are embedded in the production function:

(18) 
$$y_{it} = \alpha h_{it} + \beta k_{it} + \gamma p_{it} + \delta p s_{it} + \theta coast_i + \rho (coast_i \times p s_{it}) + \lambda_t + \varepsilon_{it},$$

where  $y_{it}$  is logged GDP (constant prices),  $k_{it}$  is the logged capital stock,  $h_{it}$  is the logged stock of human capital,  $p_{it}$  is the logged provincial stock of patents,  $ps_{it}$  is the logged stock of knowledge spillovers, *coast* is a dummy variable representing the economically most developed coastal belt, the  $\lambda_t$  are time dummies and  $\varepsilon_{it}$  is an iid error term.<sup>23</sup> The time dummies may control for macroeconomic shocks and potential endogeneity arising from transitory shocks while the "coastal effect" may pick up some omitted variables.<sup>24</sup> Previous work on regional growth across Chinese provinces has typically found a dummy variable for the coastal provinces to be positive and statistically significant. Table 1 looks at the contribution of each factor to GDP.

<sup>&</sup>lt;sup>23</sup> The coastal belt consists of the province/municipalities Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan.

<sup>&</sup>lt;sup>24</sup> Note that the stock of human capital is rather sluggish and remains fairly constant over time. This implies that fixed effects cannot be estimated because coefficients will be weakly identified. Put differently, there is not enough variation in the data for a separate identification of all coefficients.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$h_{it}$	0.291	0.286	0.282	0.278	0.328	0.321	0.328	0.321
	(20.1)	(19.6)	(18.6)	(18.4)	(19.2)	(18.5)	(19.0)	(18.5)
$k_{it}$	0.339	0.399	0.306	0.300	0.216	0.296	0.204	0.172
	(1.9)	(2.2)	(2.1)	(2.4)	(1.2)	(1.9)	(2.1)	(1.7)
$p_{it}$	0.235	0.308	0.278	0.358	0.243	0.340	0.251	0.297
	(2.5)	(3.7)	(3.5)	(4.6)	(2.9)	(4.2)	(4.1)	(4.7)
$sp_{it}$	-	-	0.066	0.020	0.049	0.020	0.025	0.010
Î			(1.9)	(0.6)	(1.4)	(0.6)	(0.9)	(0.3)
coast	-	-	-	-	0.106	0.096	0.122	0.045
					(5.1)	(4.6)	(0.9)	(0.3)
$coast \times sp_{it}$	-	-	-	-	-	-	0.094	0.085
1 "							(3.3)	(3.1)
$\lambda_t$	No	Yes	No	Yes	No	Yes	No	Yes
		(prob = 0.01)		(prob = 0.01)		(prob = 0.25)		(prob = 0.24)

#### Table 1 Olley-Pakes estimates

Notes: Olley-Pakes panel data estimation; sample period: 1993-2006; bootstrapped *t*-values are given in parentheses; the constant terms from the 1<sup>st</sup> stage of the procedure are not reported. The prob value of the joint significance of the time dummies is given in the last row.

Several results are worth noting. First, in columns 1 and 2 we start from the base specification without R&D spillovers. All estimated coefficients are significant and the resulting returns to scale are broadly in line with common sense. Thus, our results verify the role of patents as a source of growth in China.<sup>25</sup> This implies that growth in China in 1993 - 2006, contrary to the seminal work by Chow (1993) for the upstream period 1958 - 1980, was not solely brought about by capital accumulation. The results also shed doubt on Krugman's (1994) former assertion that growth in Asia is simply explained by input accumulation.<sup>26</sup>

Second, in columns 3 and 4 we have added the stocks of external accessible patents. The significant spillover variable in column 3 suggests that, to some extent, external R&D may even compensate for weak contributions of the R&D activities pursued locally. The flip side is that individual provinces may be held back, not just by their own endowment but by the endowments of their neighbouring provinces. If true, economic development would best be coordinated to be fully effective. As expected, the impacts of inter-provincial R&D spillovers are smaller than those of the own-R&D effect. The estimated magnitude of the stock of patents is 0.066, indicating that a 1 per

<sup>&</sup>lt;sup>25</sup> This province-level evidence is consistent with the firm-level evidence in Hu and Jefferson (2004) indicating substantial and significant returns to R&D.

<sup>&</sup>lt;sup>26</sup> The claim that "the newly industrializing countries of Asia, like the Soviet Union of the 1950s, have achieved rapid growth in large part through an astonishing mobilization of resources ... Once one accounts for the role of rapidly growing inputs in these countries' growth, one finds little left to explain" [Krugman (1994), p. 70] has stimulated a heated debate. This is a somewhat extreme position, and Young's (1995) careful growth accounting work indicates that, except for Singapore, there is evidence of TFP growth in the three Asian tigers: Hong Kong, South Korea and Taiwan.

cent increase in the stock of neighbouring patents would boost regional GDP by about 0.066 per cent, controlling for other variables. Column 4, however, indicates that the spillover variable turns out insignificant once the time dummies are added. The implication is that the uniform spillover variable has to be considered fragile.

Third, in column 5 and 6 we augment the specification with the coastal dummy as a further robustness check. The empirical results in Table 1 show that the dummy variable turns out to be significant, indicating a "missing element" in explaining regional growth across mainland China. Alternatively, one might say that there is something unique explaining the coastal growth process. The results in columns (5) and (6) indicate that coastal regions were on average 10 per cent more efficient than what would be predicted given their other characteristics.

Finally, we looked at the interaction of the coastal dummy variable with the R&D spillover variable. The results are documented in column 7 and 8. As can be seen, the results show that an increase in neighbouring R&D significantly boosts growth in the coastal areas while the own-patent stocks are still highly significant. Thus, other things being equal, a coastal province within an innovative neighbourhood is more advanced than one in the vicinity of less innovative provinces. When the interactive term is included, the coastal dummy variable and the overall R&D spillover variable become insignificant. This implies that the intranational R&D spillover effects matter only for the more developed eastern provinces. The results also shed new light on the omitted causes of economic growth captured by the traditional eastern belt dummy.<sup>27</sup>

### 4 Summary remarks and conclusions

If China is to sustain growth in the years ahead, it must become a more innovation-based economy. Firms need to introduce or improve products or production processes over time, first to satisfy market needs and second to cope with increased competition from diffusion phenomena.<sup>28</sup> We have therefore carefully utilized semiparametric estimation techniques and measurement to determine the impacts of innovation and diffusion on economic growth in China.

<sup>&</sup>lt;sup>27</sup> The new perspective implies that there is no need for a special "coastal theory", at least with regard to unequal growth across Chinese provinces. In other words, the suggested channel may be able to identify the cause of the "coastal enigma".

<sup>&</sup>lt;sup>28</sup> Japan, Taiwan and South Korea, which also started off by competing mainly on cheap labour, ended up challenging the West's biggest technology companies. All three countries now own a plethora of patents.

Our econometric approach enabled us to analyse different regional capacities to innovate and to assimilate innovation. Several implications can be extracted from the results of the empirical analysis. Local patents show a positive and significant relationship with provincial GDP. This implies that innovation was an important engine of growth in China over the sample period. Furthermore, external patents and neighbourhood effects turn out to be significant solely in the economically more active coastal provinces. This distinction is important for understanding the nature of the innovative landscape across China and may help us to develop regionally differentiated development policies.<sup>29</sup>

In the future one can imagine the development of even more elaborated methods and models. Nevertheless, we feel that already much is learned by applying state of the art models to the data, and seeing how well they describe reality. We hope this is what we have accomplished here.

<sup>&</sup>lt;sup>29</sup> According to the recent "*Global Innovation Index*" of the Economist Intelligence Unit, China was one of the biggest gainers. China has moved from 59th to 54th in the ranking, a gain that was expected to take five years instead of two. The index, which measures innovation performance in 82 countries, is based on the number of patents granted by patent offices in the United States, European Union and Japan. It also includes factors that help and hinder the ability to innovate, such as the amount of research and development undertaken and the technical skills of the country's workforce (see Hhttp://graphics.eiu.com/PDF/Cisco\_Innovation\_Complete.pdfH).

### References

- Ackerberg, D., Benkard, C.L., Berry, S. and A. Pakes (2007) "Econometric Tools for Analyzing Market Outcomes", in: Heckman, J.J. and Leamer, E.E. (eds) *Handbook of Econometrics*, Vol. 6A, Amsterdam (Elsevier), 4171-4276.
- Adams, J.D. and A.B. Jaffe (2002) "Bounding the Effects of R&D: An Investigation Using Matched Firm and Establishment Data", *Rand Journal of Economics* 27, 700-721.
- Audretsch, D.B. and M.P. Feldman (1996) "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review* 86, 630-640.
- Aziz, J. and C. Duenwald (2001) "China's Provincial Growth Dynamics", *IMF Working Paper No.* WP/01/3, Washington.
- Bernard, A.B., Redding, S. and P.K. Schott (2005) "Products and Productivity", *NBER Working Paper* 11575, Cambridge.
- Blundell, R. and S. Bond (2000) "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews* 19, 321-340.
- Bottazzi, L. and G. Peri (2003) "Innovation and Spillovers in Regions: Evidence from European Patent Data", *European Economic Review* 47, 687-710.
- Brock, W.A. and S.N. Durlauf (2001) "Growth Empirics and Reality", *World Bank Economic Review* 15, 229-272.
- Cheung, K.Y. and P. Lin (2004) "Spillover Effects of FDI on Innovation in China: Evidence from the Provincial Data", *China Economic Review* 15, 25-44.
- Chow, G. (1993) "Capital Formation and Economic Growth", *Quarterly Journal of Economics* 108, 809-842.
- Coe, D. and E. Helpman (1995) "International R&D Spillovers", *European Economic Review* 39, 859-887.
- Coe, D.T., Helpman, E. and A. Hoffmaister (1997) "North-South R&D Spillovers", *The Economic Journal* 107, 134-149.
- Curran, D., Funke, M. and J. Wang (2007) "Economic Growth Across Space and Time: Subprovincial Evidence from Mainland China", *Bank of Finland*, *BOFIT Discussion Paper No*. 21/2007, Helsinki..
- Doraszelski, U. and J. Jaumandreu (2008) "R&D and Productivity: Estimating Production Functions When Productivity is Endogenous", *CEPR Discussion Paper No.* 6636, London.
- Funke, M. and A. Niebuhr (2005) "Regional Geographic Research and Development Spillovers and Economic Growth", *Regional Studies* 39, 143-153.
- Greuz, L. (2003) "Geographically and Technologically Mediated Knowledge Spillovers Between European Regions", *Annals of Regional Science* 37, 657-680.
- Griliches, Z. (1990) "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature* 28, 1661-1707.

- Hausman, J., Griliches, Z, and B.B. Hall (1986) "Econometric Models for Count Data with an Application to the Patent-R&D Relationship", *Econometrica* 52, 909-938.
- He, D., Zhang, W. and J. Shek (2007) "How Efficient Has Been China's Investment? Empirical Evidence from National and Provincial Data", *Pacific Economic Review* 12, 597-617.
- Hu, A. and Jefferson, G.H. (2004) "Returns to Research and Development in Chinese Industry: Evidence from State-Owned Enterprises in Beijing", *China Economic Review* 15, 86-107.
- Hu, A. and Jefferson, G.H. (2009) "A Great Wall of Patents: What is Behind China's Recent Patent Explosion?", *Journal of Development Economics* (forthcoming).
- Islam, N., Dai, E., Sakamoto, H. (2006) "Role of TFP in China's Growth", *Asian Economic Journal* 20, 127-159.
- Izushi, H. (2008) "What Does Endogenous Growth Theory Tell about Regional Economies? Empirics of R&D Worker-based Productivity Growth", *Regional Studies* 42, 946-960.
- Krugman, P. (1994) "The Myth of Asia's Miracle", *Foreign Affairs* 73, Number 6 (November/December), 62-78.
- Kuo, C.-C. and Yang, C.-H. (2008) "Knowledge Capital and Spillovers on Regional Economic Growth: Evidence from China", *China Economic Review* 19, 594-604.
- Levinsohn, J. and A. Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies* 70, 317-341.
- Maasoumi, E. and Le Wang (2008) "Economic Reform, Growth and Convergence in China", *The Econometrics Journal* 11, 128-154.
- Olley, S.G. and A. Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry", *Econometrica* 64, 1263-1297.
- Pavcnik, N. (2002) "Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants", *Review of Economic Studies* 69, 245-276.
- Rodrik, D. (2006) "What's So Special About China's Exports?", China & World Economy 14, 1-19.
- Romer, P.M. (1990) "Endogenous Technological Change", *Journal of Political Economy* 98, S71-S102.
- Sakamoto, H. and N. Islam (2008) "Convergence Across Chinese Provinces: An Analysis Using Markov Transition Matrix", *China Economic Review* 19, 66-79.
- Schott, P.K. (2006) "The Relative Sophistication of Chinese Exports", *NBER Working Paper No.* 12173, Cambridge (Mass.).
- Solow, R.M. (1957) "Technical Change and the Aggregate Production Function", *The Review of Economics and Statistics* 39, 312-320.
- Sun, Y. (2000) "Spatial Distribution of Patents in China", Regional Studies 34, 441-454.
- Wang, Y. and Yao, Y. (2003) "Sources of China's Economic Growth 1952-1999: Incorporating Human Capital Accumulation", *China Economic Review* 14, 32-52.
- Young, A. (1995) "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience", *Quarterly Journal of Economics* 110, 641-680.
Zou, W. and H. Zhou (2007) "Classification of Growth Clubs and Convergence: Evidence from Panel Data in China, 1981–2004", *China & World Economy* 15, Number 5, 91-106.



# Uncertainty and risk analysis of the Langrun Chinese GDP Forecast: Fan Charts revisited

Hao Yu\*

Department of Economics, Hamburg University, Von-Melle-Park 5, 20146 Hamburg, Germany

(Received 1 April 2011; final version received 16 January 2012)

In this paper, we develop a fan chart methodology for Chinese economic growth to incorporate uncertainty analysis into the gross domestic product growth forecast. Using the 'Langrun Forecast' project results exclusively, we estimate the density distribution for Chinese gross domestic product growth forecasts and build corresponding fan charts for the first time. Our analysis shows that the fan chart method effectively highlights the overall uncertainty and balance of risks surrounding Chinese gross domestic product growth, especially during the past international financial crisis between 2007 and 2009. Wallis' interval forecast test is conducted to evaluate the performance of the produced fan charts, and the results indicate that our forecasts perform well for the period being investigated.

**Keywords:** Fan chart; balance of risks; economic forecast; Chinese GDP growth rate; Langrun Forecast

JEL Classification: C82; E37; O53

# 1. Introduction

As with any forecast, economic forecasts inherently suffer from the potential problem of uncertainty. According to Hendry and Ericsson (2001), such uncertainty stems from two aspects: the unavoidable uncertainty latent in unknown future events and the controllable uncertainty arising from the models and methodologies used to forecast.<sup>1</sup> The uncertainty may cause poor forecast results and, sometimes, even severe forecast failures. Although forecast technologies have improved substantially during the past decades, the uncertainty problem in economic forecasting has not been efficiently solved.

Because forecast uncertainty is inevitable, a more practical strategy is to identify it instead of trying to eliminate it. The traditional and currently prevailing point forecasts contain no uncertainty information at all, making it more vulnerable to inaccuracy due to uncertainty. In comparison, density and interval forecasts provide descriptions of forecast uncertainty to different degrees. As a result, in recent years, evaluations of these forecast methods are receiving increasing attention. At the same time, the fan chart, as one possible intuitive visual communication device that can in a single graph represent the evolvement of uncertainty situation embodied in interval or density forecasts over time, becomes increasingly popular. Fan charts were first used by the Bank of England's Monetary

<sup>\*</sup>Email: Hao.Yu@wiso.uni-hamburg.de

Policy Committee (MPC) in its quarterly *Inflation Report* in 1996. Soon after the Bank of England's original effort, the use of fan charts was rapidly extended to other aspects of the economy, including gross domestic product (GDP) growth rates and other important economic indicators. For example, GDP fan charts have appeared in each *Inflation Report* since November 1997.

Although fan chart is widely used as an illustrative device to represent the uncertainty situation in inflation or GDP density forecasts, the fan chart method itself has drawn researchers' increasing interest and the volume of literature concerning fan chart method is increasing rapidly since 1998. Blix and Sellin (1999) formally compared fan chart and several other methods representing inflation forecasts with uncertainty intervals and highlighted the fan chart's ability to demonstrate the uncertainty and asymmetric risk. Tay and Wallis (2000) surveyed the applications of density forecast in macroeconomics and finance and illustrated the usefulness of fan chart as an illustrative communication device. Wallis (2003) reviewed likelihood ratio tests of goodness-of-fit and independence of interval forecasts by evaluating the US Survey of Professional Forecasters and the Bank of England's fan charts. He found that fan chart method tends to produce better forecast in shorter horizon (for current quarter), while it might fan out too quickly for one-year-ahead forecasts. Dowd (2008) made a systematic evaluation of the Bank of England's real GDP growth fan charts. His conclusion suggests that, contrary to Wallis (2003), the shorter horizon GDP forecasts have more problems, while the performance of longer-term forecasts depend to some extent on the GDP estimates used in the assessment. More recently, Celasum, Debrun and Ostry (2007) extended the use of fan chart method to analyze the public debt sustainability in emerging market countries. They derived fan charts by marring the pattern of shocks and the endogenous response of fiscal policy at the same time. Elekdag and Kannan (2009) developed a procedure for incorporating market-based information into the construction of fan charts, which gives fan charts more power to represent the uncertainty and risk balance in a more objective way. Then they applied their approach to build fan charts for the global growth rate forecast.

Despite the numerous applications of fan charts so far, this method has not yet been applied to the Chinese economy. During the last three decades, the Chinese economy has made tremendous achievements in promoting robust and steady growth. As China becomes more important to the world economy, the demand for high-quality economic forecasts for China is growing. Furthermore, the recent financial crisis and its associated heightened levels of macroeconomic volatility have underscored the limitation of traditional point forecasts as a sufficient information resource for macroeconomic policymakers and therefore shed light on the great importance of density forecasts. As economic globalization is advancing with an accelerated pace nowadays, the transmission of market risks between different countries is much faster than any other time in history. Therefore, interval and density forecasts for the Chinese economy are becoming more appropriate and important than traditional point forecasts due to their ability to reveal uncertainty and balance of risk in Chinese economy. As a result, our main contribution in this paper is, for the first time, to produce density forecasts of Chinese GDP growth rates and apply the fan chart approach to display the forecast results. Additionally, we also make a tentative effort to evaluate the performance of the fan charts according to Wallis' (2003) interval test methodology.

The remainder of this paper is organized as follows. In section 2, we briefly introduce the empirical methodology of direct and indirect approaches to constructing fan charts. Our empirical work of estimating the distribution of Chinese economic growth forecasts and producing corresponding fan charts is implemented in section 3. Then, we conduct a simple evaluation of the prediction performance of established fan charts. The last section offers main conclusions and possible directions for future research. Some technique details are provided in the appendix.

## 2. Empirical methodology

Given intrinsic uncertainty, the forecast of future GDP growth rates and other economic indicators should, in fact, follow certain statistical distributions. The main task for forecasters is therefore to detect and reveal these distributions. In this paper I am dealing the problem of Chinese GDP growth rate density forecast.

Following the practice of MPC and other researchers, we assume that the density forecasts of Chinese GDP growth rate follow the two-piece normal distribution function (TPN). As Blix and Sellin (1999) stressed, employing the TPN function is basically because it is easy to compute and it allows skewness, which reflects the asymmetry of upside and downside risks to GDP growth rate. In fact, TPN distribution has already become a standard assumption in the researches concerning density forecast. Britton, Fisher and Whitley (1998) and Wallis (1999) are pioneers in this field. They explained the underlying principles, features and advantages of utilizing TPN distribution and fan chart in central banks' inflation forecasts. Tay and Wallis (2000) took TPN distribution as a concrete example to interpret the application of the density forecasts. They elaborated how the Bank of England utilized the TPN distribution in their inflation forecast practices and how the bank used fan charts to display their forecast results. More recently, Celasun et al. (2006) and Elekdag and Kannan (2009) further extended the applications of TPN distribution into new researching areas such as public debt sustainability analysis and the forecast of global economic growth under uncertainty.

From a mathematical perspective, a TPN distribution could be considered as a combination of two standard normal distributions with the same mode. The three key variables for the TPN distribution are mode, variance and skewness. The mode is the single most likely outcome based on current knowledge and judgment, even if the actual chance of it matching the eventual outcome is small. The variance measures the total uncertainty surrounding the target variable (in our case the GDP growth rate). The skewness is an approximate measure of whether or not the distribution is asymmetric. It is usually reasonable to presume that the possibility of the benchmark forecast to underestimate or overestimate the true value is not identical. When underestimation is more likely, the skewness is negative, and vice versa. The three key variables, which completely determine a TPN distribution, could be derived from the common mode and two standard deviations of the two component normal distributions of the TPN distribution, making the calculation practically easy.<sup>2</sup> The ease of calculation is the biggest advantage of the TPN distribution among all the candidate distributions which may fit the density of GDP or inflation forecast. In fact, the choice of TPN distribution to describe the density forecast is atheoretical, any distribution that allows skewness can actually be used to fit the density. However, other distributions that allow skewness, such as skewed-normal or skewed-t distributions, are more involved to solve and fail to increase forecast power in practice.<sup>3</sup> Moreover, the conventional null hypothesis of normality for the two component normal distributions of a TPN distribution would be unlikely to be rejected as long as the degree of asymmetry in the TPN distribution applies in practice, which is usually the case in empirical applications (Wallis 1999). As a result, in this paper, I use TPN distribution to fit the density forecast and build fan charts upon the estimation results of corresponding TPN distributions.

# 4 *H. Yu*

The fan chart, as a graphical representation of the density forecast, uses different forms of lines and shades to demonstrate the bands of different percentiles (say, for example, the 10th, 20th, 30th... and 90th percentiles) of the estimated probability distribution. As forecast uncertainty increases over time, the bands for the same percentiles become wider and spread out, making the graphical representation look like a wider fan over time. This is why such representation is named as 'fan chart'. Other possible methods to display the uncertainty situation of a forecast distribution include, for example, showing the complete distribution density function for every time spot (eight graphs for eight quarters, therefore), and calculating the probability of the outcome of the target variable falling in all possible intervals (represented in the form of a table).<sup>4</sup> Compared with other methods, fan chart has several prominent advantages, including the ability to highlight the baseline forecast, indicate the level of uncertainty surrounding the baseline forecast, and show the balance of risks, according to Elekdag and Kannan (2009). More importantly, the producing procedure of fan chart can reflect the forecast process and help people better understand the underlying principles of the forecaster. As Britton, Fisher and Whitley (1998: 36) pointed out, the fan charts produced by the Bank of England for its inflation forecasts help 'to make it clear that monetary policy is about making decisions in an uncertain world, and that the MPC does not pretend to know with certainty the exact rate of inflation in two years' time'.

As with other forecasting methodologies, the fan chart approach itself has also been constantly improved. In early applications, mainly subjective factors, such as surveyed forecast results and professional opinions, were utilized to create fan charts. Recently, Elekdag and Kannan (2009) made the first effort to incorporate some objective marketbased information, such as option trades and market-determined oil prices, into the fan chart building process. On the other hand, the increasing use of fan chart approach in forecasts has also stimulated interest in evaluating its forecasting performance. Because a fan chart is merely a representation of estimated confidence intervals of an interval or density forecast, to evaluate the performance of a fan chart is actually equivalent to evaluate confidence intervals the fan chart displays. Wallis (2003) developed a small-sample likelihood framework evaluation approach based on Christoffersen's (1998) interval forecast assessment, and then applied it to compare the forecasting performances of fan charts and the US Survey of Professional Forecasters. Elder (2005) attempted to build a synthetic and elaborate assessment system for MPC's fan chart application. Recently, Dowd (2008) designed a more precise method to evaluate the forecasting performance of fan charts for GDP growth rate density forecasts.

Before heading to the estimation approaches, one point must be noted here. There is a big difference between the density distribution for Chinese GDP growth rate forecasts and the density forecast for Chinese GDP growth rate and the two should not be confused. The latter is the purpose of this paper while the former is merely a distribution produced by piling up the existing GDP growth rate forecasts. Since no formal density forecast for Chinese GDP growth rate is released so far as I know, my estimation in this paper makes a contribution in this field. However, I have to point out that my calculation in this paper is still based on currently available point forecasts (in the applications of this paper, the quarterly forecasts surveyed by Langrun Forecast project are utilized as the solo data source, which will be introduced in detail in the next subsection).<sup>5</sup> In this case, I am actually standing on the shoulders of prominent forecasters.

Next, we briefly introduce two approaches to estimate the uncertainty intervals and construct fan charts, namely the direct and indirect approaches. The difference between the two lies in how overall uncertainty and the balance of risks for the GDP growth rate forecasts are estimated.

# 2.1. Direct estimation approach

First, we introduce the direct estimation approach. As a straightforward approach to access the risk surrounding GDP growth rate forecasts, we simply focus on the historical and current exercises of GDP growth rate forecasts. Because forecast variance reflects a divergence of attitude toward overall uncertainty of the GDP growth rate among forecasters, we can infer the change of forecasters' attitudes toward uncertainty by comparing the variance of current and historical forecasts. If we assume that the forecasting technology employed by the forecasters, including the techniques to evaluate uncertainty, remain relatively stable over the comparison period, the possible forecast error could be evaluated through comparing the variance of current forecasts and historical ones.

The main procedures of the direct estimation approach are as follows:

- (1) Calculate the variance of current GDP growth rate forecasts and the corresponding shifting historical average levels (of eight continuous periods prior to the quarter). The ratio of the two (latter divided by the former) serves as the amplification/dampening parameter.
- (2) Multiply this amplification/dampening parameter by the historical forecast error variance (also for eight previous continuous periods) of GDP growth rate forecasts to estimate the variance of assumed TPN distribution for the GDP growth rate forecast.
- (3) Calculate skewness of current GDP growth rate forecasts as the skewness of assumed TPN distribution for the GDP growth rate forecast. Furthermore, the mode of the TPN distribution is equal to the baseline forecast value. After estimating the three key variables, the TPN distribution of GDP growth rate could be completely estimated. Then, fan charts based on this estimated TPN distribution can be produced. See Appendix B for more details.

Because the TPN distribution for GDP growth rate forecasts is determined entirely from GDP growth rate forecast series, this approach is named as the direct estimation approach. The greatest advantage of this approach is its low data requirement: as long as historical GDP growth rate forecasts are available, this approach can be utilized to estimate the TPN distribution and produce fan charts of GDP growth rate forecasts.

#### 2.2. Indirect estimation approach

Although simple and straightforward, the direct estimation approach is still not satisfactory to generate density forecasts by combining the historical forecasts of GDP growth rate. The reasons are as follows. First of all, relying on information gauged only from GDP growth rate forecasts might neglect the uncertainties that originate from different aspects of the economy. Besides, as Chancellor (2010) points out, because the GDP growth rate has already become a political symbol for the Chinese government rather than a pure economic index, Chinese GDP growth rate forecasts could be potentially affected by the government's strong willingness to maintain steady and fast growth. Given these potential problems, we have to explore further to find the underlying causes of the overall uncertainty ultimately reflected by fan charts, and the benchmark GDP growth rate forecast should also be revised accordingly when necessary.

Concretely speaking, the procedure for indirect estimation is similar to the 'risk factors' method introduced by Elekdag and Kannan (2009). Our methodology can be briefly introduced as follows.

- 6 Н. Үи
  - (1) First, the underlying causes of overall uncertainty should be determined. Following Elekdag and Kannan's (2009) terminology, we call these causes 'risk factors'. The impact of each 'risk factor' on overall uncertainty should be estimated through a regression of GDP growth rates on the possible causes.
  - (2) Following Blix and Sellin (1998), for simplicity, we assume a linear relationship between the skewness of selected 'risk factors' and that of GDP growth, or

$$\gamma_Y = \sum_i \beta_i \gamma_i \tag{1}$$

where  $\gamma_Y$  and  $\gamma_i$  represent the skewness of the forecasted GDP growth rate and risk factor *i*, respectively.  $\beta_i$  is the weighting coefficient of risk factor *i*, which is actually the coefficient of the corresponding 'risk factor' obtained in step 1.

(3) The same as in direct estimation, the forecast error of future GDP growth rate is also assumed to be proportional to its historical counterpart:

$$\sigma_Y^2 = \phi \overline{\sigma}_Y^2 \tag{2}$$

where  $\sigma_Y^2$  is the measurement of the variance of assumed TPN distribution, which represents the estimated forecast error of the GDP growth rate, while  $\overline{\sigma}_Y^2$  is the shifting average of historical forecast error variance for the previous eight periods.

(4) Parameter  $\phi$  is the amplification/dampening parameter that reflects how much the uncertainty level is estimated to be enlarged compared with the historical average level. This parameter is calculated by the ratio of the weighted sum of uncertainty generated by the chosen 'risk factors' over the investigated period to their historical average level:<sup>6</sup>

$$\phi = \frac{\sum_{i} (\beta_i \sigma_i)^2}{\sum_{i} (\beta_i \overline{\sigma}_i)^2}$$
(3)

where  $\beta_i$  has the same meaning as in step 2 and is utilized as the weight of risk factor *i*.  $\sigma_i$  and  $\overline{\sigma_i}$  represent forecast deviation over the investigated period and its corresponding historical average level for risk factor *i*, respectively. In practice, once again, we take the shifting average for eight periods of variances of 'risk factor' *i* as the measurement as the historical average level  $\overline{\sigma_i}^2$ .

#### 3. Building fan charts for Chinese GDP growth rate forecast and its evaluation

Based on the empirical methodology developed so far, in this section, we carry out an empirical estimation to estimate the density forecast for the Chinese economic growth rate and produce fan charts correspondingly. The organization of this section is as follows. In the first subsection, we introduce and interpret the data set used for our empirical practice. Afterwards, we estimate the TPN distribution and produce fan charts accordingly for Chinese GDP growth rates by direct and indirect estimation approaches, respectively. Finally, we evaluate the performance of produced fan charts according to Wallis' (2003) small-sample interval forecast tests.

### 3.1. The data source: 'Langrun Forecast' project

There are two types of data sources that can be used to produce fan charts: surveyed forecasts and market information. The former is more traditional and still widely used by many researchers and institutes, such as central banks like the Bank of England, which use fan charts as intuitive devices to show inflation uncertainty. The latter is relatively new and still in an experimental phase.<sup>7</sup> Due to data availability and some practical considerations discussed later, we focus on the traditional data source – surveyed forecasts – in this paper in the initial attempt to make the density forecast and correspondingly construct fan charts for the Chinese GDP growth rate.

Our data source is from the 'Langrun Forecast' project, a quarterly forecast survey, founded and conducted by the China Center for Economic Research (CCER). This project is named after the Langrun Garden at Peking University, where CCER is located. Since May 2007, more than 20 academic and financial institutes have been invited to participate in this project. At the end of each quarter, the participating institutes provide their quarterly point forecasts of GDP growth rate and other important quarterly indexes for the following quarter on a year-to-year basis.<sup>8</sup>

The reasons for choosing the 'Langrun Forecast' project as our benchmark data source are twofold. First, this survey project provides adequate quarterly forecasts from various institutions, and the sample period is long enough (dating back from the third quarter of 2005) that we can collect enough forecasting data for estimation. Second, because all the surveyed forecasts reported in this project are offered by prestigious academic institutes or leading commercial organizations, the reliability and continuity of their forecasts are guaranteed, which is a critical prerequisite for utilizing our estimation approaches.

Several specific features and limitations of our data need to be clarified and are worth discussing. One inherent data problem is that most quarterly forecasted growth rate series, such as GDP and industrial value-added series, are only on a year-to-year basis rather than the currently prevailing period-to-period form adopted by many countries. Additionally, because the 'Langrun Forecast' project is essentially a short-run forecast (one quarter ahead), the fan charts based on this project could merely reflect the short-run uncertainty situation for the coming quarter. Intuitively, this means our fan charts could 'fan out' only one quarter, making it unable to demonstrate the development of uncertainty over time in a single fan chart. For comparison, the inflation fan charts of the Bank of England 'fan out' much further along the time axis for eight quarters, because the forecast horizon for inflation forecast is eight quarters by MPC.

Starting from the second exercise of the 'Langrun Forecast' project, a comprehensive weighted forecast based on all the individual forecasts provided by participating institutes has been reported. The principle to calculate the weighted forecast is to assign higher weights to the institutes whose historical forecasts turn out to be closer to the actual outcomes. Concretely speaking, in each period, all participating institutes are at first ranked in the reverse order according to the performance of their historical forecasts. That means that the better the performance of the historical forecasts of an institute is, the lower rank it may receive. Then the rank of a certain institute is directly used as its weighting in calculating the comprehensive weighted forecast.<sup>9</sup> The weighted forecast series and the corresponding actual GDP growth rates are shown in Figure 1. We also add the historical GDP growth rate forecast', as a comparison.

As shown in Figure 1, for most of the time of the sample period, the GDP growth rate forecast of CCER and the comprehensive weighted forecast cohere with the actual



Figure 1. Historical comprehensive weighted forecast and CCER's forecast series of GDP growth rates compared with the actual growth rates.

Source: Published Results of 'Langrun Forecast' and National Bureau of Statistics.

outcome. The only exception is between the second quarter of 2008 and the third quarter of 2009, when the global economy was undergoing a serious recession. Furthermore, for precise comparisons, we calculate the U-statistics and root mean square errors (RMSE) of 14 participating institutes who took part in all forecasting exercises and the comprehensive weighted forecasts over the sample period. The results are reported in Table 1.

According to Theil's U-statistics and the RMSEs listed in Table 1 for the sample period, the performance of the comprehensive weighted forecast of the 'Langrun Forecast' project is among the best of all forecasts. Considering that the weighted forecast was first time reported in the fourth quarter of 2005 while all participating institutes were kept anonymous till 2008, we simply chose this weighted forecast as our baseline (central projection) forecast series. From the perspective of statistics, the weighted forecast series is considered to be the most likely actual outcome. Therefore, we could directly take this series as the mode of TPN distribution function.

Table 1. An ex-post comparison of the performances for some participating institutes over the sample period from the third quarter of 2008 to the third quarter of 2010.

Institutes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Theil's	0.052	0.053	0.053	0.055	0.060	0.062	0.053	0.061	0.053	0.066	0.053	0.041	0.055	0.046	0.050
RMSE	0.930	1.027	1.024	1.077	1.163	1.133	1.011	1.160	1.021	1.290	1.025	0.805	1.027	0.896	0.949

Data Source: Historical 'Langrun Forecast' results and authors' calculation.

Note: The participating institutes are: 1: CCER; 2: Department of Economic Forecasting State Information Center; 3: Guotai Junan Securities; 4: Citibank; 5: HSBC; 6: Merrill Lynch; 7: UBS; 8: Institute of Quantitative & Technical Economics Chinese Academy of Social Sciences; 9: Shenyin Wanguo Securities; 10: China Galaxy Securities; 11: China International Capital Corporation Limited; 12: China Securities Co.; 13: CITIC Securities; 14: Bank of China International Securities Essences Securities; 15: Comprehensive Weighted Forecast.

# 3.2. Direct estimation result

In this subsection, we turn to the concrete, practical estimation result according to direct estimation approach. Because only GDP growth rate forecast series matter to the direct estimation approach, we focus on this series and ignore the forecasts for all other indexes in this section. First, we compare historical variance of GDP growth rate forecasts of participating institutes, which reflect the different attitudes of them toward the uncertainty surrounding GDP growth rate, and the actual outcomes of historical GDP growth rates in Figure 2.

Figure 2 provides a new illustration for Prati and Sbracia's (2002) finding: the degree of divergence between different forecasters could potentially indicate the steadiness level of an economy, because high dispersion is usually an indicator of a forthcoming financial crisis. For example, the surge in the forecast variance in the first three quarters of 2007 likely reflects the worry about the Chinese economy's overheating at that time. It was soon followed by a continuous economic slowdown until the first quarter of 2009. Another even higher peak of the variance appears in the first quarter of 2010, after a beyond-expectation robust and quick recovery of the Chinese economy following the toughest period of the international financial crisis. This variance hike is likely to reflect the concerns of economists about the continuity of China's robust recovery in the post-crisis involved world economic environment. Not surprisingly, afterwards, the Chinese GDP growth rate turned out to skid to around 10% from the four-year highest record of 11.9% in the first quarter of 2010.

The skewness of GDP growth rate forecasts, another crucial parameter of TPN distribution, is depicted in Figure 3 together with the actual GDP growth rate outcomes. Note that because skewness reflects the forecasters' expectation for the balance of risks, a positive skewness implies an expectation that GDP growth rate may turn upward, and vice versa.

After brief qualitative analyses of the variance and skewness of GDP growth rate forecasts, now we turn to the empirical estimation of TPN distribution by the direct



Figure 2. The historical dispersion of forecasted GDP growth rates. Source: Historical 'Langrun Forecast' results and authors' calculation. Deviation is calculated as the forecast variance.



Figure 3. The historical skewness of forecasted GDP growth rates. Source: Forecast results from 'Langrun Forecast' project and authors' calculation.

Crucial parameters	Amplification/ dampening parameter for variance	Estimated deviation of GDP growth rate forecasts	Estimated skewness of GDP growth rate forecasts 0.66 Upper limit	
	3.58	1.80		
Confidence intervals	Level of confidence (%)	Lower limit		
	50 70 90	10.48 9.99 9.16	12.88 13.67 15.01	

Table 2. Crucial parameters of TPN distribution and corresponding confidence intervals for the first quarter of 2010 by direct estimation method.

Source: 'Langrun Forecast' results and authors' calculations.

estimation approach. We use the fan chart for the second quarter of 2010 as the initial illustration. According to the procedures described in Appendix A and B, the values of the three determinant variables are calculated and shown in Table 2, and the corresponding fan chart is depicted in Figure 4.

In Figure 4, different shades represent the ranges of different confidence intervals. The inner darkest band, covering the mode value (which is equal to benchmark forecast level and represented by the blue line), covers 50% of the probability. The adjacent lighter band covers 70% of the probability, and the outer lightest band covers 90% of the probability. The actual outcome of the GDP growth rate is identified by another line located within the 50% probability band. With this fan chart, we can observe the expected total uncertainty (the width of fan chart) and the risk of balance (the relative height of the bands at both sides of the mode level) of the GDP growth rate forecast for the first quarter of 2010 is shown intuitively.



Figure 4. Fan chart of Chinese GDP growth forecast for the first quarter of 2010 by direct estimation approach.

Source: 'Langrun Forecast' project and authors' calculations. This graph is produced with the help of the spreadsheet software for constructing fan chart designed by Elekdag and Kannan (2009).

As shown in figure 1, the 'Langrun Forecast' forecasts turned out to be inaccurate during the global economic crisis. Therefore, we construct the fan charts for some quarters during this period (from the second quarter of 2008 to the first quarter of 2009) and compare their forecast performance with that of traditional point forecasts. These fan charts are integrated into Figure 5.

During the chosen periods, the evidence supporting the fan chart as a better uncertainty communication device is, at most, mixed. In only two of the four fan charts (for the first quarter of 2008 in the upper left panel and the first quarter of 2009 in the lower right panel) the actual GDP growth rate falls within the 50% probability band, and the balance of risks is correctly predicted because the change direction of GDP growth rate in the following quarter is consistent with the skewness of the forecast reflected by the fan chart.<sup>10</sup> However, in the other two quarters, the performance of the predictions, shown by the fan charts, are not so satisfactory. The fan chart in the upper right panel for the third quarter of 2008 was abnormal as the mode value (baseline forecast) does not lie within any given band of the fan chart. In the fourth quarter of 2008, the Chinese economy suffered the most severe slowdown in a decade. The GDP growth rate dove to 6.8% from 9% of the preceding quarter. Such a rapid decrease in the GDP growth rate was entirely unexpected and was not even covered by the outer 90% probability band of the fan chart for that quarter, making the expected balance of risk incorrect.



Figure 5. Fan charts constructed by direct estimation approach for several chosen periods during global economy crisis.

Source: 'Langrun Forecast' results and authors' calculations.

There are several reasons for the unsatisfactory performance of the fan charts constructed using the direct approach. An objective reason is that, during the global economic crisis, Chinese GDP growth rates experienced an unexpected, serious slowdown. Because the variance of the TPN distribution is calculated with historical performance of forecasts as the reference, the unexpected worsening economic environment causes the performance of forecast during the crisis to be considerably worse than previous ones, making the estimated variance of the TPN distribution rather large, which would in turn cause the density forecast result fails to be inaccurate. The more important reason is likely due to the way that uncertainty situation is evaluated by the direct estimation approach. As discussed earlier, the information explored from the GDP growth rate forecast series is usually insufficient for a complete description of the total uncertainty situation because the specific causes of overall uncertainty are not considered in this approach. For this reason, we should further develop the indirect estimation approach, which is the main topic of next section.

# 3.3. Indirect estimation result

As mentioned previously, the main difference between the direct and indirect estimation methods lies in the way how uncertainty surrounding the GDP growth rate is estimated. According to the indirect estimation approach, the variance and skewness of the TNP

distribution are estimated from chosen 'risk factors' instead of directly from GDP growth rate forecasts.

To implement indirect estimation, the 'risk factors' must first be identified. From these 'risk factors,' the overall uncertainty for the entire economy is derived. In Elekdag and Kannan's (2009) paper, three sets of variables are chosen as 'risk factors', including oil price risk and inflation risk. Although these choices are reasonable for estimating the uncertainty surrounding global GDP growth rate, they might not be suitable for investigating the Chinese economy.<sup>11</sup> Because the Chinese financial market is still far from mature and has undergone frequent interventions by Chinese government, China's financial condition is not a suitable candidate for our 'risk factors' because it does not reflect fully the expectation of market participants but does reflect, to a great extent, the will of the Chinese government. Furthermore, given the Chinese government's firm control over the domestic price of energy and important natural resources, the international oil price risk also might not be a proper candidate for analysis of Chinese economy uncertainty.<sup>12</sup>

The 'risk factors' for the Chinese economy should be chosen according to specific features of the Chinese economy. As a starting point, we turn our look to some famous economic activity indexes concerning Chinese economy, as uncertainty situation should be properly considered when these indexes are made. We selected two indexes to evaluate domestic economic circumstances: The Leading Index for Macroeconomy Economy Climate (LMCI) and the Conference Board Coincident Indicator for China (CBCI). The LMCI is a predictive indicator released by the National Bureau of Statistics to describe the overall business environment and help define the coming business cycle's turning point. The CBCI is a comprehensive indicator produced by The Conference Board that includes information from industries and enterprises, which could be considered as a measurement of the activity of China's real economy.<sup>13</sup> Since China joined the World Trade Organization (WTO) in 1999, Chinese economic growth has increasingly benefited from globalization and depended more heavily on foreign trade.<sup>14</sup> Given the close relationship between the Chinese economy and the global economy, the uncertainty embedded in global economic growth should also be an important source of uncertainty over Chinese economy. For this reason, we take Standard & Poor's Global BMI indicator (SPGB) as another 'risk factor'. The SPGB is a comprehensive indicator, which is basically estimated by artificially composing the performance of worldwide stock market, including both developed and emerging markets. Because stock market performance is generally acknowledged as the barometer of economic health and also because the stock market is essentially forwardlooking and reflects the expectation of the economic conditions to some extent, the SPGB could well reflect the uncertainty of the global economic growth.

Next, we determine the impact of the chosen 'risk factors' on the total uncertainty surrounding GDP growth rates. For this purpose, we simply carry out an ordinary least squares (OLS) regression of the GDP growth rate against the growth rates of chosen 'risk factors', where all regressors are standardized in advance. In a standardized form regression, the coefficient of an explanatory variable (so-called beta coefficient) measures how much the dependent variable's deviation changes when that explanatory variable changes by a standard deviation, holding other conditions constant. Because deviation/variance is a common approximate measurement for uncertainty, the beta coefficients of this standardized form regression can be directly used to gauge the impact of 'risk factors' on the overall uncertainty surrounding GDP growth rates. Concretely speaking, these coefficients are taken as the value of  $\beta_i$ s in Equation (3). In order to deal with potential autocorrelation problem in time series, the first-order lagged GDP growth rate is introduced as another explanatory variable. The sample period is from the first quarter of 2001 and the first quarter of 2010, a total of 37 quarters. The regression results are shown in Table 3.

### 14 *H. Yu*

Source of risk (corresponding risk factor)	LMCI	SPGB	CBCI	GDP_1 (first-order lagged GDP growth rate to control for potential auto-correlation)
Coefficient	0.151	0.261	0.288	0.427
t-statistic	1.76	2.00	2.08	3.34
p-value	0.087	0.053	0.045	0.002

Table 3. Regression result for estimating the impact of 'risk factors' to the total uncertainty surrounding GDP growth rate.

F-statistic for Breusch-Godfrey LM tests: 1st-order: 3.13 (p = 0.10); 4th-order: 1.47 (p = 0.24) Jarque-Bera statistic for normality test: 4.45 (p = 0.11) F-statistic for Ramsey RESET test: 0.55 (p = 0.71)

Source: 'Langrun Forecast' results and authors' calculations. Note:

1.All regressors are in the form of standardized growth rates. All explanatory variable series are original monthly data, and the quarterly observations used in this regression are calculated by taking a simple arithmetic average value of the three months in the same quarter.

2. The adjusted  $R^2$  of this equation is 0.87. The F-statistic for the hypothesis that all coefficients are jointly equal to zero is 55.4 (p-value is lower than 0.001).

The adjusted  $R^2$  is 0.87. Because the adjusted  $R^2$  of the simple regression of the standardized GDP growth rate on its own first-order lagged series is 0.68, approximately 20% of fluctuation in the variance of the Chinese GDP growth rate could be explained by these 'risk factors'.<sup>15</sup> Along with routine t-tests, three additional test results are reported in Table 3. The Breusch-Godfrey LM test results show that the possibility for autocorrelation problems is very slim (no higher than the 10% critical level). The normality test suggests that we have enough confidence (also higher than the 10% confidence level) to believe that the regression residuals are well modeled by normality distribution. The Ramsey RESET test result also suggests that the regression equation is correctly specified. Because the coefficient of CBCI is larger than those of the other two, the change of a standard deviation of CBCI would have greater impact on overall uncertainty than other two 'risk factors'. Notice that CBCI represents the prosperity condition of firms and enterprises, this means that the uncertainty the Chinese firms and enterprises face may convert to uncertainty over the whole economy. This is not a surprising finding, as manufacturing industries have already played a very important role in the Chinese economy. We also observe that the uncertainty originated from the global economy environment is another important source of overall uncertainty, since the coefficient of SPGB is only slightly lower than that of CBCI, which is also accordance with our expectation.

As discussed previously and shown in Figure 1 and Figure 5, the 'Langrun Forecast' GDP growth rates turned out to be too optimistic during the global economic crisis in 2008 and early 2009, making the direct estimation results not satisfactory. According to the indirect estimation approach, the overall uncertainty of GDP growth rate forecast could, to a great extent, be determined by the uncertainties of the three chosen 'risk factors'; therefore, we can revise the benchmark GDP growth rate forecast for this period accordingly. For this purpose, we simply regress the actual GDP growth rate outcomes against the chosen 'risk factors' growth rates and the first-order lagged GDP growth rate for the entire sample period (from first quarter of 2001 to the first quarter of 2010). The benchmark forecasted GDP growth rate forecasts during the economic crisis period should thus be rebuilt based on the regression result:

Forecasted GDP growth rate<sub>t</sub> =  $4.565 + 0.375 * \text{Actual GDP growth rate_{t-1}} + 0.076 * \text{LMCI growth rate}_t + 0.018 * \text{SPGB growth rate}_t + 0.148 * \text{CBCI growth rate}_t$ , where t

Quarter	Revised forecasted growth rate	'Langrun Forecast' growth rate	Actual GDP growth rate	Forecast error of revised rate	Forecast error of 'Langrun Forecast' rate
2008Q2	10.0	10.5	10.2	$0.2 \\ -0.1 \\ 0.7$	-0.3
2008Q3	9.1	10	9		-1
2008Q4	7.5	9.1	6.8	-0.7	-2.3 -0.5
2009Q1	6.6	6.6	6.1	-0.5	

Table 4. Comparison of revised GDP growth rate forecasts and 'Langrun Forecast' predictions for the period between 2008Q2 and 2009Q1.

Source: 'Langrun Forecast' results and authors' calculations.

represents year t. The adjusted  $R^2$  is 0.88 and the F statistic value for the hypothesis that all coefficients are jointly equal to zero is significant at 0.001.

The renewed forecasted benchmark GDP growth rate forecasts are shown in Table 4. 'Langrun Forecast' weighted average forecasts and actual GDP growth rates are also presented in the same table for comparison.

It is obvious that the revised GDP growth rate forecasts in the chosen period are closer to the actual growth rates, which confirms our belief that the indirect estimation approach could lead to better predictions of GDP growth rate, because the chosen 'risk factors' clearly explain the causes of overall uncertainty surrounding GDP growth rate. According to Equation (3), the variance and skewness of the forecasts for each factor in a single quarter contribute to the variance and skewness of TPN distribution, respectively. Moreover, the shifting average of the estimated variance for eight consecutive quarters prior to that quarter is used as the baseline measure of forecast uncertainty, or  $\bar{\sigma}_Y^2$  in Equation (2). After calculating Equations (1)–(3), the TPN distribution for GDP growth rates can be determined and then fan charts could be produced accordingly. Again, the fan chart for the first quarter of 2010 is shown in Figure 6 as a primary illustration, the key parameters of which are presented in Table 5.



Figure 6. Fan chart of Chinese GDP growth forecast for the first quarter of 2010 by indirect estimation approach.

Source: 'Langrun Forecast' results and authors' calculations. This graph is produced with the help of the spreadsheet software designed according to Elekdag and Kannan (2009).

### 16 *H. Yu*

Key influential parameters	Weights (β coefficients) for calculating variance and skewness	Deviation of the estimated deviation of the variable's forecasts	Skewness of the estimated skewness of the variable's forecasts		
LMCI	0.15	0.36	1.15		
SPGB	0.26	5.71	1.57		
CBCI	0.29	2.06	-1.44		
GDP growth rate	N/A	0.42	0.17		
Confidence intervals	Level of confidence (%)	Lower limit	Upper limit		
	50	11.19	11.75		
	70	11.08	11.94		
	90	10.89	12.25		

Table 5. Fan chart of Chinese GDP growth forecast for the first quarter of 2010 by indirect estimation approach.

Source: 'Langrun Forecast' results and authors' calculations.

Compared to the fan chart in Figure 4 by direct estimation approach, this fan chart performs better, because it is more positively skewed and narrower, with the actual GDP growth rate lying just in the 70% probability band. More significant improvements can be seen when we turn to the specially chosen periods between the second quarter of 2008 and the first quarter of 2009, which are shown in Figure 7 below.

For these four quarters, the actual outcome of GDP growth rates fell within the bands of each fan chart (at least in the 90% probability bands). In the fourth quarter of 2008, when the Chinese economy underwent an unprecedented slowdown, Chinese GDP grew only at 6.8% compared to the same quarter of the previous year, which was its worst quarterly performance since 1999. The actual GDP growth rate in that quarter is barely covered by





Source: 'Langrun Forecast' results and authors' calculations.

the outer 90% probability band. In the other three quarters, however, the 50% probability intervals successfully predicted the actual output growth rates. Furthermore, the forecasted balance of risk turns out to be consistent with the actual outcomes in all four quarters. As a result, compared to the fan charts produced by the direct estimation approach in Figure 5, the indirect estimation approach indeed improve the performance of density forecasts and, therefore, the corresponding fan charts.

#### 3.4.. Evaluating interval forecasts and fan charts

So far, we have constructed fan charts according to both direct and indirect estimation approaches, which naturally leads to a question that all kinds of forecasts may face: how accurate is our estimated TNP distribution and the correspondingly constructed fan charts? In other words, how close are our forecasts to the actual outcomes?<sup>16</sup>

To answer this question, we need to evaluate our constructed fan charts. As mentioned previously, the fan charts are, in essence, an intuitive representative of confidence intervals derived from assumed TPN distribution for GDP growth rate forecasts. As a result, the density forecast tests could be used to precisely assess these estimated TPN distributions, such as Edler (2005) and Dowd (2008) have discussed. However, because of lacking data, the density forecast tests cannot be effectively conducted for the fan charts produced over the sample period.<sup>17</sup> In this case, small-sample interval forecast tests may be more proper for our purpose.

Therefore, we employ Wallis' (2003) chi-square test method for small samples, which is a concrete and extensive application of Christoffersen's (1998) interval test methodology. The Wallis (2003) test is in fact a Markov chain approach to evaluate both the coverage and independence of an interval forecast sequence. The basic idea of this test is that if the 'hit frequency' of constructed confidence intervals (the chance of outcomes falling into a certain confidence interval) is approximately consistent with the *ex post* confidence level (coverage) and the bulks of zeros and ones do not appear in clusters (independence), then the interval forecast should be considered accurate. More technical details could be found in Wallis' (2003) paper. In Figure 8, the estimated interval forecasts of GDP growth rate series by two methods as well as the actual outcomes are depicted intuitively.

Following Wallis' (2003) small-sample methodology, we conducted a likelihood ratio test and Pearson's chi-square test for the joint assumption of coverage and independence at 50%, 70% and 90% confidence intervals. The results are represented in Table 6.

The p-values reported in the parentheses next to the test statistics indicate how much confidence we have in our null hypotheses that the coverage of fan chart bonds is both correct and independent. The smaller the p-value is, the more likely that the null hypothesis may be rejected. As shown in Table 6, except for Person's chi-square test of the 90% confidence level from the direct estimation approach, the null hypothesis for both approaches cannot be rejected at the 10% significant level because the corresponding p-values are significantly higher than the 10% threshold. The relatively low p-values for both tests of the direct estimation approach for the 90% confidence bands for direct estimation approach are inaccurate. Generally, the overall uncertainty estimated through direct estimation tends to be inaccurate, especially for the outer band of the fan chart which corresponds to the confidence intervals of higher confidence level.<sup>18</sup> Another observation from Figure 8 is the expected overall uncertainty, which is expressed by the width of the fan chart, tends to follow a strong anti-cyclical pattern. This is likely because the expected overall uncertainty associated with the GDP growth rate forecast tends to be higher during a recession, reflecting the fact that forecasters are



Figure 8. Chinese GDP growth interval forecasts and realizations.

Note: Along the vertical direction, round brackets indicate the confidence intervals calculated by direct estimation, while the square parentheses represent the confidence intervals from indirect estimation. All intervals are 50% confidence intervals. Point forecast results of baseline forecasts, or central projection forecasts, are denoted with round dots, while the revised baseline forecasts are marked with solid squares. By comparison, actual GDP growth rates are marked with small triangles.

more likely to differ widely from each other facing a changing economic environment in the economy downturn.<sup>19</sup>

As shown in Figure 8, in six out of 10 times for direct estimation results, compared with five out of 10 times for indirect estimation results. Although the covering ratios are similar, the advantage of the fan chart methodology is better illustrated when we compare the ability of revealing the risk balance of the GDP growth rate forecast. This is represented by skewness and displayed intuitively by the relative lengths of the two tails of the fan chart: if the upper tail of the central prediction is longer, the forecasted risk is upward, and vice versa. In Figure 8, in five out of 10 estimations by direct approach, the skewness of fan chart turns out to reflect the changing direction of GDP growth rate correctly; while

Estimation method	Direct	estimation ap	proach	Indirect estimation approach			
Confidence level	50%	70%	90%	50%	70%	90%	
Likelihood ratio	1.019	0.047	3.623	1.248	NA	NA	
test (p-value)	(0.600)	(0.977)	(0.163)	(0.534)			
Person's	1.000	0.048	5.444	1.200	2.429	0.167	
chi-square test (p-value)	(0.607)	(0.976)	(0.066)*	(0.549)	(0.297)	(0.920)	

Table 6. The results of joint test of coverage and independence for main confidence intervals.

Source: 'Langrun Forecast' project and author's calculation.

Note: the p-values are shown in the parentheses behind the corresponding test statistics. \*: p-value is lower than 10% significant level. \*\*\*; p-value is lower than 5% significant level. \*\*\*; p-value is lower than 1% significant level.

for indirect estimation results, the number of correct predictions increases to eight out of 10 times.

## 4. Conclusion

In this paper, we make an initial attempt to estimate the density forecast of Chinese economic growth and use fan charts as the visible communication device to demonstrate the forecast results. The estimation methodology is adapted from Elekdag and Kannan's (2009) method, given the specific characteristics of the Chinese economy. The fan chart is employed as the representation of the forecast results, so as to intuitively reflect uncertainty condition in the forecasting process. As a first attempt and due to data availability, we focused on the one-quarter-ahead 'Langrun Forecast' project and several complementary indicators for the empirical application of our fan chart approaches. The empirical results show that our fan charts do convey more valuable information about the uncertainty situation for Chinese GDP growth. The advantages of our fan chart approaches are more obvious during the global financial crisis, when traditional point forecasts appeared to produce overly optimistic predictions. Comparatively, indirect estimation approach performs better because it reveals the uncertainty situation surrounding GDP growth rate with more details. The direct estimation approach, although outperformed by indirect approach, has lower requirement for data.

Our estimation approaches could be further developed. The primary idea is to introduce more objective elements into the estimation. These elements include market-based factors, which, according to Elekdag and Kannan (2009), refer to the information obtained from market behaviors, such as call option prices in financial markets and international oil option prices because the market is considered as forward-looking. However, because similar markets in China do not exist or are still far from mature, we simply ignore such factors in our analysis. Nevertheless, because the government regulations are doomed to be relaxed and finally abandoned as the Chinese economy becomes more mature, open and international, such market-based information might be sufficient and available to be used in fan chart studies in the future.

#### Notes

- 1. More precisely, Clements and Hendry (1998) elaborately categorized five sources of modelbased forecast error, which can be emerge into these two main aspects as concluded by Hendry and Ericsson (2001).
- 2. See Appendix A for more details of the relation between a TPN distribution and the two component normal distributions.
- 3. As for the skewed normal distribution, for example, except for mean and variance, two other parameters (a location parameter and a scale parameter) have to be estimated in order to determine the distribution completely. Comparatively, only three parameters (mode and two variances for the two component normal distributions) need to be estimated to determine a TPN distribution.
- 4. Blix and Sellin (1999) have applied the three methods mentioned here to display the same inflation density forecasts made by the Riksbank.
- 5. Individual density forecasts could also be aggregated to form comprehensive density forecast according to designed schemes. In a recent paper, Kenny, Kostka and Masera (2011) proposed a framework to investigate the information content of subjective expert density forecasts and corresponding aggregated density forecast using micro data from the ECB's Survey of Professional Forecasters. Due to lack of density forecast results, only point forecasts are used in this paper to estimate the density forecast for Chinese GDP growth rate.
- 6. Compared with the original formula introduced by Elekdag and Kannan (2009), we simply drop the covariance between 'risk factors', not only because of the difficulty to determine their

20 *H. Yu* 

value but also due to the fact that they are simply dropped out when the authors calculated  $\phi$  in practice.

- 7. Because an important feature of market participants is forward-looking, much information for future uncertainty might be exploited from market behaviors. The research in this field is growing rapidly in recent years. For example, Bahra (1997) summarized the early efforts to extract implied risk-neutral probability density functions from option prices, Ält-Sahalia and Duarte (2003) developed a more advanced nonparametric approach for the same purpose.
- The regularly forecasted indices include quarterly growth rates of GDP, consumer price index 8. (CPI), industrial value-added, investment, consumption, export and import, interest rate and exchange rate. The list of participating institutes has grown from 14 members in the first exercise to 22 members in the 24th exercise (forecast for the second quarter of 2011) and might probably continue to increase. Participating institutes of this project (containing those which have once participated but later quitted) include Bank of China International Securities Essences Securities, Bank of Communications, Guotai Junan Securities, Blue Oak Capital, HSBC, BNP Paribas, Industrial and Commercial Bank of China, China Center for Economic Research Peking University, Institute of Quantitative & Technical Economics Chinese Academy of Social Sciences, China Galaxy Securities, Merrill Lynch, China International Capital Corporation Limited, Morgan Stanley, China Merchants Securities, Nomura Securities China Securities Co., Shenyin Wanguo Securities, Citibank, Standard Chartered Bank, CITIC Securities, UBS, Department of Economic Forecasting State Information Center, Unirule Institute of Economics, Essence Securities, Greatwall Securities and China Securities Co. The historical forecasting data could be found at http://www.nsd.edu.cn/cn/list.asp?classid=634# (in Chinese).
- 9. Despite being easy to handle with, this weighting method is fairly rough and has potential to be improved. Plenty of work has been done in this field. For example, Clark and McCracken (2009) use Monte Carlo experiments to decide whether the recursive and rolling schemes or a scalar convex method should be used for combination, especially when linear predictive models are subject to structural change. Hsiao and Zhao (2000) explored the usefulness of opinion surveys with time-series data. For general principal for forecast combination and a brief review of early forecast combination methods, see Palm and Zellner (1992).
- 10. Note that, quantitatively, a negative skew indicates that the tail on the left side of the probability density function is longer than the right side and the bulk of the values (including the median) lie to the right of the mean. A positive skew indicates an exactly opposite situation.
- 11. Elekdag and Kannan (2009) focused on the term spread and the Standard and Poor's (S&P) 500 index in order to estimate the impact of financial conditions on world economy growth. Given the high representativeness of the stock market in United States and high correlation of the financial markets among major developed countries, such financial condition is a good indicator for uncertainty estimation. Considering that major developed countries contribute more than 70% of world GDP but are relatively vulnerable to shocks in oil market, oil price risk should also be highly relevant to the total uncertainty surrounding global economy growth.
- 12. As an illustration of price control, the National Development and Reform Commission (NDRC) has long been regulating the domestic price of petroleum product, and Chinese domestic oil price adjustment is usually lagging the international price fluctuation in time and less so much in scale.
- 13. According to the description of The Conference Board, CBCI is a composite index, the components of which include Value Added of Industrial Production, Retail Sales of Consumer Goods, Electricity Production, Volume of Passenger Traffic and Manufacturing Employment. More information of this indicator could be found at the official website www.conference-board.org.
- 14. The importance of net export to Chinese GDP could be measured by the ratio of total foreign trade value to GDP. This ratio has increased steadily during the last decade, from 36% in 1999 to approximately 50% in 2010.
- 15. The incremental contribution of the 'risk factors' is also tested to be significant. The p-value of F-test that such 'risk factors' has no effect on the fluctuations of GDP growth rate is lower than 0.001.
- 16. Since late 1990s, many scholars have raised increasing doubts about the quality of China's official statistic data, especially the key indicators like GDP and its growth rate, as well as CPI inflation. Some scholars believe that Chinese governments, especially provincial and local governments, tend to exaggerate the GDP level and its growth rate to show off their political achievements. Krugman (2011) even claimed that China's statistic numbers 'are more fictional

than the most boring form of science fiction'. The criticisms of Chinese official statistic quality mainly focus on the inconsistency of GDP growth rate and other key indicators which are believed to relate closely to and move at a similar pace with GDP growth rate, such as energy consumption and railway traffic volumes (Rawski 2001a, b; Sinton 2001). On the other hand, however, some scholars still have cautious faith in China's official statistic data. In a very influential empirical study, Klein and Özmucur (2002) examined some strategic indicators that are suggested by basic social accounting principles and concluded that principal components of these indicators indeed reflect the movement of official estimates of the Chinese economy. They further indicated that, in fact no one knows the correct estimate, and the key point lies in the way how the estimate is calculated, which is true not only to China but the world wide over. Following Klein and Özmucur's (2002) opinion. I leave this controversial problem aside and still use official data as benchmark to evaluate the accuracy of forecasts. Another reason for using official data is the self-adaption of forecasters of Chinese economy. It is reasonable to believe that the inherent sources which may cause measurement error in official GDP growth rates might probably also affect forecasters of Chinese economy in a similar pattern, because the forecasters have to make their forecast results as close to the official data as possible to show their prediction ability when no more reliable statistic data are available.

- 17. Two of the classical methods of testing density forecast, namely the likelihood ratio and Pearson chi-squared tests, are conducted through dividing the range of the variable into several mutually exclusive classes and then comparing the probabilities of actual outcomes that fall into these classes with the theoretical value. By doing so the density forecast tests are essentially degraded to interval forecast tests.
- 18. Notice that our tests here are only for the overall uncertainty of constructed fan charts, leaving the accuracy of forecasted balance of risks untouched.
- 19. For example, during the whole 2009 and the first quarter of 2010, when was the main period of the ongoing world financial crisis, the intervals are obviously wider than other time. The confidence interval by indirect estimation approach in the first quarter of 2010 is considerably wide, reflecting the huge dispersion of forecasts for key influential variables due to the turbulence triggered by the unprecedented GDP growth rate slowdown from 9% in the third quarter of 2009 to 6.8% in the fourth quarter of 2009.
- 20. This appendix is adapted from Box 1 and section II of Elekdag and Kannan's (2009) paper.
- 21. The criterion for real roots of the quadratic Equation (B.1) is  $(4 3/2\pi) \gamma_Y^2 + \sigma_Y^2 \ge 0$ , which can be satisfied in a normal case so we ignore the possibility of complex roots, although it is not totally impossible theoretically.
- 22. Wallis (1999) has discussed another approach for constructing confidence interval, that is, to make equal tail probabilities of the interval: Pr(x < a) = Pr(x > b) = (1-p)/2. Interestingly, he argued that this constructing method is theoretically superior to the shortest interval method we applied because the loss function implied in the former approach is more reasonable than the 'all-or-nothing' loss function implicit in the latter.

#### References

- Ält-Sahalia, Y., and J. Duarte. 2003. Nonparametric option pricing under shape restrictions. *Journal of Econometrics* 116: 9–47.
- Bahra, Bhupinder. 1997. Implied risk-neutral probability density functions from option prices: theory and application. Bank of England Working Paper, No.66.
- Blix, M., and P. Sellin. 1998. Uncertainty bands for inflation forecasts. Sveriges Riksbank Working Paper, No.65.
- Blix, M., and P. Sellin. 1999. Inflation forecasts with uncertainty intervals. *Quarterly Review* 2: 12–28.
- Britton, E., P. Fisher, and J. Whitley. 1998. *The Inflation Report Projections: Understanding the Fan Chart*. Bank of England Quarterly Bulletin, February.
- Celasun Oya, Xavier Debrun, and Jonathan D. Ostry. 2006. Primary surplus behavior and risks to fiscal sustainability in emerging market countries: a 'fan-chart' approach. *IMF Working Paper*, 06/67.
- Chancellor, Edward. 2010. China's red flags. GMO White Paper, March.
- Christoffersen, Peter F. 1998. Evaluating interval forecasts. *International Economic Review* 39, no. 4: 841–62.

- Clark, Todd. E., and Michael W. McCracken 2009. Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review* 50: 363–95.
- Clements, M.P., and D.F. Hendry. 1998. Forecasting economic time series. Cambridge: Cambridge University Press.
- Dowd, Kevin. 2008. The GDP fan charts: an empirical evaluation. *National Institute Economic Review*, No.203, January.
- Elder, Robert. 2005. Assessing the MPC's fan charts. Bank of England Quarterly Bulletin, Autumn.
- Elekdag, Selim, and Prakash Kannan. 2009. Incorporating market information into the construction of the fan chart. IMF Working Paper, August.
- Hendry, F.D., and R.N. Ericsson. 2001. Understanding economic forecast. Cambridge, MA: The MIT Press.
- Hsiao Cheng, and Zhao Zhongyun. 2000. Combining opinion surveys with time-series data to forecast the Japanese economy. *The Japanese Economic Review* 51: 155–69.
- John, S. 1982. The three-parameter two-piece normal family of distributions and its fitting. Communications in Statistics – Theory and Methods 11: 879–85.
- Johnson, N.L., S. Kotz, and N. Balakrishnan. 1994. *Continuous univariate distributions*, Vol. 1. 2nd ed. New York: John Wiley.
- Kenny, Geoff, Thomas Kostka, and Federico Masera. 2011. How informative are the subjective density forecasts of macroeconomists? CESifo Working Paper, No.3671.
- Klein, L.R., and S. Özmucur. 2002. The estimation of China's economic growth rate. Journal of E conomic and Social Measurement 28, no. 4: 187–202.
- Krugman, P. 2011. Will China break? The New York Times, December 18.
- Palm, F.C., and Zellner, A. 1992. To combine or not to combine? issues of combining forecasts. *Journal of Forecasting*, Vol. 11(8), pp. 687–701.
- Prati, Alessandro, and Massimo Sbracia. 2002. Currency crises and uncertainty about fundamentals. IMF Working Papers 02/3.
- Rawski, T. 2001a. What is happening to China's GDP statistics? *China Economic Review* 12, no. 4: 347–354.
- Rawski, T. 2001b. China's GDP statistics: a case of caveat lector? *China Economic Quarterly* 12, no. 5: 18–22.
- Sinton, Jonathan E. 2001. Accuracy and reliability of China's energy statistics. *China Economic Review* 12, no. 4: 373–83.
- Tay, Anthony S., and Kenneth F. Wallis. 2000. Density forecasting: a survey. *Journal of Forecasting* 19: 235–54.
- Wallis, Kenneth Frank. 1999. Asymmetric density forecasts of inflation and the Bank of England's fan chart. *National Institute Economic Review*, No. 167.
- Wallis, Kenneth Frank. 2003. Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts. *International Journal of Forecasting* 19, no. 2: 165–75.

# Appendix<sup>20</sup>

### A. The two-piece normal distribution

The two-piece normal distribution (TPN) was invented by John (1982) and is now used extensively for fan charts due to its relative ease for calculation and explanation. Johnson, Kotz and Balakrishnan (1994) made an early briefly introduction and discussion of the key features of this distribution. This represents a necessary and simple summarization of TPN distribution, which may help our readers to follow and develop our concrete approaches of constructing fan charts mainly discussed in section 2.

In principle, a TPN distribution is formed by the two halves of normal distributions. For example, suppose the two component normal distributions have a common mean,  $\mu$ , while their variances are  $\sigma_1$  and  $\sigma_2$ , respectively. Then, the density function of the two-piece normal distribution is:

$$f(x) = \begin{cases} A \exp\left[-(x-u)^2 / (2\sigma_1^2)\right] & \text{for } x \le \mu \\ A \exp\left[-(x-u)^2 / (2\sigma_2^2)\right] & \text{for } x > \mu \end{cases}$$
(A.1)

where  $A = (\sqrt{2/\pi} (\sigma_1 + \sigma_2))^{-1}$  is a constant, which ensures that the distribution is continuous and integrates to 1 when  $\sigma_1 > \sigma_2$ , the two-piece normal distribution has positive skew and when  $\sigma_1 < \sigma_2$  has negative skew. If  $\sigma_1 = \sigma_2$ , the TPN degrades to the normal distribution.

The first two moments and the skew of the two-piece normal distribution are as follows:

$$E(x) = \mu + \sqrt{\frac{2}{\pi}} (\sigma_2 - \sigma_1)$$
 (A.2)

$$V(x) = \sigma_1 \sigma_2 + \left(1 - \frac{2}{\pi}\right)(\sigma_2 - \sigma_1)^2$$
(A.3)

$$\gamma(x) = \sqrt{\frac{2}{\pi}} (\sigma_2 - \sigma_1) \left[ \left( \frac{4}{\pi} - 1 \right) (\sigma_2 - \sigma_1)^2 + \sigma_1 \sigma_2 \right]$$
(A.4)

where E(x), V(x) and  $\gamma(x)$  represent the mean, variance and skewness of the two-piece normal distribution, respectively. According to the last two equations, the variance parameters of the two component normal distributions  $\sigma_1$  and  $\sigma_2$  could be thoroughly determined from the variance and skew of the TPN distribution. Further, as discussed in Blix and Sellin (1998), the skewness of TPN distribution could be approximately calculated as:

$$\gamma(x) = \sqrt{\frac{2}{\pi}} \left(\sigma_2 - \sigma_1\right). \tag{A.4}$$

In our practice of constructing confidence intervals and fan charts, this simplified form of TPN distribution skewness simplifies, to a great extent, our calculations for  $\sigma_1$  and  $\sigma_2$ . In the next section, we will briefly introduce our tactics for this calculation and the corresponding construction of fan charts from estimated and known information, such as the mean, variance and skew of the two-piece normal distribution.

#### B. Drawing fan chart based on information of TPN

Once we have estimated the skew and variance (according to direct or indirect approach) of the overall economic growth rate distribution, which is assumed to be the TPN distribution, we can solve the two variance parameters  $\sigma_1$  and  $\sigma_2$  from Equation (A.3) and Equation (A.5). By substituting  $\sigma_2$  in Equation (A.3) with skewness  $\gamma(x)$  and  $\sigma_1$ , we get the following quadratic equation:

$$\sigma_1^2 + \sqrt{\frac{\pi}{2}} \gamma_Y \sigma_1 - \left[ \left( 1 - \frac{\pi}{2} \right) \gamma_Y^2 + \sigma_Y^2 \right] = 0 \tag{B.1}$$

where  $\gamma_{Y}$  and  $\sigma_{Y}$  represent the estimated skew and variance of the TPN distribution, respectively. Typically, the quadratic Equation (B.1) has two roots, one positive and one negative. We chose the one with the highest real value as the value of  $\sigma_1$ .<sup>21</sup> Then, the value of  $\sigma_2$  could be calculated immediately according to Equation (A.5).

After estimating the two key variance parameters of  $\sigma_1$  and  $\sigma_2$  with the known common mean,  $\mu$  (the predicted economy growth rate of the baseline forecast), we have the information needed for the two-piece normal distribution of the overall growth rate and can, therefore, construct confidence intervals. Suppose a confidence interval for a given probability *p* is (*a*, *b*), the following equation can be satisfied:

$$\Pr(a \le x \le b) = F(b) - F(a) = p \tag{B.2}$$

where F(.) is the cumulative distribution function of the TPN distribution.

Due to the asymmetry of the two-piece normal distribution, there are several possible methods for constructing the confidence interval. One intuitively easy approach is to choose the shortest interval b-a, which can be obtained by keeping the height of the TPN distribution function at the two ends of the interval identical, or<sup>22</sup>

$$f(a) = f(b). \tag{B.3}$$

# 24 Н. Үи

Combining Equations (B.2) and (B.3) and the TPN density function Equation (A.1), the confidence interval can be calculated as

$$b = \mu + \sigma_2 \Phi^{-1} \left(\frac{1+p}{2}\right) \tag{B.4}$$

$$a = \mu - \frac{\sigma_1}{\sigma_2} \left( b - \mu \right) \tag{B.5}$$

where  $\Phi^{-1}(\bullet)$  is the inverse of the standard normal distribution. This strategy could be shown visually as Figure B1.



Figure B1. Construction of confidence interval of two-piece normal distribution (the figure is produced with  $\mu = 0$ ,  $\sigma_1 = 1$ ,  $\sigma_2 = 2$ ).