Measuring Convergence using Dynamic Equilibrium Models: Evidence from Chinese Provinces*

Lei Pan[a], Olaf Posch[b,c] and Michel van der Wel[c,d]
[a]Wageningen University [b]Aarhus University, [c]CREATES,
[d]Erasmus University Rotterdam

May 30, 2012

Abstract

We propose a model to study economic convergence in the tradition of neoclassical growth theory. We employ a novel stochastic set-up of the Solow (1956) model with shocks to both capital and labor. Our novel approach identifies the speed of convergence directly from estimating the parameters which determine equilibrium dynamics. The inference on the structural parameters is done using a maximum-likelihood approach. We estimate our model using growth and population data for China's provinces from 1980 to 2009. We report heterogeneity in the speed of convergence both across provinces and time. The Eastern provinces show a higher tendency of convergence, while there is no evidence of convergence for the Central and Western provinces. We find empirical evidence that the speed of convergence decreases over time for most provinces.

JEL classification: C13; E32; O40

Keywords: Economic convergence; Dynamic stochastic equilibrium models; Solow model; Structural estimation

*Olaf Posch and Michel van der Wel appreciate financial support from the Center for Research in Econometric Analysis of TimE Series (CREATES), which is funded by the Danish National Research Foundation, and the Danish Social Science Research Council. Michel van der Wel is grateful to Netherlands Organisation for Scientific Research (NWO) for a Veni grant. Contact author: Lei Pan, Development Economics Group, Wageningen University, Hollandseweg 1, 6706 KN, The Netherlands; Phone: (+31)317484354; Email: lei.pan@wur.nl.

1 Introduction

An important issue in economic growth is whether poor regions grow faster than richer ones and eventually will catch up. Literature has devoted enormous enthusiasm in exploring this question in the 1990's and generally finds evidence of economic convergence in the world (e.g., Mankiw, Romer, and Weil, 1992; Barro and Sala-i-Martin, 1992), United States (e.g., Barro and Sala-i-Martin, 1992), and Japan and European nations (e.g., Sala-i-Martin, 1996a,b). Despite the effort, the economic growth in the new century seems to provide dubious evidence to confirm economic convergence. For example, the lagged behind Sub-Sahara Africa does neither show predicted rapid economic growth nor a high potential to catch up with the rich nations in the short run. However, while the new economies (e.g., China, India, Brazil) are emerging and catching up, only a few provinces seem to be the driving forces. In this paper we study convergence within China based on a novel stochastic Solow set-up.

Traditionally, standard tests of convergence are derived from a deterministic Solow model which involves investigating the relation between the growth rate and initial income level (see, e.g., Mankiw, Romer, and Weil, 1992; Barro and Sala-i-Martin, 1992). A negative relation between the two variables implies (absolute) economic convergence. The Solow model implies conditional convergence, which means that the growth rate is larger for countries with lower initial income level conditional on the structural characteristics. Cross-section and panel data regression approaches have been applied to explore the relation (see, e.g., Mankiw, Romer, and Weil, 1992; Barro and Sala-i-Martin, 1992; Islam, 1995). More recent literature shows that taking the stochastic feature of the Solow model into consideration the convergence estimates from these traditional cross-section or panel data regressions are biased (see, e.g., Lee, Pesaran, and Smith, 1997, 1998; Hauk and Wacziarg, 2009).[1]

We propose a model to study economic convergence in the tradition of neoclassical growth theory. In our stochastic set-up of the Solow (1956) model, we include shocks to capital and labor, and obtain equilibrium dynamics for output and population growth. We first derive a testable implication from our stochastic set-up following the traditional approach by assuming that the economy is close to its (stochastic) steady state. We then relax this assumption and fully explore the dynamic structure and obtain the speed of conditional convergence directly from the equilibrium dynamics.

From a modeling perspective we contribute to recent advances in the literature on using a continuous-time formulation to estimate the structural parameters in dynamic equilibrium models (see, e.g., Posch, 2009; Christensen, Posch, and van der Wel, 2011). Our approach

[1] Recent methodological advances on economic growth and convergence include Maasoumi, Racine, and Stengos (2007); Durlauf, Kourtellos, and Tan (2008) and Henderson, Papageorgiou, and Parmeter (2011).

builds on the seminal papers in the economic growth and convergence literature in continuous time (Merton, 1975; Mankiw, Romer, and Weil, 1992; Barro and Sala-i-Martin, 1992). In contrast to the traditional literature on convergence, however, we emphasize the stochastic nature of these models. This has three main advantages for our purpose. First, it enables us to conduct a structural estimation approach. Despite the rich nature of our models, we show that inference on the structural parameters of the model can be done using a maximum-likelihood approach. Second, the estimation results for the parameter measuring (conditional) convergence based on the deterministic model are likely to be upward-biased. In fact, in case of deterministic growth rates the convergence measure in our stochastic model collapses into the well-known deterministic measure. Third, we do not need to assume that the economy is close to its steady state in order to study local equilibrium dynamics, which clearly reflects a conceptional problem when analyzing the economy's transition. Our measure of the speed of convergence is specific for each economy and varies over time.

China is particularly important for the study of economic convergence. With its rapidly rising economy, it contributed to the majority of poverty reduction in the world in the past 30 years. However, its economic growth is also characterized by a high inequality between the coastal and inland areas and between urban and rural areas. Whether all regions in China will eventually converge to the same steady state and at which rate the convergence will happen is not only essential for economic growth in China, but also for the world given China's growing importance in the world economy. The traditional approaches for testing absolute and conditional convergence within deterministic neoclassical growth model have been applied to pre-reform and post-reform Chinese provincial data (e.g., Chen and Fleisher, 1996; Raiser, 1998; Weeks and Yao, 2003). These studies find convergence in the pre-reform period and mixed evidence for the post-reform period (from 1978 to the early 1990's). From theoretical and methodological aspects, we contribute to these studies by exploring the convergence in China in the framework of a stochastic Solow growth model, and allow for estimating province specific steady states and convergence rates. From an empirical aspect, China was still at the beginning of its economic growth in the early 1990's and the provinces were probably still far away from their steady states. Our approach allows us to study the equilibrium dynamics for this transitional period, in particular, when the speed of convergence is expected to be substantially higher than its value around the steady state. In this paper we use data for China's provinces from 1978 to 2009, a period which captures the rise of the Chinese economy and also provides a clearer picture on convergence (divergence) of its regional economies.

Our empirical results report heterogeneity in the speed of convergence both across provinces and time. We estimate our model using cross-section data, as in Mankiw, Romer, and

Weil (1992), and Barro and Sala-i-Martin (1992), panel data, as in Islam (1995) and Lee, Pesaran, and Smith (1997), and by exploiting the full dynamic structure of our stochastic Solow model. The cross-section and panel approaches provide mixed results. The cross-section set-up indicates significant absolute convergence in China, but the panel estimate is far lower. A salient feature of a disaggregation to the results to three regions in China is heterogeneity in convergence across the provinces in China, though the cross-section and panel approaches differ on which regions actually converge. With our novel approach to estimate the full dynamic structure of the stochastic Solow model we can examine these findings in more detail. Our approach provides an idiosyncratic and time-varying speed of convergence, while benefiting from the available data. The results exploiting the dynamic structure indicate significant convergence for all provinces and confirm the heterogeneity in convergence across provinces, and in addition highlight in line with theory that the speed of convergence decreases over time for most provinces.

The remainder of the paper is organized as follows. In Section 2 we detail our stochastic Solow model, and derive the convergence measures implied by the model. Section 3 introduces and discusses our data set for Chinese provinces, provides summary statistics, and reports the results using the traditional approaches. In Section 4 we discuss our convergence estimates for China using our new model. Section 5 concludes.

2 The stochastic Solow model

In this section we discuss our set-up to examine convergence. We first discuss in Section 2.1 the basic set-up, which is a stochastic Solow model closely following Merton (1975). Then, in Section 2.2, we derive empirically testable implications of the model.

2.1 The basic set-up

Consider a set of provinces, $i = 1, 2, ..., N$, with each T annual observations, $t = 1, 2, ..., T$. We assume that time is continuous. Output in each province and date, Y_{it}, is produced according to the technology

$$Y_{it} = K_{it}^{\alpha} L_{it}^{1-\alpha}, \tag{1}$$

where K_{it} is the physical capital stock, L_{it} is labor employed. Accounting for technological progress poses no conceptional difficulties but introduces an inherently latent variable to the problem formulation. Assuming that observed variables are in efficiency units, it would only increases the parameter space without adding too many insights. Our baseline model thus abstracts from technological progress. Capital accumulation follows

$$dK_{it} = (I_{it} - \delta K_{it})dt + \sigma_i K_{it} dZ_{it}, \tag{2}$$

in which $I_{it} = s_i Y_{it}$ where δ denotes the mean and σ the volatility of the of the instantaneous stochastic depreciation rate, driven by a standard Brownian motion Z_{it}.[2] Economically, we assume that physical depreciation of the capital stock is stochastic where σ_i governs the dispersion around a rate of δ percent (annual) depreciation. This assumption makes the capital stock and thus capital returns risky (cf. Merton, 1975).

We have some freedom to specify the labor dynamics. Assume that labor obeys

$$dL_{it} = n_i L_{it} dt + \eta_i L_{it} dB_{it}. \tag{3}$$

Similar to the shocks to capital depreciation, the rate of population growth is stochastic, driven by a standard Brownian motion B_{it} (and thus for simplicity independent from Z_{it}).

2.2 Deriving empirically testable implications

To answer our research questions on convergence of Chinese provinces, this section derives empirically testable implications to test the data for the convergence patterns. We provide two approaches: one based on the traditional approaches which were discussed in the Introduction, and one fully exploiting the dynamic structure of the model.

2.2.1 Using the traditional approach

An important measure is the speed of the transitional dynamics. Our approach to measure convergence closely follows Barro and Sala-i-Martin (1992). It is well known that the model implies *conditional convergence*, which means that convergence tends to be more rapid the lower the initial level of output in efficiency units conditional on structural characteristics.

In contrast to the previous analysis, however, below we fully specify the stochastic nature of the model and show how it feeds into the measure of convergence. Defining the capital stock in efficiency units, $k_{it} = K_{it}/L_{it}$, it follows

$$\begin{aligned} dk_{it} &= (s_i y_{it} - k_{it}\delta)dt + \sigma_i k_{it} dZ_{it} - k_{it} n_i dt - \eta_i k_{it} dB_{it} + \eta_i^2 k_{it} dt \\ &= (s_i k_{it}^{\alpha-1} - (\delta + n_i - \eta_i^2))k_{it} dt + (\sigma_i dZ_{it} - \eta_i dB_{it})k_{it}, \end{aligned} \tag{4}$$

in which we used $y_{it} = Y_{it}/L_{it} = k_{it}^{\alpha}$.

[2] Since Z_t is a standard Brownian motion, $Z_0 = 0$, $Z_{t+\Delta} - Z_t \sim \mathcal{N}(0, \Delta)$, $t \in [0, \infty)$.

We define the *speed of convergence* such that it measures by how much the growth rate declines as the effective capital stock increases in a proportional sense, i.e.,[3]

$$\beta_t = \beta(k_t) \equiv -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(1/k_t)dk_t]}{\partial \ln k_t}\right]. \tag{5}$$

As shown in the Appendix A, the speed of convergence for the growth rate of output per capita is the same as for capital in efficiency units. Hence, the speed of convergence in (4) is

$$\begin{aligned}
\beta_{it} = \beta(k_{it}) &= -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(s_i k_{it}^{\alpha-1} - (\delta + n_i - \eta_i^2))dt + (\sigma_i dZ_{it} - \eta_i dB_{it})]}{\partial \ln k_t}\right] \\
&= -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(s_i e^{(\alpha-1)\ln(k_{it})})dt]}{\partial \ln k_t}\right] \\
&= -\frac{1}{dt}\mathbb{E}_t\left[(\alpha-1)(s_i k_{it}^{\alpha-1})\right] \\
&= (1-\alpha)s_i k_{it}^{\alpha-1}.
\end{aligned} \tag{6}$$

Obviously, the speed of convergence is *not* constant but declines monotonically as the capital stock increases towards its steady-state value. Around the stochastic steady state where $dk_{it} = 0$ we typically observe that $s_i k_{it}^{\alpha-1} = \delta + n_i - \eta_i^2$. Hence, in the absence of shocks in the neighborhood of the stochastic steady-state value the speed of convergence equals

$$\beta_i^* = (1-\alpha)(\delta + n_i - \eta_i^2). \tag{7}$$

Note that setting $\eta_i = 0$, our measure of convergence (more precisely $\beta-$convergence) at the steady-state collapses to the traditional measure used in Mankiw, Romer, and Weil (1992).

Typically, the literature is considering a log-linear approximation of the deterministic version of (4) around the deterministic steady state. Keeping the stochastic nature of the problem at hand but log-linearizing (4) around the stochastic steady state (14), gives

$$(1/k_{it})dk_{it} \cong -\beta_i^* \ln(k_{it}/k_i^*)dt + \sigma_i dZ_{it} - \eta_i dB_{it}.$$

It follows from Itô's formula that

$$\begin{aligned}
(1/y_{it})dy_{it} &\cong -\alpha\beta_i^* \ln(k_{it}/k_i^*)dt + \alpha\sigma_i dZ_{it} - \alpha\eta_i dB_{it} + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2 + \eta^2)dt \\
&= -\beta_i^* \ln(y_{it}/y_i^*)dt + \alpha\sigma_i dZ_{it} - \alpha\eta_i dB_{it} + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2 + \eta^2)dt \\
&= -\beta_i^* \ln(y_{it})dt + \alpha\beta_i^* \ln k_i^* + \alpha\sigma_i dZ_{it} - \alpha\eta_i dB_{it} + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2 + \eta^2)dt,
\end{aligned}$$

[3]Alternatively, we could define the speed of convergence as (cf. Posch and Waelde, 2011),

$$\tilde{\beta}_t = \tilde{\beta}(k_t) \equiv -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial dk_t}{\partial k_t}\right].$$

Hence, the speed of convergence in (4) is $\tilde{\beta}_{it} = \delta + n_i - \eta_i^2 - \alpha s_i k_{it}^{\alpha-1}$. In the absence of shocks and around the stochastic steady state it takes the value $\tilde{\beta}_i^* = (1-\alpha)(\delta + n_i - \eta_i^2)$. It measures the absolute decrease in the capital stock for an infinitesimal increase in the effective capital stock. Since it is not a monotonically decreasing function, we use β in (5). Both measures converge the the same value at the steady state.

where we may use the definitions $\beta_i^* = (1-\alpha)(\delta + n_i - \eta_i^2)$ and $k_i^* = [s_i/(\delta + n_i - \eta^2)]^{1/(1-\alpha)}$. Using Itô's formula, we obtain

$$\begin{aligned} d\ln y_{it} &= (1/y_{it})dy_{it} - \tfrac{1}{2}(\sigma_i^2 + \eta_i^2)\alpha^2 dt \\ &\cong -\beta_i^* \ln(y_{it})dt + \alpha\beta_i^* \ln k_i^* dt + \tfrac{1}{2}(\alpha^2 - \alpha)(\sigma_i^2 + \eta^2)dt \\ &\quad +\alpha\sigma_i dZ_{it} - \alpha\eta_i dB_{it} - \tfrac{1}{2}(\sigma_i^2 + \eta_i^2)\alpha^2 dt \\ &= (-\beta_i^* \ln(y_{it}) + \alpha\beta_i^* \ln k_i^* - \tfrac{1}{2}\alpha(\sigma_i^2 + \eta^2))dt + \alpha\sigma_i dZ_{it} - \alpha\eta_i dB_{it}. \end{aligned}$$

It follows that the solution to this SDE is

$$\begin{aligned} \ln y_{it} &= e^{-\beta_i^* t}\left(\ln y_{i0} - \int_0^t (-\alpha\beta_i^* \ln k_i^* + \tfrac{1}{2}\alpha(\sigma_i^2 + \eta_i^2))e^{\beta_i^* s}ds + \alpha\sigma_i \int_0^t e^{\beta_i^* s} dZ_{is} - \alpha\eta_i \int_0^t e^{\beta_i^* s} dB_{is}\right) \\ &= e^{-\beta_i^* t}\ln y_{i0} + (1 - e^{-\beta_i^* t})\alpha \ln k_i^* - \tfrac{1}{2}\alpha(\sigma_i^2 + \eta_i^2)(1 - e^{-\beta_i^* t})/\beta_i^* \\ &\quad +\alpha\sigma_i \int_0^t e^{\beta_i^*(s-t)} dZ_{is} - \alpha\eta_i \int_0^t e^{\beta_i^*(s-t)} dB_{is}. \end{aligned}$$

After substituting k_i^* and β_i^* into the equation, we have

$$\begin{aligned} \ln y_{it} - \ln y_{i0} &= (1 - e^{-\beta_i^* t})\frac{\alpha}{1-\alpha}\left(\ln(s_i) - \ln(\delta + n_i - \eta^2) - \tfrac{1}{2}\frac{\sigma_i^2 + \eta_i^2}{\delta + n_i - \eta_i^2}\right) \\ &\quad -(1 - e^{-\beta_i^* t})\ln y_{i0} + \varepsilon_{it}. \end{aligned} \tag{8}$$

Comparing our result to the existing deterministic approaches gives two main insights. First, we have a complete specification of the implied error structure in a cross-sectional regression framework. Second, the parameter specifying the variance of the labor dynamics enters the convergence measure, i.e., the lower the variability of the growth rate of labor in province i, the more rapid is the speed of convergence in that province.[4] Estimation results for the convergence parameter based on the deterministic model are likely to be upward-biased.

2.2.2 Fully exploiting the dynamic structure

The traditional approach uses an approximation where provinces are assumed to be close to their steady-state values, which may or may not be a good approximation. This section fully exploits the dynamic structure of the model in order to estimate the structural parameters of the model, which finally can be used to provide answers to the convergence question. In this case, no log-linear approximation around the steady state is necessary.

It is well known that during the transition to the steady state value, the convergence rate β_{it} exceeds β_i^* as long as $k_{it} < k_i^*$ but tends to revert back to β_i^*. Thus, by making use of the dynamic structure of the model we may gain further insights into the model implications for

[4] The well known regression equations in the deterministic models is obtained for $\eta_i = \sigma_i = 0$.

convergence. Previous papers use panel data and the investment to output ratio together with a regression-based approach to capture these effects (e.g., Islam, 1995). Our approach builds on that approach and exploits the stochastic properties of the Solow growth model together with its moment conditions.

Using Itô's formula, the stochastic differential for output Y_{it} reads

$$\begin{aligned} dY_t &= \alpha \frac{Y_{it}}{K_{it}} dK_{it} + \tfrac{1}{2}\alpha(\alpha-1)\frac{Y_{it}^2}{K_{it}^2}\sigma_i^2 K_{it}^2 dt + (1-\alpha)\frac{Y_{it}}{L_{it}} dL_{it} - \tfrac{1}{2}(1-\alpha)\alpha \frac{Y_{it}}{L_{it}^2}\eta_i^2 L_{it}^2 dt \\ &= \alpha Y_{it}\left((I_{it}/K_{it}-\delta)dt + \sigma_i dZ_{it}\right) + \tfrac{1}{2}\alpha(\alpha-1)Y_{it}(\sigma_i^2+\eta_i^2)dt \\ &\quad +(1-\alpha)Y_{it}\left(n_i dt + \eta_i dB_{it}\right), \end{aligned} \tag{9}$$

or equivalently,

$$\begin{aligned} d\ln Y_{it} &= \alpha(I_{it}/K_{it}-\delta)dt - \tfrac{1}{2}(\alpha^2\sigma_i^2 + (1-\alpha)^2\eta_i^2)dt + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2+\eta_i^2)dt \\ &\quad +(1-\alpha)n_i dt + \alpha\sigma_i dZ_{it} + (1-\alpha)\eta_i dB_{it} \\ &= (\alpha I_{it}/K_{it} - \alpha\delta + (1-\alpha)n_i - \tfrac{1}{2}((1-\alpha)\eta_i^2 + \alpha\sigma_i^2))dt \\ &\quad +\alpha\sigma_i dZ_{it} + (1-\alpha)\eta_i dB_{it}. \end{aligned} \tag{10}$$

Similarly, the growth rate of labor employed from (3) follows

$$d\ln L_{it} = (n_i - \tfrac{1}{2}\eta^2)dt + \eta_i dB_{it}. \tag{11}$$

Equations (10) and (11) form the basis for our estimation. Our novel approach uses the structure for the estimation of the structural parameters. From the structural parameters we make inferences on β-convergence. The speed of convergence implied by the model is given in equation (6). This speed of convergence varies over time and differs across province, due to the variation of the stock of capital over time and across province and the variation of the saving rate per province. In addition, the speed of convergence near the stochastic steady state for each province can be obtained from (7).

3 Data and traditional approaches

In this section we describe the data we use and provide summary statistics. In addition, we show evidence of convergence in China using the traditional measures.

3.1 Province-level data for China

The data we use are the provincial level data from China Statistical Year Book 1978-2009 published by the National Bureau of Statistics of China (NBS).[5] The year books provides

[5]Besides the papers on convergence in China mentioned in the Introduction, others have used this data (for different time periods). For example, Fan and Zhang (2002) model production and productivity growth

basic information on population, labor market, investment, agriculture, industry, service business, public finance, education, culture and education. In this paper, we use the GDP and population data to calculate GDP per capita which we use as the index of economic growth. We approximate saving using investment in fixed assets and, for the traditional regressions, human capital level by the ratio between the number of students enrolled in high schools and the total population. We focus on the 1980-2009 data due to the noise in the population data at the end of the 1970's.

[insert Table 1]

In Table 1 we list the means and growth rate of real GDP per capita and population. We can see that China's population increases around 1% per year and the GDP per capita increases at an amazing rate of 9% in the past 30 years. There is evidence of geographic imbalance in China's economic growth in Table 1. The fastest growing provinces (with the growth rate bigger than 10%) are all in the coastal region except Inner Mongolia. Regional inequality is obvious from the mean GDP per capita. The average GDP per capita in the richest (city) province Shanghai is 13,888 yuan (in 1990 prices), almost ten times that of the poorest province Guizhou in Western China. However, there is some evidence of absolute economic convergence. Even though some of the coastal provinces show very high growth rates (e.g., Jiangsu, Zhejiang, Fujian, Guangdong), some traditionally developed regions such as Beijing, Tianjin and Shanghai grow on average slower than the country average and some of the Western provinces such as Sichuan, Yunnan and Ningxia. To examine dispersion of convergence across China, in some of our results we consider regions within China consisting of a subset of all provinces. Specifically, we split the provinces in China in three groups: East, Central, and West. Moreover, we consider a subgroup of the Eastern provinces that excluding the North-Eastern provinces. The last column of Table 1 lists the group each of the considered province falls under.

[insert Figure 1]

We plot the province level data in Figure 1. For expositional purposes we show the average per region over time. We find an upward trend of GDP per capita (Panel A). We do see a few jumps in the plots. For all provinces GDP per capita decreased during 1989, which was mostly due to the large level of inflation in that year (real GDP generally increased that year). For population data (Panel B), the jumps always happen in the census years in which the NBS use the census information to correct for the population data, which are obtained

in Chinese agriculture, Jin, Qian, and Weingast (2005) investigate provincial government's fiscal incentives and market development, and Fleisher, Li, and Zhao (2010) study growth patterns in China.

by a 1% population sample survey in non-census years. Also the jumps happen in cities such as Beijing and Shanghai, where there is high mobility of population and the estimate of total population is often with bigger standard error compared to other provinces. For the GDP per capita data there are jumps for a few provinces in 2001. We analyze the raw data, but we performed robustness analyses and verified our results do not depend on these jumps qualitatively.[6]. Finally, in Panel C we show the real investment data (top plot). We use this data to construct our capital measure,[7] which we plot in the middle part. Of relevance for our estimation is the ratio of investment to capital (bottom plot), which varies between 0.1 and 0.3.

3.2 Traditional approaches to measure convergence

A few concepts of economic convergence have emerged in the literature, as was pointed out in the Introduction. The first pair is absolute and conditional convergence. Absolute, or unconditional, convergence means that all economies are converging to the same steady state. Since the poor economies are further away from the steady state, absolute convergence also implies that poor economies grow faster. Conditional convergence means that economies approach different steady states determined by different structural characteristics, which are hence conditioned on, and the growth rate falls as the economy approaches its steady state (Galor, 1996; Sala-i-Martin, 1996a). The second pair is β-convergence and σ-convergence. Absolute and conditional convergence are also called absolute and conditional β-convergence (and it is also this type of convergence we have examined in Section 2.2 and we focus on in this paper). σ-convergence means that the dispersion of real per capita GDP levels tend to decrease over time and it is a necessary condition for β-convergence (Sala-i-Martin, 1996a). The concepts and empirical studies of economic convergence mentioned above are within the framework of deterministic neoclassical growth models. Before we set out to estimate the stochastic Solow system, we first use our data to explore these traditional concepts and approaches.

[6]For the robustness analyses, we looked at four alternatives. First, we smoothed the jumps out of the data: If, for population data, the growth rate is bigger than 5% or smaller than 5% and the nearby years do not show such big jumps, we replace the population data with the average of the two nearby years. As a second alternative, we apply a Hodrick-Prescott smoother on GDP and population variables. Third, we only smooth Guangdong's jump in the 1986 population number by replacing it with the average of 1985 and 1987. Fourth, we analyze raw data but use dummies for either only 1990 and 2000, or dummies for 1990, 1991, 2000, 2001 and 2005. For all four variations, our results are qualitatively similar to using the raw unsmoothed data.

[7]The main measure of capital is constructed as follows. First, we use the 1978 capital stock estimates from Zhang (2008). Then, for each subsequent period, we add the real investment to the previous period's capital stock, which we multiplied by 1-δ to capture depreciation (with $\delta = 0.05$). As alternatives, we have tried the Hall and Jones (1999) perpetual inventory method to estimate the initial capital stock, and varied the depreciation rate δ. Results when using these alternatives were qualitatively similar.

[insert Figures 2 and 3]

In Figure 2 and 3, we take a look at the absolute β-convergence and σ-convergence in Chinese provinces. Figure 2 plots the log-growth rate of real GDP per capita against the initial level of log real GDP per capita. Absolute $\beta-$convergence is consistent with a decreasing line: provinces with a low starting point display higher growth. We can see from Figure 2 that the poor provinces, the ones with lower GDP per capita in 1978, indeed grew fast in the past 30 years. This provides evidence of absolute β-convergence. In addition, we examine the convergence for the different regions in China. We find substantial heterogeneity across these different regions. The convergence is most obvious in the (richer) Eastern provinces, while the convergence is less pronounced in the Western provinces, and there is almost no convergence in the Central provinces. Figure 3 shows σ-convergence: the cross-sectional dispersion of real GDP per capita, scaled by the cross-sectional mean, over time. In case of σ-convergence, there would be less variation of the GDP per capita and thus a decreasing line. For China as a whole there is clear σ-convergence from 1978-1990. From 1990-2000 there is σ-divergence and after 2000 the pattern reversed again. The convergence mainly comes from the Eastern region. The Central provinces show clear divergence and the Western provinces do not show clear pattern of convergence or divergence.

[insert Table 2]

Table 2 provides the convergence measures of the traditional regressions approaches introduced in Barro and Sala-i-Martin (1992), Mankiw, Romer, and Weil (1992) and Islam (1995) for the province-level China data. In the table we distinguish absolute and conditional convergence. In the first two columns the convergence measures are estimated from the coefficients of regressing the log GDP per capita growth rate on the log of the initial GDP per capita using cross-section and panel data respectively. Columns three and five are their conditional convergence counterparts controlling for investment, population and technology growth rate.[8] Following Mankiw, Romer, and Weil (1992), we also control for human capital in column four by including a variable capturing the ratio between the number of high school students and total population. The results show a difference between the cross-section approaches of Barro and Sala-i-Martin (1992) and Mankiw, Romer, and Weil (1992), and the panel data approach of Islam (1995) (which we implement using data at a 5-year frequency and the LSDV estimator). Column one provides a significant convergence estimate of 0.01. As the regression does not condition on other factors, this is evidence of

[8]We implement the traditional approaches with the theoretical restriction that the coefficients of log investment rate and log of the sum of depreciation rate, technology growth rate and population growth rate are equal. The results are robust without this restriction.

absolute convergence: all provinces in China will converge to the same steady state. After we condition on the variables and parameters of the Solow model, we also find evidence of convergence. The estimate of 0.010 is similar (though no longer significant), and when there is also conditioned on human capital the estimate is 0.037, indicating even more rapid convergence of the provinces to their own steady state. Contrary to these consistent results, the panel data approaches provide significant estimate of absolute divergence but conditional convergence.

In addition, to get a clearer picture of the heterogeneity across the provinces we examine the convergence measures for the different subsets of provinces. The lower rows in Table 2 provide the speed of convergence for each of the four groups using the traditional growth regressions. The cross-section results show that the convergence is quicker in the East than the West. In fact, for both cases convergence is only significant in the East (consistent with Figure 2), which is the most developed part of China. The panel data analysis provides significant evidence of absolute divergence for the Central and Western provinces, but surprisingly significant evidence of conditional convergence for precisely these same provinces. It is these mixed findings which we examine further benefiting from estimating a full dynamic system.

4 Convergence evidence in the stochastic Solow model

We now turn to the results of convergence in China using our stochastic Solow model. As a starting point, in Section 4.1 we take the traditional approach of log-linearizing around a steady state to obtain our testable implications. Then, in Section 4.2, we fully exploit the dynamic system to study convergence in China.

4.1 Regression evidence

Before we turn to the full dynamic estimation, we first study the stochastic Solow model inspired by the traditional approaches. In this case, our measure of the speed of convergence is based on equation (8). We estimate two broad variations of this equation.

[insert Table 3]

First, we follow the traditional cross-section approaches of Barro and Sala-i-Martin (1992) and Mankiw, Romer, and Weil (1992) and consider for each province only data from the begin- until the end-point (in our application from 1980 to 2009). That is, we estimate

$$\ln y_{i,2009} - \ln y_{i,1980} = \kappa_0 + \kappa_1 \ln y_{i,1980} + \kappa_2(\ln(s_i) - \ln(\delta + n_i)) + u_i,$$

where $y_{i,2009}$ denotes the real GDP per capita of province i in 2009, s_i is estimated using the average investment over GDP ratio over 1980-2009, n_i is the average annual growth rate of labor force over the sample for each province and δ is fixed at 0.05.[9] The intercept, represented with κ_0, captures the covariance terms, which structure we do not consider in this first specification. As a slight modification, we also consider a specification where we in a first stage estimate the average annual growth rate of the labor force n_i and the variance η_i^2 of this growth rate for each province i using a regression, and in a second stage plug this in equation (8). This gives

$$\ln y_{i,2009} - \ln y_{i,1980} = \kappa_0 + \kappa_1 \ln y_{i,1980} + \kappa_2(\ln(s_i) - \ln(\delta + \hat{n}_i - \hat{\eta}_i^2)) + u_i.$$

In both cases, only a country/group-wide convergence measure can be obtained. For example, through $\kappa_1 = -(1 - \exp(-29 \times \beta))$ the speed of convergence β is obtained.

The first two columns in Table 3 provide the output for the cross-sectional analysis of the stochastic Solow model. The speed of convergence is significant for both the sample average and estimated labor growth. The estimate of 0.01 in the case with the in-sample estimate of labor growth is similar to the cross-section conditional convergence approach of the previous section. When using the estimated population growth and the standard deviation thereof, the convergence estimate drops slightly. Both estimates are, however, insignificant.

Second, in the spirit of Islam (1995) we estimate (8) using panel data for each province. Dividing the full sample from 1980 to 2009 in steps of length L, a panel variation of the speed of convergence equation can be written as

$$\ln y_{i,t} - \ln y_{i,t-L} = \kappa_{0,i} + \kappa_{1,i} \ln y_{i,t-L} + \kappa_{2,i}(\ln(s_{i,t}) - \ln(\delta + n_{i,t})) + u_{i,t},$$

where $y_{i,t}$ denotes the GDP of province i in year t, L is the lag in the specification, $s_{i,t}$ is estimated using the average investment over GDP ratio over the period from $t - L + 1$ up to t, $n_{i,t}$ is the average annual growth rate of labor force and δ is fixed at 0.05. Also here, as a slight modification, we first estimate the average annual growth rate of the labor force $n_{i,t}$ and the variance $\eta_{i,t}^2$ of this growth rate over the relevant period from $t - L + 1$ up to t for each province i using a regression, and plug this in the to get

$$\ln y_{i,t} - \ln y_{i,t-L} = \kappa_{0,i} + \kappa_{1,i} \ln y_{i,t-L} + \kappa_{2,i}(\ln(s_{i,t}) - \ln(\delta + \hat{n}_{i,t} - \hat{\eta}_{i,t}^2)) + u_{i,t}.$$

Similar to before, from the $\kappa_{1,i}$ estimates we obtain in both cases the convergence measure β_i which is now province specific.

[9] Mankiw, Romer, and Weil (1992) fix $g + \delta$ at 0.05, with g the growth rate of technology. We fix δ such that our regression is comparable to their setting, as in their case the last term in the regression is $\ln(g+\delta+n_i)$ compared to $\ln(\delta+n_i)$ in our case. In this way this regression collapses to the Mankiw, Romer, and Weil (1992) regression (column three in Table 2) with the restriction on the coefficients mentioned in footnote 8. Results are however robust to setting δ at different levels.

The last two columns in Table 3 report the speed of convergence for each of the provinces obtained using this panel approach. In the implementation, we take $L = 3$ as this gives some observations to estimate the labor growth rate in each interval but still ensures we have enough time series observations in the panel to estimate the convergence measure. There is great variation in the conditional convergence estimates for each province using both the sample average and estimated labor growth. Interestingly, except for Tianjin all significant estimates point to convergence, and are generally in relatively rich and developed provinces. Thus, contrary to the cross-section evidence, the panel data approach does give rise to significant conditional convergence estimates for some provinces.[10] The analysis highlights the heterogeneity of convergence in the Chinese provinces, further motivating the use of our dynamic system with a time-varying province specific convergence estimate that is not obtained by log-linearization around the steady state.[11]

4.2 Fully exploiting the dynamic structure

We implement the stochastic Solow model using the dynamic structure from Section 2.2.2. By exploiting the dynamic structure we obtain structural parameters without making the approximation that provinces are close to their steady-state. We estimate the model based on equations (10) and (11). In order to estimate the structural parameters using the discrete-time structure of the data and accounting for the dynamic structure of the model, we integrate over $s < t$ for our discrete-time empirical specification:

$$\begin{aligned}
\ln Y_{it} - \ln Y_{is} &= \int_s^t \alpha I_{iu}/K_{iu} du - (\alpha\delta - (1-\alpha)n_i + \tfrac{1}{2}((1-\alpha)\eta_i^2 + \alpha\sigma_i^2))(t-s), \\
&\quad + \alpha\sigma_i(Z_{it} - Z_{is}) + (1-\alpha)\eta_i(B_{it} - B_{is}), \\
\ln L_{it} - \ln L_{is} &= (n_i - \tfrac{1}{2}\eta_i^2)(t-s) + \eta_i(B_{it} - B_{is}).
\end{aligned}$$

We consider annual data, such that in the above general case we have $s = t - 1$, and take an Euler approximation of the integral $\int_s^t I_{iu}/K_{iu} du \approx (t-s) I_{is}/K_{is}$. Using our constructed

[10] These results do depend to some extend on the length of the interval L. While for many values of L results are qualitatively similar, for $L = 5$ the only significant estimates for the provinces all point to divergence. This mixed evidence is an additional motivation for employing our novel set-up of the next section.

[11] An alternative way of showing this heterogeneity is to extend the traditional panel data approach of Islam (1995). With panel data it is possible to estimate a province specific speed of convergence from the time series of log real GDP per capita of each province. Note that a heterogeneous speed of convergence is fact quite natural. The original speed of convergence as in Mankiw, Romer, and Weil (1992) of $\lambda = (1-\alpha)(n + g + \delta)$ depends on the labor growth rate n. As in the traditional regressions is taken province specific, this already implies a province specific speed of convergence.

capital stock from investment data (see Section 3.1) we obtain the system:

$$\ln\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = -(\alpha\delta - (1-\alpha)n_i + \tfrac{1}{2}((1-\alpha)\eta_i^2 + \alpha\sigma_i^2)) + \alpha\frac{I_{i,t-1}}{K_{i,t-1}} + \varepsilon_{it}^Y, \qquad (12)$$

$$\varepsilon_{it}^Y \sim N(0, \alpha^2\sigma_i^2 + (1-\alpha)^2\eta_i^2),$$

$$\ln\left(\frac{L_{it}}{L_{i,t-1}}\right) = (n_i - \tfrac{1}{2}\eta_i^2) + \varepsilon_{it}^L, \qquad \varepsilon_{it}^L \sim N(0, \eta_i^2), \qquad (13)$$

with $Cov(\varepsilon_{it}^Y, \varepsilon_{it}^L) = (1-\alpha)\eta_i^2$. This system is in fact a SUR system, which we estimate with maximum-likelihood using data on real GDP and labor force (proxied by population data).

[insert Table 4]

Table 4 reports estimates for the parameters of stochastic depreciation σ_i, population growth n_i, and labor volatility η_i. Consistent with our earlier analyses, we fix δ at 0.05. Based on our estimates, stochastic depreciation is largest in Beijing, Shanxi, Inner Mongolia, Liaoning, Shanghai, Jiangsu, Guangdong and Sichuan. The population growth estimates all roughly correspond to the summary statistics reported in 1. For example, Beijing's average annual population growth rate is 2.36%, and our estimate indicates 2.43% annual growth. Population growth is most volatile in Beijing, Shanghai and Guangdong. The output elasticity of capital is estimated at 0.714.[12] The estimates in the table are almost all significant at the 1% level.

From estimating the full dynamic system we obtain two types of convergence results. First, the last column in Table 4 provides the province specific convergence measure around the steady state from equation (7). This measures denotes the speed with which a province converges back to its own steady state after it was near the steady state and got hit with some shock. The convergence estimates are roughly in the same order of size as those reported earlier using the traditional methods in Table 2, and the cross-sectional regressions based on our stochastic Solow model in Table 3. By exploiting the full dynamic structure of our model, we show that convergence is significant for all provinces. Nevertheless, these still is great dispersion in the speed of convergence across the different provinces: the speed of convergence ranges from 0.0078 to 0.0132. The provinces that showed significant convergence using the panel approach in Table 3 also generally rank relatively high on the list of convergence in this case.

[insert Figure 4]

[12]From the regression approach to the stochastic Solow model that we pursue in Section 4.1, it is possible to get an implied value of α. The estimate of κ_2 in the regressions represents $(1-e^{-\beta_i^* t})\frac{\alpha}{1-\alpha}$ from equation (8). Solving for α by using the implied β from the κ_1 estimate gives an output elasticity of capital of 0.67.

The second type of convergence results exploits the full dynamic nature of our system. Figure 4 provides the time-varying province specific convergence measure from equation (6), where s_i as the average ratio of real investment to real GDP I_{it}/Y_{it}, and we use $k_{it}^{\alpha-1} = Y_{it}/K_{it}$. The speed of convergence in this case represents the speed with which a province converges to its own steady state, and thus not rely on any approximation of the province already being near the steady state. The results show that there is a generally decreasing pattern of convergence over time. For some provinces this pattern is more pronounced (such as Zhejiang), while for other provinces the pattern is flatter (such as Gansu). We compare these speed of convergences to the steady state convergence of Table 4. In all cases these differ significantly (we have also plotted 95% confidence bounds). From this we can conclude that all of Chinese's provinces are not near their steady state yet. For some provinces the difference between the steady state speed of convergence and the time-varying speed of convergence differs more than for others. Particularly the Eastern provinces Shanghai, Jiangsu and Zhejiang converge rapidly to the steady state convergence speed. The picture is different for Gansu. This relatively poor province (GDP per capita of 1,963 yuan, compared to 13,388 yuan for the richest province Shanghai) is already relatively close to its own steady state speed of converge. The province is thus already near its steady state growth path, leaving little room for serious catching up with the richer provinces.

[insert Figure 5]

The difference between the steady-state and the time-varying measures of the speed of convergence also confirms the importance of the dynamic convergence analysis. In Figure 5, we plot the steady-state speed of convergence and the speed of convergence in 2009 against initial GDP per capita. A very different conclusion can be drawn from these two types of convergence measures: the steady-state convergence rate has a negative relation with the initial GDP per capita level, but the actual convergence rate in 2009 shows the opposite. The 2009 convergence rates are also much more heterogeneous compared to the steady-state measures. Ignoring the dynamic feature of convergence will lead to an incomplete view of convergence in China.

Finally, in the framework we can test whether the provinces converge to the same steady state or a different one. We can estimate a restricted version where we set $n_i = n$ and $\eta_i = \eta$ and perform a likelihood-ratio (LR) test against the unrestricted model.[13] The statistics is 713.8, which is significant at the 1% level. As a second alternative we further also restrict $\sigma_i = \sigma$. Also in this case the LR test rejects the null that the coefficients are equal is

[13]Such a test is similar in spirit to testing whether the conditioning variables in the traditional regressions are significant. They generally are, pointing to conditional rather than absolute convergence.

reject (significant at 1% level). Both analyses point to conditional rather than absolute convergence.

5 Conclusion

We propose a model to study the convergence process in the tradition of neoclassical growth theory. We employ a novel stochastic set-up of the Solow (1956) model with shocks to capital and labor and obtain equilibrium dynamics for output. We obtain the speed of conditional convergence directly from the equilibrium dynamics.

We study our model using growth and population data for China's provinces from 1980 to 2009. Our results show heterogeneity in the speed of convergence both across provinces and time. Our novel approach to estimate the full dynamic structure of the stochastic Solow model provides an idiosyncratic and time-varying speed of convergence, while benefiting from all available data *and* without requiring an approximation of dynamics around steady states. Convergence is significant for most provinces and differs across provinces, and the speed of convergence decreases over time for most provinces as predicted by theory.

A Appendix

A.1 The speed of convergence for output

Consider the stochastic growth rate of output $y_{it} = k_{it}^{\alpha}$

$$\begin{aligned} (1/y_{it})dy_{it} &= \alpha(s_i k_{it}^{\alpha-1} - (\delta + n_i - \eta_i^2))dt + \alpha(\sigma_i dZ_{it} - \eta_i dB_{it}) + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2 + \eta^2)dt \\ &= \alpha(1/k_{it})dk_{it} + \tfrac{1}{2}\alpha(\alpha-1)(\sigma_i^2 + \eta^2)dt \end{aligned}$$

Defining the speed of convergence as

$$\begin{aligned} \lambda_t &= \lambda(k_t) \equiv -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(1/y_t)dy_t]}{\partial \ln y_t}\right] \\ &= -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(1/k_{it})dk_{it} + \frac{1}{2}(\alpha-1)(\sigma_i^2 + \eta^2)dt]}{\partial \ln k_t}\right] \\ &= -\frac{1}{dt}\mathbb{E}_t\left[\frac{\partial[(1/k_{it})dk_{it}dt]}{\partial \ln k_t}\right] \end{aligned}$$

shows that $\lambda_t = \beta_t$.

A.2 Log-linearization in the stochastic Solow model

The advantage of the log-linearization method is that it provides a closed-form solution for the convergence coefficient, but it applies only as an approximation around a stationary point. In this section we show how we log-linearize (4) of the stochastic Solow model,

$$dk_{it} \quad = \quad (s_i k_{it}^{\alpha-1} - (\delta + n_i - \eta_i^2))k_{it}dt + (\sigma_i dZ_{it} - \eta_i dB_{it})k_{it}.$$

We define the stochastic steady state value as the value of the effective capital stock where $dk_{it} = 0$ and $dB_{it} = dZ_{it} = 0$, that is,

$$k_i^* = \left(\frac{s_i}{\delta + n_i - \eta_i^2}\right)^{\frac{1}{1-\alpha}}. \tag{14}$$

We rewrite the equation

$$(1/k_{it})dk_{it} - (\sigma_i dZ_{it} - \eta_i dB_{it}) = (s_i k_{it}^{\alpha-1} - (\delta + n_i - \eta_i^2))dt.$$

Log-linearization of the right-hand side around k_i^* yields

$$\begin{aligned} &(s_i e^{(\alpha-1)\ln k_i^*} - (\delta + n_i - \eta_i^2) + (\alpha - 1)s_i e^{(\alpha-1)\ln k_i^*}(\ln k_{it} - \ln k_i^*))dt \\ &= (s_i e^{\ln((\delta+n_i-\eta_i)/s_i)} - (\delta + n_i - \eta_i^2) + (\alpha - 1)s_i e^{\ln((\delta+n_i-\eta_i^2)/s_i)}(\ln k_{it} - \ln k_i^*))dt \\ &= (\alpha - 1)(\delta + n_i - \eta_i^2)(\ln k_{it} - \ln k_i^*)dt \end{aligned}$$

Thus, we may write

$$\begin{aligned} (1/k_{it})dk_{it} &\approx (\alpha - 1)(\delta + n_i - \eta_i^2)\ln(k_{it}/k_i^*)dt + \sigma_i dZ_{it} - \eta_i dB_{it} \\ &= -(1 - \alpha)(\delta + n_i - \eta_i^2)\ln(k_{it}/k_i^*)dt + \sigma_i dZ_{it} - \eta_i dB_{it} \end{aligned}$$

which shows that the stochastic growth rate approximately decreases at the rate β_i^*.

References

Barro, R. J., and X. Sala-i-Martin (1992): "Convergence," *J. Polit. Economy*, 100(2), 223–251.

Chen, J., and B. M. Fleisher (1996): "Regional Income Inequality and Economic Growth in China," *J. of Comparative Econ.*, 22(2), 141–164.

Christensen, B. J., O. Posch, and M. van der Wel (2011): "Estimating Dynamic Equilibrium Models using Macro and Financial Data," *CREATES WP*, (2011-21).

Durlauf, S. N., A. Kourtellos, and C. M. Tan (2008): "Are Any Growth Theories Robust?," *Economic Journal*, 118, 329–346.

Fan, S., and X. Zhang (2002): "Production and Productivity Growth in Chinese Agriculture: New National and Regional Measures," *Econ. Development and Cultural Change*, 50, 819–838.

Fleisher, B., H. Li, and M. Q. Zhao (2010): "Human capital, economic growth, and regional inequality in China," *J. of Development Econ.*, 92, 215–231.

Galor, O. (1996): "Convergence? Inferences from Theoretical Models," *Economic Journal*, 106(437), 1056–4069.

Hall, R. E., and C. I. Jones (1999): "Why do some countries produce so much more output per worker than others?," *Quart. J. Econ.*, 114, 83–116.

Hauk, W. R., and R. Wacziarg (2009): "A Monte Carlo study of growth regressions," *J. Econ. Growth*, 14, 103–147.

Henderson, D. J., C. Papageorgiou, and C. F. Parmeter (2011): "Growth Empirics Without Parameters," *Economic Journal*, 122, 125–154.

Islam, N. (1995): "Growth Empirics: A Panel Data Approach," *Quart. J. Econ.*, 110(4), 1127–1170.

Jin, H., Y. Qian, and B. R. Weingast (2005): "Regional decentralization and fiscal incentives: Federalism, Chinese style," *J. Public Econ.*, 89, 1719–1742.

Lee, K., M. H. Pesaran, and R. Smith (1997): "Growth and Convergence in A Multi-Country Empirical Stochastic Solow Model," *J. Appl. Econometrics*, 12, 357–392.

——— (1998): "Growth Empirics: A Panel Data Approach- A Comment," *Quart. J. of Econ.*, 113(1), 319–323.

Maasoumi, E., J. Racine, and T. Stengos (2007): "Growth and convergence: A profile of distribution dynamics and mobility," *Journal of Econometrics*, 136, 483–508.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quart. J. Econ.*, pp. 407–437.

Merton, R. C. (1975): "An asymptotic theory of growth under uncertainty," *Rev. Econ. Stud.*, 42(3), 375–393.

Posch, O. (2009): "Structural estimation of jump-diffusion processes in macroeconomics," *J. Econometrics*, 153(2), 196–210.

Posch, O., and K. Waelde (2011): "On the link between volatility and growth," *J. Econ. Growth*, (16), 285–308.

Raiser, M. (1998): "Subsidising inequality: Economic reforms, fiscal transfers and convergence across Chinese provinces," *The Journal of Development Studies*, 34(3), 1–26.

Sala-i-Martin, X. (1996a): "The classical approach to convergence analysis," *Economic Journal*, 106, 1019–1036.

——— (1996b): "Regional cohesion: Evidence and theories of regional growth and convergence," *European Economic Review*, 40, 1325–1352.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quart. J. Econ.*, 70, 65–94.

Weeks, M., and J. Y. Yao (2003): "Provincial Conditional Income Convergence in China, 1953-1997: A Panel Data Approach," *Econometric Reviews*, 22(1), 59–77.

Zhang, J. (2008): "Estimation of China's provincial capital stock (1952-2004) with applications," *J. of Chinese Economic and Business Studies*, 6, 177–196.

Table 1: China Province-level Data – 1980-2009

The table reports the Gross Domestic Product (GDP) per capita and Population for each of the provinces in our data set. We consider 28 provinces, where we include Chongqing in Sichuan and Hainan in Guangzhou, and ignore Tibet due to lack of data. We show the real GDP per capita in 1990 prices, which is obtained by dividing the nominal GDP by the CPI index with baseyear 1990, and the province population. For both we show average over 1980-2009, and average percentage growth which is obtained as 100 times the difference of the log-series. In the last column we specify the region each province falls under: East (with East* the subset of Eastern provinces excluding the North-East), Central and West.

Province-level Data – Real GDP per capita (in yuan, 1990 prices) and Population						
		Real GDP per capita		Population (millions)		
Province		Mean	%-Growth	Mean	%-Growth	Region
1	Beijing	10,674.4	7.69	12.0	2.36	East*
2	Tianjin	8,514.2	7.70	9.3	1.71	East*
3	Hebei	3,465.5	8.62	62.9	1.06	Central
4	Shanxi	3,032.5	8.04	30.1	1.12	Central
5	Inner Mongolia	3,923.1	10.85	22.2	0.88	Central
6	Liaoning	5,156.9	7.66	40.0	0.74	East
7	Jilin	3,513.1	8.73	25.2	0.74	East
8	Heilongjiang	3,847.8	6.64	35.8	0.61	East
9	Shanghai	13,388.2	6.22	14.3	1.78	East*
10	Jiangsu	5,720.0	9.86	68.9	0.91	East*
11	Zhejiang	6,187.3	10.32	44.3	1.04	East*
12	Anhui	2,314.9	8.57	58.0	0.78	Central
13	Fujian	4,657.8	10.43	31.5	1.26	East*
14	Jiangxi	2,398.5	8.18	39.1	1.05	Central
15	Shandong	4,579.9	10.12	85.8	0.90	East*
16	Henan	2,650.5	9.04	88.0	0.91	Central
17	Hubei	3,081.6	8.34	55.1	0.69	Central
18	Hunan	2,647.7	8.52	61.7	0.67	Central
19	Guangdong[a]	5,538.6	10.09	78.1	2.06	East*
20	Guangxi	2,219.6	8.64	43.9	1.09	West
21	Sichuan[b]	2,424.5	8.58	109.4	0.41	West
22	Guizhou	1,430.0	7.94	34.0	1.08	West
23	Yunnan	2,121.0	8.18	39.2	1.26	West
24	Tibet[c]					
25	Shaanxi	2,595.8	9.00	33.9	0.99	Central
26	Gansu	1,963.4	6.73	23.5	1.10	West
27	Qinghai	2,750.3	7.45	4.7	1.35	West
28	Ningxia	2,719.3	8.17	5.1	1.77	West
29	Xinjiang	3,334.0	7.96	16.8	1.79	West
30	Chongqing[b]					
31	Hainan[a]					
	China Total	3,658.9	8.65	1,166.3	1.01	

[a] *Hainan is included in Guangdong, though it was separated in 1988*

[b] *Chongqing is included in Sichuan, though it was separated in 1997*

[c] *Omitted as data starts in 1984*

Table 2: Convergence in China – Evidence using Traditional Approaches

The table provides various speed of convergence estimates for different regions in China. We split the provinces in China in three groups: East (with as subset East* the Eastern provinces excluding the North-East), Central, and West. We show the results of both absolute and conditional convergence regressions. For absolute convergence regressions, we regress the log-growth rate of real GDP per capita on a constant and the initial log real GDP per capita. We do this regression based on cross-section data (column CS), where we regress the log-growth rate of real GDP per capita over 1980-2009 of the provinces on 1980 log real GDP per capita, and panel data (column PD), where we look at the log-growth rate of real GDP per capita during each 5-year interval compared to the log real GDP at the beginning of each interval. For the conditional convergence regressions we also use cross-section and panel data, and additionally condition on the log of real investment over real GDP, and the log of the sum of the depreciation rate δ, labor growth n_i of province i and technological growth g. Consistent with Mankiw, Romer and Weil (1992), we set $g+\delta = 0.05$. In addition, for conditional convergence we follow Mankiw, Romer and Weil (1992) and also condition on schooling (column CS-HC). In all cases, we only report the speed of convergence estimate, which is obtained from the estimated coefficient on the initial log real GDP per capita. The cross-section regression results are obtained using OLS, the LSDV estimator is used for the panel models. Values below estimates denote standard errors. Asterisks (*/**/***) indicate significance at 10%/5%/1% level.

Convergence Estimates for China					
	Absolute convergence		Conditional convergence		
	CS	PD	CS	CS-HC	PD
China	0.011**	-0.015***	0.010	0.030**	0.015**
	(0.005)	(0.004)	(0.012)	(0.013)	(0.007)
Convergence Estimates for Regions in China					
East	0.028***	-0.007	0.027***	0.041***	0.008
	(0.010)	(0.006)	(0.008)	(0.013)	(0.008)
East*	0.031***	-0.004	0.033***	0.038***	0.010
	(0.003)	(0.008)	(0.002)	(0.005)	(0.010)
Central	0.014	-0.023***	0.022	0.004	0.034*
	(0.032)	(0.007)	(0.038)	(0.130)	(0.022)
West	0.011	-0.020***	0.012	0.045	0.035**
	(0.012)	(0.007)	(0.048)	(0.039)	(0.017)

Table 3: Convergence in China – Regression Evidence from Stochastic Solow Model

The table provides speed of convergence estimates for China using cross-section and panel regressions based on our stochastic Solow model linearized around the steady state. We show output for four different variations, two cross-section and two panel settings. For both cross-section and panel settings, we regress the log-growth rate of real GDP per capita over 1980-2009 of the provinces on 1980 log real GDP per capita and the the log savings rate minus the log of δ plus the labor growth rate n. For comparison with the traditional regressions, we set $\delta = 0.05$. We do this regression based on cross-section data (column CS), where we regress the log-growth rate of real GDP per capita over 1980-2009 of the provinces on 1980 log real GDP per capita and the controls, and panel data (column PD), where we look at the log-growth rate of real GDP per capita during each 5-year interval compared to the log real GDP at the beginning of each interval and the controls. In specifications (A) we use the sample average of n, while in specifications (B) we estimate these first using regressions. In all cases, we only report the speed of convergence estimate, which is obtained from the estimated coefficient on the initial log real GDP per capita. For the cross-section analyses this is a country-wide speed of convergence. For the panel approach we obtain a province specific speed of convergence. The cross-section regression results are obtained using OLS, the LSDV estimator is used for the panel models. Asterisks (*/**/***) indicate significance at 10%/5%/1% level.

		Speed of Convergence β_i			
		Cross-section		Panel Data	
		(A)	(B)	(A)	(B)
	Country-wide	0.0103	0.0101		
1	Beijing			-0.0020	-0.0019
2	Tianjin			-0.0042**	-0.0041**
3	Hebei			0.0017	0.0017
4	Shanxi			-0.0013	-0.0013
5	Inner Mongolia			0.0038	0.0038
6	Liaoning			0.0039	0.0039
7	Jilin			0.0094**	0.0097**
8	Heilongjiang			-0.0016	-0.0016
9	Shanghai			-0.0032	-0.0030
10	Jiangsu			0.0052*	0.0056**
11	Zhejiang			0.0057*	0.0058*
12	Anhui			0.0096**	0.0097**
13	Fujian			0.0055**	0.0057**
14	Jiangxi			0.0038	0.0038
15	Shandong			0.0019	0.0019
16	Henan			0.0063*	0.0063*
17	Hubei			0.0042	0.0041
18	Hunan			0.0032*	0.0028
19	Guangdong			0.0008	0.0009
20	Guangxi			0.0008	0.0008
21	Sichuan			0.0039	0.0032
22	Guizhou			0.0042	0.0040
23	Yunnan			0.0024	0.0023
25	Shaanxi			0.0063	0.0063
26	Gansu			0.0104	0.0105
27	Qinghai			0.0001	0.0002
28	Ningxia			0.0037	0.0036
29	Xinjiang			0.0032*	0.0032*

Table 4: Stochastic Solow Model – Dynamic System Estimates and Convergence around Steady State

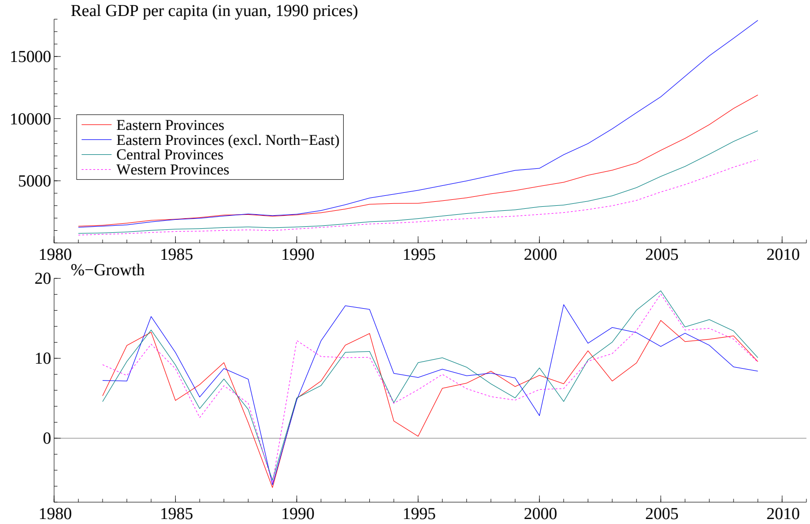
The table reports estimates of the Stochastic Solow model. We consider 28 provinces in China, where we include Chongqing in Sichuan and Hainan in Guangzhou and ignore Tibet due to lack of data. For province i we show estimates of the capital volatility σ_i, instantaneous population growth n_i, and labor volatility η_i. In addition, we show the estimate of the production function parameter α and fixed value of depreciation rate $\delta = 0.05$, which we both take constant across provinces. The final column (St St Conv) reports the speed of convergence for each province around its steady state, β_i^*, which is implied by the parameter estimates. The parameter estimates are obtained using maximum-likelihood for the SUR system of real GDP in 1990 prices and population. Asterisks (*/**/***) indicate significance at 10%/5%/1% level.

Full Dynamic System Estimates of Stochastic Solow Model							
		Province i estimates			Country		St St Conv
		σ_i	n_i	η_i	α	δ	β_i^*
1	Beijing	0.1233***	0.0243***	0.0388***			0.0078***
2	Tianjin	0.0890***	0.0172***	0.0182***			0.0095***
3	Hebei	0.0718***	0.0107***	0.0074***			0.0113***
4	Shanxi	0.1065***	0.0112***	0.0086***			0.0111***
5	Inner Mongolia	0.1372***	0.0088***	0.0055***			0.0118***
6	Liaoning	0.1006***	0.0074***	0.0071***			0.0122***
7	Jilin	0.0824***	0.0075***	0.0129***			0.0122***
8	Heilongjiang	0.0867***	0.0061***	0.0068***			0.0126***
9	Shanghai	0.1097***	0.0188**	0.0445***			0.0095***
10	Jiangsu	0.0962***	0.0091***	0.0109***			0.0117***
11	Zhejiang	0.0851***	0.0105***	0.0100***			0.0113***
12	Anhui	0.0744***	0.0079***	0.0180***			0.0121***
13	Fujian	0.0723***	0.0126***	0.0108***			0.0107***
14	Jiangxi	0.0656***	0.0105***	0.0036***			0.0113***
15	Shandong	0.0754***	0.0090***	0.0062***			0.0117***
16	Henan	0.0805***	0.0092***	0.0131***			0.0117***
17	Hubei	0.0915***	0.0070***	0.0131***			0.0124***
18	Hunan	0.0704***	0.0068***	0.0139***			0.0124***
19	Guangdong	0.0947***	0.0214***	0.0411***			0.0087***
20	Guangxi	0.0755***	0.0111***	0.0179***			0.0112***
21	Sichuan	0.0800***	0.0042*	0.0157***			0.0132***
22	Guizhou	0.0646***	0.0109***	0.0174***			0.0113***
23	Yunnan	0.0875***	0.0126***	0.0107***			0.0107***
25	Shaanxi	0.0751***	0.0099***	0.0058***			0.0115***
26	Gansu	0.0802***	0.0110***	0.0068***			0.0112***
27	Qinghai	0.0844***	0.0136***	0.0155***			0.0105***
28	Ningxia	0.0744***	0.0177***	0.0063***			0.0092***
29	Xinjiang	0.0778***	0.0181***	0.0160***			0.0092***
	Country-wide				0.714***	0.05	

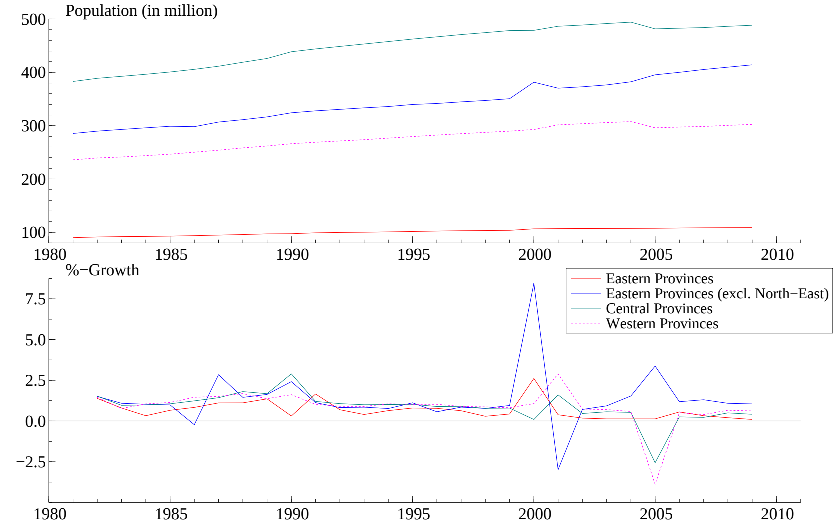
Figure 1: China Provincial GDP Data

The figure reports plots of our data of the provinces in our data set over 1980-2009. We consider 28 provinces, where we include Chongqing in Sichuan and Hainan in Guangzhou and ignore Tibet due to lack of data. For expositional purposes we aggregate the data for these figures in three regions: East (with as subset East* the Eastern provinces excluding the North-East), Central and West. We show the time series of real GDP per capita (in 1990 prices) and growth thereof (both in Panel A), and population (in millions) and population growth (Panel B). In Panel C we show the real investment (top plot), capital (middle plot) as constructed from the real investment series (both in billions), and the ratio of investment to capital (bottom plot).

(A) Real GDP per capita



(B) Population (millions)



(C) Real Investment and Capital

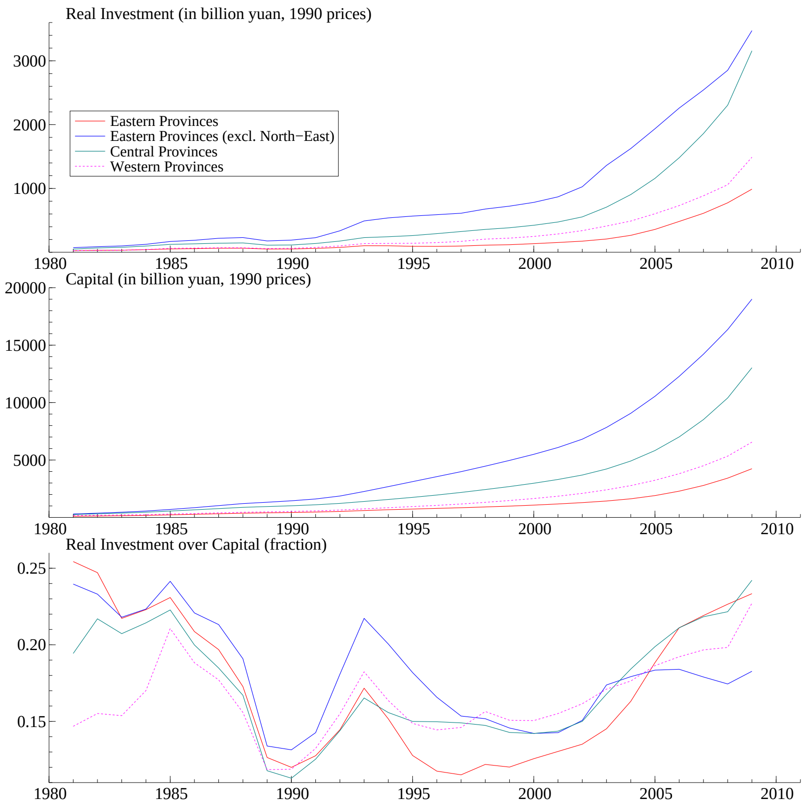


Figure 2: Absolute β-convergence

The figure plots absolute/unconditional β-convergence in China. On the y-axis is the log-growth rate of real GDP per capita from 1980-2009, and on the x-axis is log real GDP per capita in 1980. In addition to the top plot for all Chinese provinces, we split the provinces in three groups: East (with as subset East* the Eastern provinces excluding the North-East), Central, and West.

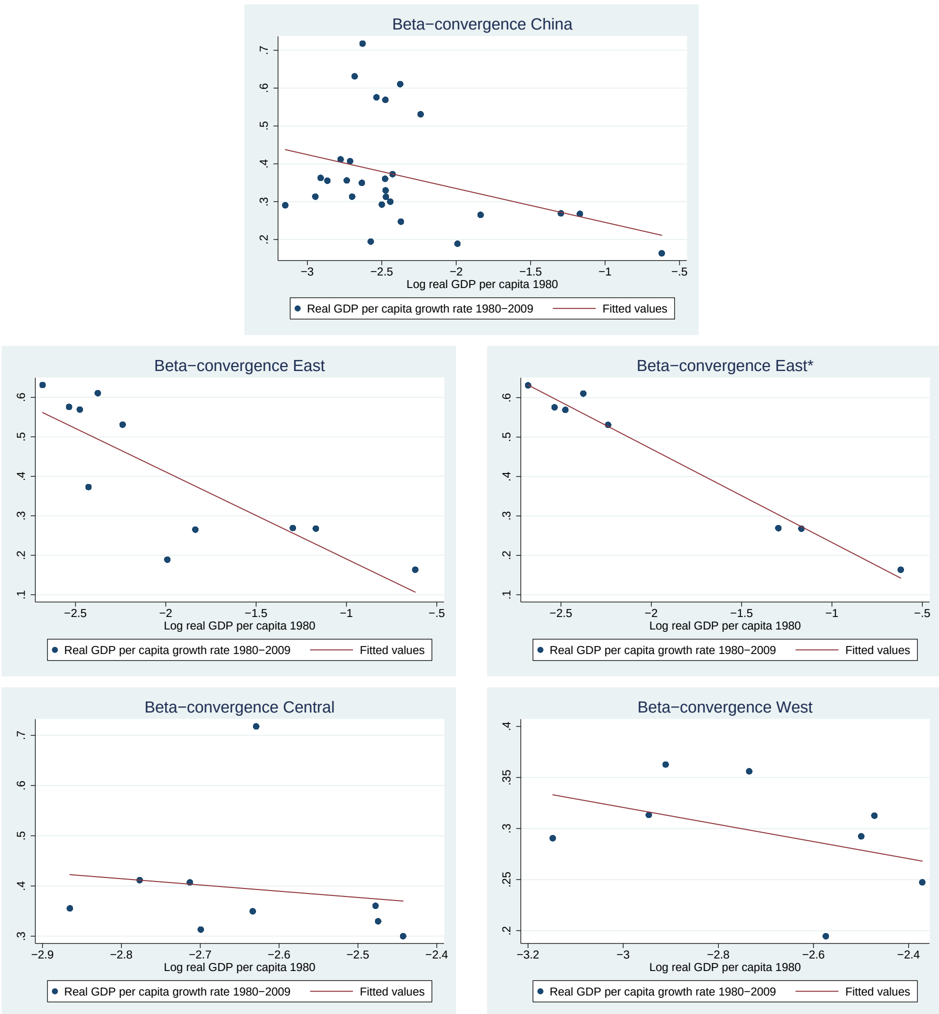


Figure 3: σ-convergence

The figure plots σ-convergence in China. For each year, we calculate the cross-sectional standard deviation of real GDP per capita divided by the cross-sectional mean of real GDP per capita. On the y-axis we plot this series, and on the x-axis time. In addition to the top plot for all Chinese provinces, we split the provinces in three groups: East (with as subset East* the Eastern provinces excluding the North-East), Central, and West.

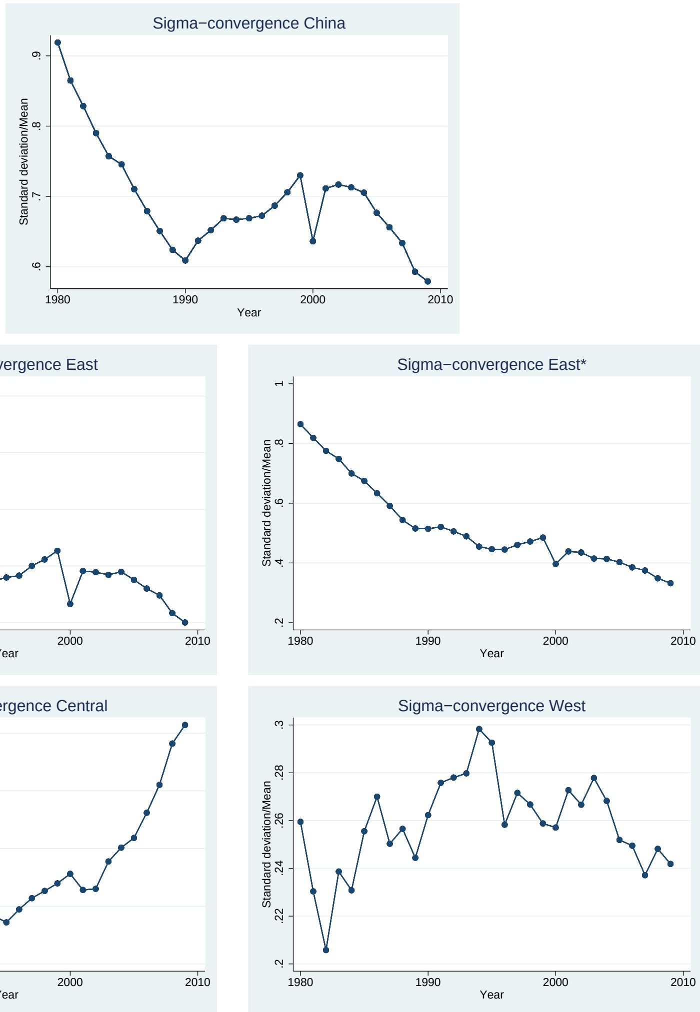


Figure 4: Speed of Convergence exploiting Full Dynamic System

The figure reports evidence of convergence in China exploiting the full dynamic structure of the stochastic Solow model. We consider 28 provinces, where we include Chongqing in Sichuan and Hainan in Guangzhou and ignore Tibet due to lack of data. For each province we plot the time series of the speed of convergence β_{it} as implied by the parameter estimates. In addition, we report the speed of convergence for each province around its steady state, denoted with β_i^*. For both we also show 95% confidence bounds.

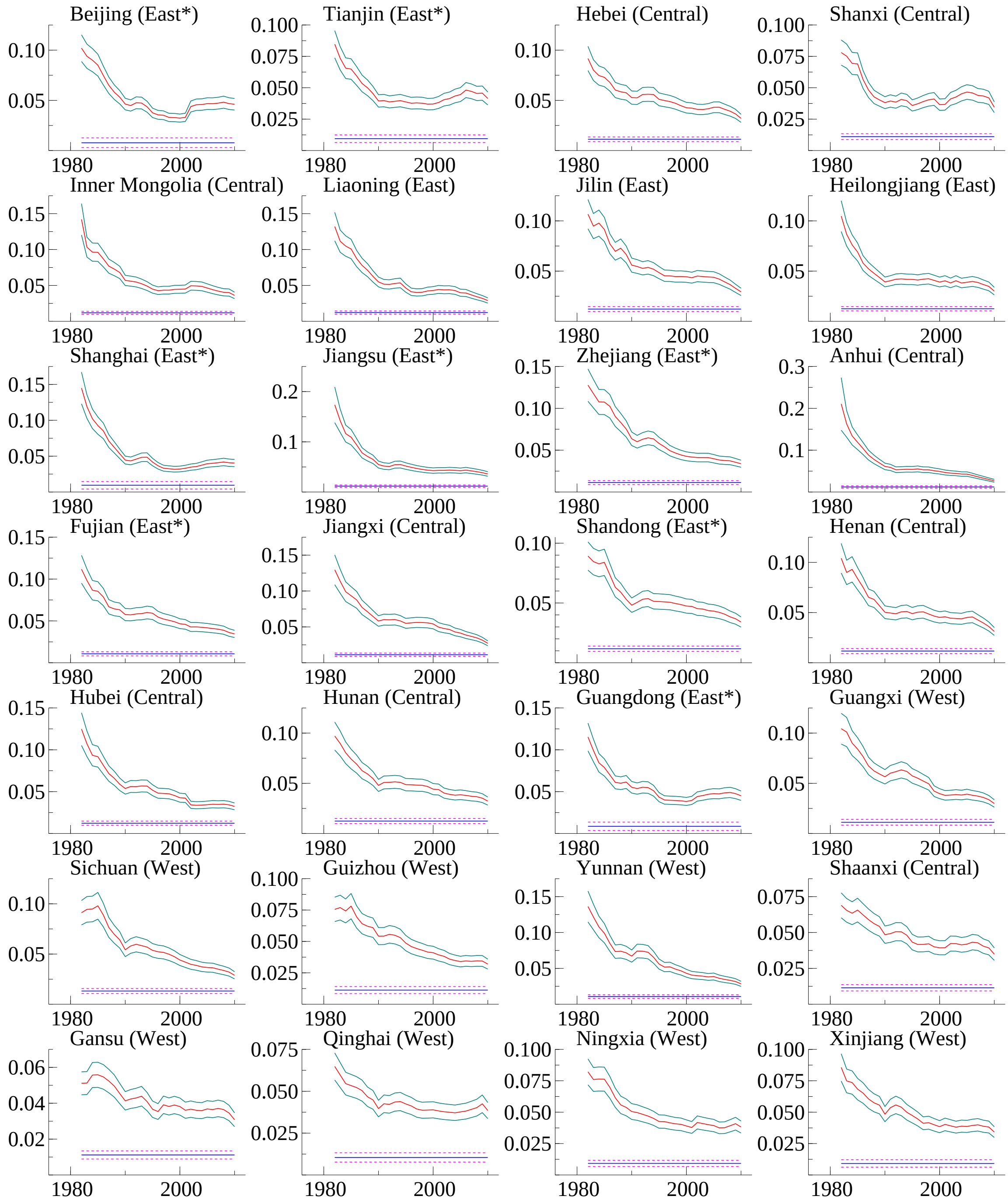


Figure 5: Convergence Rate in 2009 and Steady-State Convergence Rate

The figure shows the differences between the convergence rate in 2009 and the steady-state convergence rate for the Chinese provinces. We consider 28 provinces, where we include Chongqing in Sichuan and Hainan in Guangzhou and ignore Tibet due to lack of data. For each province we plot the real GDP per capita in 1980 against both the convergence rate in 2009 from our stochastic Solow model (triangles) and the steady-state speed of convergence (dots). A fitted line is included for both.