Is China’s Rising Service Sector Leading to Cost Disease?

Duo Qin *

Economics and Research Department
Asian Development Bank &
Department of Economics
Queen Mary, University of London, UK

December 2003

ABSTRACT
The focal issue of this investigation is whether the rising service sector would contract cost disease and stagnate economic growth in People’s Republic of China (PRC), as diagnosed generally of tertiarization in Baumol’s unbalanced growth model. Using provincial panel data for 1985-2001, the paper examines how labour productivity of the service sector has been associated with GDP growth and whether the service sector has shown signs of cost disease. The key findings are: the currently positive contribution of the service sector to growth is largely due to shifts of labour from agriculture into services; however, signs of cost disease are discernible from demand for service products, wage determination and labour input demand of the service sector.

JEL: O41, O53, L80, E24, C23, P23

Keywords: tertiarization, cost disease, labour productivity

* Contacting address: Macroeconomics and Finance Research Division, ERD, Asian Development bank, 6 ADB Avenue, Mandaluyong City, 0401 MM, PO Box 789, 0980 Manila, Philippines; Fax: 632 6322444; Email: duoqin@adb.org
I. Introduction

The last two decades have witnessed rapid economic development in the People’s Republic of China (PRC). The development is accompanied by significant structural transformation in various aspects of the economy. In particular, industrialization and urbanization has stimulated enormous demand for services. Growth in services has been phenomenal since 1990, as the average income level soared from around $200(US) per annum to $1,000(US) in 2002 in terms of per capita GDP. The real take-off of a services economy is expected, however, to occur in PRC in the near future on two accounts. First, tertiarization or emergence of a ‘services economy’ is widely recognised as a natural process for a developing economy to evolve into a developed or post-industrial economy (see, for example, Illeris 1996 and Sundrum 1990). Second, promotion of “tertiary industries” has recently been designated to a prominent place in the Tenth 5-year Plan (2001-05) of the PRC.¹

The government strategy of promoting tertiary industries is primarily aimed at alleviating the acute labour supply pressure at relatively low expenses of material and capital resources in the PRC. Moreover, it is felt that growing service industries would help strengthening the non-state-owned business sector and therefore increase the overall economic efficiency, and that they would also help accelerating trade and technological progress, as generally found in developing economies, (e.g. see Riddle 1986 and Mattoo et al 2001). Yet, little has been discussed about the possible downside of the strategy.

One influential theory about the negative effect of tertiarization on growth is Baumol’s unbalanced growth model (1967), see also (Baumol et al 1985). The model maintains that tertiarization is likely to decelerate overall productivity improvement of an economy and hence stagnate economic growth. This is because growth of the service sector is typically characterised by faster expansion of service employment than of service output, due mainly to the labour-intensive nature of the sector, relative to the agricultural and the manufacturing sectors. Such effect is often referred to as ‘cost disease’, as tertiarization is seen as dominantly demand-driven to the extent that demand for services becomes price inelastic and encourages the wage level of the service sector to grow faster than its labour productivity growth.

¹ According to the plan, the share of the service sector in GDP is to rise to 36% by 2005 from 33% in 2000; employment of the service sector is to increase 4% per annum on average, see (Li 2001).
The present study investigates empirically how much and in what ways the service sector absorbs labour and contributes to the overall economic growth, and whether there are discernible signs of cost disease in the PRC. Panel data of 30 provinces\(^2\) for the period of 1985-2001 are used for the purpose. The investigation starts from an overview, in the following section, of the Chinese service sector in relation to the overall economic growth. In section 3, two types of decomposition methods are employed to examine how labour productivity of the service sector has been associated with GDP growth via its association to the growth of the primary and the secondary sectors, frequently known as the agricultural sector and the manufacturing sector. Econometric analysis is carried out in Section 4 to detect signs of cost disease characterised in Baumol's unbalanced growth model. Section 5 concludes. It is well-known that statistical measures of the service sector are full of difficulties and there is significantly downward mis-measurement in service output and productivity (e.g. see Wolff 1999; Triplett and Bosworth 2001). However, the present study shall not go into these measurement issues.

II. An Overview\(^3\)

The service sector of the PRC expanded at an average annual growth rate of over 9% during the 1990s. The expansion is accompanied by rapid income growth, as reflected in the average annual rate of 8.3% in terms of per capita GDP. The rising service sector has particularly helped to provide employment opportunities for the abundant Chinese labour market, with the current service share being 28% of the total employment, more than doubled the 1980’s figure of 13%, as shown in Figure 1. The increase apparently comes mostly from the agricultural sector, which has dwindled from nearly 70% in 1980 to 50% at the present in terms of its employment share. One of the most striking aspects, however, lies with the output share changes of the service sector. The output share of the service sector in nominal terms rises significantly from one fifth roughly in 1980 to one third currently, but the output share in real terms remains almost constant, as shown in Figure 1. The reverse is found from the secondary sector. Its nominal output share remains relatively

\(^{2}\) Beijing, Tianjin and Shanghai are counted as provinces, but Chongqing, the new autonomous municipality, is still regarded as part of Sichuan in our sample. A list of the province abbreviations is given in the appendix.

\(^{3}\) In order to focus on the role of services to the whole economy in aggregate, we ignore the compositional changes, or the heterogeneous development, within the service sector in this paper.
unchanged while its share in constant prices rises. The output share of the primary sector has decreased in both nominal and constant terms. These phenomena suggest that there is marked disparity in the movement of the sectoral prices.

[Figure 1 about here]

In Figure 2, the service sector shares in nominal GDP and in employment are plotted at the provincial level. The provinces are ranked by per capita GDP of the 2001 statistics. The richer a province, the higher its shares of the service sector. The nominal output shares are below 40% for most of the provinces. Only Beijing and Shanghai exceed 50%, just about comparable to the level of advanced economies. Employment shares are remarkably low for inland and poor provinces. It is clear that China is still at the very beginning of tertiarization and that there is great potential for employment expansion in the service sector.

[Figure 2 about here]

Figure 3 presents a summary view of the labour productivity, wage rates and labour costs of the three sectors. The real output per labour in both the primary and the tertiary sectors is shown to have hardly grown over the last two decades, in great contrast with the secondary sector (the top left panel). The output prices of the primary and the tertiary sectors have significantly outgrown that of the secondary sector, with the price of the tertiary sector getting in the lead towards the end of the 1990s (the bottom left panel). The nominal wage level of the service sector has grown in pairs with that of the secondary sector, gradually outpaced it since mid 1990s, and induced a rising gap in real unit labour cost between the two sectors (the right panels).4 These phenomena remind us clearly of cost disease. However, it is the agricultural sector, rather than the manufacturing sector, which supplies abundant labour to the service sector, as shown in Figure 1 and confirmed by the persistently high level of real unit agricultural labour cost in Figure 3. Notice that the wage data used here may have significantly overestimate the per worker income in the agricultural sector. Nevertheless, low agricultural labour productivity is a widely acknowledged fact in the PRC, see, e.g. Yang and Zhou (1999). There are discernible signs of slowdown in the secondary sector growth, as shown from the stagnant shares of its employment and nominal output since the mid 1990s (see Figure 1).

4 If \( Y \) denotes real output, \( P \) output price, \( L \) employment and \( w \) wage rate, real unit labour cost is calculated by: \( \frac{wL}{PY} \).
To further examine the extent of the gap in real unit labour cost between the service and the manufacturing sectors, we plot the ratio of the two labour cost series using provincial data in Figure 4. Interestingly, there is no obviously upward trend for the three large cosmopolitan cities, i.e. Beijing, Shanghai and Tianjin, but the trend is discernible for most provinces. In terms of the absolute gap, real unit labour cost in the service sector has reached about twice of that of the manufacturing sector in those relatively developed provinces, whereas it has exceeded twice and even reached three times in some inland and poor provinces. These suggest that the gap is likely to decrease in the long run as the economy further develops, but that the real unit labour cost of the service sector is likely remain higher that that of the manufacturing sector.

The graphed evidence shows us that growth of the service sector in the PRC occurs mainly in its nominal output share and employment share at the expense of the agricultural shares. The gap between nominal and real output shares of the service sector is apparently induced by rapidly rising prices of services as well as the rising wage bills of the sector. As labour cost is closely related to labour productivity, we shall focus our attention on the contribution of services to GDP via labour productivity in the next section.

III. Labour productivity and Economic Growth

In order to assess how the service sector contributes to the national level of labour productivity, two decomposition methods are applied to the panel dataset in this section. The first follows the spirit of the decomposition method proposed by Berman et al (1994). The second method is developed recently by Nordhaus (2002a; 2002b).

Denote real GDP by $Y$ and its three sectoral real output by $Y_i$, i.e.:

$$Y_i = \sum_{i=1}^{3} Y_{it}$$

where sector one is the primary sector, sector two the manufacturing sector and sector three the service sector. Denote labour productivity by $q = \frac{Y}{L}$ where $L$ represents employment. We have:

$$q_t = \frac{Y_{1t} + Y_{2t} + Y_{3t}}{L_t} = \sum_{i=1}^{3} y_i = \sum_{i=1}^{3} q_i l_{it} \quad \sum_{i=1}^{3} l_{it} = 1$$
where \( y_i = \frac{Y_i}{L} \) can be viewed as representing approximately per capita real sector output,

\[ q_i = \frac{Y_i}{L} \]

denotes sector labour productivity and \( l_i = \frac{L_i}{L} \) denotes employment share. The changes or growth of \( q_t \) in (2) should be:

\[ \Delta q_t = \sum_{i=1}^{3} \Delta q_{it} l_{it} + \sum_{i=1}^{3} \Delta l_{it} q_{it}, \quad \Delta q_t = q_t - q_{t-1} \], or

\[ q_t = \sum_{i=1}^{3} q_{it} l_{it} \left( \frac{q_t}{q_{t-1}} \right) + \sum_{i=1}^{3} \Delta l_{it} \frac{q_{it}}{q_{it-1}}, \quad q_t = \frac{q_t}{q_{t-1}} - 1 \]

Let us refer to as ‘within effects’ the components in the first summation term on the right-hand side of (3) or (3′) and as ‘between effects’ the components in the second summation term, following the terminology of Berman et al (1994). Figure 5 illustrates the time series of the six components of (3′) under the heading of ‘within’ and ‘between’ effects. The time series are calculated at both the national and regional levels. The national data set covers 1980-2001 whereas the regional set covers 1985-2001.5

5 The categorisation of all the provinces into three regions follows mainly that by the National Bureau of Statistics of China. See the Appendix for the categorisation.
In order to focus on the role of the service sector, let us evaluate the changes of $q_t$ by the changes of $y_3$, which can be regarded approximately as per capita consumption or demand of services. Utilising $\sum_{i=1}^{3} l_{it} = 1$, we have:

\[
\frac{\Delta q_t}{\Delta y_{3t}} = \sum_{i=1}^{2} \frac{\Delta q_{it}}{\Delta y_{3t}} l_{it} + \sum_{i=1}^{2} \frac{\Delta l_{it}}{\Delta y_{3t}} q_{it} + 1
\]

Again, we refer to as ‘within effects’ components in first the summation term and as ‘between effects’ components in the second summation term on the right-hand side of (4). Figure 6 shows the time series of these four components under the heading of ‘within’ and ‘between’ effects, together with the time series of $\frac{\Delta q_t}{\Delta y_{3t}}$, denoted as the total effect.

It is reassuring that the service sector has made continuously positive contribution to the national labour productivity growth, especially during the first half of the 1990s. The contribution is again dominantly through the ‘within’ effects, especially through the within effect on the manufacturing sector. This serves as supporting evidence to the theory by Oulton (2001), namely that the dooming effect of cost disease may disappear if the contribution of services, used as vital intermediate inputs, to the sectors of high productivity growth is taken into consideration. The present evidence shows that intermediate services input helps to improve the productivity of the manufacturing industries. The regional evidence on the ‘within’ effects also shows a significantly leading role of the coastal region and a catching-up trend of the inland regions during the latter part of the 1990s. Similar to Figure 5, ‘between’ effects are very small and turned to negative on the manufacturing sector in the late 1990s. This calls our attention to the question of what the overall effect is of a rapidly growing service sector on the national labour productivity via employment redistribution.

An intuitive way of measuring such effect is developed by Nordhaus (2002a; 2002b). Instead of decomposing the growth effects of labour productivity by the accounting identity of GDP in (1) and (2), Nordhaus proposes to decompose labour productivity in accordance with welfare theory. Specifically, he decomposes the aggregate productivity growth into four types of effects: the pure productivity effect, the Baumol effect (i.e. interaction between productivity changes and output share changes), the Denison effect (i.e. the effect of changing employment shares on productivity growth) and the fixed-weight drift term.
The four effects are denoted in sequence by the four terms on the right-hand side of the following equation:

\[
q_t = \sum_{i=1}^{3} q_is_i + \sum_{i=1}^{3} q_is_i(s_{it} - s_{i0}) + \sum_{i=1}^{3} L_i(s_{it} - l_{it}) + \sum_{i=1}^{3} Y_i(s_{it} - z_{it})
\]

where \( s_i = \frac{P_i Y_i}{P_Y} \) denotes the nominal output share of sector \( i \) (\( P_i \) denotes output price of sector \( i \)) and \( z_i = \frac{Y_i}{Y} \) the real share. Nordhaus shows that the sum of the first two terms, i.e. the pure productivity effect plus the Baumol effect, embodies the welfare gain in the aggregate productivity growth, \( \hat{q}_t \). Figure 7 gives the time series of \( \hat{q}_t \), the pure productivity effect and the Baumol effect calculated using the national and the regional data sets. The sectoral components of these effects are given in Figure 8.

[Figure 7 and Figure 8 about here]

It is discernible from Figure 7 that most of the total productivity growth, \( \hat{q}_t \), comes from welfare gain and that most of the welfare gain is brought about by pure productivity growth. The Baumol effect is very small in comparison and exhibits continuously negative values only for the western region. There is also a slight overall improvement in the Baumol effect towards the end of the 1990s. Detailed sectoral results in Figure 8 show that the manufacturing sector remains the major contributor to the pure productivity growth, and that the contribution by the service sector is on the rise as the contribution by the manufacturing sector levels off during the late 1990s. Again, the coastal region plays the leading role in the sectoral changes. Surprisingly, the service sector turns out to be the saviour of the negative Baumol effects caused mainly by the agricultural sector, especially during the late 1990s, as shown in Figure 8. These results seem to suggest that fast economic growth and rise of the service sector in the PRC have apparently been a cure for the cost disease.

What we find from Nordhaus’ decomposition method actually reinforces what we find from the first decomposition method, namely the service sector has contributed positively to GDP growth mainly via absorbing excess labour from the agricultural sector. However, we should be cautious in making inference from the above result to cost disease, as the decomposition methods are not designed to identify price elasticity of demand for services and cost effectiveness in service production, or to discover behavioural patterns relating to
IV. Signs of Cost Disease?

This section is devoted to econometric tests of three key propositions underlying Baumol's unbalanced growth model. The first proposition states that, as average income level rises in an economy, the income elasticity of aggregate demand for services increases and may well exceed unity making the price elasticity insensitive. The second states that the labour input of the service sector is primarily driven by the aggregate demand for services and is scarcely curbed by labour cost. The third states that the wage level of the service sector follows closely that of the sector with the fastest productivity growth, rather than being determined by its own relatively low productivity level. The three propositions can be formed into three testable structural equations: a demand equation for services, a wage equation of the service sector and a factor demand equation for labour input in the service sector.

However, simple estimation of these equations derived directly from the Baumol’s model has not produced satisfactory results, e.g. see (Curtis and Murthy 1998). This is mainly because the model contains a number of simplifying assumptions that are too stringent to fit data directly. Here, we generalise the model in several ways while maintaining the spirit of the three propositions. The first generalisation is to include all the three sectors of an economy, instead of the manufacturing and the service two sectors in the original model. The second is to adopt more general model forms than what was originally used. The third is to regard the theoretical equations as the long-run states of dynamically specified econometric equations in view of the time-series properties of data, and to choose the estimation methods accordingly. The final generalisation is to utilise regional differences in the panel data for the inference of future development trend.

Let us start with the demand equation for services by using a standard demand model to explain $y_3$, which can be regarded approximately as per capita demand for services:

$$y_3 = \Psi(t)P_1^\psi P_2^\psi P_3^\psi W^\psi \sum_{j=1}^{4} \psi_j = 0$$
where $W$ is nominal wage rate and represents approximately per capita income, $\Psi_0$ is a constant parameter and $\Psi(t)$ a time effect parameter.\(^6\) The condition of homogeneous of degree zero is assumed in the prices, $P_i$, and income, $W$ in equation (6). The first proposition amounts to hypothesizing $\psi_3 \geq 1$ and $\psi_3 \approx 0$ in (6).

Next, we propose a simple wage equation for $W_3$:

\begin{equation}
W_3 = W_1^{\beta_1} W_2^{\beta_2} (P_3 q_3)^{\beta_3}, \quad \sum_{j=1}^{3} \beta_j = 1
\end{equation}

where $W_i$ denotes nominal wage rate of sector $i$ and $B_0$ is constant parameter. Equation (7) approves of the second proposition if $\beta_2 = 1$ and $\beta_1 = \beta_3 = 0$. The equation becomes a standard wage equation under perfect competition when $\beta_1 = \beta_2 = 0$, and approves of Cobb-Douglas production technology if $\beta_3 = 1$.

As for the labour input demand equation, a linear homogeneous production function is adopted with constant elasticity of substitution (CES) technology. Moreover, constant returns to scale is assumed for the long-run equilibrium state. We thus have:

\begin{equation}
Y_3 = \Lambda(t) \left[ \alpha_L L_3^{\delta} + \alpha_K K_3^{\delta} \right]^{1/\delta}
\end{equation}

where the elasticity of substitution is $\sigma = \frac{1}{1-\delta}$ (when $\sigma = 1$, (8) converges to Cobb-Douglas technology), $\Lambda(t)$ is commonly used as an index of technological progress to represent total factor productivity (TFP), and $\alpha_L$ and $\alpha_K$ are the input share parameters for labour $L$ and capital $K$. The demand for labour input conditional upon cost minimisation of (8) should be:

\begin{equation}
L_3 = \alpha_L^{\sigma} \Lambda(t)^{-1} Y_3 w_3^{-\sigma} \quad \text{or} \quad l_3 = \alpha_L^{\sigma} \Lambda(t)^{-1} Y_3 w_3^{-\sigma}
\end{equation}

where $w_3 = \frac{W_3}{P_3}$ is often referred to as real wage rate (see, for example, Hamermesh 1986). The bottom equation in (9) has the advantage of representing the employment share of the service sector. Notice that $w_3$ is absent and $\Lambda(t)$ is assumed to grow very slowly in Baumol's labour demand equation, which is purely demand driven.

\(^6\) Baumol’s (1967) assumes $\Psi(t) = e^{At}$ to capture the growth rate at which services are commercialised.
In regard to the time-series properties of the panel data set, two types of econometric considerations are taken into here. One is on model specification and the other on estimation. On model specification, the principle of dynamic specification approach is adopted (see Hendry 1995, see also Nickell 1986 for dynamic models of labour demand). The approach enables us to regard equations (6), (7) and (9) as long-run equilibrium-correction models (ECM) embedded in dynamically adequate representations of the data generating processes. In view of the relatively short time series in the panel (17 observations), a first-order auto-regressive, distributed-lags model is used here as the general representation. This model can be easily reparameterised into an ECM to facilitate our focus on the long-run parameter estimates. The general format of a two-variable ECM using panel data is as follows:

\[ \Delta \ln X_{1jt} = \theta_0 + \theta_{0j} + \theta_t + \theta_1 \ln X_{1j(t-1)} + \theta_2 \ln X_{2j(t-1)} + \theta_3 \Delta \ln X_{2jt} + \epsilon_{jt} \]

(10)

\[ = \theta_0 + \theta_{0j} + \theta_t + \theta_1 \left[ \ln X_{1j} + \frac{\theta_2}{\theta_1} \ln X_{2j} \right]_{t-1} + \theta_3 \Delta \ln X_{2jt} + \epsilon_{jt} \]

\[ j = 1, \ldots, n \]

where \( \theta_{0j} \) denotes individual effect and \( \theta_t \) the time effect. It is expected \(-1 \leq \theta_t < 0\) for the embedded equilibrium-correction (EC) term (i.e. the term in the squared brackets) to take effect. The EC term implies a long-run equilibrium equation:

\[ \ln X_v = -\frac{\theta_2}{\theta_1} \ln X_{2j} \]

with \( -\frac{\theta_2}{\theta_1} \) being the key parameter of interest in the estimation of (10). To accommodate to (10), equations (6), (7) and (9) are re-written into the following long-run forms:

(6') \[ \ln y_{3j} = \ln \Psi(t) + \psi_2 \ln \left( \frac{P_3}{P_1} \right) + \psi_3 \ln \left( \frac{P_3}{P_1} \right) + \psi_4 \ln \left( \frac{W}{P_1} \right) \]

(7') \[ \ln W_{3j} = \beta_1 \ln W_{1j} + \beta_2 \ln W_{2j} + \beta_3 \ln (P_{3j}q_{3j}) \]

(9') \[ \ln \frac{l_{3j}}{y_{3j}} = \sigma \ln \alpha_0 - \ln \Lambda(t) - \sigma \ln w_{3j} \]

However, it is well-known that the parameter estimates of a long-run EC equation are likely to be afflicted by collinearity when the corresponding variables are significantly trended. This makes us concerned particularly about equation (6) since the strong trend in \( (W/P_1) \) would be collinear with \( \Psi(t) \). On reflection, it is questionable whether it is appropriate to use \( \Psi(t) \) in the long-run EC term since the rate of commercialisation of the service sector, which \( \Psi(t) \) is to represent, should be a transitory feature after all. In the
context of a dynamic model, such rate is already embodied in short-run growth variables, such as \( \Delta \ln X_{2,jt} \) in (10). Therefore, the term \( \Psi(t) \) in (6’) is to be ignored in the dynamic model specification. On the other hand, we also try to incorporate the transitory contribution of services to other sectors as intermediate inputs, as postulated by Oulton (2001), into the dynamic ECM based on (6’). More precisely, we introduce \( \Delta \ln Y_{2t} \) as an additional short-run variable.

In regard to panel data information, we can also exploit regional differences to get a longer term perspective than the 16-year sample period. In other words, we can estimate different regions separately such that the results from more advanced regions can be used as indicating future developments for the poorer regions. This amounts to relaxing the homogeneous parameter restrictions in (10). On the basis of the three-region division, we have:

\[
\Delta \ln X_{1,jt} = \theta_0 + \theta_{0j} + \theta_j + \sum_{k=1}^{3} \theta_{jk} \ln X_{1,jk} + \frac{\theta_{2k}}{\theta_{jk}} \ln X_{2,jk} + \sum_{k=1}^{3} \theta_{3k} \Delta \ln X_{2,jkt} + \varepsilon_{jt}
\]

\( j = 1, \cdots, n \)

We refer, thereafter, to (10) as the full panel model and (10’) as the regional model.

As for estimation, our key objective is to estimate the long-run parameters of interest in equations (6’), (7’) and (9’). The parameters are estimated by three methods, considering the fact that the time series of some variables in these equations are likely to contain roots near to unity. The first is simply to estimate static regression models based on (6’), (7’) and (9’) by the feasible GLS (Generalised Least Squares) method. This method should give us fairly consistent estimates of the long-run parameters when the variables in the static models are nonstationary and cointegrated, as the case is similar to the OLS (ordinary least squares) estimation of a static model of time-series variables which are nonstationary and cointegrated, see (Engle and Granger 1987). The second method is to derive the long-run solution from the GLS estimates of the first-order auto-regressive, distributed-lags model of equations (6’), (7’) and (9’). The third method is to estimate the above dynamic model in the ECM format of (10) by the GMM (Generalised Method of Moments) or combined GMM method, see (Arellano and Bover 1995) and also (Blundell and Bond 1998). The estimation is normally carried out on the first line of (10) and the long-run parameters are then derived from the estimates of \( \theta_j \) and \( \theta_{2j} \). GMM estimates should be more efficient in principle, but they may suffer from poor robustness and weak instruments if some of the regressors exhibit evident nonstationarity, see (Arellano and Honoré 2000).
and (Phillips and Moon 2000). The three sets of parameter estimates are considered in combination in our inference.

In order to check parameter constancy, estimation is run for four different sample sizes: full sample (i.e. 1985-2001), 1988-2001, 1990-2001 and 1992-2001. The three sets of the long-run parameter estimates of \((6')\), \((7')\) and \((9')\) are reported respectively in Tables 1, 3 and 5. Short-run variables remaining significant through the dynamic specification are listed in these tables but their parameter estimates are not reported. Key diagnostic test results of the GMM estimation of the three equations are reported in Tables 2, 4 and 6 respectively. These include Sargan test for the validity of over-identifying instruments and residual autocorrelation tests. It is extremely difficult, as expected and shown in tables 4 and 6, to get valid instruments when the equation in estimation contains variables with evident nonstationarity, such as the nominal wage series in \((7')\). The GMM estimates vary considerably with different instruments under such circumstance. We should hence interpret these estimates with caution. Notice that the residual autocorrelation test results can be seen as validating the first-order dynamic model.\(^7\)

Let us now look at the key estimation results. We see, from the demand equation for \(y_3\) in Table 1, that the hypothesis \(\psi_3 \approx 0\) is confirmed, as this parameter drops out of the model for being statistically insignificant. Chinese consumers are indeed insensitive to prices of service products, in spite of the rapid price hike, as shown in Figure 3. Meanwhile, the significantly negative cross price elasticity \(\psi_2\) indicates a strong price effect of manufactured goods. However, this price effect would not help much to dampen demand for \(y_3\) since the growth of \(P_2\) has remained the slowest of the three sectors (see Figure 3). In fact, \(P_2\) is likely to stimulate \(y_3\) as the prices of many manufactured goods have been coming down. On the other hand, we see from Figure 3 that \(P_1\) has been rising fast in line with \(P_3\). But its cross price elasticity \(\psi_1\) is found to be small and decreasing. If we look at the estimates of the income elasticity, we fail to find enough evidence for the hypothesis \(\psi_4 \geq 1\). The elasticity looks likely to be around unity. More interestingly and contrary to common expectation, the elasticity estimates decrease slightly as the economy develops, as shown from both the sub-sample and the regional results. The full sample estimates are larger than the sub-sample estimates. The estimates of the poorest region of the west are

\(^7\) Significant first-order serial correlation is an expected feature of the GMM method, see Doornik and Hendry (2001, Chapter 7, vol. 3) for details of the residual autocorrelation test.
larger than those of the richest coastal region. A feasible explanation is that severely suppressed demand for services under the old economic regime was rapidly released by economic reforms in the PRC and that the long-run elasticity converges to unit once this transitory regime shift was over. The dynamic estimation results also confirm the earlier argument that commercialisation of services is virtually transitory by nature. Two positive short-run factors are identified in this respect: an income effect represented by the income growth rate, $\Delta \ln W_t$ (its parameter estimates are in the range of 0.11 ~ 0.19), and a between-sector effect represented by the growth rate of the manufacturing sector, $\Delta \ln Y_{2t}$ (its parameter estimate is around 0.2). The latter is in supportive of Oulton’s model (2001) to some extent. In short, income is found to accelerate demand for services only in a short-run manner, whereas the price factors are found to be the main culprit for failing to curb the rising demand in the long run.

[Tables 1 and 2 about here]

Let us now turn to the wage equation, as shown in Table 3. Much of the evidence here is in supportive of Baumol’s third proposition. The parameter of the productivity of the service sector, $\beta_3$, is insignificant throughout. $W_3$ is dominantly determined by $W_2$ in the long run. The estimates of $\beta_2$ are fairly uniform across regions as well as over time with $\hat{\beta}_2 \approx 1.8$ In addition, we find that the growth rate, $\Delta \ln W_{2t}$, exerts a strong and positive impact on $W_3$ (its parameter estimates are in the range of 0.7~1.5). The finding tells us that there is a high degree of wage rigidity irrespectively of labour productivity in the service sector. This is not very surprising since output prices of service products do not form a serious constraint to deter consumers.

[Tables 3 and 4 about here]

Finally, let us examine the result of the service labour demand equation (9'). Notice that constant-returns-to-scale is imposed in (9). This enables us to estimate the technological progress factor, $\Lambda(t)$, without worrying about collinearity between $\Lambda(t)$ and $y_3$. The imposition unfortunately makes it impossible to perform the second estimation method. Table 5 reports the rest two types of estimation results. During the GMM estimation, the significance of the negative feedback parameter, $\theta_1$, for variable $\ln(l_3 / y_3)_{t-1}$ can serve as

---

8 It is noticeable from Tables 3 and 4 that parameter estimates based on the dynamic specification are not very robust and the GMM lacks good instruments. This is due to the problem of evident nonstationarity in nominal wage series, as stated earlier.
an indirect verification of the constant-returns-to-scale assumption (this parameter takes the value around –0.17 in the first two samples and around –0.37 in the last two samples in the full-panel model estimation). As for the estimated Λ(t), the GLS estimates and the GMM estimates deliver considerably different patterns, as shown from the two plots in Table 5. But both patterns show fairly slow and cyclical movement, confirming Baumol’s proposition that technical progress in the service sector is too slow to help improving labour productivity.9 What contradicts Baumol’s model is the labour cost variable, via the significance of σ, albeit not very robust as some of the GMM estimates have fairly low significance levels, as shown in Table 5. Nevertheless, real wage rate exerts a small role in curbing labour input. Interestingly, its parameter estimates show a slight declining tendency both over time as well as from the more developed region to the less developed region. This seems to reflect a mixture of development effects. On the one hand, economic development via marketisation appears to gradually shape the service sector to become receptive to labour cost signals from the supply side; on the other, such a cost constraint appears to be diluted by the rapidly growing need for service products from the demand side.

[Tables 5 and 6 about here]

V Reflections and Summary

The PRC is still at the very beginning of tertiarization. A ‘services economy’ is yet to occur with respect to rapidly rising income, commercialisation of services, abundant excess labour supply, and adequately developed capacity in both manufacturing and agricultural sectors at present in comparison to many service industries.

The shift of labour shares from agriculture into services has so far kept the service sector in positive contribution to the overall labour productivity and growth of the Chinese economy. But plenty of evidence is found in this investigation to show that pricing in the service sector, in terms of both output and labour input, already poses significant risk of cost disease to the whole economy. The rapidly growing economy has stimulated enough demand for services to disable the own price effect to keep the demand in check. This

9 It is shown in some literature that technical progress has recently accelerated in the service sector with the rapid expansion of the ‘new economy’, i.e. the information, communication and computing industries (e.g. see Temple 2002). However, it is not yet clear that the progress with make net contribution to labour cost reduction, since the new economy normally enjoys a relatively high wage rates.
encourages nominal wage rigidity and has somehow allowed the wage level of the service sector to outpace the most productive sector of manufacturing industries. As a result, real wage cost forms too weak a constraint for the Chinese service producers in optimising their production.

Our findings carry important policy implications. While it is a viable strategy to encourage development of the service sector and a service-led growth path in future, the strategy is unlikely to materialise without more carefully designed policies in place. The positive contribution of the service sector to overall growth may not sustain by itself. The net labour productivity gain will disappear as soon as the shift of labour shares from the agricultural sector to services is replaced by the shift from the secondary sector to services, as is expected normally of substantial tertiarization in a developed economy. It is thus imperative to introduce policies to induce the service sector to be economically more efficient before it takes off in surpassing the secondary sector.

Our findings suggest that the key guidance of such policy design should be encouragement of competition and market efficiency of the service industries. A recent five-country study by Van Ark et al (1999) has verified that deregulation of markets plays an important role in promoting productivity in services. Since there are not enough pricing signals to check demand for services and for service labour, policies are particularly wanted in the realm of wage and employment determination. Specifically, these policies should aim to remove those institutional arrangements in protection of nominal wage rigidity and to stimulate real wage cost response from service producers. Considering the abundant labour supply in the PRC, there is no reason for service producers not to be receptive to such policies. In reality however, design and implementation of such policies can be extremely challenging politically, as a sizeable and the leading part of the service sector is still dominantly state-owned, such as banking and finance, education and culture, social services, telecommunication, health care and government administration.
Appendix

Data sources:

Variable definition:

\[ Y_{ij}: \] value-added output of sector \( i \) in province \( j \), adjusted to constant price by \( P_{ij} \)

\[ (i = 1,2,3 \quad j = 1,\ldots,30) \]

\[ Y_j: \] constant price GDP of province \( j \)

\[ P_{ij}: \] output price indices of sector \( i \) in province \( j \) (the indices are derived from indices of annual GDP growth data and nominal GDP data by sector)

\[ L_{ij}: \] employment of sector \( i \) in province \( j \)

\[ W_{ij}: \] average wage per worker in sector \( i \) in province \( j \) from the table of Average Wage of Staff and Workers by Sector (\( W_1 \) uses the category of farming, forestry animal husbandry and fishery; \( W_2 \) uses the category of manufacturing; \( W_3 \) is the employment weighted average of nine categories: transport and telecommunication, trade and catering services, banking and insurance, real estate trade, social services, health care, education and culture, scientific research, government agencies)

Abbreviation of provinces by region:

<table>
<thead>
<tr>
<th>Coastal region</th>
<th>Central region</th>
<th>Western region</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ Beijing</td>
<td>SX Shanxi</td>
<td>SC Sichuan</td>
</tr>
<tr>
<td>TJ Tianjin</td>
<td>NM Inner Mongolia</td>
<td>GZ Guizhou</td>
</tr>
<tr>
<td>HB Hebei</td>
<td>JL Jilin</td>
<td>YN Yunnan</td>
</tr>
<tr>
<td>LN Liaoning</td>
<td>HLJ Heilongjiang</td>
<td>XZ Tibet</td>
</tr>
<tr>
<td>SH Shanghai</td>
<td>AH Anhui</td>
<td>SHX Shaanxi</td>
</tr>
<tr>
<td>JS Jiangsu</td>
<td>JX Jiangxi</td>
<td>GS Gansu</td>
</tr>
<tr>
<td>ZJ Zhejiang</td>
<td>HN Henan</td>
<td>QH Qinghai</td>
</tr>
<tr>
<td>FJ Fujian</td>
<td>HUB Hubei</td>
<td>NX Ningxia</td>
</tr>
<tr>
<td>SD Shandong</td>
<td>HUN Hunan</td>
<td>XJ Xinjiang</td>
</tr>
<tr>
<td>GD Guangdong</td>
<td>GX Guangxi</td>
<td></td>
</tr>
<tr>
<td>HAN Hainan</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


Nordhaus, William D. (2002a) Alternative methods for measuring productivity growth including approaches when output is measured with chain indexes, Yale University.
Figure 1. Output, employment and their composition

Per capita GDP (yuan, top curve) and its composition in nominal price (Top area: service sector; Middle: industrial sector; Bottom: primary sector)

Sector shares in GDP in nominal price (Top area: service sector; Middle: industrial sector; Bottom: primary sector)

Per capita GDP (yuan, top curve) and its composition in 1980 constant price (Top area: service sector; Middle: industrial sector; Bottom: primary sector)

Sector shares in GDP in 1980 constant price (Top area: service sector; Middle: industrial sector; Bottom: primary sector)

Employment (10,000) and its composition (Top area: service sector; Middle: industrial sector; Bottom: primary sector)

Employment share (Top area: service sector; Middle: industrial sector; Bottom: primary sector)
Figure 2. Shares of the service sector in GDP (solid line) and in employment (dotted line) by province

Note: The provinces are ordered by per capita GDP in 2001 by row. Notice that units of the vertical axis are varied.
Figure 3. Labour productivity, wage rate and labour cost

Note: Real unit labour cost is defined as the product of nominal wage rate and employment divided by nominal output.
Figure 4. Ratio of real unit labour costs of the third sector to the second sector

Note: The horizontal axis marks at 1, where the unit costs of the two sectors are identical; the units of the vertical axes may differ across provinces. See the note in Figure 3 for the definition of the real unit costs; See the appendix for definitions of province acronyms.
Figure 5. Decomposition of labour productivity growth by (3')

The whole economy

Coastal region

Central region

Western region
Figure 6. Growth effects of the service sector by (4)

The whole economy

Within Effects
- Total  S1  S2

Between Effects
- Total  S1  S2

Coastal region

Within Effect
- Total  S1  S2

Between Effect
- Total  S1  S2

Central region

Within Effect
- Total  S1  S2

Between Effect
- Total  S1  S2

Western region

Within Effect
- Total  S1  S2

Between Effect
- Total  S1  S2
Figure 7. Labour productivity decomposition

The whole economy

Coastal region

Central region

Western region
Figure 8. Pure and Baumol effects by sector total productivity growth

The whole economy

Coastal region

Central region

Western region
Table 1. Main estimates of demand for services (equation (6') without time effects)

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\psi_1 = -(\psi_2 + \psi_4)$</th>
<th>$\psi_2$ (t value)</th>
<th>$\psi_4$ (t value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS for S. M.</td>
<td>-0.033</td>
<td>-0.101</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR GLS for D. M.</td>
<td>-0.457</td>
<td>-0.492</td>
<td>-0.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.241</td>
<td>0.364</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: S. M. stands for static model and D. M. for dynamic model; LR GLS for D. M. means long-run solution of a dynamic model using GLS estimates. The dynamic model has the form:

$$\Delta \ln y_{it} = f \left( \ln y_{it-1}, \ln \left( \frac{P_{it}}{P_{i,t-1}} \right), \ln \left( \frac{W_{it}}{P_{it}} \right), \Delta \ln W_{it}, \Delta \ln y_{it-1} \right)$$

Table 2. Diagnostic tests associated with GMM estimation of the services demand equation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity of over-identifying instrument tests (upper: full panel model; lower: regional model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>359.7 [0.495]</td>
<td>252.4 [0.195]</td>
<td>159.5 [0.605]</td>
<td>72.59 [0.458]</td>
</tr>
<tr>
<td></td>
<td>167.4 [1.00]</td>
<td>173.4 [0.999]</td>
<td>12.41 [1.00]</td>
<td>12.06 [1.00]</td>
</tr>
<tr>
<td>Residual non-autocorrelation tests (upper: full panel model; lower: regional model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1) [p value]</td>
<td>-3.263 [0.001]</td>
<td>-2.846 [0.004]</td>
<td>-2.297 [0.022]</td>
<td>-2.562 [0.010]</td>
</tr>
<tr>
<td>N(0, 1)</td>
<td>-3.268 [0.001]</td>
<td>-2.824 [0.005]</td>
<td>-2.384 [0.017]</td>
<td>-2.98 [0.003]</td>
</tr>
<tr>
<td>AR(2) [p value]</td>
<td>0.059 [0.953]</td>
<td>0.83 [0.407]</td>
<td>0.973 [0.331]</td>
<td>1.047 [0.295]</td>
</tr>
<tr>
<td>N(0, 1)</td>
<td>0.032 [0.975]</td>
<td>0.657 [0.511]</td>
<td>0.945 [0.345]</td>
<td>1.324 [0.186]</td>
</tr>
</tbody>
</table>

Note: $\chi^2$ test statistics have different degrees of freedom mainly due to different sample sizes.
Table 3. Main estimates for the wage equation (7’)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GLS for S. M.</td>
<td>-0.117</td>
<td>-0.1408</td>
<td>-0.14</td>
<td>-0.0841</td>
<td>1.117</td>
<td>1.1408</td>
<td>1.140</td>
<td>1.0841</td>
</tr>
<tr>
<td>LR GLS for D. M.</td>
<td>-0.1272</td>
<td>-0.1476</td>
<td>-0.1272</td>
<td>-0.0071</td>
<td>1.1272</td>
<td>1.1476</td>
<td>1.1272</td>
<td>1.0071</td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.1906</td>
<td>0.1129</td>
<td>0.4797</td>
<td>-0.2566</td>
<td>0.8094</td>
<td>0.8871</td>
<td>0.5203</td>
<td>1.2566</td>
</tr>
</tbody>
</table>

| Coastal region | GLS for S. M. | 0.386   | 0.141    | 0.226    | 0.149    | 1.117     | 1.1489    | 1.1573    | 1.1287    |
|----------------|---------------|---------|----------|----------|----------|-----------|-----------|-----------|
| LR GLS for D. M. | -1.138 | -1.694    | -1.157   | -0.416   | 1.1321    | 1.1626    | 1.1622    | 1.0315    |
| GMM for D. M.  | -0.405   | -0.429    | 1.45     | -1.398   | 1.405     | 1.429     | -0.45     | 2.398     |

| Central region | GLS for S. M. | -0.168  | -0.114   | -0.08    | -0.011   | 1.1262    | 1.1371    | 1.1202    | 1.0333    |
|----------------|---------------|---------|----------|----------|----------|-----------|-----------|-----------|
| LR GLS for D. M. | -0.181 | -0.251    | -0.361   | -0.4     | 1.1282    | 1.1288    | 1.082     | -0.61     |
| GMM for D. M.  | 0.6397  | 0.7056    | 1.042    | 0.5161   | 0.3603    | 0.2944    | -0.042    | 0.4839    |

| Western region | GLS for S. M. | -0.269  | -0.178   | -0.218   | -0.08    | 1.1066    | 1.1311    | 1.1342    | 1.077     |
|----------------|---------------|---------|----------|----------|----------|-----------|-----------|-----------|
| LR GLS for D. M. | -0.343 | -0.205    | -0.295   | -0.225   | 1.1093    | 1.1347    | 1.1085    | -0.881    |
| GMM for D. M.  | -0.2204 | -0.3609   | -0.4705  | -2.2298  | 1.2204    | 1.3609    | 1.4705    | 3.2298    |

Note: S. M. stands for static model and D. M. for dynamic model; LR GLS for D. M. means long-run solution of a dynamic model using GLS estimates. The dynamic model has the form:

$$\Delta W_{it} = f \left( \ln \left( \frac{W_{it}}{W_{i-1}} \right), \ln \left( \frac{W_{it}}{W_{i-1}} \right) \right)$$

But the short-run variable $\Delta \ln W_{it}$ is significant in LR GLS estimation but not in GMM estimation.

Table 4. Diagnostic tests associated with GMM estimation of the wage equation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity of over-identifying instrument tests (upper: full panel model; lower: regional model)</td>
<td>$\chi^2$</td>
<td>302.4 [0.000]</td>
<td>231.1 [0.000]</td>
<td>139.5 [0.000]</td>
</tr>
<tr>
<td></td>
<td>302.3 [0.000]</td>
<td>193.1 [0.000]</td>
<td>89.68 [0.000]</td>
<td>4.838 [0.436]</td>
</tr>
<tr>
<td>Residual non-autocorrelation tests (upper: full panel model; lower: regional model)</td>
<td>AR(1) [p value]</td>
<td>N(0, 1)</td>
<td>-2.687 [0.007]</td>
<td>-3.161 [0.002]</td>
</tr>
<tr>
<td></td>
<td>-2.588 [0.01]</td>
<td>N(0, 1)</td>
<td>-3.429 [0.001]</td>
<td>-2.692 [0.007]</td>
</tr>
<tr>
<td></td>
<td>AR(2) [p value]</td>
<td>N(0, 1)</td>
<td>-0.1882 [0.851]</td>
<td>-0.8854 [0.376]</td>
</tr>
<tr>
<td></td>
<td>-0.7407 [0.459]</td>
<td>N(0, 1)</td>
<td>-1.937 [0.053]</td>
<td>-0.3663 [0.714]</td>
</tr>
</tbody>
</table>

Note: $\chi^2$ test statistics have different degrees of freedom mainly due to different sample sizes.
Table 5. Main estimates for labour demand of equation (9’)

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\sigma$ (t value)</th>
<th>$\Lambda(t)$ (Top: GLS; Bottom: GMM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full panel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS for S. M.</td>
<td>0.5744 (12.0)</td>
<td>0.4665 (8.95)</td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.4486 (1.98)</td>
<td>0.519 (1.59)</td>
</tr>
<tr>
<td>Coastal region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS for S. M.</td>
<td>0.645 (13.7)</td>
<td>0.5569 (10.6)</td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.7746 (1.63)</td>
<td>1.1331 (1.64)</td>
</tr>
<tr>
<td>Central region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS for S. M.</td>
<td>0.493 (8.36)</td>
<td>0.4152 (6.31)</td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.2692 (0.359)</td>
<td>0.7588 (1.88)</td>
</tr>
<tr>
<td>Western region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLS for S. M.</td>
<td>0.4084 (7.96)</td>
<td>0.3336 (6.18)</td>
</tr>
<tr>
<td>GMM for D. M.</td>
<td>0.2873 (1.82)</td>
<td>0.3275 (1.31)</td>
</tr>
</tbody>
</table>

Note: S. M. stands for static model and D. M. for dynamic model. The dynamic model has the form:

$$\Delta \ln l_t = f\left(\ln \left(\frac{l_t}{y_3}\right)_{t-1}, \ln w_{3t-1}, \ln \Lambda_{t}\right)$$

Table 6. Diagnostic tests associated with GMM estimation of the labour demand equation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity of over-identifying instrument tests (upper: full panel model; lower: regional model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>212.1 [0.049]</td>
<td>169.6 [0.011]</td>
<td>113.7 [0.034]</td>
<td>76.58 [0.023]</td>
</tr>
<tr>
<td></td>
<td>205.3 [0.46]</td>
<td>152.7 [0.053]</td>
<td>102.9 [0.079]</td>
<td>72.66 [0.02]</td>
</tr>
<tr>
<td>Residual non-autocorrelation tests (upper: full panel model; lower: regional model)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1) [p value]</td>
<td>-2.549 [0.011]</td>
<td>-2.256 [0.024]</td>
<td>-2.527 [0.012]</td>
<td>-2.633 [0.008]</td>
</tr>
<tr>
<td>N(0, 1)</td>
<td>-2.733 [0.006]</td>
<td>-2.532 [0.011]</td>
<td>-2.652 [0.008]</td>
<td>-2.902 [0.004]</td>
</tr>
<tr>
<td>AR(2) [p value]</td>
<td>1.642 [0.101]</td>
<td>1.398 [0.162]</td>
<td>1.57 [0.116]</td>
<td>1.648 [0.099]</td>
</tr>
<tr>
<td>N(0, 1)</td>
<td>1.807 [0.071]</td>
<td>1.549 [0.121]</td>
<td>1.8 [0.072]</td>
<td>1.809 [0.071]</td>
</tr>
</tbody>
</table>

Note: $\chi^2$ test statistics have different degrees of freedom mainly due to different sample sizes.