

NBER WORKING PAPER SERIES

HOW MIGRATION RESTRICTIONS LIMIT  
AGGLOMERATION AND PRODUCTIVITY IN CHINA

Chun-Chung Au  
Vernon Henderson

Working Paper 8707  
<http://www.nber.org/papers/w8707>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2002

We gratefully acknowledge support for this work from a grant from the Research Committee of the World Bank and by a Mellon Foundation grant to the Population Studies and Training Center at Brown University. The opinions expressed in the paper do not necessarily represent those of these organizations. Remy Prud'homme provided early inspiration and insight in formulating some of the issues in the paper. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Chun-Chung Au and Vernon Henderson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Migration Restrictions Limit Agglomeration and Productivity in China  
Chun-Chung Au and Vernon Henderson  
NBER Working Paper No. 8707  
January 2002  
JEL No. J600, O100, O400, P200, R100, R200, R300

### ABSTRACT

China strongly restricts rural-rural, urban-urban, and rural-urban migration. The result which this paper documents is a surplus of labor in agriculture. However, the paper argues that these restrictions also lead to insufficient agglomeration of economic activity within both rural industrial and urban areas, with resulting first order losses in GDP. For urban areas the paper estimates a city productivity relationship, based on city GDP numbers for 1990-97. The effects of access, educational attainment, FDI, and public infrastructure on productivity are estimated. Worker productivity is shown to be an inverted U-shape function of city employment level, with the peak point shifting out as industrial composition moves from manufacturing to services. As far as we know this is the first paper to actually estimate the relationship between output per worker and city scale, as it varies with industrial composition. The majority of Chinese cities are shown to be potentially undersized - below the lower bound on the 95% confidence interval about the size where their output per worker peaks. The paper calculates the large gains from increased agglomeration in both the rural industrial and urban sectors. It also examines the effect of capital reallocations, where the rural sector is grossly undercapitalized.

Chun-Chung Au  
Brown University

Vernon Henderson  
Department of Economics  
Box B  
Brown University  
Providence, RI 02912  
and NBER  
[j\\_henderson@brown.edu](mailto:j_henderson@brown.edu)

China restricts internal migration of its population between urban and rural areas, within the rural sector, between big and small cities, and between regions. The prime instrument of control as will be described later is the household registration, or hukou system. The usual view is that the primary effect of the restrictions is to enforce a surplus of labor in the rural sector. As China's economic growth has slowed in recent years, an explicit issue articulated by policy makers in formulating the most recent five year plan is whether permitting greater migration to cities would increase overall productivity and help sustain growth by allowing the rural surplus labor to gain from urban employment (see Prud'homme (2000)). A related policy concern is national income inequality, a serious domestic social and political issue. In 1978 China's Gini coefficient was similar to India's and urban real incomes were 2.0 times those in the rural sector. Since 1978, China's rural-urban income disparity has increased; and the Gini coefficient has risen and now far exceeds countries such as India, Bangladesh and Pakistan, not to mention Korea and Japan.<sup>1</sup>

This paper argues that the issue as posed, that rural-urban migration would enhance productivity, is not straightforward. Given the estimates we obtain, it will indeed be the case that, under the current allocations of resources, labor productivity in the urban sector exceeds that in the rural sector. However it turns out that for physical capital the reverse is the case: the return to capital in the urban sector is much lower than in the rural sector. Capital allocations in China are subject to strong controls, with investment directed towards urban industry through quantity controls and discriminatory taxation. As a result, apart from our findings, Jefferson and Singhe (1999, Table 1.9) find in 1995 and 1996 that the pre-tax return on capital for rural industry was 25% higher than in the state owned urban sector. Moreover education is very unequally spread among sectors, partly because migration policies direct the educated rural towards larger cities, while the less educated are held in place. How do we then proceed with productivity assessments?

We could hypothetically redistribute capital and perhaps education more efficiently across sectors and then compare (having held labor in place) the hypothetical rural versus urban marginal products of labor. We do that and still find support for the idea that rural labor productivity is less than urban. However these exercises face two problems. First is that they are partial equilibrium and large resources shifts between sectors will have impacts on relative prices and evaluations of output of different sectors. Second is that the approach overlooks a crucial aspect of the problem, namely scale (dis)economies in spatial agglomeration within sectors.

---

<sup>1</sup>In 1995 the World Bank estimated China's Gini to be .42 compared to about .30 for the noted South Asian countries (WDR (1998)). Korea and Japan are viewed as having among the lowest Gini's in the world.

We will argue that migration restrictions on rural to rural migration and restrictions on the movement of rural capital strongly inhibit spatial agglomeration of rural industry. We will show empirically that rural industry has enormous unexploited economies of scale from spatial agglomeration. Similarly migration restrictions across cities, as well as rural-urban migration restrictions, have left most cities significantly undersized, with unexploited economies of scale. In terms of migration, while the issue can be couched as a rural-urban problem, in fact the problem is more general. Migration restrictions cause productivity losses simply because they stop desired levels of spatial agglomeration in both the rural and urban sectors.

In order to do this analysis, it is essential then to document the magnitudes of scale economies and diseconomies from spatial agglomeration. Urban economists and geographers (Henderson (1974), Fujita, Krugman and Venables (2000) Duranton and Puga (2001)) postulate that spatial agglomerations occur to exploit scale economies in production and distribution of goods, but are limited by diseconomies that also arise with agglomeration, as well as the spatial extent of an agglomeration's market. Systems of cities models postulate an inverted U-shape function to utility, or to real output per worker in city population size. That is, there is a city size at which output per worker is maximized, where that peak point will vary with city industrial composition or city characteristics. This fact limits city sizes and gives us multiple cities in an economy, where individual city sizes vary with city type, or industrial composition.

To evaluate the impact of migration restrictions we must estimate the effect on productivity of insufficient agglomeration. To accomplish that, we will estimate real output per worker as an inverted U-shape function of city scale, that shifts with industrial composition. To do so, we need to know GDP figures for cities. China, unlike any other country we know of, calculates and publishes such figures. This is the first study on any country to estimate this inverted U-shape function. China's migration restrictions turn out to be fortuitous from a research perspective. Since cities are found to both the left and right of the peak of the inverted-U, that helps identify the shape of the curve. As such this is a fundamental research contribution of the work, in addition to the analysis of the costs of migration restrictions.

## 1 Migration Restrictions and Urbanization Policy in China

The hukou system in China is similar to an internal passport system (see Chan (1994) for a detailed description). A person's local "citizenship" and residence is initially defined for a child as a

birth right, traditionally by the mother's place of legal residence. The entitlements and details of the system differ for urban and rural residents. Legal residence in a city entitles one to local access to permanent jobs, regular housing, public schooling, and public health care (where almost all health care is public). Until the early 1990's, it also entitled urban people to "grain rations" – rations of essentials such as grain and kerosene.

Legal residence in a village or rural township entitles residents to land for farming, township housing, job opportunities in rural industrial enterprises, and local health and schooling facilities. Residents also have some degree of "ownership" in local enterprises, although distributed profits all go to the local public budget, which may be used to finance township housing and infrastructure. Again, until recent years, legal residence in a township also entitled a "peasant" to some share in locally produced (or allocated from the outside) grain and other essentials.

How does a person change their local citizenship? There are several common mechanisms. First is through the academically competitive education system. A smart rural youth may persist through village schooling and competitive exams to get a spot in a nearby township or county town high school ("senior middle school") and then to get a spot in a college in a city. Upon college graduation, the rural youth will be hired into an urban job, with an urban hukou. It is this aspect of migration policy which potentially skews the allocation of educational people towards larger cities. Second, the state at times can open the gates, permitting factories to hire permanent workers from rural areas, permitting family reunification, or permitting legal migration from rural areas to nearby small cities. These transfers confer some type of urban hukou upon these formerly rural residents. Permanent migration is not without costs. Permanently leaving a village means abandoning ownership claims without compensation to agricultural land that one's family may have farmed for decades and to the profits of local rural industries which are distributed in-kind, as for example with township housing.

People can migrate to an area without local hukou, either illegally ("unregistered") or legally as a temporary worker or visitor or even as a "permanent resident" on a long-term permit. For example, a rural person may be hired as a "contract worker" in industry or services, for a term of three years. Or a rural person may get permission to work temporarily in another local area in construction, food services, domestic services and the like. Or they may move illegally, without registering in the new location, and work in the informal sector for low pay, under poor conditions, with risk of deportation. Despite these possibilities and despite some recent relaxations of restrictions in particular provinces, the restrictions on migration remain tight.

Temporary migrants to larger cities typically have no, or very high priced access to health care

and schooling facilities and regular, "legal" housing. In fact cities have strict national guidelines on conversion of agriculture to urban land; and institutional difficulties in housing markets in expanding supply which makes it particularly difficult for migrants to find decent housing. All this means living and social conditions for migrants are difficult; and it also means migration of families is extremely difficult, since children face no or very high priced access to schooling and health care. Migrants also face restrictions on job choices. Legal migration requires getting a permit from the city of in-migration and cities can impose various hurdles to getting a permit — permission from the home location, proof of a guaranteed job and specific housing, and the like. Cities also publish job lists, citing jobs for which migrants are not eligible; in 2000 Beijing listed over 100 occupations as non-eligible ones. Migrants may still have to pay taxes to their rural home village for services they don't consume and on land left fallow. Finally migrants face direct fees which can be quite arbitrary (Cai (2000)). There is a license fee to work outside the home township paid to the township that can be equivalent to a month's wages. At the destination there can be fees for city management, for a license to work as a "foreign" worker, for construction management, for crime fighting, for temporary residence, and even for family planning if the migrant is female. All these restrictions sharply reduce the benefits and raise the costs of migration, particularly into large cities. Migration is limited and most migration is short-term, or "return" migration. Return migration has the added feature of reducing the incentive to invest heavily in on-the-job training in urban jobs.

The application of migration restrictions in part reflects the urbanization provisions in the Sixth Five Year Plan (1981-1985) (Chan (1994), Lin (1992b), and Henderson (1988)). Several features of national urban planning in the 1980's strongly influence the urban sector today. For most of the 20 years prior to 1982, urbanization in China was strongly restricted, with the urban sector's share of national population declining, reflecting in part the bias against urbanization and its association with capitalism, Western moral pollution, and potentially counterrevolutionary forces. The Sixth Five Year Plan, while planning to constrain the growth of larger cities, promoted population growth of smaller cities, through flexibility in transferring full or limited hukou to small cities from rural areas. Thus one expects to see much more rapid population growth amongst smaller cities and towns, than larger cities, as well as transformation of towns to cities. An issue is whether such differential growth is efficient.

While the Sixth Five Year Plan permitted growth of smaller cities, it is still the case that overall urbanization was to be restricted. Given that, a key issue was what to do with the surplus rural population. A policy of "leave the land but not the village" promoted development of rural industry. Rural industrial development became a key engine of Chinese growth through the mid-1990's. But the

restriction of not leaving the village makes rural agglomeration of industry difficult, as does an inability to transfer capital ownership across villages as described later.

A final aspect of planning in the 1980's was the strict spatial hierarchy that is still in place. The large ("sophisticated") lead the small. So only the largest coastal cities were to have initial access to new technologies and FDI; and in turn were supposed to transfer their older technologies to smaller and hinterland cities. Capital allocations, tax policies and governance structure favor the large over the small; and this hierarchical structure may influence productivity differentials across cities.

## 2 The Urban Sector

This section starts with an overview of the urban sector, both looking cross-sectionally at its key features and looking at dynamic tendencies from 1990 to 1996 or 1997, the time frame of the data. This paper doesn't focus on the dynamic development per se, but understanding some features is important to modeling productivity of individual cities and key econometric issues in estimating the relevant parameters.

### 2.1 China's Urban System

We have data for 1990-1997 on about 225 prefecture and provincial level cities, the larger cities at the top of the urban hierarchy. For 1990, 1991 and 1994-1996 we have data on about 435 other cities, generally urban places over 100,000 population. Those latter cities are county cities, either traditional ones or new ones. In 1990, at all levels, there are about 465 cities; by 1996 that number grows to about 665 cities, so there is entry of an additional 200 cities. Virtually all new cities are county level, or non-prefecture cities.

Tables 1 and 2 we present some basic facts about the urban system in China. Columns in Table 1 give figures for all cities present in 1990, for prefecture level cities present in 1990, for county cities present in 1990, and for new county cities that enter the set of cities after 1990. The first four rows show high overall non-agricultural employment growth from 1990 to 1996, over 4.3% per year for all cities, but much slower population growth, about 1.9% per year. The differential reflects two things. In some cases in 1990, cities contained significant rural-agricultural populations which during the decade moved into non-agricultural employment. This is urbanization from within. Second and more critically, population numbers may exclude certain types of immigrants while employment figures better reflect them. In particular are shorter-term or longer-term migrants working in the city, who can't obtain

housing within the city but can obtain (often "illegal") housing just beyond the boundaries of urban districts. As a result, in the paper, our city size measure will always be in terms of non-agricultural employment.

Table 1 also shows that real output per worker grew at an incredible rate during the period; for all cities, the average annual rate was about 7.8% a year. Finally over time the manufacturing to service ratio declines. The decline involves some redefinition of manufacturing activity as service activity around 1993-1994, a redefinition which is relevant to later analysis. For example, some service activities owned by state-owned industrial enterprises may have been defined as services after 1994, but as industry before 1994. In the data, over the 1990-93 period the ratio declines modestly in total by 4%; between 1993 and 1994 it declines by 24% in what we assume are mostly definitional changes prompted by economic reforms; and from 1994 on it declines by 4-5% a year, as restrictions on private service sector development are loosened and the service sector starts to boom.

Looking across the urban system, prefecture level and above cities are larger and have much greater human capital. Among below prefecture level cities, the big story are new cities, which are larger than traditional county cities, have more industry, and are nearer the coast. These cities reflect rapid urbanization and industrialization of former rural towns and townships.

The period 1990-1996 is one of rapid industrial reforms, removing some of the props under state owned industry and exposing them to increasing competition. Mostly heavily hit were interior and northern heavy industry cities, especially under the reforms in 1993-94. These reforms attempted to further force state-owned enterprises to operate on a market basis and represent a break point in the data in terms of how outputs are evaluated, as well as diminishing the role of planning. Along with the rapid growth of business and financial service activity, the result, in this very short period of time, is to dramatically shake up the urban system. To get a sense of the overturning of the old industrial order, we look at transition matrices. In Table 2, Part A is the transition matrix for real output per worker. The first row shows into what 1996 quartile new cities, or entrants from 1990 to 1996, would enter. Interestingly, new cities would enter fairly evenly, across states. The second through fifth rows show transitions for 1990 cities. So of the 116 cities in the lowest quartile in 1990, 58 remain in the lowest quartile in 1996, while 35 move to the second, 10 to the third, and 12 to the fourth. In addition 1 city in the first quartile in 1990 disappears from the data set by 1996.

What is stunning about Table 2 Part A is the incredible mobility. The diagonal measuring the propensity for a city to stay in its original quartile is small, as indicated by the bold numbers. The average probability across states of remaining in the own state is only .41. As noted, 12 of 116 cities

jump from the lowest to highest state in just 6 years, while 18 of 116 fall from the highest state to one of the two lowest states. Consistent with the picture of turmoil is the notion of lack of persistence. The correlation between growth in output per worker from 1990 to 1991 and growth between 1995 and 1996 is zero. A city's features which gave it a comparative advantage in 1990 may not do so in 1996.

In contrast the transition matrix for total employment in Part B is more "regular". The average probability of staying in the own state is .74; and cities in the top state in 1990 have a very high probability of staying there – .92. Although there is more motion in terms of relative size than in a developed country, cities only move to adjacent states and big cities persist in remaining big, just as one would find in other countries (Eaton and Eckstein (1996), Black and Henderson (2001)). The turmoil lies in the huge relative jumps up and down in output per worker across cities. That gives us lots of change over time in the data to analyze productivity determinants but, as we will see, can make the analysis less straightforward.

## 2.2 Productivity in the Urban Sector

We model productivity in the urban sector, with a standard urban model. Suppose countries are composed of a variety of different types of cities, where each type produces a particular good for export from the city, in an urban hierarchy with an endogenous number of cities of each type. In any city type, final city output is produced with local inputs from manufacturers and from service producers. So at the bottom of the hierarchy we might have heavy industry with, say, steel type hinterland cities that combine rolled steel production with local finance, personnel, and market services to sell rolled steel to other cities. At the top in a coastal city we could have headquarters activity for manufacturers whose services (to hinterland plants) are produced mostly with local export, R&D, and management inputs.

More generally for city  $i$  of type  $j$ , value-added of the city,  $Y_j^i$ , is a function of production amenities in city  $i$ ,  $A^i$ , local manufacturers' output  $M_j^i$ , and a function with intermediate non-traded service inputs,  $\tilde{g}_j(x_1^i, x_2^i, \dots, x_s^i)$ .  $x_l$  is output of service firm  $l$  and there are  $s^i$  such firms in city  $i$ . Usually symmetry is imposed (i.e.,  $x_l$  firms have identical technology and enter  $g$  identically), so  $\tilde{g}_j(\cdot)$  is written as  $g_j(s^i, x^i)$  where  $x^i$  is the output of the representative service firm in city  $i$ . The  $g_j(\cdot)$  function represents the Dixit-Stiglitz-Ethier notion that diversity,  $s^i$ , plays a role in enhancing productivity in  $Y_j^i$  production. Thus value added of city  $i$  in the simple case of Hicks' neutrality of  $A^i$  is

$$Y_j^i = A^i f_j(M_j^i, g_j(s^i, x^i)) \quad (1)$$

Each of the  $x^i$  is produced according to technology

$$x^i = x(k^i, n^i) \quad (2)$$

where  $k^i$  and  $n^i$  are per firm inputs of capital and labor in the service sector in city  $i$ . Manufactures,  $M_j^i$ , are produced with inputs of capital and labor, or

$$M_j^i = M_j(K^i, N^i) \quad (3)$$

City endowments of capital and effective labor are  $\bar{K}^i$  and  $\bar{N}^i$  where effective labor is defined below, so

$$K^i + s^i k^i = \bar{K}^i \quad (4a)$$

$$N^i + s^i n^i = \bar{N}^i \quad (4b)$$

Capital and labor are observed but not their allocations between the final output sector and the intermediate input service sector. Accordingly, if we impose internal allocation rules, we can reformulate eq. (1) as a "meta-production" function in terms of  $\bar{K}^i$  and  $\bar{N}^i$ . For example, there could be optimal allocation rules, where after substituting (2)-(4) into (1) and optimizing with respect to  $K, N$  and  $s$  (dropping sub and superscripts) we have

$$s f_M M_K = f_g g_x x_k \quad (5a)$$

$$s f_M M_N = f_g g_x x_n \quad (5b)$$

$$s g_s = g_x (x_n n + x_k k) \quad (5c)$$

Utilizing these to solve out  $k, n$ , and  $s$ , allows us to write (1) as

$$Y_j^i = A^i \tilde{F}_j(\bar{K}^i, \bar{N}^i) \quad (6)$$

The allocation rules which are used to derive a meta-production function need not be optimal. So, for example, we could assume marginal products of capital or labor are only equalized across sectors within cities up to a factor of proportionality,  $\phi_K$  or  $\phi_N$ , where the  $\phi$ 's might depend on human capital in the city or the extent of economic liberalization. Second there are local scale externalities in cities. The role of diversity,  $s^i$ , has been identified. However  $M_j^i$  can also be produced with scale externalities,

dependent on local manufacture's scale (measured by, say, either  $M_j^i$  or  $N^i$ ). Such scale externalities are based on local information spillovers and intra-industry specialization (Marshall (1980); see Eberts and McMillen (1999) for a review of the literature on micro-foundations and empirical evidence). Allocation rules reflect the extent to which these externalities are successfully locally internalized.

As we move up the urban hierarchy, based on empirical evidence (Kolko (1999) on the USA, our raw data on China) and on theory (Duranton and Puga (2001)), we expect cities to become increasingly service oriented, in terms of the composition of local observed production of,  $M_j$  versus  $g_j(\cdot)$ . For this to come out of a theoretical model we must account for the juxtaposition of parameters as they vary by type  $j$ , that describe the intensity of service versus manufacturing inputs in  $Y_j$ , the importance of service input diversity to  $Y_j$ , and the degree of scale economies in the manufacture's that are inputs to  $Y_j$ . For the hierarchy we have described to play-out, it is sufficient in simple models, for example, that the benefits of service diversity increase as we move up the urban hierarchy.<sup>2</sup> To see this we must discuss the rudiments of how city sizes are determined.

### 2.2.1 The Inverted U-Shape to Output Per Worker

So far we have discussed the existence of scale economies in cities which are the basis of local agglomeration. However with undiminished scale economies, we would have no limit to the benefits of agglomeration and city sizes. Apart from limits on a city's market potential (Fujita et al (1999)), the key ingredient in most urban models involves the notion of "effective labor". In a city, the down-side to local agglomeration which limits city sizes is represented by commuting/congestion costs (Henderson (1974)). People live in residential areas surrounding a central business district (CBD); as city population, or employment size increases the per person time devoted to commuting (to claim more land for housing at the expanding city edge) increases on average.<sup>3</sup> The net disadvantages of city size increases can be

<sup>2</sup>For example, assume

$$Y_j = \tilde{A}_j \left( \sum_{t=1}^s x_t^{\rho_j} \right)^{\gamma/\rho_j} M^{(1-\gamma_j)}, \quad M = A_m N_m^{1-\alpha} K_m^\alpha, \quad x = (n-f)/c \quad (\text{based on Dixit-Stiglitz technology with fixed}$$

cost,  $f$ , and marginal cost  $c$ ), and  $A_m = N_m^{\varepsilon_j}$  (or  $M^{\varepsilon_j}$ ). With optimal allocation decisions (before dealing with the issue of effective labor), we have  $N_m/N = (1-\gamma_j) \rho_j (1-\alpha-\varepsilon_j) [(1-\alpha+\varepsilon_j) (1-\gamma_j) \rho_j + \gamma_j]^{-1}$ . Then  $d \ln (N_m/N) / d \ln \rho_j > d \ln (N_m/N) / d \ln \varepsilon_j > 0$ . Second overall scale economies (the exponent of  $N$ ) are  $(1-\alpha+\varepsilon_j) (1-\gamma_j) + \gamma_j / \rho_j$ . If  $\rho_j$  declines (diversity becomes more important in enhancing productivity) as we move up the city hierarchy, then  $N_m/N$  will decline and scale economies (and hence the peak to  $Y_j/L$ ) will rise. Manufacturing scale economies,  $\varepsilon_j$  can also rise, as we move up the hierarchy accentuating the increase in city size, as long as the percent decline in  $\rho_j$  is greater (absolutely), so  $N_m/N$  declines. Increases in  $\gamma_j$  lower than  $N_m/N$  but the effect on scale is ambiguous. Greater scale occurs if  $1/\rho_j > 1-\alpha+\varepsilon_j$ .

<sup>3</sup>The sum of rents plus commuting costs increases for everyone; but rents are redistributed, say, equally back to residents as Arrow-Debreu shareholders in their city. Even allowing for polycentric cities, per person commuting costs relative to agglomeration benefits rise with city size.

represented as lost working time. So if  $L$  is total city population (= labor force), the fraction available for work is  $\tau(L)$  where  $\tau < 1$  and  $\tau' < 0$ . That is

$$\bar{N} = \tau(L)L, \quad \tau < 1, \tau' < 0 \quad (6)$$

so effective labor,  $\bar{N}$ , is a fraction of the labor force,  $L$ , where that fraction declines with  $L$ .

Incorporating commuting costs, we rewrite (5) as

$$Y_j^i = A^i F_j(\bar{K}^i, L^i) \quad (7)$$

The shape of  $F(\cdot)$  in (7) reflects internal city (mis)allocation rules and any marginal products are “social” marginal products (given allocation rules). Empirically we will estimate a version of (7) where value-added per worker will be viewed as some type of inverted U-shaped function of scale,  $L$ . As city scale increases from a low level initially scale economies dominate, but then commuting cost considerations kick-in, so there is a city scale which maximizes value-added per worker. One issue will be how to capture these non-linear effects. A second issue is how to incorporate notions of hierarchy so the inverted  $U$ -shape to value-added per worker in local total employment varies with industrial competition. In footnote 2, we showed that if the importance of service diversity grows as we move up the urban hierarchy we have both an increasing degree of scale economies and a decline in the local manufacturing to service ratio. The former will shift the peak point to the inverted  $U$ -shape of value-added per worker to the right (higher total local employment). To capture the idea that peak points shift right as industrial composition changes, in estimation we interact the manufacturing to service ratio with city scale effects, to allow the inverted  $U$ -shape to value-added per worker to vary with industrial composition, reflecting how the underlying constellation of structural parameters varies over the urban hierarchy (Henderson (1988)).

In a market economy with perfect migration, free capital markets, and developers and/or local government involved in formation of new cities, any city type  $j$  operates near its peak point to value-added per worker, which will be proportional to real compensation per worker. All cities face the same horizontal national supply curve of labor (as viewed by an individual city). As we move up the urban hierarchy, bigger cities have their peak points to real compensation per worker shifted right, cresting near the supply curve. In particular, with perfect divisibility of cities, many cities of each type, and all cities having identical amenities,  $A^i$ , each inverted  $U$  is tangent to the supply curve at its peak point. If  $A^i$ 's vary, say, then those with high  $A^i$ 's operate to the right of their peak points in stable equilibria.

We have two remaining issues to deal with, before turning to empirics. One is the measure of capital. The other concerns how city sizes are determined in China.

### 2.2.2 Capital Stock in Cities

The capital stock data we have is for “independent accounting units”. In 1997 we know the share of independent accounting units in value-added, where they account for 35-40% of value-added in prefecture level cities. In 1996 and earlier we can estimate their share in gross output (basically sales, not value-added); and they account for 70% of gross output. In the mid-1970’s independent accounting unit numbers cover virtually the whole state owned sector but less than half of urban collectives and private firms. Non-independent accounting units tend to be smaller units. If both sectors have the same value-added production function and face the same opportunity cost of capital, and if the ratio of total value-added (or gross output) by all units to independent accounting units is the same across cities, then total capital would simply be our observed capital times a constant factor of proportionality (total value-added to independent unit value-added). However this factor of proportionality is not constant across cities and is negatively correlated with observed capital by construction. Thus it may be critical to control for the factor of proportionality. We can do that precisely for 1997, but for other years we only have a proxy based on gross output (rather than value-added) discussed in the data Appendix. While we will control for a factor of proportionality, we want to interpret that factor; and to assess whether independent accounting units, which are dominated by state owned firms, face a lower opportunity cost of capital than non-independent accounting units as suggested in Jefferson and Singhe (1999).

For interpretation we need structure and we consider the following simple framework. Consider two sectors,  $I$  for independent and  $N$  for non-independent, where total city output is  $A [K_N^\alpha L_N^{1-\alpha} + K_I^\alpha L_I^{1-\alpha}]$ . Both sectors have identical technologies drawing on underlying manufacturing and service production, so we have meta-functions for each sector. Inside the square brackets are technologies for producing  $Y_N$  and  $Y_I$ . If labor is internally efficiently allocated in the city (equalized private marginal products) then we can solve  $L_N = L (Y_N/Y)$  and  $L_I = (Y_I/Y)L$ . For capital, assume sector  $I$  faces a cost  $r_I$  and sector  $N$  a cost  $r_N$ , where we expect  $r_I/r_N < 1$ . Then markets will yield  $K_N = (r_I/r_N) K_I (Y_N/Y_I)$ . Substituting for  $L_N$ ,  $L_I$  and  $K_N$  back in the basic production function we get

$$Y = AK_I^\alpha L^{1-\alpha} (Y/Y_I)^\alpha \left[ \left(\frac{r_I}{r_N}\right)^\alpha + \left(\frac{Y_I}{Y}\right) \left(1 - \left(\frac{r_I}{r_N}\right)^\alpha\right) \right] \quad (8)$$

If  $r_I/r_N = 1$ , then eq. (8) reduces to

$$Y = A (K_I(Y/Y_I))^\alpha L^{1-\alpha} \quad (8a)$$

That is, if  $r_I/r_N = 1$  then total capital is simply  $K_I$  inflated by  $Y/Y_I$ . In estimation we will generally adopt a form of (8a) where  $\alpha$  is the coefficient for  $K_I$  but  $Y/Y_I$  is assigned a coefficient  $\sigma$ . If (8a) is correct, then in estimation  $\alpha$  will equal  $\sigma$ . However by differentiating (8) we see that the coefficient for  $d(\ln Y)/d\ln(Y/Y_I)$  is

$$\alpha - (1 - (r_I/r_N)^\alpha) (1 + (r_I/r_N)^\alpha Y_N/Y_I)^{-1} \quad (9)$$

If  $r_I/r_N < 1$  then the coefficient ( $\sigma$ ) is less than  $\alpha$ , whereas if  $r_I/r_N > 1$ , the opposite is the case. Finally, at least for 1997, we can and will directly estimate eq. (8) by non-linear methods.

### 2.2.3 Determining City Population

The final consideration is how city populations are determined. Having some understanding of that is useful for relating output per worker in cities to welfare of residents and equilibrium city sizes. It is also important in estimation to understanding choice of instrumental variables. The usual framework is one of free migration across cities; here the situation is rather different. Each city offers a utility level to a resident which is a function of its real wage and local quality of life,  $Q_t^i$ . Real wages are related to the city's allocation of labor and capital and scale (dis)economies, determining marginal products. For the moment we treat capital as given. Thus city utility that can be offered to migrants is some function

$$U_t^i = U(\bar{K}_t^i, L_t^i, A_t^i, Q_t^i) \quad (10)$$

In (10) we expect  $U_K, U_A, U_Q > 0$  but the sign of  $U_L$  is ambiguous (given the inverted  $U$ -shape to real wages). Focusing on the determinants of city population,  $L_t^i$ , in China growth comes primarily from migration, particularly the rural counties surrounding the city. The underlying basis of the paper is that for similar skill workers within a region, the utility from rural residence is less than from urban residence and migration is very costly, or there are market frictions.

To model market frictions, suppose city  $i$  draws from a surrounding rural population, where utility in the rural sector is given by

$$R_t^i = R(K_t^{Ri}, \bar{L}_t^i - L_t^i, A_t^{Ri}, Q_t^{Ri}) \quad (11)$$

$\bar{L}_t^i$  is the fixed population of the whole region of  $i$ , encompassing the urban and rural sectors, so the

rural sector population is this total less the urban population,  $\bar{L}_t^i$ . As noted in the data Appendix for China,  $\bar{L}_t^i$  is fairly well defined for the 225 largest cities.  $K_t^{Ri}$ ,  $A_t^{Ri}$ , and  $Q_t^{Ri}$  are respectively the capital, production amenities (e.g., soil quality, rainfall) and quality life measures in the rural sector.  $\partial R_t^i / \partial L_t^i$  may be greater than zero, implying diminishing real wages in the rural sector, so  $R_t^i$  rises with outmigration. If there were no migration restrictions  $L_t^i$  would adjust to equalize  $R_t^i$  and  $U_t^i$ . In estimation of an urban production relation, if  $L_t^i$  is viewed as endogenous, then  $Q_t^i$ ,  $\bar{L}_t^i$ ,  $A_t^{Ri}$ , and  $Q_t^{Ri}$  could be instruments in IV estimation, since they influence  $L_t^i$  but arguably don't directly affect city productivity.

In China, we presume  $U_t^i > R_t^i$ , a differential sustained by two considerations. First are migration restrictions, as discussed earlier, imposed by cities and nationally. Second are on-going technological shocks and quality of life shocks which differentially favor the urban sector, given the on-going and even accelerated disparities in average education, FDI, and public service provision in urban versus rural areas. We view migration restrictions operating as frictions that have the per person cost of immigration rising as the rate of net in-migration to the city,  $\dot{L}_t^i \equiv (dL_t^i/dt)/L_t^i$ , rises. At any instant the gap between urban and rural utility equals the cost of migration  $m(\cdot)$ , or

$$U_t^i - R_t^i = m(\dot{L}_t^i), \quad m, m' > 0 \tag{12}$$

Of course, absent naive expectation,  $U_t^i$ , and  $R_t^i$ , should represent the present values of utility in the urban versus rural sectors. Equation (12) (with the LHS in present value terms) is the specification of migration frictions in the USA, used in Mueser and Graves (1995) and Rappaport (2000).

The differential equation in (12) provides a link between present and past values of city employment. So, for example, in 1990 there will be a rural-urban utility gap based on historical allocations within the municipality in question and 1990 conditions in the rural versus urban areas. 1997 city employment will depend on the 1990 conditions, and how parameters  $A^i$ ,  $Q^i$ , and so on have evolved. Second the employment growth in any period will be related to prior period conditions defining past employment levels in rural versus urban areas. These notions are important for instrumentation strategies that use historical values (1990) that are independent of current productivity shocks to instrument for current inputs or that use past level variables which describe city employment and urban versus rural conditions to instrument for current changes. While we have focused on labor, capital allocations are subject to frictions and quantity controls, as well as an accumulation process, that also may make historical factors good instruments for current values and changes.

## 2.3 Econometric Implementation

In formulating estimating equations, first we specify the full set of covariates. What is observed for each city is value-added in the industrial and service sectors. If  $\hat{Y}^i$  is “value-added”, evaluated in national markets at uniform prices, then what we observe is what is the local evaluation  $Y^i$ . The link is transport costs; purchasers of products of cities further from national markets will pay higher transport costs and pay producers in those cities less for any amount of output. We model this as  $\ln \hat{Y}^i = \ln Y^i + T^i \phi$  where  $\phi$  is a vector of parameters and  $T^i$  a set of variables including distance from city  $i$  to the coast and whether the city is on a national highway or not. In addition to access determinants, we need to specify technology variables,  $A_t^i$ , such as education of the adult population in 1990 and share of FDI in the local capital stock, which shift productivity. It may also be the case that access variables influence effective technology. Finally there is public capital, which we control for by a measure of the surface area of roads per capita within the city.

For cities, we will present two types of estimates. The first, as a benchmark and to use in comparisons with the rural sector, ignores city output composition effects and non-linear scale effects (the inverted- $U$ ). Then based on eq. (7) and (8a) the estimating equation for city  $i$  in province  $p$  is

$$\ln (Y_t^{ip}/L_t^{ip}) = A_t^{ip} \gamma - T^{ip} \phi + \alpha \ln (K_t^{ip}/L_t^{ip}) + \sigma \ln (Y_t^{ip}/Y_{It}^{ip}) + \beta \ln L_t^{ip} + u_t^p + \delta_t + \varepsilon_t^i \quad (13)$$

In (13)  $\alpha$  is the capital share coefficient and  $\beta$  is any degree of undiminished (dis)economies of scale, where the social marginal product of labor is  $(1 - \alpha + \beta)$ . Note estimates of production functions reflect underlying (mis)allocation rules imposed on cities; and marginal products are “social” marginal products, reflecting underlying scale externalities. The social marginal product interpretation is consistent with firms having constant returns to scale technology with a labor share coefficient of  $(1-\alpha)$ , so  $\beta$  represents scale externalities (as opposed to internal scale economies). However as we saw in footnote 2, it is possible to “mix and match” given internal allocation rules, so  $\beta$  represents overall scale effects, without requiring  $(1-\alpha)$  to be labor’s compensation coefficient.

In (13), the term  $\ln (Y_t^{ip}/Y_{It}^{ip})$  controls for the ratio of total activity to that of independent accounting units, to account for the fact that capital is only observed for the latter units. If both sectors face the same opportunity cost and capital intensity then  $\sigma = \alpha$ .<sup>4</sup>  $A_t^{ip}$  is a vector of variables influencing technology and  $T^{ip}$  a vector influencing accessibility. This specification investigates for Chinese cities

---

<sup>4</sup>We also need to assume in ordinary estimation that the error structure affects  $Y_t^{ij}$  and  $Y_{It}^{ij}$  in proportion equally, so the error structure doesn’t affect the ratio.

whether overall there are economies ( $\beta > 0$ ) or diseconomies ( $\beta < 0$ ) to urban production. However we believe ultimately that real output per worker for a given  $Y/L$  has an inverted  $U$ -shape relationship to  $L$  which varies with industrial composition. This leads to the second specification.

A determinant city size requires first net scale economies and then diseconomies. Additionally as discussed earlier, we want to allow for urban specialization with different types of cities with differing equilibrium or efficient sizes. We use the ratio of secondary to tertiary sector activity, which we call manufacturing and services to indicate how the constellation of scale economy effects varies as we move up the urban hierarchy (see earlier discussion and footnote 2). To capture the ideas that (1) real value-added per worker in cities is an inverted  $U$ -shaped function of city size and (2) that the shape of the function and its peak point are related to industrial composition, we specify

$$\begin{aligned} \ln(Y_t^{ip}/L_t^{ipj}) &= A_t^{ip}\gamma - T^{ip}\phi + \alpha \ln(K_t^{ip}/L_t^{ip}) + \sigma \ln(Y_t^{ip}/Y_{It}^{ip}) \\ &+ \beta_0 MS_t^{ip} + \beta_1 L_t^{ip} + \beta_2 (L_t^{ip})^2 + \beta_3 (L_t^{ip} MS_t^{ip}) \\ &+ u_t^p + \delta_t + \varepsilon_t^i \end{aligned} \quad (14)$$

where  $MS$  is the manufacturing-service ratio (entered separately to control for composition effects on evaluation of output). We generally use the ratio of value-added in the two sectors but also experiment with the employment ratio.

In (14), scale effects are expressed as a quadratic in an exponential function (explaining  $Y/L$ ), where the linear term varies with the manufacturing to service ratio,  $MS$ . Output per worker peaks at an employment of

$$L_t^{*ip} = \frac{\beta_1 + \beta_3 MS_t^{ip}}{-2\beta_2}, \quad \beta_2 < 0, \quad \beta_1 + \beta_3 MS > 0 \quad (15)$$

We will discuss results on more complex forms to (14) and (15) where there are more interactive terms. In (15) the  $MS$  ratio affects scale economies, but the allocation of factors within the city between manufacturing and services is hidden. However if  $\alpha$  is not affected by  $MS$  (as will be the case), then that ratio just affects the degree of scale economies all firms experience in the city.

### 2.3.1 Error Structure

In the error structure in eqs. (13) and (14),  $\delta_t$  are national shocks and a control for inflation and national price reevaluations.  $u_t^p$  is either a provincial time invariant ( $u^p$ ) fixed effect or a province-wide effect that varies from period to period. The provincial fixed or time-fixed effects represent unobservables

such as the extent of market liberalization in the province, provincial natural resources, and provincial know-how that is common to all cities within a province. China under Mao was a very decentralized country with even the basic planning functions effectively assigned to the provinces. That provincial autonomy remains to some extent. So economic regimes may effectively vary by province; hence the strong role for provincial fixed effects. City effects  $\varepsilon_i^j$  might be assumed to be i.i.d. That is, there are no city unobservables that persist in influencing productivity from period to period, given the turbulence in city output per worker noted earlier. This idea is not supported by Bruesch-Pagan tests which overwhelmingly reject no covariance of within city errors over time for any models we estimate.

To explore the issue of changes in underlying city productivity with the regime change in 1993-1994, we estimated casually a fixed effect in the data for 1990-1993 and then again for 1994-1997 for prefecture level cities, with the 1993-94 split representing the regime shift in service sector evaluation and liberalization. To get a fixed effect, we assigned capital, a priori, a coefficient of .4 and labor one of .6 (no scale economies), based on international results and Chinese data. Using these shares we calculate TFP per city and regress that on time dummies and exogenous values of access for the two time periods. We then average the residuals for the years in each of the two time periods to get estimates of the two fixed effects. The correlation coefficient between the 1990-93 and the 1994-97 fixed effects are always around 0.8, and are robust to using different capital and labor share coefficients. Thus at a minimum, standard errors should be estimated allowing for within city correlation of errors over time; this we do.

However one would like to experiment with fixed effect methods which allow for endogeneity of covariates to city time invariant unobservables. There are problems with regular city fixed effects, as we will see, due to attenuation bias from measurement error (Griliches and Hausman, (1986)). The measurement error seems to arise from several sources. First as noted above, there appear to be "regime" switches for certain cities typically between 1993 and 1994, where magnitudes of  $Y/L$ ,  $K/L$ , and manufacturing to services jump by about 100% with no corresponding change in urban scale, indicating a shift in the way quantities are evaluated. Second, growth is rapid in the time interval and there appear to be lags in updating, say, employment figures compared to output figures, so magnitudes which should move in unison between periods don't. Third, city expansions of defined boundaries in the data are discontinuous, with employment and other growth recorded in some cities as occurring disproportionately in one time period, when a city updates its definition (indicating before that time period there was undercounting in terms of what was the city operationally). Finally, there is regular measurement error, where numbers for a city appear to just bounce around, as though officials are unsure as to whom to count for employment or how to value output from year-to-year.

To deal with econometric issues there are several possibilities. We focus on simple IV estimation of a 1996-97 production function using appropriate instruments from 1990 that reflect historical allocation decisions under planning but are independent of unobserved influences on city productivity in 1996-97. The intent is to deal with both measurement error and endogeneity of right hand side variables to city errors. The precise set of instruments used is discussed below as is the choice of years (from specification tests). There is also instrumental variables estimation with first differenced equations using lagged level values as instruments. This would deal both with measurement error and with endogeneity of RHS variables to the  $\varepsilon_t^i$ , as well as removing fixed effects; it is a procedure we experiment with but reject for reasons discussed later.

## 2.4 Results

We present three sets of results. One is for a representative city model for prefecture level cities, where city size is not explicitly modeled, employing a standard production function approach. Our discussion of those results is brief, but we elaborate on econometric issues that affect all estimations. The second set contains the key results and employs a model where there could potentially be city types with varying equilibrium or efficient city sizes among prefecture level cities. In both cases we start with baseline ordinary regression results; and then proceed to IV estimators. Finally we look at some combined results for non-prefecture level cities with their limited years of data and more limited set of variables.

### 2.4.1 Representative City Model: Prefecture Level Cities

In Table 3 are a set of basic results for the pooled data on prefecture level cities from 1990-97. Given the noise in the data especially in the early years and problems with outliers in the pooled sample, we tend to rely on least absolute deviation [LAD], or median regression when pooling earlier and later years, with 200 bootstrap repetitions for the calculation of standard errors. The basic results control for separate time and provincial fixed effects. F-tests reject any improvement in estimation for OLS results from allowing province-time fixed effects (i.e. separate provincial dummies for each year, to control for annual provincial level shocks); and results are almost identical across the two specifications. So we stick with separate time and provincial fixed effects. In this base model, OLS and LAD results are almost the same; but in later specifications they can differ. Column (1) gives OLS results and column (2) LAD results. Standard errors in column (1) are calculated based on within city clustering of errors in OLS, so errors may be heteroskedastic across cities and correlated over time within cities. But the

clustering structure seems conservative, so, for example, standard errors under clustering are generally greater than under bootstrapping in LAD which imposes no specific structure.

The LAD results in column (2) are the focus of initial discussion. In column (2), the capital coefficient is .420, and the labor coefficient is .60. Overall constant returns to scale can't be rejected, a point we return to later. For other variables, knowledge accumulation and access seem very important. A one-standard deviation (8.5) increase in the percent adults with high school leads to a 8.9% increase in output per worker. Similarly FDI seems important as a source of technology transfer. The covariate is accumulated FDI from 1990 to the current year divided by capital stock in the current year, effectively trying to measure the ratio of capital from foreign investment to all capital. A one-standard deviation (.069) increase in this ratio raises output per worker by 5.1%, indicating the benefits of technology transfer. For access, a one-standard deviation (6.1) increase in access to the coast leads to a 14% increase in value-added per worker, while being on a highway raises value-added per worker by 6.9%. Public capital also plays a role with a one standard deviation (1.3) in roads leading to a 5.8% increase in output per worker.

Finally there are the controls for the ratio of total activity to that of independent accounting units. The effect of controls is always to raise the capital coefficient, given in fact that  $\ln K$  and the ratio are both conceptually and empirically negatively correlated. In this case the rise is modest (without a control the coefficient on capital is .39); but in some other situations such as IV estimation the capital coefficient can double with the inclusion of the control. For the years 1990-96, the control is the estimated ratio of total gross to gross output of independent accounting units, a variable that at times is clearly poorly measured (see data Appendix). For 1997 the control is what we want – the recorded ratio of total value-added to value-added of independent accounting units. The coefficient on either variable in column (3) is much less than capital's share coefficient, so  $\alpha \neq \sigma$  in equation (13), based on equation (8a). This could indicate that, for prefecture level cities, the two sectors face different opportunity costs of capital as in eq. (8). This is a point we will return to below.

The results ignore endogeneity issues, beyond provincial fixed effects. City fixed effects estimates of the pooled data suggest that attenuation bias due to measurement error is a major problem here. In Table 3, column (3), we present city fixed effect results (which do not include controls for the ratio of total to independent accounting units activity, because the variable is not consistently defined across years). For standard fixed effects, while the capital coefficient only drops from .39 to .35, the labor coefficient drops from .61 to .39 and overall severe decreasing returns to scale would be indicated.<sup>5</sup>

---

<sup>5</sup>The capital coefficient without control for the ratio of total output to that of independent accounting units is .389 and

If the only problem was serially uncorrelated measurement error, a fix-up would be fixed effects IV estimation using lagged values of covariates as instruments.<sup>6</sup> However measurement errors will be serially correlated given our earlier discussion of the data; and, most critically, lagged covariates are unlikely to be exogenous to both the fixed effects and all contemporaneous error terms. We need to move beyond standard fixed effects methods.

A more sophisticated version of what is attempted in column (3) would difference years to eliminate the fixed effect and then instrument in each differenced equation year with lagged values of variables. This would potentially deal with both measurement error and endogeneity of covariates to the contemporaneous error term, as desired. Using GMM (Hansen (1982), Arellano and Bond (1991)) we would also account for serial correlation across differenced equation years and within year heterogeneity. We will footnote one of those results later; but for the current specification they are similar to fixed effects. The problem is one of weak instruments. The idea from eqs. (10) to (12) that past values of employment are correlated with current changes is supported by first stage regressions of (differenced) covariates on (level) instruments. However for other variables such as  $\ln K/L$ , first stage regressions are extremely bad. First stages regressions of  $d \ln (K/L)$  on past levels have adjusted  $R^2$ 's (and  $F$ 's) for 1997, 1996, ..., and 1992 of respectively -.014 (.58), .087 (4.04), .088 (4.02), -.014 (.57), 0 (.92), and .081 (3.82).<sup>7</sup> Accordingly we don't focus on IV estimation for differenced equations in the paper.

**IV Estimation.** One way in this context to deal with both measurement error and endogeneity is to estimate 1996-97 levels equations by IV methods, using 1990 variables as instruments. We could potentially add 1995 or, even 1994, as estimating years as well. However Sargan specification tests (see later) indicate this is not a good idea. Using 1990 instruments requires two assumptions. First and quite reasonably, measurement error in 1996-97 is not correlated with 1990 measurement error, given the move to strict market accounting and the long time lag. Second is the assumption that 1990 instruments are uncorrelated with errors affecting productivity in 1996-97. We have two kinds of 1990 variables. First are ones that, from the framework of eqs. (10) to (12), affect city employment and population for example, but not productivity, such as consumer amenities (telephones per capita, books

---

the labor coefficient is .014.

<sup>6</sup>In that case the capital coefficient relative to regular fixed effects bounces up to .7, an extraordinary capital share coefficient. Labor's share remains unchanged, but overall there would be strong increasing returns to scale (the sum of factor coefficients rises from 0.76 to 1.1).

<sup>7</sup>Instruments are  $t-2$  values of  $\ln K/L$ ,  $\ln L$ , the manufacturing to service ratio,  $\ln (Y/L)$ , share of agriculture in local GDP, FDI/capital stock, and the ratio of total output to output of independent accounting units. Adding  $t-3$  variables as instruments only worsens the  $F$ 's.

per capita, and doctors per capita), and factors affecting the supply of labor to the city.<sup>8</sup> Second are city variables that were determined in a planning context historically and which we assert are not influenced by market unobservables affecting productivity after 1995.<sup>9</sup>

The instruments form a strong set, with high first stage  $R^2$ 's (typically .5 to .6) and high  $F$ 's as well. The question is whether they are exogenous-determined in a planning context and uninfluenced by unobservables influencing productivity after the market reforms of 1993-95. We excluded as an instrument 1990 employment or population since early migration may have responded to a city's inherent productivity factors and we excluded 1990 education attainment for similar reasons. As one test of exogeneity, we examined the simple correlation coefficient between the 1994-97 fixed effect calculated earlier<sup>10</sup> and the 1990 instruments. The instruments generally have simple correlation coefficients with the fixed effect of less than .14; the two worst exceptions are telephones per capita and FDI/capital stock but they only have simple correlation coefficients around 0.2.<sup>11</sup> Specification tests (see later) on the over-identifying restrictions indicate we have a solid specification.

In IV estimation, we estimate the model as two separate equation years (with hence separate influences of instruments), instrumenting with 1990 variables<sup>12</sup> and allowing for within year heteroskedasticity and across year serial correlation, using DPD 98 (Arellano and Bond (1991)). (In this case first and second step estimates produce similar standard errors). Without heteroskedasticity and an unbalanced panel, the results would be 3SLS ones. Simple 2SLS results on the key parameters are similar.

In Table 3 are OLS and IV results for 1996-97 in columns (4) and (5). Regional, rather than provincial, dummy variables are used given the limited sample size. Compared to the LAD results in column (2), in the IV results, in column (5), capital's coefficient is modestly lower. Again, overall scale economies are rejected. The coefficient on the control for the ratio of total gross output to gross output of independent accounting units in 1996 rises to .442. and the 1997 control for the value-added ratios also rises. IV results on education, access to the coast, and being on a highway are similar to results in

---

<sup>8</sup>These are the ex-urban (rural) population outside the city proper but within the municipal district (which is typically the size of a USA state like Connecticut) and the share of agriculture in city value-added.

<sup>9</sup>These include for 1990, land area of the city, the ratio of manufacturing to services, the ratio of gross output of independent to non-independent accounting units, whether the city had 1990 FDI (with policy at the center designating cities eligible for FDI), the city capital to labor ratio, and the 1990 ratio of FDI to capital stock.

<sup>10</sup>This fixed effect is the average residual of a regression of  $\ln(Y/N) - .4 \ln(K/N)$  on the access variables.

<sup>11</sup>With a sample size of 207, the 95% confidence interval around zero for a simple correlation coefficient between "variables" is  $\pm .14$  (but the fixed effect is estimated). It is interesting to note for example that 1990  $\ln(K/L)$  has a negative correlation (-.07) with the 1994-97 fixed effect. However the 1997  $\ln(K/L)$  has a positive correlation (.13) with the 1990-93 fixed effect and the percent change in  $K/N$  from 1994-97 has a correlation coefficient of .11 and .17 with the 1990-93 and 1994-97 fixed effects respectively. This suggests investment since 1994 is more oriented to high inherent productivity cities.

<sup>12</sup>Experimenting with using '91 instruments for 1997 and '90 instruments for 1996 yielded very similar results, but specification tests suggested '91 instruments perhaps shouldn't be used (the Sargan p-values drop noticeably).

column (2). The impact of FDI is strengthened, as is public infrastructure under IV estimation. Finally we note that regional dummies in Table 3 and next in Table 4 are never significant, indicating we may have already captured determinants of regional differences.

In terms of specification tests on over-identifying restrictions, the Sargan  $p$ -value is an acceptable .102, for the column (5) IV estimation. In the estimation to follow allowing for non-linear scale effects the  $p$ -value rises to .98 (see Table 4, column (4)), indicating a very good specification. In that Table if we add 1995 to the IV model the Sargan  $p$ -value drops to .078, indicating that adding 1995 is not a great idea. The 1990 instruments may not be valid for 1995.

#### 2.4.2 Non-Representative City Model: Prefecture Level Cities

The previous section generally suggested overall constant returns to scale in the urban sector. It is interesting to plot the relationship between output per worker,  $Y/L$ , and city scale  $L$ . Figure 1 does this for real output per worker for cities for 1995-97 data. Figure 1 shows the data points and then draws a curve through the median values of  $Y/L$  for ten intervals of  $L$ . Whether we look at the entire range of  $L$  or truncated versions to see more detail, the curve itself is rather flat, although there is evidence in the 0-200 range of  $L$  (in 10,000's) of an inverted U-shape. However, in a free migration context, even with scale (dis)economies we would be looking for a flat line. As discussed earlier, with different kinds of cities, the utility, or real income curve each city can offer varies with the manufacturing to service ratio [ $MS$ ], so that with an inverted  $U$ -shape, the peak point to  $Y/L$  (and hence to real compensation per worker) in terms of  $L$  shifts as  $MS$  changes. Then in equilibrium under free migration, cities would operate, each near their peak point, to offer roughly the same real output per worker.

To quantify these ideas, we now turn to the specification in eqs. (14) and (15). Results are in Table 4. Columns (1) and (2) give OLS and LAD estimates for the pooled 1990-97 sample. (Now coefficients under LAD are rather different from OLS given the more complex formulation.) Columns (3) and (4) give OLS and IV results for 1996 and 1997. We start by focusing on the non-linear scale effects as interacted with the manufacturing to service ratio. First we note that the only interactive term that has a consistent significant effect is  $MS$  interacted with  $L$ . If we enter a term  $MS * \ln(K/L)$ , the coefficient is always small and insignificant indicating capital intensity doesn't vary with industrial composition. Similarly, interacting education, access to the coast, or public capital (roads) with employment or  $MS$  with employment squared in all Table 4 formulations produces insignificant coefficients. The most promising in terms of quantitative effects is education, but its effects weaken statistically in later years, perhaps reflecting the fact that 1997 education is not that well measured by

1990 education.

For the non-linear scale effect in Table 4, column (2), we have  $(.00334 - .00121 MS) L - .00000420 L^2$  where coefficients are all statistically strong. For these coefficients, we solve for the peak point to the inverted  $\ln(Y/L)$  in Table 5, as it varies with the manufacturing to service ratio. The solution is only relevant if  $(.00334 - .00121 MS) > 0$  or if  $MS < 2.76$ , which is about a standard deviation above its mean (or median) value for the pooled sample. Beyond this value, the functional form suggests undiminished scale diseconomies. In Table 5 we see that the employment where output per worker is maximized increases as the manufacturing to service ratio declines. As in the USA, bigger cities are better sites for modern service activity. So for Table 4, column (2) results, at a manufacturing to service ratio of .75, the peak is at 3.0 million workers, while for a ratio of 2.5 it is at 500,000 workers. City populations tend to be double employment.

Once we move from the pooled sample to later years, the results in columns (3) and (4) in Table 4 on the shape of the inverted -U differ from those in columns (1) and (2). We focus on the IV results in column (4) and the solution to how the peak varies with  $MS$ , as reported in Table 5.<sup>13</sup> Generally peak points for any  $MS$  occur at smaller employment levels under estimation for the '96-'97 years, and the relevant range of  $MS$  for calculating a peak ends at 2.11. However for 1996-97, 2.11 is still one standard deviation above its then mean value in the estimating sample. The two rows on peak points in Table 5 are not entirely inconsistent. The explanation is that by definition (how workers and output were classified) the manufacturing to service ratio fell by about 30% from 1993-1995, so the '96-'97 estimates reflect the modern definition.

In Table 5, the peak point for the IV estimates varies from 3.1 m. workers at the lowest  $MS$  ratio of 0.4 to .21 m. workers at the highest  $MS$  ratio of 2.0 for which we do calculations. To confirm the approximate quadratic shape, as it varies with  $MS$ , we did two other experiments. First we estimated the quadratic interacted with  $MS$  for each year separately in OLS and LAD. There is a split at 1994 with coefficients from 1994 on looking similar and very different from those for earlier years but the quadratic relationship holds in all the different years. Second, for 1996-97 in OLS and LAD we broke  $MS$  into six equal size groups and estimated coefficients on  $L$  and  $L^2$  for each group. Results for lower  $MS$  intervals are quite strong and they support those in Table 4 and 5. In the highest  $MS$  group, as in Tables 4 and 5, there are undiminished diseconomies to size.

---

<sup>13</sup>GMM results on differenced equations using lagged level values as instruments have a strong quadratic form to  $L$ , with a peak at 4.0 m to  $\ln(Y/L)$ . However there is no variation with  $MS$ . But given very weak instruments this is not a surprise.

Do Chinese cities operate to the left or right of the peak point to this inverted-U shape to  $Y/L$ ? To determine this, we calculated (by the delta-method) the 95% confidence interval about each peak point for each  $MS$  value. These are reported in the last two rows of Table 5. At the tails to  $MS$ , the intervals are large. So at  $MS = .4$ , the interval is 2.1 m. to 4.2 m; while, at  $MS = 1.6$ , the interval is .11 to 1.8 m. and at higher  $MS$  the lower bound is negative. For typical  $MS$  values around 1.2 the band is tighter (1.1 m. to 2.3 m.) – but appropriately still wide, given the challenge of the exercise.

Despite these wide bands, many Chinese cities in 1997 fall outside them. In particular 58% of cities have employment below the lower bound on the 95% confidence interval. About 13% of cities have  $MS$  values too high to calculate a peak. About 10% of cities have sizes beyond their peak point, although only 3% significantly so. But these variations allow us to trace out this shifting (with  $MS$ ) inverted U-shape to  $Y/L$ .

The fact that 58% of prefecture level cities have employment below the 95% confidence interval would indicate a severe tendency to having undersized cities. But we must interpret this with care. First we note that with free migration and city formation, we would want cities to operate at their peaks (if all  $A^i$ 's in eq. (7) are the same). In each city we want to maximize  $(Y - rK)/L$  where  $r$  is the return paid to capital; or we want to maximize the residual return to labor per worker.<sup>14</sup> If  $K/L$  is fixed or capital is allocated across cities so  $\alpha Y/K = r$ , then by maximizing  $Y/L$ , the residual return per worker is maximized. This analysis presumes labor gets the residual either directly from firms (if they are CRS) in the form of  $(1-\alpha) Y/L$ , or some directly from firms and the rest in compensation from city provided housing and services. Regardless, from eq. (10) utility of workers is not strictly this real income per worker. We don't have city amenities, including public services, pollution, climate, and diversity of consumer products. A second point is that, theoretically, neither equilibrium nor optimal city sizes will be at peak points, even ignoring the issue of internal misallocation roles and imperfect migration, because models indicate that cities with better amenities will be larger than cities of the same type (industrial composition) with lower amenities. The latter will operate near their peaks and the former with shifted up inverted  $U$ -shape curves will operate beyond their peaks (Henderson (1988)), under either equilibrium or efficient sizes. Of course, this fact simply bolsters the case of severe undersizing of cities in China.

Finally our calculations are based on a simple exercise, where peak points are (1) calculated from

---

<sup>14</sup>When  $Y/L$  is maximized then the social marginal product of labor equals  $Y/L$ , which defines "optimal" city size (for invariant  $A^i$ 's); see Henderson (1988).

just a quadratic, hence creating a range for high  $MS$  where the function doesn't fit,<sup>15</sup> (2) insensitive to technology factors such as education, (3) based on a crude industrial composition measure, and (4) based on GDP measures that may not miss some increasing costs associated with increased city sizes.<sup>16</sup> The wide error bands may only partially account for these factors. Still the results are suggestive of a severe lack of agglomeration, occasioned by migration restrictions.

Finally we note that, in this model with a non-linear formulation to scale, the role of certain other variables changes. Results for other variables in columns (1) and (2) in Table 4 are similar to those in Table 3. However results in columns (3) and (4) of Table 4 differ. Education's role is reduced, although that could occur because 1990 education is a weak measure of education relevant in 1996-97 in rapidly growing cities. IV estimation in column (4) in Table 4 has insignificant coefficients for being on a highway and for public infrastructure (roads per capita). However FDI and access continue to play very strong roles. A one-standard deviation in the FDI ratio raises output per worker by 8%, while a one-standard deviation increase in access to the coast raises output per worker by 15%. Finally IV estimation weakens the control for the ratio of total gross output to that of independent accounting units in 1996 and, in general, the results suggest  $\alpha$  and  $\sigma$  are not equal in eqs. (8a), (13) and (14). The effects of IV estimation of the non-linear scale model on coefficients for roads, education, and the ratios of total activity to that of independent accounting units suggest that these variables in Table 3 may have picked up some of the non-linearities in city scale that are modeled in Table 4. Note that the non-linear scale formulation in Table 4 has a much better Sargan value than the formulation in Table 3, indicating a better overall specification (for the same set of instruments).

**Capital Allocation.** Independent accounting units which are disproportionately state-owned enterprises are viewed as facing a lower effective price of capital compared to non-independent accounting units. Table 4, column (4) results suggest an  $r_I/r_N < 1$ , given estimates suggest  $\sigma < \alpha$ , based on eq. (9). This conclusion is consistent with results in section 4, below. There we will observe an overall marginal product of capital in prefecture level cities of .33 and in rural industry of .48. If we assume non-independent accounting units in our prefecture level cities face the same opportunity cost of capital as rural industry cities so  $r_N = .48$  we can calculate  $r_I$  from  $r_N Y_N/Y + r_I Y_I/Y = .33$  (overall return in prefecture cities). In 1996, based on gross output data,  $Y_N/Y = .3$ , in which case  $r_I = .27$

<sup>15</sup>Results using the ratio of employment for  $MS$  (rather than value-added) are weaker and indicate peaks at large employment levels which don't vary much with  $MS$  (the three coefficients are .0055, -.000011 and -.00059 where the last, interactive one is insignificant). We believe manufacturing employment is poorly measured given employees in state-owned firms may be carried on the books after retirement or layoffs.

<sup>16</sup>That is, we do not have purchasing power parity numbers. However given the more uniform pricing of housing and other non-traded goods for China this may not be the issue it would be for the USA.

or  $r_I/r_N \approx .56$ . What is interesting is that the result is consistent with the coefficient on  $\ln(Y/Y_I)$  in Table 4, column (4) for 1996 of .28. Why? If we plug in  $r_I/r_N = .56$ ,  $\alpha = .48$ ,  $Y_I/Y = .7$ , in eq. (9) the coefficient on  $(Y/Y_I)$  has a value of .29, which is close to .28. For 1997, if  $Y_N/Y$  based on value-added data equals .6, we can solve  $r_I = .11$  so  $r_I/r_N = .23$ . That corresponds to a coefficient in eq. (9) of .18, which is similar to the coefficient for 1997 of .16 in Table 4. However the unanswered question is why  $r_I/r_N$  is greater in 1996 than 1997, which suggests our ability to uncover its value is limited.

In theory, we can directly estimate  $r_I/r_N$ , by estimating eq. (8) by non-linear methods, or by estimating

$$\begin{aligned} \ln(Y_t^{ip}/L_t^{ip}) &= A_t^{ip}\alpha - T^{ip}\phi + \alpha \ln(K_t^{ip}/L_t^{ip}) + \beta \ln L_t^{ip} + \alpha \ln(Y_t^{ip}/Y_{It}^{ip}) \\ &+ \ln(Y_{It}^{ip}/Y_t^{ip} + (r_I/r_N)^\alpha Y_{Nt}^{ip}/Y_t^{ip}) + w_t^p + \delta_t + \varepsilon_t^i \end{aligned} \quad (16)$$

Since  $Y/Y_I$  is only measured accurately (value-added) for 1997, given we are constraining coefficients, we can only legitimately do this for 1997. The basic non-linear result yields an  $r_I/r_N$  value of .159 (with a standard error of .090), which is consistent with the other 1997 results. Unfortunately the critical estimates are IV ones and the IV results under non-linear estimation are unstable and  $r_I/r_N$  coefficients are always insignificant.

### 2.4.3 Non-Prefecture Level Cities

We now turn to non-prefecture level cities where we have more limited data. Rather than eight years, there are five years of data. We have less information on base characteristics of cities (to use as instruments). Finally there is extensive entry of cities, so for many cities we have only the later years.

The results for non-prefecture level cities are in Table 6. Columns 1-3 cover the base case where complex scale effects are not modeled; and columns (4) and (5) cover the more complex formulation. Columns (1), (2) and (4) pool all years ('90, '91, '94-'96) while columns (3) and (5) contain IV estimation for '95 and '96. We discuss details of IV estimation momentarily. We start with some general comments that hold across columns. Relative to prefecture level cities in Table 3 in particular but also in Table 4, the first obvious difference is that the role of access and knowledge externalities is sharply reduced. Education, distance to the coast and being on a highway have weaker and often statistically insignificant effects, compared to prefecture level cities. This suggests these cities may more service local markets with less sophisticated products. Only FDI/capital has very strong effects in both non-IV and IV estimation. The capital coefficient moves around a bit from column to column, but seems a little less

than in prefecture level cities. Finally, for these cities, the role of the ratio of total gross output to that of independent accounting units seems more stable. Its inclusion generally increases the coefficient on  $\ln(K/L)$  by 50%; and its magnitude suggests  $\sigma = \alpha$ , so there is no discrimination across sectors in these small cities.

The big issue in Table 6 is scale economies. OLS and quantile regressions suggest overall diseconomies of scale in columns (1) and (2). In column (4) with the more complex scale formation, it looks like there is also decreasing returns to scale. For an  $MS$  value of 1.5,  $Y/N$  declines up to an employment of 370,000. This is about two-standard deviations above mean employment in these cities, so again this suggests overall decreasing returns to scale.

However this result on (dis)economies is dramatically reversed under IV estimation; but IV estimation is problematical. First, as instruments we no longer have historical values on city amenities (books, doctors, telephones) or surrounding rural populations for these cities. Second, estimating an equation for the last year with early year instruments means we lose entrants and the base sample for IV estimation is 236 cities, rather than 432. Finally the assumption of exogeneity of early period instruments is now much more questionable. We use 1990 variables as instruments. If we calculate a fixed effect for 1994-1996 for non-prefecture level cities as for prefecture level cities,<sup>17</sup> now fixed effects are strongly correlated with instruments. For variables like  $\ln(K/L)$ ,  $\ln(\text{ratio total gross output to that of non-independent accounting units})$ , and whether a city has FDI or not, we have simple correlation coefficients between the fixed effect and these variables of .33 to .43 in absolute magnitude. Also given a lack of variables for instruments we need also to use employment variables, with simple correlation coefficients of around .22.<sup>18</sup> As a result specification tests are problematical as we will see.

For the simple form to scale economies in column (3), rather than (dis)economies, estimates suggest strong economies with an elasticity of .08.<sup>19</sup> However the model (marginally) fails the Sargan test, indicating perhaps the problem with instruments. In the complex scale form in column (5), the model does pass the Sargan test and unexhausted scale economies are indicated also. For a manufacturing service ratio of 1.6, all scale terms are positive. For a ratio of, say, 1.4 after a scale of .16 m., cities experience unexhausted scale economies.

---

<sup>17</sup>We regress  $\ln(Y/L) - .4 \ln(K/L)$  on access variables and take the average residual for 1994-96. Similar results apply if we use  $\alpha = .3$ , rather than .4.

<sup>18</sup>For column (5) in Table 7, the instrument list is distance to coast, on highway, regional dummies, distance to provincial capital, and 1990 values of  $\ln(K/L)$ ,  $MS$ ,  $L$ ,  $L^2$ ,  $MS * L$ , no FDI in 1991, share of agriculture in local GDP, ratio of total to ind. acct. unit output,  $FDI/K$ ,  $\ln(\text{roads})$ .

<sup>19</sup>Accounting for heteroskedasticity in GMM estimation raises the standard errors. For 2SLS, ordinary standard errors leave the coefficient strongly significant. Playing with instruments under 2SLS can raise the coefficient as high as .12 with a high level of significance.

The IV results suggest (1) unexhausted scale effects and (2) sizes of non-prefecture level cities are far too small to trace out the inverted U-shape function. The results are suggestive but not conclusive given the problem in finding valid instruments.

### 3 Productivity in the Rural Sector

The rural sector in China has key features noted in the introduction. Overall it appears that there is a surplus of labor within the sector due to restrictive migration policies. While under the “leave the land but not the village” policy, workers have moved out of farming to non-farm agricultural activity and, most critically, to rural industrial production, there remains a surplus of labor in agriculture, probably felt most strongly in more remote regions where rural industrialization is not very competitive and residents can’t migrate to richer rural and agricultural areas. Second, as we will discuss momentarily, the rural industrial, or town and village enterprise [TVE] sector in more competitive locations, faces constraints that restrict TVE development.

The shift to the TVE sector has been dramatic. In 1981 TVE’s accounted for 6% of all industrial gross output, while the urban state owned and collective sector accounted for the balance. By 1997, TVE’s accounted for 28% of all gross industrial production with the urban state owned and collective share dropping from 93% in 1981 to 40%. Of course apart from the TVE sector’s share growing from 6 to 28%, the absolute growth is astronomical. Despite this relative and absolute growth of TVE’s and blurred ownership distinctions, the rural industrial sector faces two major constraints. First, as noted earlier, the sector is capital constrained. Second, and, as we will argue, more critically, the TVE sector is constrained by forced factor immobility within the rural sector, not only across provinces, but also, in particular, within provinces.

Rural industrial activities can’t relocate to appropriately spatially agglomerate to exploit the local scale externalities inherent in agglomeration, whether they are local information spillovers or firm and intra-industry local specialization in production and exchange of parts, components, and services. Even today, TVE’s cannot readily be bought and sold, nor shares transferred among individuals, so a township in one location cannot readily buy and relocate to their township TVE’s in another location. Similarly if workers leave their township permanently they lose their rights to agricultural land and the in-kind (e.g., housing) distribution of profits of local TVE’s, with a limited ability to garner such rights elsewhere. While provinces such as Zhejiang recognize the problem and have rural industrial cluster policies, little physical clustering and relocation is involved.

If TVE's could spatially relocate and cluster, of course, that would transform the rural sector and invalidate the analytical approach we will take below. With agglomeration the "rural" sector would become two distinct parts, a rural agricultural sector with supporting services and industry and an urbanized sector with some of the many townships transformed into cities. In fact, as noted earlier, the rapid entry of new cities noted above consists in part of rural townships that have grown into cities, so although constrained, some transformation has occurred.

Following Yang and An's (1999) insightful work, we analyze the rural sector as it largely operated in the 1990's – local townships with integrated agricultural and non-agricultural sectors. A town's function in part is to allocate resources to farm activity (farming), to non-farm agricultural activity which is mostly animal husbandry and fisheries, and to TVE and related activities covering manufacturing, services, construction, transportation and so on. Our data are limited. We have provincial level data for three years, 1995, 1997, and 1998, on value-added in the agricultural and in the TVE sector. We have specific TVE employment numbers, but agricultural employment is difficult to measure, being simply defined as persons "primarily engaged" in agriculture. While we have capital figures (original value of assets, with similar results for horsepower) for TVE and agricultural production, we are not confident of the accuracy of their time variation in agriculture. And time variations for land in agriculture are poorly measured. We start with a meta-production function approach for value-added in the rural sector, as Yang and An (1997) do, assuming allocation rules corresponding to those in equation (4) for the urban sector. But given the data (see next) for the TVE sector appear to be better than for the overall rural sector, we examine also the TVE sector separately.

Why do we do our own examination at all? Jefferson and Singhe (1999) have robust results on productivity parameters based on extensive micro data on the TVE sector, that are consistent with other results on TVE's in China. There are three reasons. First, it is helpful to confirm their basic results on TVE's. Second, we want to reformulate the estimation and nature of scale effects<sup>20</sup> for TVE's. Third, estimates for agriculture and the overall rural sector in China are all over the place. Labor share coefficients range from .60 (Yang and An (1997)) to .14 (Lin (1992a)) to .06 (Putterman (1993)), with the vast variation in the literature remaining even when units of analysis (households, production teams, provinces) are the same (Putterman and Chiacu (1994)). While our estimates of labor share will correspond to those of Yang and An (1997), they generally find zero returns to land and almost zero returns to capital. We believe our results and the formulation of scale effects, to be

---

<sup>20</sup>Not only do Jefferson and Singhe have output on both the LHS and RHS of the estimating equation which raises issues of error structure, but they don't examine township scale effects per se but rather related plant scale effects.

more plausible, despite our inferior data and very limited sample size. Fortunately for the parameters of interest, the private and social marginal product of labor, our results on TVE's and on the overall rural sector correspond to those in Jefferson and Singhe (1999) and Yang and An (1997) respectively.

For the overall rural sector, value-added in province  $p$ ,  $Z^{Rp}$ , is postulated to be a function of capital  $K^{Rp}$ , sown land  $S^{Rp}$ , labor  $N^{Rp}$ , and education  $E^{Rp}$  in the rural sector. We estimate

$$\begin{aligned} \ln (Z_t^{Rp}/N_t^{Rp}) &= \alpha \ln (K_t^{Rp}/N_t^{Rp}) + \gamma \ln (S_t^{Rp}/N_t^{Rp}) \\ &+ \beta \ln N_t^{Rp} + \theta E_t^{Rp} + \delta_t + \mu^p + e_t^p \end{aligned} \quad (17)$$

In (17), given a meta-production function with underlying labor allocation rules, the "social return" to labor is  $[(1 - \alpha - \gamma) + \beta] Z^{Rp}/N^{Rp}$ , assuming scale effects derive from employment scale (although other interpretations can be readily utilized<sup>21</sup>) and that scale effects derive from adding a worker to the rural sector in a province. While adding labor to the provincial rural sector with a fixed number of townships indicates the scale benefits of adding a worker to some township somewhere, it is a little vague. For TVE's, it will be possible to be more specific.

In estimation of (17), the key issues are error structure and data quality. Input choices are not exogenous to provincial fixed effects, representing long term weather conditions, soil fertility, local political climate and extent of economic and social liberalization. We do have regional dummies, distinguishing the coast, central and western regions. However at a minimum, we would like to employ provincial fixed effects, but annual variations in all inputs (especially land) are poorly measured and fixed effect results are insignificant with a zero return to land. We rely on two types of estimates. First are time and region fixed effect estimates, with provincial random effects for the three years of data which are close together. Second are IV estimation results using instruments from 1981. Instruments are values of 1981 variables that meet two criteria. They are determined by political and historical forces under Russian style industrialization, planning, the political chaos from the mid-1960's on, and the military industrial sector, which are independent of current unobserved influences on rural sector productivity; but they have a strong impact on influencing TVE development and agriculture populations today.<sup>22</sup> First stage

<sup>21</sup>For example, as in Jefferson and Singhe, the formulation can be  $Z = Z^\beta K^\alpha L^\gamma N^{1-\alpha-\gamma}$  so that a proper estimating equation would be  $\ln (Z/N) = \frac{\alpha}{1-\beta} \ln (K/N) + \frac{\gamma}{1-\beta} \ln (L/N) + \frac{\beta}{1-\beta} \ln N$ .

<sup>22</sup>Instruments are (1) the number of communes, and their population prior to any free migration reflecting the base from which populations are adjusting today; (2) provincial industrial output and the share of light industry in that outside the rural sector, that formed an early basis for local demand for TVE industrial parts and components; (3) enrollments in regular secondary schools and institutions of higher learning that influence adult educational attainment today; and (4) arable land defined under Mao's "grain everywhere" policy that defines land that can be planted with grain but not necessarily economically viable today (again a base from which China is adjusting in determining what is productive land

regressions indicate a strong set of instruments, although sample sizes for IV results are limited.

**The Overall Rural Sector.** The results for the meta-production function for the rural sector are in Table 7. In the random effect results, the residual labor share coefficient is .53 (=1 -.379 -.094) and the labor scale effect is .10, although the latter is not statistically strong. The two sum to .63, which is within the realm of Yang and An's results on the social marginal product of labor (.52 to .61). The main difference between this result and Yang and An's is that we have significant and reasonable coefficient on capital. The IV results are stronger, and yield an overall labor coefficient of .66, with very significant scale effects of .16. In both cases, education affects productivity very, very strongly where a one standard deviation in percent in the rural population completing high school (2.5) results in a 27% increase in output for IV results. This result supports Yang and An's hypothesis that education improves overall productivity, by improving allocation decisions between the urban and rural sectors.

**TVE'S** Random effects and IV results on TVE's are in Table 8. There are two sets of results. The first is for a generalized scale effect – adding labor to the TVE sector in a province. The second is more precise – the effect of increasing average local township TVE sector size in a province, or the degree of local industrial agglomeration. In general, as we move from random effect to IV results, capital's coefficient rises, but the IV results vary by scale formulation. Results are consistent with Jefferson and Singhe (1999, Table 7.4) who find a capital coefficient of .29, mid-way between the two IV numbers in Table 8. For scale effects, the generalized IV formulation has a scale elasticity of .13 and a social marginal product of labor coefficient of .77 to be compared with Jefferson and Singhe's of .97. Under the more specific township scale effect, IV results suggest a scale elasticity of .30 and a social marginal product coefficient of 1.04. What is generally apparent is the very high level of scale effects in our results, as well as in Jefferson and Singhe. This would suggest that the costs of constraining agglomeration in the TVE sector are very high.

Finally we note that rural education has no direct positive impact on TVE productivity. That combined with a positive overall effect of education on rural productivity suggest the education's effect works through improved allocation decisions for rural sector activities, as Yang and An (1997) suggest. In all results, regional dummies indicate productivity in the west is much lower than in coastal or central areas. We now turn to a general comparison of labor productivity across sectors and regions.

---

today).

## 4 Productivity in the Urban And Rural Sectors

This summary section considers the allocation of resources across sectors in China. The analysis is divided into two parts. First we look at the issue of agglomeration within sectors, holding relative capital allocations (the capital to labor ratio in each sector) constant. The issue is: what are the magnitudes of costs of under-agglomeration within sectors? Second we examine the issue of sectoral allocation, where the sectors are prefecture level cities, non-prefecture level cities, TVE's and the overall rural sector. Given current levels of agglomeration and factor allocations, does the marginal product of labor follow the perceived hierarchy – highest in prefecture level cities, then non-prefecture cities, then TVE's, and then the overall rural sector. If so, does this imply that rural-urban migration would enhance overall productivity?

### 4.1 Agglomeration Gains Within the Spatial Hierarchy

The different sections in this paper argue that spatial agglomeration in China is insufficient. TVE industry and non-prefecture level cities are all undersized, with as yet unexhausted scale economies. Table 9 shows the gains from doubling local scale such as the sizes of non-prefecture level cities and average TVE employment in townships, according to estimates (of  $\beta$ ) from Tables 6 and 8. For TVE's the effects are enormous – a 23% increase in output per worker holding input ratios fixed. It is not clear what the local scale of "rural industry" or non-prefecture level cities should ultimately be. In particular allowing spatial agglomeration of TVE's would transform townships into non-prefecture level cities; and non-prefecture level cities as they enter the data set are still so small as to have unexhausted scale economies as well. It would seem reasonable to presume that, if non-prefecture level cities were permitted to grow, they would become subject to the same econometric representation as prefecture level cities.

For prefecture level cities where there is a huge size range, we are able to estimate the size where output per worker peaks as industrial composition varies. Based on IV results in Tables 4 and 5, Table 9 tells us the percentage gain in output per worker (holding input ratios and other city characteristics fixed) of moving from a size that is 50, 40, 30, and then 20% below the size where  $Y/L$  peaks, to the size where it peaks. The gains are independent of industrial composition (although the peak sizes differ). If a city is half its peak size, the gain is 35%. But the gains drop off quickly as the gap shrinks. So if a city is only 20% below its peak size the gain to moving to peak size is only 4%. However the majority of Chinese prefecture cities are at least 40% below their estimated peak size.

In summary, results suggest that the losses within sectors from underagglomeration are enormous. Doubling sizes of many current rural industrial or urban agglomerations would substantially increase national output.

## 4.2 Allocation Gains Across the Spatial Hierarchy

A focus of Chinese policy has been the gain from permitting increased rural to urban migration. The idea is that rural-urban migration restrictions leave a gap between the marginal product of labor in rural versus urban areas. This gap does appear in the data. For prefecture cities, non-prefecture cities, the rural sector overall, and TVE's, based on results in Tables 3, 6, 7, and 8, we can calculate the (weighted) average social marginal product of labor by region or the nation by multiplying the overall labor coefficient by  $Y/L$  where  $Y$  and  $L$  are the aggregate output and employment in the sector by region or the nation. For these sectors the overall labor coefficient is  $1 - \alpha + \beta$  in eq. (13) or (18). For prefecture level cities we use IV results from Table 3 (column (5)) since it is the simpler formulation that corresponds to what we use for other sectors. For prefecture cities, non-prefecture cities, rural sector and TVE's we use a  $(1 - \alpha + \beta)$  of .60, .73, .65, and .90 respectively.<sup>23</sup> The last represents a compromise between the .78 and 1.04 implied by columns (3) vs. (6) in Table 8 and is near the magnitude in Jefferson and Singhe (1999).

The calculations of  $Y$  and  $L$  are based on 1995 data, the last year data are available across all sectors. Results are in Table 10, Part A. Nationally and within regions, in part A it is apparent that the social marginal product of labor in prefecture level cities is the highest, although it only exceeds that of non-prefecture level cities by a modest 3.6%. This lack of differential could reflect the lack of scale economies estimated in Table 3 for prefecture level cities (compared to Table 4) so we are understating labor's social productivity in those cities. In turn the social marginal product of labor in cities nationally and by region exceeds that in TVE's. Overall the excess over TVE's for prefecture level cities is 40%. The productivity of labor in the overall rural sector is depressing and indicates the enormous surplus of labor in agriculture.

The comparison in Table 10 Part A suggests rural-urban migration would substantially increase national output. However things are not so simple. The ordering of social marginal products of labor in Table 10, Part A is at least in part due to differences in access to capital. Prefecture level cities overall operate with a capital to labor ratio in production that is 130% higher than in non-prefecture

---

<sup>23</sup>The .73 for non-prefecture cities uses an  $\alpha$  of .35 and a  $\beta$  of .08. The reason for raising  $\alpha$  from .31 to .35 in column (3) of Table 7 is discussed below.

level cities and 300% higher than in TVE's. Of course that could be efficient, but Table 10, Part B indicates otherwise. There we calculate the weighted average of the marginal product of capital by sector and region, assigning  $\alpha$ 's to prefecture cities, non-prefecture cities TVE's and the rural sector of respectively 0.40, 0.35, 0.3, and .40.<sup>24</sup> For prefecture level cities and the rural sector these are the estimated  $\alpha$ 's in columns (5) and (2) of Tables 3 and 7 respectively. For non-prefecture cities and TVE's they are a compromise between results in columns (3) and (5) of Table 6 and columns (3) and (6) of Table 8 respectively. The results in Table 10, Part B are consistent with common perceptions (e.g., Jefferson and Singhe (1999)). The marginal product of capital of prefecture level cities is about 35% lower than in non-prefecture level cities and TVE's, overall nationally. Even the productivity of capital in the overall rural sector exceeds that in prefecture level cities. Regionally the same pattern holds with some special details. Traditional, stagnant non-prefecture level cities in the west have very low returns to both capital and labor. TVE's in the coast seem to have easier access to capital, through "partnerships" with firms in prefecture level cities.

A natural question is whether, if we allocated capital on a more efficient intersector basis, would rural-urban migration still produce the same gains? Consider an experiment where we hold labor in place and reallocate capital so as to equalize the marginal product of capital across space.<sup>25</sup> But then there is an issue of adjusting for the ratio of total value-added to that of independent accounting units. To do this we give each city the same ratio. Then for each city we calculate a hypothetical value-added which is actual value-added adjusted for capital allocated on a common cost,  $r$ , and a common ratio of activity of independent accounting units to all units. The result is<sup>26</sup>

$$\tilde{Y}^i = (Y^i)^{\frac{\sigma}{1-\alpha}} (K^i)^{-\frac{\alpha}{1-\alpha}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} \left(\frac{\bar{Y}/\bar{Y}_I}{Y^i/Y_I^i}\right)^{\frac{\sigma}{1-\alpha}} \quad (18)$$

Given these adjustments, the weighted average social marginal product of labor in a group of cities is the labor coefficient multiplied by  $\Sigma \tilde{Y}^i / \Sigma L^i$ . For TVE's and the rural sector, the only adjustment is

<sup>24</sup>Note we are implying industry in prefecture level cities is more capital intensive than in non-prefecture cities and TVE's. This could represent industrial composition differences. But it could also be that prefecture level cities have been so over-endowed with capital for so many decades that their technological development and choices are simply different!

<sup>25</sup>One could of course do the experiment the other way - reallocate labor to equalize its social marginal product and then examine capital allocations. The problem is that mixes in agglomeration (of labor) gains.

<sup>26</sup>Actual city output  $Y^i$  is  $Y^i = J^i (\tilde{K}^i)^\alpha (Y^i/Y_I^i)^\sigma$  where  $J^i$  contains all scale, education access, and error term variables. What we want to calculate is a hypothetical output  $\tilde{Y}^i$  where  $\tilde{Y}^i = J^i (\tilde{K}^i)^\alpha (\bar{Y}/\bar{Y}_I)^\sigma$  where  $\bar{Y}/\bar{Y}_I$  is constant across cities.  $\tilde{K}^i$  is determined by a marginal condition where  $r = \alpha \tilde{Y}^i / \tilde{K}^i$  for  $r$  the national opportunity cost of capital. Substituting for  $\tilde{K}^i$  and  $J^i$  into the first equation yields (18).

for capital. For these sectors

$$\tilde{Z}^p = (Z^p)^{\frac{1}{1-\alpha}} (K^p)^{-\frac{\alpha}{1-\alpha}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} \quad (19)$$

To perform comparisons based on these equations, we assign capital an opportunity cost of .42, between the calculated current marginal product of capital in prefecture level cities and other geographic units in Table 10. We set the ratio of all to independent accounting units (for gross output) equal to the urban average of 1.50 in 1995. For prefecture level cities, non-prefecture level cities, TVE's and the rural sector we assign  $\alpha$ 's of .40, .35, .30 and .40 as before. For non-prefecture level cities, given the results in Table 6 on the coefficient of the ratio of total sales to those of independent accounting units (for 1996), we set  $\alpha = \sigma$ . For prefecture level cities in Table 3 IV results suggest  $\sigma = \alpha$ , but results in Table 4 and other discussion suggests  $\sigma = .28$ . Fortunately that has little effect on outcomes, lowering the returns to labor in prefecture level cities by 2-3%. We go with the lower figures and  $\sigma = .28$ .

Table 11 gives the revised social marginal products of labor "under a level playing field". Nationally, despite a high labor coefficient due to scale externalities and capital reallocation, the social marginal product of labor in TVE's remains 12% or so below that in cities. But now the notion that labor reallocations from TVE's to cities would be desirable is weakened and in the central region is reversed. However we note that we have assigned prefecture level cities no scale externalities (despite Table 4 results). Prefecture level cities still have labor marginal products in excess of those in non-prefecture level cities, but the differences remain very modest. One might have thought capital reallocation would favor non-prefecture level cities, but the adjustment to equalize the ratio of activity of independent to non-independent accounting units removes that advantage.

Finally the rural sector results overall indicate an enormous surplus of labor (in agriculture). Even if we inflate the rural sector numbers in Table 11 by giving the rural sector national average education (12.5% high school) versus its current level (7.4%), while that raises the national marginal product of labor to 5700, it is still less than half that in the urban sector.<sup>27</sup> There is no story to tell that can justify the surplus agricultural labor in China.

Surplus labor in agriculture remains a first order cost in terms of lost GDP. However migration restrictions also produce first order GDP losses by forcing insufficient spatial agglomeration of resources within both the urban and rural industrial sectors.

<sup>27</sup>To do this we inflate  $\tilde{Z}$  in (21) by exp.  $[(.106/(1-\alpha)) (12.5-7.4)]$  where  $\alpha = .40$ , based on an education coefficient of .106 in Table 7.

## References

- Arellano, M. and S. Bond (1991), "Some Tests of Specification in Panel Data: Monte Carlo Evidence and Application," Review of Economic Studies, 58, 537-559.
- Black, D. and V. Henderson (2001), "Urban Evolution in the USA," University of California, Irvine mimeo.
- Cai, Fang (2000), Zongguo Liudong Renkou Wenti (The Mobile Population Problem in China), Henan People's Publishing House: Zhengzhou.
- Chan, H.W. (1994), Cities With Invisible Walls, Oxford University Press: Hong Kong.
- Duranton, G. and D. Puga (2001), "Nursery Cities," American Economic Review, (forthcoming).
- Eaton, J. and Z. Eckstein (1997), "Cities and Growth: Theory and Evidence from France and Japan," Regional Science and Urban Economics, 27, 443-474.
- Eberts, R. and D. McMillen (1999), "Agglomeration Economies and Public Infrastructure", Handbook of Regional and Urban Economics, Vol. 3, P. Cheshire and E. Mills (ed.) North Holland.
- Hansen, L. (1982), "Large Sample Properties of Generalized Methods of Moments," Econometrica, 50.
- Henderson, J.V. (1974), "The Size and Types of Cities," American Economic Review.
- Henderson, J.V. (1988), Urban Development: Theory, Fact and Illusion, Oxford University Press.
- Jefferson, G. and I. Singhe (1999), Enterprise Reform in China: Ownership Transition and Performance, Oxford University Press: New York.
- Kolko, J. (1999), "Can I Get Some Service Here? Information Technology, Service Industries and the Future of Cities," Harvard University mimeo.
- Lin, J.Y. (1992a), "Rural Reforms and Agricultural Growth in China," American Economic Review, 82, 34-51.
- Lin, J.Y. (1992b), The China Miracle: Development Strategy and Economic Reform, Hong Kong: Chinese University Press.
- Prud'homme, Remy (2000), "Estimating the Benefits of Urbanization in China," Report to World Bank on China's Urbanization Policy.

- Putterman, L. (1993), Continuity and Change in China's Rural Development: Collective and Reform Eras in Perspective, Oxford University Press: New York.
- Putterman, L. and A. Chiacu (1994), "Elasticities and Factor Weights for Agricultural Growth Accounting: A Look at Data for China," China Economic Review, 5, 191-204.
- Marshall, A. (1890), Principles of Economics, London: MacMillan.
- Mueser, P. and P. Graves (1995), "Examining the Role of Economic Opportunity and Amenities in Explaining Population Redistribution," Journal of Urban Economics, 37, 176-200.
- Rappaport, J. (2000), "Why Are Population Flows So Persistent?," Federal Reserve Bank of Kansas City mimeo.
- Yang, D.T. and M.Y. An (1997), "Human Capital, Entrepreneurship and Farm Household Earnings," Duke University mimeo.

**Table 1. The Chinese Urban System**

	All cities present in 1990	Prefecture level cities present in 1990	County (non-pref.) cities present in 1990	New (non- pref.) cities
Avg. pop. 1996 (thousands)	811	1050	589	706
Pop. growth 1990-96	12%	15%	7.9%	n.a.
Avg. non-agri. employ. 1996 (thousands)	327	511	158	183
Non-agri. employ. growth 1990-96	29%	26%	40%	n.a.
% population (over 6) completing high school in 1990	14.6%	21.7%	12.8%	8.8%
Distance to coast	7.1	6.3	9.4	5.4
Avg. output per worker in 1996 (1990 yuan)	8942	10198	7802	9229
Growth in real output per worker 1990-96	57%	53%	62%	n.a.
Avg. manu. service ratio 1996	1.52	1.59	1.47	1.59
Change in manu. serv. ratio 1990-96	-25%	-32%	-8.2	n.a.
N	459	220	240	193

**Table 2. Transition Matrices**  
(all 669 cities ever present)

**A. Output per Worker – 1996**

	not present	first quartile	second quartile	third quartile	fourth quartile
not present	6	46	47	51	55
<b>1990</b> first quartile	1	<b>58</b>	35	10	12
second quartile	1	46	<b>33</b>	28	8
third quartile	3	13	33	<b>39</b>	28
fourth quartile	0	2	16	37	<b>61</b>

**B. Employment – 1996**

	not present	first quartile	second quartile	third quartile	fourth quartile
not present	5	56	65	51	28
<b>1990</b> first quartile	1	<b>91</b>	21	4	0
second quartile	2	17	<b>65</b>	24	7
third quartile	2	0	13	<b>79</b>	22
fourth quartile	0	1	1	7	<b>107</b>

**Table 3. Basic Results on Urban Sector**

	(1)	(2)	(3)	(4)	(5)
	OLS	LAD <sup>a</sup>	fixed effects	OLS	IV est.
	('90 - '97)	('90 - '97)	('90 - '97)	('96, '97)	('96, '97)
ln (capital/ employment)	.426** (.0285)	.420** (.0193)	.351** (.0235)	.448** (.045)	.394** (.0532)
ln (employment)	.0117 (.0251)	.0166 (.0154)	-.261** (.0310)	.0438 (.0278)	.00708 (.0356)
% high school 1990	.00963** (.00297)	.0105** (.00170)		.00394 (.00350)	.00907** (.00384)
distance to coast	-.0190** (.00963)	-.0225** (.00477)		-.0199** (.00925)	-.0236** (.00672)
on highway	.0617 (.0424)	.0693** (.0181)		.126** (.0551)	.104** (.0504)
ln (out. all units/out. ind. acct. units) [90-96]	.218** (.0602)	.214** (.0401)		.442** (.100)	.436** (.105)
ln (VA all units/VA ind./acct. units) [97]	.112** (.0561)	.111 (.077)		.169* (.0882)	.227** (.065)

<sup>a</sup>) Standard errors are obtained through bootstrapping with 200 repetitions.

**Table 3. Basic Results on Urban Sector – Continued**

	(1)	(2)	(3)	(4)	(5)
	OLS	LAD	fixed effects	OLS	IV est.
	('90 - '97)	('90 - '97)	('90 - '97)	('96, '97)	('96, '97)
FDI/ capital	.729** (.260)	.744** (.232)		1.00 (.651)	1.33** (.441)
ln (roads per capita)	.0612** (.0212)	.0451** (.0164)		.110** (.0469)	.220** (.0653)
time effects	yes	yes	yes	yes	yes
province effects	yes	yes	n.a.	no	no
regional effects	n.a.	n.a.	n.a.	yes	yes
R <sup>2</sup> [pseudo]	.827	[.643]	.948	.563	n.a.
Sargan p-value	n.a.	n.a.	n.a.	n.a.	.102
N [cities]	1751 [225]	1751 [225]	1762 [234]	417 [212]	417 [212]

<sup>a)</sup> Standard errors are obtained through bootstrapping with 200 repetitions.

**Table 4. Modeling City Scale Effects**

	(1) OLS ( '90 -'97)	(2) LAD ( '90 -'97)	(3) OLS ( '96, '97)	(4) IV estimation ( '96, '97)
ln (capital/ employ.)	.406** (.0388)	.415** (.0244)	.441** (.0556)	.480** (.0804)
employ.	.00272** (.00119)	.00334** (.000658)	.00327* (.00175)	.0125** (.00406)
employ. sq.	-.00000201 (.00000217)	-.00000420** (.00000135)	-.00000230 (.00000216)	-.0000163** (.00000768)
(MS) * employ.	-.00116** (.000534)	-.00121** (.000287)	-.0181 (.00129)	-.00592** (.00171)
MS	.0548 (.0377)	.0398** (.0179)	.0672 (.0875)	.0724 (.0886)
% high school	.00805** (.00284)	.00760** (.00176)	.00317 (.00370)	.00489 (.00475)
distance to coast	-.0179* (.0101)	-.0218** (.00484)	-.0203** (.00936)	-.0240** (.00703)
on highway	.0671 (.0431)	.0690** (.0185)	.137** (.0588)	.0683 (.0586)
ln (output/output of ind. acct. units) 1990-96	.202** (.0627)	.194** (.0416)	.433** (.104)	.282** (.095)

**Table 4. Modeling City Scale Effects – Continued**

	(1)	(2)	(3)	(4)
	OLS	LAD	OLS	IV estimation
	('90-'97)	('90-'97)	('96-'97)	('96, '97)
ln (VA/VA of ind.acct. units 1997)	.105** (.0535)	.0982 (.0700)	.164* (.0929)	.158** (.0583)
FDI/Capital	.736** (.255)	.586** (.203)	.959 (.644)	1.17** (.356)
ln (roads per capita)	.0644** (.0218)	.0432** (.0166)	.105** (.0479)	.0290 (.0864)
time effects	yes	yes	yes	yes
province effects	yes	yes	no	no
region effects	n.a.	n.a.	yes	yes
R <sup>2</sup> [pseudo]	.832	[.648]	.571	n.a.
Sargan p-value	n.a.	n.a.	n.a.	.982
N [cities]	1751 [225]	1751 [225]	417 [212]	417 [212]

**Table 5. Employment a Peak Point of Value-Added Per Worker**

peak point employment (1000's)	manufacturing to service ratio									
	.40	.60	.80	1.0	1.2	1.4	1.6	1.8	2.0	2.6
median regression '90-'97	3400	3110	2820	2540	2250	1960	1670	1380	1100	230
IV '96-'97	3110	2750	2390	2030	1660	1300	940	580	210	
<b>IV 95% confidence interval – lower</b>	2090	1890	1690	1430	1070	630	110			
– upper	4170	3610	3080	2620	2250	1970	1770	1600	1450	

**Table 6. Non-Prefecture Level Cities**

	(1) OLS (’90, ’91, 94-’96)	(2) LAD (’90, ’91, 94-’96)	(3) IV estimation (’96, ’95)	(4) LAD (’90, ’91, 94-’96)	(5) IV estimation (’96, ’95)
ln (capital/ employ	.362** (.0300)	.420** (.0224)	.314** (.0559)	.428** (.0290)	.425** (.0931)
ln (employ.)	-.0987** (.0260)	-.0908** (.0206)	.0783* (.0430)		
employ.				-.0150** (.00475)	-.0319** (.0140)
employ.sq.				.000150** (.0000704)	.000122 (.000183)
MS * employ.				.00259 (.00185)	.0200** (.00746)
MS				-.0500 (.0372)	-.287** (.144)
% high school	-.00359 (.00364)	-.00621** (.00311)	.000040 (.0110)	-.00346 (.00279)	-.0101 (.0115)
distance to coast	-.0149** (.00598)	-.0100** (.00429)	-.00197 (.00337)	-.00675* (.00412)	-.00269 (.00386)

**Table 6. Non-Prefecture Level Cities Continued**

	(1) OLS ( <sup>'90, '91, '94-'96</sup> )	(2) LAD ( <sup>'90, '91, '94-'96</sup> )	(3) IV estimation ( <sup>'96, '95</sup> )	(4) LAD ( <sup>'90, '91, '94-'96</sup> )	(5) IV estimation ( <sup>'96, '95</sup> )
on highway	.0663** (.0336)	.0327 (.0246)	.0567 (.0510)	.0180 (.0224)	.0979* (.0530)
ln (output/output of ind. acct. units)	.396** (.0410)	.398** (.0279)	.550** (.111)	.424** (.0370)	.509** (.132)
FDI/capital	.660** (.233)	.961* (.498)	2.68** (.634)	1.05** (.526)	3.01** (.823)
ln (roads per capita)	.0475** (.0179)	.0579** (.0154)	.115 (.0769)	.0587** (.0136)	.0638 (.0847)
time effects	yes	yes	yes	yes	yes
provincial effects	yes	yes	no	yes	no
regional effects	n.a.	n.a.	yes	n.a.	yes
R <sup>2</sup> [pseudo]	.749	[.548]	n.a.	[.546]	n.a.
Sargan p-value	n.a.	n.a.	.024	n.a.	.134
N	1707 [432]	1707 [432]	472 [236]	1707 [432]	472 [236]

**Table 7. Productivity in the Rural Sector****Generalized provincial scale effects**

	<u>random effects</u>	<u>IV estimation</u>
ln (capital/labor)	.379** (.0884)	.396** (.0544)
ln (sown land/labor)	.0942 (.133)	.102 (.0937)
ln (labor)	.101 (.0662)	.156** (.0179)
% completing high school in rural sector	.0648** (.0303)	.106** (.0250)
central region	-.169 (.134)	-.0767 (.0476)
west region	-.560** (.155)	-.392** (.0458)
time dummies	yes	yes
adj R <sup>2</sup>	.849	n.a.
Sargan value	n.a.	.605
N	87	87

**Table 8. Productivity in TVE's**

	Generalized Scale Effect		Township Scale Effect	
	<u>random effects</u>	<u>IV estimation</u>	<u>random effects</u>	<u>IV estimation</u>
ln (capital/ labor)	.229** (.0672)	.355** (.0412)	.211** (.0664)	.256** (.0271)
ln (labor)	.119** (.0410)	.128** (.0158)	n.a.	n.a.
ln (avg. town- ship TVE size)	n.a.	n.a.	.158** (.058)	.300** (.0346)
% completing high school	.011 (.0077)	-.00494 (.0118)	-.0180 (.0173)	-.0120** (.00582)
central region	-.0625 (.105)	.120** (.043)	-.0102 (.111)	.057* (.033)
west region	-.515** (.119)	-.335** (.172)	-.504** (.144)	-.258** (.0623)
time dummies	yes	yes	yes	n.a.
R <sup>2</sup>	.786	n.a.	.781	n.a.
Sargan value	n.a.	.760	n.a.	.846
N	87	87	87	87

**Table 9. Gains from Increased Within-Sector Agglomeration**

**A. Gain in output per worker from doubling sizes  
of local employment agglomerations**

Non-prefecture level cities 5.6%

TVE's (township TVE employment) 23%

**B. Gain from moving to size where output per  
worker peaks for prefecture level cities**

Percent size is below peak

50   40   30   20

**Percent gain in output per  
worker**   35%   20%   9.5%   4.1%

**Table 10.**

**Part A. Social Value of Marginal Product of Labor**

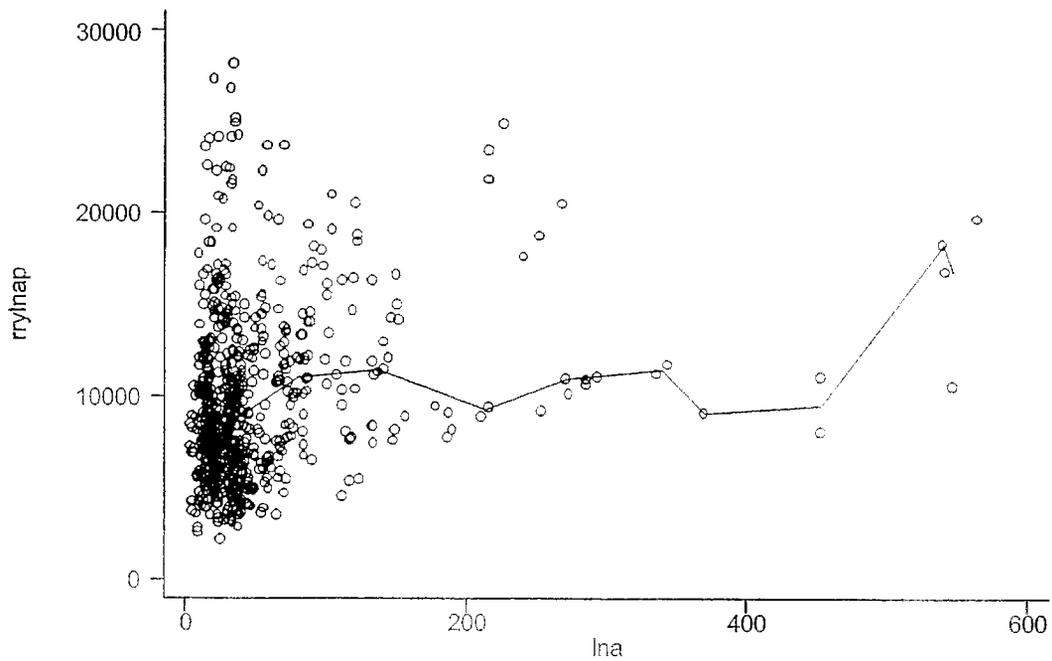
	Prefecture Level Cities	Non-Prefecture Level Cities	TVE's	Rural Sector
Nation	14300	13800	10200	2200
East	16600	15900	11600	2900
Central	11300	11100	9800	2100
West	12000	9800	6800	1300

**Part B. Marginal Product of Capital**

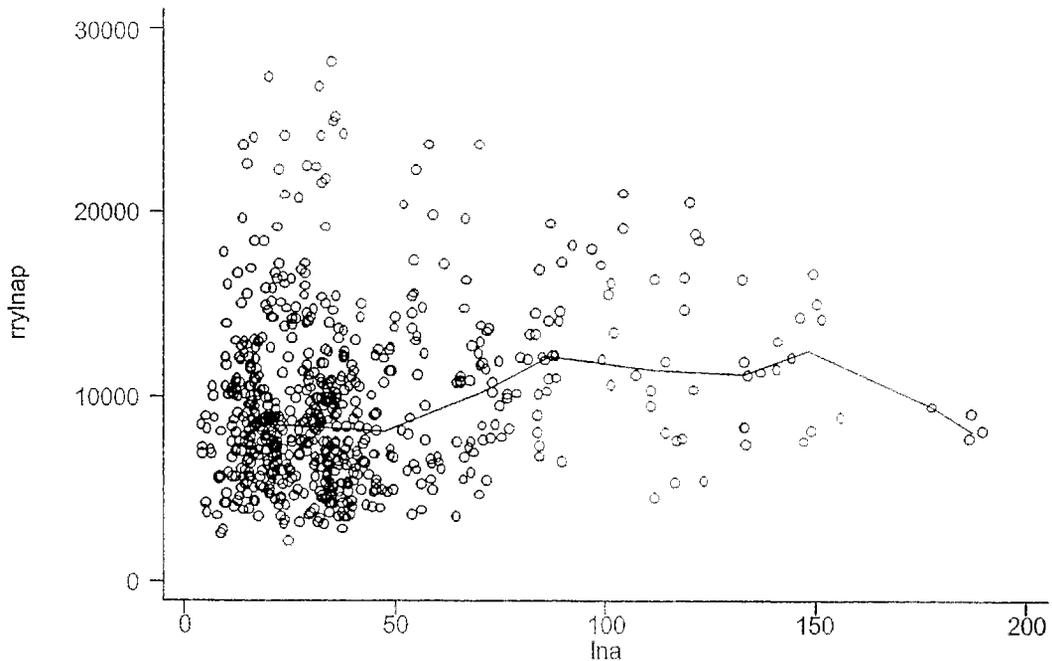
	Prefecture Level Cities	Non-Prefecture Level Cities	TVE's	Rural Sector
Nation	.33	.51	.48	.41
East	.34	.58	.39	.35
Central	.29	.49	.77	.52
West	.34	.29	.61	.42

**Table 11. Social Marginal Product of Labor With a Level Playing Field**

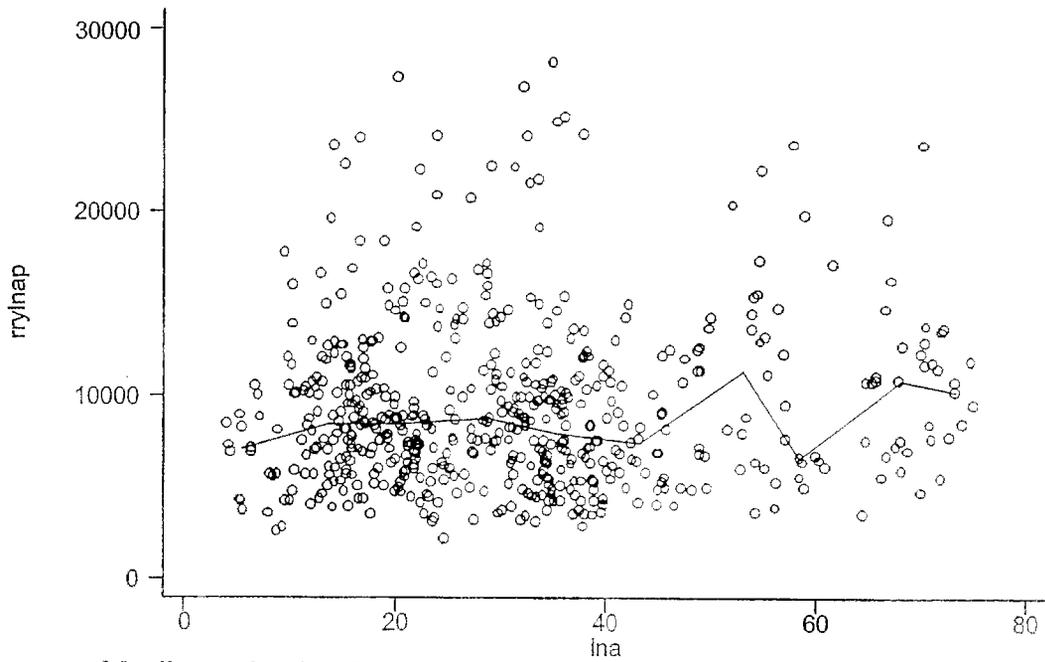
	Prefecture Level Cities	Non-Prefecture Level Cities	TVE's	Rural Sector
National	12800	12400	11400	2300
East	15300	14900	11600	2800
Central	8600	9600	12900	2500
West	11700	6900	8300	1300



**FIGURE 1a. OUTPUT PER WORKER AS A FUNCTION OF CITY SIZE (WHOLE SAMPLE)**



**FIGURE 1b. OUTPUT PER WORKER AND CITY SIZE (SIZE < 2M EMPLOYMENT)**



Median rrylnap/p: Prefecture level cities 1995-1997 (portion)

**FIGURE 1C. OUTPUT PER WORKER AND CITY SIZE (EMPLOYMENT < .75 M)**

## Data Appendix

### *The Urban Sector*

City level data used in our analysis come from several sources. Most economic and amenity variables were taken from the 1991 to 1998 annual volumes (for data years 1990 to 1997) of the *Urban Statistical Yearbook of China* (hereafter *Yearbook*)<sup>1</sup>, and *Cities China 1949-1998*. The former contains data for both prefecture level and county level cities. The latter includes a compilation of selected data in 1990 to 1997 for prefecture level cities from the *Yearbook* volumes and a complete history of new city establishment and changes in administrative area of all cities during the period. Distance proxies are measured with a ruler from *Map of China*. Highway access is read directly from the same map (occasionally with help from a more detailed map). Educational attainment is aggregated from the *China County-Level Data on Population (Census) and Agriculture, Keyed to 1:1M GIS Map, 1990*.

GDP figures are not available for 1992 and 1993 from the *Yearbook*, but are documented for prefecture level cities in *Cities China 1949-1998*. County level cities are no longer included in the *Yearbook* starting from the 1998 volume (containing 1997 data). A change in definition of labor force (number of employed person) was effected starting from data year 1998. Therefore, the panel data set that we use in our analysis consists of prefecture level cities each year between 1990 to 1997, and county level cities in 1990, 1991 and 1994 to 1996. Table A shows the number of cities with available data in each year. It should be noted that all city level data that we use are those of the more confined city proper (shi qu) rather than the municipal district (di qu). The city proper corresponds to an "urbanized area" in the USA, or the urbanized portion of a metropolitan statistical area. In our analysis of prefecture level cities, we have excluded three oil-dominant cities<sup>2</sup>, and a minimal set of outlying city-years<sup>3</sup> based on extraordinary year-to-year change in output per capita, capital-labor ratio or manufacturing to service ratio which are likely the results of misdocumentation. The identified outliers constitute less than 2.5% of the available sample of prefecture level city-years. Brief descriptions of the variables used in our analysis are in Table B and C.

Three issues should be noted here. First, capital is original book value of capital of industrial enterprises with independent accounting systems. Constructed capital data are used for prefecture level cities in 1992 and 1993. These are the only two years in which actual data for capital are not available in the *Yearbooks*. In these years we observe the ratio of gross output to capital for independent accounting units, but we don't observe gross output for these independent accounting units. However the gross output value of industrial enterprises with independent accounting system has a very close correspondence with the gross output value at township and higher levels for prefecture level cities in 1990, 1991, 1995 and 1996 (a simple correlation of 0.997 with sample size 902)<sup>4</sup>. We assume that the ratio of the two gross output values remains stable in 1992 and 1993, and is equal to the average of the ratio over the four years of 1990, 1991, 1995 and 1996 for each prefecture level city. We have data for gross industrial output value at township and higher levels in 1992 and 1993 and based on the assumed ratio, we estimate the gross industrial output value of independent accounting units in each city. Dividing this estimate by the industrial output value realized per 100 yuan of fixed assets at book value of industrial enterprises with

<sup>1</sup> A combined volume was published for 1993 and 1994.

<sup>2</sup> Daqing, Dongying and Karamay in all years.

<sup>3</sup> Chuzhou in 1992, Maanshan in 19193, Mianyang in 1993, Mundanjiang in 1996, Yingtan in 1992, Haikou in 1994, Wanxian in 1992, Jining of Shandong province in 1994 and 1997, Fangchenggang in 1995 and Songyuan in 1992-1997.

<sup>4</sup> The 1994 figures for industrial output value realized per 100 yuan of fixed assets at book value (from which the gross output values are calculated) are abnormally high as compared to that of the other years. In 1994, the mean value among prefecture level cities is 245. The mean value in the other years are: 136 (1990), 134 (1991), 113 (1995), 108 (1996).

independent accounting systems, we arrive at an estimate of the capital for each prefecture level city in 1992 and 1993. There is no significant change in our regression results for prefecture level cities with the inclusion of these estimated capital stocks.

Second, since only capital of independent accounting units is measured, we need to control for the share of independent accounting units among production units in a city. Ideally we would want to use the share of value-added of independent accounting units in the city economy. The data needed to derive this share is available only in 1997. For 1990-1996, a proxy based on the share of industrial output value (i.e. sales value) of independent accounting units among the industrial output value of a city is used. We allow the effects of the actual share and the proxy share to be different in all regressions of prefecture level cities. For county level cities, only the proxy share is available in the relevant years. As noted in footnote 4, the data required to derive the industrial output value is probably not very reliable.

Third, for comparison of real growth of output (GDP), we use the provincial level urban resident consumer price index to deflate nominal GDP's. The index is taken from the Price Indices section of the annual *China Statistical Yearbook* in the relevant period. To compare the real output across cities, we have to assume comparability based on nominal prices in a certain year (1990 in our case).

### *The Rural Sector*

Provincial level data used in our analysis of the rural sector come mainly from two sources: annual volumes of the *China Statistical Yearbook* in 1996, 1998 and 1999 (data year 1995, 1997 and 1998). Provincial educational attainment in rural areas is aggregated from relevant variables of counties not included in urban areas from the *China County-Level Data on Population (Census) and Agriculture, Keyed to 1:1M GIS Map, 1990, 1990*. Data year of 1995, 1997, and 1998 are chosen because these are the data years when value-added (instead of gross output value) of TVEs is available. Table D shows brief description of the variables of the rural sector.

Table A. Number of Cities with Available Data

<b>Year</b>	<b>Number of Prefecture Level Cities</b>	<b>Number of County Level Cities</b>	<b>Total Number of All Cities</b>
1990	225	239	464
1991	225	251	476
1992	227	/	
1993	225	/	
1994	226	393	619
1995	226	411	637
1996	225	433	658
1997	223	/	

Table B. Description of Variables (Urban Sector)

<b>Variable</b>	<b>Description</b>	<b>Source(s)</b>
population	- population at the end of the year	S1, S2
output of a city	- GDP of city in 2nd and 3rd sectors at current prices	S1, S2
manufacturing to service ratio (MS)	- ratio of GDP in 2nd sector to GDP in 3rd sector	S1, S2
employment	- number of persons employed in 2nd and 3rd sectors	S1, S2
capital	- original value of capital of industrial enterprises with independent accounting system	S1
output (value-added) of all units	- gross industrial output value (GDP in 2nd and 3rd sector of city) at current prices	S1, S2
output (value-added) of independent accounting units	- gross industrial output value (value-added of industry) of industrial enterprises with independent accounting system at current prices <sup>5</sup>	S1
FDI	- accumulated sum of foreign direct investment (foreign capital actually used) since 1990	S1, S2
roads per capita	- paved area of all roads with width greater than 3.5 meters	S1, S2
% high school	- percentage of population aged 6+ that has completed senior middle school or above	S3
distance to coast	- shortest horizontal distance from coast, measured in centimeters from map S4	S4, S5
distance to provincial capital	- horizontal distance from capital of province in which a city is located, measured in centimeters from map S4	S4, S5
on highway	- dummy for cities with access to highway (the highest category of all roads on map)	S4, S5
area (1990)	- built-up area in city proper	S2
doctors per capita (1990)	- number of medical doctors per capita	S2
books per capita (1990)	- number of books in public library per capita	S2
telephone per 100 persons	- number of telephones per 100 persons	S2
ratio of municipal agriculture to city value-added	- ratio of total GDP in 1st sector in municipal area to total non-agricultural GDP in city proper	S1, S2
ln (ex-urban population)	- log of (1+ population living in municipal district excluding city proper)	S1, S2

<sup>5</sup> Calculated from industrial output value realized per 100 yuan of fixed assets at book value (value-added realized per 100 yuan of fixed Assets at book value) and fixed assets at book value of industrial enterprises with independent accounting system

Table C. Urban Variable Means and Standard Deviations

	<u>Prefecture Level Cities in 1997</u>		<u>Non-Prefecture Level Cities 1996</u>	
	<u>mean</u>	<u>standard deviation</u>	<u>mean</u>	<u>standard deviation</u>
output per worker	23079	11077	16836	9429
capital per worker	30335	18197	15030	13794
employment (1000's)	510	661	158	104
total value added/value added of independent acct. units	3.33	2.05	n.a.	n.a.
total value sales/sales of independent acct. units 1996	1.42	.522	2.07	1.30
roads per capita	4.72	3.56	2.75	3.38
ln (roads per capita)	1.36	.675	.668	.771
% high school	21.9	8.48	12.8	6.41
distance to coast	6.09	6.06	9.36	10.5
on highway (highest road category)	.66	.47	.271	.445
manufacturing to service ratio (GDP)	1.46	.739	1.47	.809
accumulated FDI since 1990 (\$)/capital (yuan)	.0374	.0685	.0160	.0411

Table D. Description of Variables (Rural Sector)

Variable	Description	Source(s)
output of TVE	- added-value of township and village enterprises at current price	S6
labor of TVE	- total number of staff and workers employed in TVE	S6
capital of TVE	- original value of fixed assets (year-end)	S6
average township TVE size	- labor of TVE divided by number of township and town governments	S6
output of rural sector	- output of TVE plus GDP in 1st sector, deflated to 1990 constant prices <sup>6</sup>	S6+S7
labor of rural sector	- labor of TVE plus rural labor force in farming, forestry, animal husbandry and fishery	S6
capital of rural sector	- capital of TVE plus total original value of productive fixed assets of rural households <sup>7</sup>	S6
sown land	- total sown areas of farm crops	S6
% completing high school	- percentage of population aged 6+ that has completed senior middle school or above in rural areas	S3
persons in the communes (1982)	- number of persons in the communes	S6
number of communes (1982)	- number of people's communes	S6
provincial industrial output(1982)	- gross industrial output value of industrial enterprises in province	S6
share of light industry (1982)	- share of light industry in gross industrial output value of industrial enterprises	S6
arable land (1979)	- arable land area	S6
enrollment in higher institutions (1981)	- student enrollment in institutions of higher learning	S6
enrollment in secondary schools (1981)	- student enrollment in regular secondary schools or above	S6
rainfall	- annual total precipitation in capital city of province	S6
hp (1981)	- total horse power of farm machinery (year-end)	S6

<sup>6</sup> Deflation is done with the ratio of provincial total output value of agriculture, forestry, animal husbandry and fishery calculated at current prices to that calculated at 1990 constant prices.

<sup>7</sup> Calculated from original value of productive fixed assets per rural household times number of rural households.

## Sources of Data

S1. State Statistical Bureau, Urban Social and Economic Survey Team [Guojia Tongjiju Chengshi Shehui Jingji Diaocha Zongdui], *Urban Statistical Yearbook of China* [Zhongguo Chengshi Tongji Nianjian], Beijing: China Statistics Press, 1991 to 1998 (annual volumes).

S2. State Statistical Bureau, Urban Social and Economic Survey Team [Guojia Tongjiju Chengshi Shehui Jingji Diaocha Zongdui], *Cities China 1949-1998* [Xin Zhongguo Chengshi Wushi Nian], Beijing: Xinhua Press, 1998.

S3. China in Time and Space (CITAS), *The China County-Level Data on Population (Census) and Agriculture, Keyed to 1:1M GIS Map, 1990*. Data can be downloaded from:  
[ftp://ftpservers.ciesin.org/pub/data/China/CITAS/census\\_agr/](ftp://ftpservers.ciesin.org/pub/data/China/CITAS/census_agr/)

S4. *Map of China* [Zhongguo Quantu], Haerbin Map Press, 3rd ed., February 1999. #1280529-158

S5. *Transportation Map of China* [Zhongguo Jiaotong Yingyun Licheng Tuji], Beijing: People's Communication Press, 2000. ISBN 7-114-03553-5.

S6. State Statistical Bureau, *China Statistical Yearbook* [Zhongguo Tongji Nianjian], Beijing: China Statistical Publishing House, 1996, 1998 and 1999 (annual volumes) and other relevant years.

S7. State Statistical Bureau, Department of Rural Social and Economic Statistics [Guojia Tongjiju Nongcun Shehui Jingji Tongji Si], *Rural Statistical Yearbook of China* [Zhongguo Nongcun Tongji Nianjian], Beijing: China Statistics Press, 1996, 1998 and 1999 (annual volumes).