

# Export Variety, Revenue, and Productivity in

## China

Bo Chen\*

July 22, 2010

### Abstract

In this paper I argue through empirical testing that it is the underlying export variety that helps to explain the strong correlation between China's provincial export revenue and productivity. The empirical model maps export varieties into provincial GDP functions with multiple sectors by price index theory. By employing panel data that covers all 31 executive districts of mainland China from 1998 to 2005, I find that export varieties, via export revenue, significantly affect productivity: it accounts for 44.1% of interprovincial TFP differences and 36.6% of within-province TFP growth; a 10% increase in export variety leads to a 1.4% productivity growth in China.

---

\*School of International Business Administration, Shanghai University of Finance and Economics. 777 Guo Ding Rd., Shanghai, China 200433. E-mail: chen.bo@mail.shufe.edu.cn. I am indebted to Richard Harris, David Jacks, Kenneth Kasa, Nicolas Schmitt and the participants of IEFS-China Inaugural Meeting and APTS(Hong Kong) for their helpful comments. I also gratefully acknowledge the financial supports of the Ministry of Education of P.R. China (Grant No.: 09YJC790180) and the Institute of International Business of SIFT.

*Keywords:* Export Variety; Price Index; Total Factor Productivity (TFP)

*JEL Classification:* F14; O47; R11

## 1 Introduction

Since the "Open Door Policy" in 1983, China's exports have experienced a rapid (real) growth at an average rate of 20% per year. During this period China has also emerged to be one of the fastest growing economies. As a result, economists usually characterize China as a so called "export-led" economy. At the same time, the interprovincial productivity difference in mainland China<sup>1</sup> has been strikingly enlarged since 1983. Interestingly, provinces that export more (relative to its regional GDP) usually have a higher productivity. Figure 1 plots the export share (provincial export revenue to national export revenue) and TFPS<sup>2</sup> in all Chinese provinces in 2005. Evidently, it shows that there exists a significantly positive correlation between exports and productivity. This paper analyzes the effects of China's exports, by examining provincial export variety and revenue, on its interprovincial productivity variation and growth, which has been relatively lightly researched.

Although traditional trade literature has discussed intensively the posi-

---

<sup>1</sup>They include the four municipalities directly under the Central Government (i.e. Beijing, Shanghai, Tianjin, and Chongqing).

<sup>2</sup>TFPs are calculated as the Solow residuals which are the results of provincial GDP net of the contributions of land, capital, and labor. That is,  $TFP = \log(GDP) - s_T \bullet \log(T) - s_L \bullet \log(L) - (1 - s_T - s_L) \bullet \log(K)$ , where  $T$ ,  $L$ ,  $K$  refers to the (arable) land, labor, and capital stock, respectively.  $s_T$  denotes the land share for which I use the agriculture share of GDP (ratio of value added in agriculture sector to GDP) as a proxy (I also borrow the land share obtained from Feenstra and Kee (2008), and the strong positive correlation still exists).  $s_L$  denotes the labor share which is the ratio of labor income to GDP.  $K$ , the capital stock, is obtained from a perpetual inventory method defined in the text.

tive correlation between exports and productivity, they seem far from enough to fully explain such phenomenon in China's provinces for two reasons. First, productive factors (i.e. labor and capital) are supposed to be mobile across provinces whereas in traditional literature that focuses on country-wise studies they are assumed to be immobile. Second, traditional literature typically deems productivity development as the cause of export boom. However, in the "export-led" economies such as China, the causality may actually be the reverse. In this paper, I argue by empirical test that it is the underlying export variety that helps to explain the positive correlation between exports and productivity.

The hypothesis that expansion of export varieties could boost an exporting country's productivity is based on the assumption of diminishing technical rate of substitution (i.e. concavity of production possibility frontier (PPF)) or intuitively, economies of scope. Since marginal productivity of productive factors in general is diminishing, spreading productive factors in the production of more varieties will have a greater production gain than only concentrating on producing existing varieties as the average marginal productivity in latter case is lower. Intuitively, having more varieties available to produce enables producers to allocate the productive factors in a more efficient way which generates a production gain. Empirically, the link between export variety and productivity has been found by Feenstra et al (1999) for South Korea and Taiwan, by Funke and Ruhwedel (2001, 2002, 2005) for OECD, East Asian countries, and East European transition economies, and by Feenstra and Kee (hereafter FK) (2008) for 48 countries (developed and developing) in the world.

The theoretical interaction between export variety and productivity also leads to another interesting question: can we identify the direction of causality between export variety and productivity in various cases? In fact, many economists have analyzed, although not explicitly proposed, possible answers. In a standard MC model, whether a variety (firm level) can be exported depends on two key factors: the firm's own productivity and its initial irreversible investment (costs from trade barriers as well as investment cost for marketing, promotion, etc.). Holding the initial investment constant, productivity increases export variety in an obvious way: if more firms' productivity exceeds the threshold where the operational profit just covers the initial investment, more varieties will be exported. Likewise, holding its own productivity or productivity distribution constant, a decrease in trade barriers can also lead to more export varieties. Consequently, due to the effect of economies of scope, more varieties could contribute to higher aggregate productivity. For example, Melitz (2003) develops a dynamic industry model with heterogeneous firms and shows that more exposure to trade will cause market reallocation in favor of the more productive firms and thus contribute to an increase in productivity (also see Bernard and Jensen (1999), Delgado et al (2002)). On the other hand, based on the monopolistic competition model with endogenous technology, FK (2008) analyzed panel data containing 48 countries across 20 years. They found that the total increase in export variety accounts for a 3.3% average productivity growth in the exporting countries from 1980-2000. However, the causality between export varieties and productivity may be different in different countries. A plausible conjecture could be that in developed open economies with intensive

R&D activities, productivity would more likely be one of the fundamental sources of export variety expansion as well as economic growth; whereas in developing economies in the process of liberalizing trade, exports may lead economic and productivity growth.

This paper has three major contributions and improvements to existing literature. First of all, exploring the provincial data enables me to obtain a more accurate estimation for the country-specific effect of trade variety on productivity: a key (implicit) assumption in existing panel studies which use country level data is that the export elasticities of substitution between varieties within the same sector are the same across countries. For example, it is assumed that the output elasticities of substitution in the agricultural sector should be the same in both US and China. Apparently, it is too strong an assumption. While using all Chinese provincial data, I can obtain "China-specific" output elasticities of substitution without assuming identical elasticities across countries. Second, by Hausman test, I confirm the hypothesis that China is an "export-led" economy: the export variety *exogenously* accounts for interprovincial productivity difference and growth rather than the endogenous variety assumption that is typically made in existing literature (see, amongst many others, Arkolakis, Demidova, Klenow, Rodríguez-Clare (2008)). This finding is consistent with the results of Kwan and Kwok (1995), which found that exports are exogenous in explaining China's economic growth (and thus productivity). Third, a lot of empirical literature uses the U.S. data, but employing Chinese export data makes the results more reliable: an important assumption in the empirical price-factor GDP function is that the prices and production factors are all given,

i.e., exogenous. For example, FK (2008) employed the U.S. import data to approximate the exports of 48 countries including the major developed ones. However, it is very likely that the prices are actually endogenous in some industries where the developed countries may have monopolistic power (for instance, electronic products from Japan and machines from Germany). Furthermore, the U.S. itself is the largest open economy in the world suggesting that it has monopsonistic power over many of its imports. As a result, if monopoly dominates, estimates of the elasticities of productivity on export varieties tend to be overstated; while if monopsony dominates, the estimates tend to be understated. On the contrary, though China is a large economy, most of its exports are characterized as low value-added and easy-to-substitute.<sup>3</sup> In other words, Chinese exports mainly face competitive markets with very elastic demands. As a result, most Chinese exporters, to a large extent, are price takers.

The rest of the paper is organized in three parts. Section 2 summarizes the mechanism of how the underlying export variety, via export revenue, affects productivity, mainly by borrowing from the methods discussed in FK (2004).<sup>4</sup> Then I describe the dataset and estimate a system of equations relating sectoral shares and adjusted total factor productivity (TFP) to export variety. In section 3, I present the results of estimation as well

---

<sup>3</sup>For example, by 2005, the top three markets in which Chinese exports occupied the largest world shares were: Textiles (35%), Footwear (60%), and Toys (40%), which are easily found in Wal-Mart but not in luxury franchise stores.

<sup>4</sup>Although FK (2008) is the revised version of FK (2004), the 2008 version is based on a "Melitz-type" model where the export variety is assumed to be endogenously affected by productivity. However, China is an "export-led" economy which suggests export may exogenously explain productivity. Therefore, I borrow FK's method in the 2004 version where variety endogeneity is not assumed (though it is tested).

as hypothesis tests to show the validity of the estimation method and the robustness of these results that are based on various constraints imposed in the model. Finally, section 4 concludes that export variety accounts for 44.1% of the variation of interprovincial productivity in level and 36.6% of within-province productivity growth. Overall, at the sample mean, a 10% increase in export varieties of all exporting sectors leads to a 1.4% increase in China's productivity.

## 2 The Empirical Model

### 2.1 Price Index Theory and Product Variety

Feenstra (1994) derives an exact price index from a CES (constant elasticity of substitution) aggregate good allowing both variety and taste to change in existing varieties. This index can also apply for several goods or even industries as long as they are still CES aggregates.

Suppose there exist  $i = 1, \dots, R$  regions. Each region  $i$  can produce  $I_t^i$  set of product varieties at time  $t$ . The quantity of each type of variety produced in region  $i$  in period  $t$  is denoted by  $q_t^i$ . The aggregate output of region  $i$ ,  $Q_t^i$ , is characterized by a CES function of the output of each specific variety produced in that region:

$$Q_t^i = f(q_t^i, I_t^i) = \left( \sum_{j=I_t^i} a_j (q_{jt}^i)^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \quad a_j > 0, \quad i = 1, \dots, R. \quad (1)$$

where  $a_j$  is the unknown quality parameter for variety  $q_{jt}^i$ , and  $\sigma < 0$  is the

elasticity of substitution between variety and thus the PPF is concave.

As demonstrated in Feenstra (1994) the ratio of aggregate price levels associated with the CES production function can be evaluated by the product of the Sato(1976)-Vartia(1976) price index of goods that are common,  $I_t \equiv (I_t^a \cap I_t^b) \neq \emptyset$ , and by terms reflecting the revenue share of “unique” goods. The log form is given by:

$$\ln \frac{P_t^a}{P_t^b} = \sum_{i \in I_t} w_j(I_t) \ln \left( \frac{P_{jt}^a}{P_{jt}^b} \right) + \left( \frac{1}{\sigma - 1} \right) \ln \left( \frac{\lambda_t^a(I_t)}{\lambda_t^b(I_t)} \right), \quad a, b = 1, \dots, m. \quad (2)$$

where  $P_t^i$  is the aggregate price level ( $i = a, b$ ), and the weights  $w_j(I_t)$  are constructed from the revenue shares in the two regions:

$$w_j(I_t) = \left( \frac{s_{jt}^a - s_{jt}^b}{\ln s_{jt}^a(I_t) - \ln s_{jt}^b(I_t)} \right) / \sum_{i \in I_t} \left( \frac{s_{jt}^a - s_{jt}^b}{\ln s_{jt}^a(I_t) - \ln s_{jt}^b(I_t)} \right) \quad (3)$$

$$s_{jt}^i(I_t) \equiv p_{jt}^i q_{jt}^i / \sum_{j \in I_t} p_{jt}^i q_{jt}^i, \quad \text{for } i = a, b, \quad (4)$$

$$\lambda_t^i(I_t) = \frac{\sum_{j \in I_t} p_{jt}^i q_{jt}^i}{\sum_{j \in I_t^i} p_{jt}^i q_{jt}^i} = 1 - \frac{\sum_{j \in I_t^i, j \notin I_t} p_{jt}^i q_{jt}^i}{\sum_{j \in I_t^i} p_{jt}^i q_{jt}^i} \quad \text{for } i = a, b, \quad (5)$$

where  $s_{jt}$  measures the revenue share of variety  $j$  relative to all the varieties that are common in county  $a$  and  $b$  at period  $t$ , and  $\lambda_t^i(I_t)$  is the revenue share of common varieties ( $j \in I_t$ ) to total varieties ( $j \in I_t^i$ ).

The first term on the right hand side of (2) is the traditional price index, which only captures the weighted average of the price ratios for varieties in the common set  $I_t$ , in other words, it ignores the effect of variety and quality change. The second term is Feenstra's correction term, which reflects changes in product variety given that the quality of the same variety is invariant. In fact,  $\lambda$  is a variety variable since it can be deemed as a weighted variety ratio where the weight is determined by the revenue of each variety. Therefore, a new variety with sales revenue of one million is 100 times as important as a new variety with sales revenue of only ten thousand. Given (2) and (5), we can easily tell that a new product variety produced in region  $a$  will decrease  $\lambda^a$  and thus increase the price ratio (since the coefficient of the lambda ratio term,  $1/(\sigma - 1)$ , is negative). Intuitively, the new product variety induces a more efficient allocation among productive factors (gains from production) such that the output price index  $P^a$ , measured as (maximized) the unit revenue of the aggregate goods, increases.

## 2.2 Revenue and Variety

The ideal definition for a variety is a firm-brand such as Honda Civic and Ford Focus. In micro-level studies researchers often use the market-based survey data to study the variety effect on welfare or productivity. (see, for example, Blonigen and Soderbery (2009)). However, survey data has a serious limitation in data coverage: it can only cover one or a few industries for a few years. Therefore survey data can not satisfy macro-level studies which need data on a much broader scope of economy (i.e. export industries). Researchers usually rely on trade data (such as SITC or HTS data system)

to carry out their macro-level studies and typically adopt the Armington definition that a variety is a country-good pair (see, amongst many others, Broda and Weinstein (2006)). For example, the beer produced in France and in Britain is treated as two varieties of the product "beer". Since this paper uses provincial level data, a natural modification of the Armington type variety is that a variety is a province-good pair. That is, Shanghai-beer and Beijing-beer are two varieties of beer.

In fact, defining a variety as a province-good pair can help to greatly simplify the eq(2). Assume  $b$  as (mainland) China and  $a$  is a province. Then it is immediate that  $I_t = I_t^a \cap I_t^b = I_t^a$ . Furthermore, the prices of common varieties produced in both province ( $a$ ) and China ( $b$ ) should be the same since they are the prices of the identical varieties. For example, a common variety, Shanghai-beer that is produced in both Shanghai and China, refers to the same variety since no other provinces in China can produce Shanghai-beer. Hence, the first term on the right hand side of eq(2) disappears (it is zero!) and  $\lambda_t^a(I_t) = 1$ . Then eq(2) can be simplified as:

$$\ln \frac{P_t^a}{P_t^b} = \left( \frac{1}{1 - \sigma} \right) \ln \left( \lambda_t^b(I_t) \right), \quad a = 1, \dots, R. \quad (6)$$

where  $R=31$  is the total number of provincial executive regions in mainland China.

Eq(6) reveals that the exact price ratio in an industry/sector of a province, with respect to the counterpart of the whole nation, is a simple function the ratio of their corresponding sales revenues, as the variety effect on exact price can be fully captured by its sales revenue. Therefore a new variety in-

troduced in province  $a$  will lead an increase in  $\lambda_t^b(I_t)$  and thus an increase in the provincial exact price level (to its national counterpart). Furthermore, eq(6) has a very good empirical implication: to study the variety effect on industry/sector exact price, we do not need highly disaggregate variety data (i.e. HTS-8 or HTS-10 level data) in that industry/sector; what we need is just the revenue data of that industry/sector (i.e. HTS-2 or HTS-4). This implication is very important: China only has HTS-4 level export data (provincial level) prior to 2002, but this approach allows me to empirically study the export variety effect, even for years prior to 2002, without highly disaggregate data.

### 2.3 GDP Function with Export Variety

As is common in empirical GDP work, I assume the GDP function of a region follows a translog functional form where there are  $N+1$  output sectors ( $N$  export sectors and one non-traded sector) and  $M$  types of production factors:

$$\ln G_t^i(P_t^i, V_t^i) = \alpha_0^i + \beta_0^t + \sum_{n=1}^{N+1} \alpha_n \ln P_{nt}^i + \sum_{k=1}^M \beta_k \ln v_{kt}^i + \frac{1}{2} \sum_{m=1}^{N+1} \sum_{n=1}^{N+1} \gamma_{mn} \ln P_{mt}^i \ln P_{nt}^i$$

$$+ \frac{1}{2} \sum_{k=1}^M \sum_{l=1}^M \delta_{kl} \ln v_{kt}^i \ln v_{lt}^i + \sum_{n=1}^{N+1} \sum_{k=1}^M \phi_{nk} \ln P_{nt}^i \ln v_{kt}^i, \quad i = 1, \dots, R. \quad (7)$$

Notice that in a panel data regression setting,  $\alpha_0^i$  and  $\beta_0^t$  refer to region and time fixed effects, respectively. To satisfy the properties of homogene-

ity in prices and endowments as well as symmetry, I impose the following restrictions:

$$\begin{aligned} \gamma_{mn} = \gamma_{nm}, \sum_{n=1}^{N+1} \alpha_n = 1, \sum_{n=1}^{N+1} \gamma_{mn} = \sum_{n=1}^{N+1} \phi_{nk} = 0, \\ \delta_{kl} = \delta_{lk}, \sum_{k=1}^M \beta_k = 1, \sum_{k=1}^M \delta_{kl} = \sum_{k=1}^M \phi_{nk} = 0, \end{aligned} \quad (8)$$

From eq(7), the share of sector  $n$  is given by the derivative of  $\ln G_t^i(P_t^i, V_t^i)$  with respect to  $\ln P_{nt}^i$ :

$$s_{nt}^i = \alpha_n + \sum_{m=1}^{N+1} \gamma_{mn} \ln P_{mt}^i + \sum_{k=1}^M \phi_{nk} \ln v_{kt}^i, \quad n = 1, \dots, N + 1. \quad (9)$$

Taking the difference of any province  $i$  with that of the comparison region  $*$  (i.e. the whole country), we can map export variety variable  $\lambda$  into the empirical GDP function as well as the share equations with minimal computation:

$$\begin{aligned} s_{nt}^i - s_{nt}^* &= \sum_{k=1}^M \phi_{nk} (\ln v_{kt}^i - \ln v_{kt}^*) + \sum_{m=1}^N \frac{\gamma_{mn}}{(1 - \sigma_m)} \ln \lambda_{mt}^{i*} \\ &+ \gamma_{N+1,n} (\ln P_{N+1,t}^i - \ln P_{N+1,t}^*). \end{aligned} \quad (10)$$

and

$$\begin{aligned}
& \ln G_t^i(P_t^i, V_t^i) - \ln G_t^i(P_t^*, V_t^*) - \sum_{k=1}^M \frac{1}{2}(s_{kt}^i + s_{kt}^*)(\ln v_{kt}^i - \ln v_{kt}^*) \\
& - \frac{1}{2}(s_{N+1t}^i + s_{N+1t}^*)(\ln P_{N+1t}^i - \ln P_{N+1t}^*) \\
= & \alpha_0^i + \sum_{n=1}^N \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \sigma_n)} \tag{11}
\end{aligned}$$

where the price difference of tradable goods (sector 1 to N) has been substituted by the variety variable  $\ln \lambda_{nt}^*$  according to eq(6). Without the loss of generality, I normalize  $\alpha_0^* = 0$ . The left hand side of (11) can be interpreted as the productivity difference between province  $i$  and the country  $*$  (the average productivity of all provinces): it is the difference of GDP net of the differences in factor endowments and prices in nontraded goods. The remaining difference is the productivity differences in export sectors due to export variety.

## 2.4 Data and Estimating Equations

The dataset covers all 31 provincial level executive districts in mainland China. However, due to data availability, it is an unbalanced panel dataset with 17 of the provinces starting from 1998 while the rest start no later than 2002 and all series end in 2005. This dataset contains 193 observations for each regression and 1544 for the whole system. All of the data was obtained from corresponding national or provincial statistical yearbooks.

Since variety in this paper is defined as a province-good pair,  $\lambda_{nt}^*$  is then the ratio of provincial export revenue to its national counterpart in export

sector  $n$  at time  $t$ .

I assume there are three factors of production: Labor, Capital, and (arable) Land. Labor and Land (as well as nominal GDP) are directly reported by the China Statistical Yearbooks (1999-2006). Capital is constructed by the perpetual inventory method using real investment of the whole nation as well as the 31 provinces across the 8 years. Real investment is obtained by deflating the gross domestic capital formation of the whole nation as well as 31 provinces with their respective GDP deflators. All the data on gross capital formation and GDP deflators are also obtained from the China Statistical Yearbooks. In addition, I construct the base year capital stock using an infinite sum of series of investment prior to the first year (1998) by assuming that the average growth rate of investment of the first 7 years is a good proxy for the investment prior to the first year. Note that all the three productive factors are reported as year-end input rather than initial endowment in the statistical yearbooks. Therefore, there is no interprovincial factor mobility problem.

I aggregate up all the export goods into 7 sectors: agriculture, wood & paper, textile & clothing, chemicals & plastics, mining & metals, machinery<sup>5</sup>, and food & beverage<sup>6</sup>. The value-added of these sectors is obtained from the corresponding provincial statistical yearbooks (1999–2006), which are used to compare with the corresponding regional GDP to construct the sector shares. The nontraded goods price is obtained by taking an equally-weighted average of the Education, Health Care & Child Care, and Rental

---

<sup>5</sup>Including transportation equipment and electronic products.

<sup>6</sup>Including tobacco.

for Housing price indices. The regional labor share in GDP,  $s_{Lt}^i$ , is constructed by comparing the labor income to the corresponding regional GDP.

Here I use each of the 31 provinces as a specific "region", i.e.  $i = 1, \dots, 31$ , and China as a whole as the comparison region \*. By applying the homogeneity constraints on productive factors (i.e.  $\sum_{n=1}^3 s_{nt}^i = 1$ ) and prices (i.e.  $\sum_{m=1}^8 \gamma_{mn} = 0$ ), we can rewrite (10) and (11) as follows:

$$\begin{aligned} s_{nt}^i &= \delta_{nt} + \phi_{Ln} (\ln \ell_t^i - \ln \ell_t^*) + \phi_{Kn} (\ln k_t^i - \ln k_t^*) \\ &\quad + \sum_{m=1}^7 \gamma_{mn} \left( \frac{\ln \lambda_{mt}^{i*}}{(1 - \sigma_m)} - (\ln P_{8,t}^i - \ln P_{8,t}^*) \right) + \varepsilon_{nt}^i, \\ n &= 1, \dots, 7. \end{aligned} \tag{12a}$$

$$\begin{aligned} Adj.TFP_t^i &\equiv \ln G_t^i(P_t^i, V_t^i) \\ &\quad - \frac{1}{2} (s_{Lt}^i + s_{Lt}^*) (\ln \ell_t^i - \ln \ell_t^*) - \frac{1}{2} (1 - (s_{Lt}^i + s_{Lt}^*)) (\ln k_t^i - \ln k_t^*) \\ &\quad - \frac{1}{2} (\ln T_t^i - \ln T_t^*) - \sum_{n=1}^7 \frac{1}{2} (s_{nt}^i + s_{nt}^*) (\ln P_{8t}^i - \ln P_{8t}^*) \\ &= \alpha_t^* + \alpha_0^i + \beta_k (\ln k_t^i - \ln k_t^*) + \beta_8 (\ln P_{8t}^i - \ln P_{8t}^*) \\ &\quad + \sum_{n=1}^7 \frac{1}{2} (s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \sigma_n)} + \varepsilon_t^i \end{aligned} \tag{12b}$$

where  $\ln(\ell_t) = \ln(L_t/T_t)$  and  $\ln(k_t) = \ln(K_t/T_t)$ . If the homogeneity constraint in prices is not violated,  $\beta_8$  in (12b) should be equal to unity, and  $\beta_k$  in (12b) represents the negative value of the share of Land in GDP.<sup>7</sup> It

<sup>7</sup>In fact,  $\beta_k$  should be  $\beta_{kt}^i = 1 - \frac{1}{2}(s_{Lt}^i - s_{Lt}^*) - \frac{1}{2}(s_{Kt}^i - s_{Kt}^*)$ , which is a random parameter. For the sake of simplicity, we assume the relative labor share and the relative capital share do not change for different regions and across periods so that we treat  $\beta_k$  as

should be noted that the sector shares of the comparison country (China), appearing as  $s_{nt}^*$  in (10), are measured as year fixed effects,  $\delta_{nt}$  in (12a); whereas China's national (log) GDP,  $\ln G(P_t^*, V_t^*)$ , in equation (11) is treated as a year-fixed effect,  $\alpha_t^*$ , in the regression function (12b). In summary, I will regress the panel data for seven sectoral share equations (with year-fixed effects only) and a TFP equation (with both region and time fixed effects).

With equations (12a) and (12b), we can estimate the parameters of interest such as the elasticity of substitution ( $\sigma_n$ ) between different varieties in export sector  $n$ , the relative price effects on the sector shares ( $\gamma_n$ ), as well as the effects of relative endowments effects ( $\phi_{Ln}$ ,  $\phi_{Kn}$ , and  $\beta_k$ ) in the regressions of sector shares and productivity. Table 0 summarizes the main variables and key estimates in this regression system.

Furthermore, with the estimated parameters, the regional estimated productivity is given by (13):

$$\begin{aligned} Est.TFP_t^i &= Adj.TFP_t^i - \hat{\alpha}_t^* - \hat{\beta}_k(\ln k_t^i - \ln k_t^*) + \hat{\beta}_g(\ln P_{8t}^i - \ln P_{8t}^*) \\ &= \hat{\alpha}_0^i + \sum_{n=1}^N \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)} + \hat{\varepsilon}_t^i \end{aligned} \quad (13)$$

Due to the cross equation restrictions on  $1/(1 - \sigma_n)$  and  $\gamma_{mn}$ , and the multiplicative nature of these parameters, I use nonlinear system estimation for the seven share equations (12a) and one TFP equation (12b). The optimal estimates for these parameters are derived by minimizing the variance-  


---

a time- and region-invariant parameter.

covariance matrix of the residuals in the full system of the regression equations.

### 3 Estimation Results and Hypothesis Tests

#### 3.1 Estimation Results

##### 3.1.1 Nonlinear Least Square Regressions

Table 1 presents the results of the nonlinear system of share equations (12a) with the TFP equation (12b), estimated by iterative Nonlinear Least Square Regressions (NLS). All the homogeneity constraints on prices and endowments as well as the symmetric constraints on cross-price effects are imposed on the share equations, and the last column shows the estimated coefficients of the provincial productivity equation.

In table 1, the upper part of column (1) to (7), reports  $\gamma_{mn}$ , which are the partial price effects on the share of sectors in the columns due to export variety changes in the rows. Particularly, the diagonal shows the own-price effects ( $\gamma_{nn}$ ) which are all positive and significant. That is to say, the underlying supply curves of these sectors are positively sloped. The middle part shows the Rybczynski effects of endowments on the industry shares (i.e.  $\phi_{Ln}$  for labor and  $\phi_{Kn}$  for capital). For example, an increase in capital relative to land (significantly) hurts the agriculture and food & beverage industries but benefits the chemical & plastics, mining & metals, and machinery industries. On the other hand, an increase in the labor endowments relative to land (significantly) benefits agriculture, wood & paper but hurts chemical

& plastics, mining & metals, and machinery.

Given the structure of (12a) and (12b), the elasticities of substitution (i.e. the sigmas:  $\sigma_n$ ) are nonlinear coefficients of the variety variables ( $\lambda_n$ ). To be convenient, the upper part of column (8) presents the NLS point estimates of  $1/(1 - \sigma_n)$  for each industry in the row and the sigmas are then derived from these estimates. According to the assumption, the sigmas among outputs should be strictly negative. In other words, the expected estimates of  $1/(1 - \sigma_n)$  should be between zero and one. Furthermore, the higher degree of differentiation between varieties in a sector, (in absolute terms) the smaller (larger) is the sigma ( $1/(1 - \sigma_n)$ ) of that sector. As shown in the upper part of column (8), all the estimates are significant and fall in the range of zero to one. The ranking of export sectors according to their implied elasticities of substitution are: wood & paper (-0.235), machinery (-0.748), textile & garments (-1.322), chemical & plastics (-1.390), agriculture (-1.621), mining & metals (-4.882), and food & beverage (-8.025).

The middle part of column (8) presents the effects of the capital-land ratio and nontraded goods prices on adjusted TFP. As predicted in the model, the coefficient on the capital-land ratio should be the negative value of the land share in GDP. That is, our estimate,  $\beta_k = 0.083$ , implies that the estimated share of land in China's GDP is about 8.3%. However, the estimated coefficient on nontraded goods price is significantly less than unity, which suggests a violation of the homogeneity assumption on prices.<sup>8</sup>

However, our NLS method might have two potential problems. First of

---

<sup>8</sup>The violation of the homogeneity constraint in prices in the TFP equation does not affect the rest of the estimations since I did not impose it in TFP equation.

all, for a system of simultaneous regression equations there may be correlation between the error terms of these equations. In particular, the error terms of the seven sectoral share equations (12a) might be correlated, as these export sectors may compete for the same resources. For example, the Rybczynski theorem suggests that an increase in one productive factor may cause an expansion in some sectors that use this factor intensively but a contraction in some other sectors. Ignoring the correlation problem will cause the estimates to be less efficient as the variance of estimation is not at a minimum. If there is cross-equation correlation among the error terms, a seemingly unrelated regression (SUR) can yield more efficient estimates (Zellner, 1962). Secondly, endogeneity might be a problem since in our adjusted TFP equation (12b) productivity can also affect export variety. That is, a simultaneity problem may arise since productivity growth may help some products gain comparative advantage over their international counterparts so that they become new exported varieties (for example, Melitz (2003), Ghironi and Melitz (2005)). Ignoring this potential endogeneity problem, if any, will cause the estimates to be biased. To overcome the endogeneity problem, I use a nonlinear two stage least square (N2SLS) method to derive unbiased estimates based on sufficient valid instrumental variables (IVs) which affect productivity only via export variety.

To check the validity of our estimates of NLS, I compare them to those of SUR and N2SLS respectively in the following two subsections.<sup>9</sup>

---

<sup>9</sup>Though the SUR and N2SLS can be combined into a N3SLS (so called "full information estimation" in theory), we choose to check for these two problems separately for the following reasons: First of all, if the estimates of NLS were not robust, I wish to find out the main causes: SUR, endogeneity, or both. Secondly, when only one or none of the problem seriously exists in our regression system, the empirical estimates of N3SLS may

### 3.1.2 Iterative Seemingly-Unrelated Regressions

Table 2 presents the results estimated by SUR. The results reveal that both the partial price effects on the share of industries and the Rybczynski effects of endowments are very similar to those estimated by NLS. The similarity also appears in column (8) except that the coefficients on Textile & Garments and Wood & Paper are above unity, but not significantly.

Even though SUR is a theoretically more efficient estimation method than OLS, a number of empirical studies reveal that this superiority is ambiguous in practice (for example, Srivastava and Giles, 1987; Maeshiro, 1980; Heien, et al., 1998). Indeed, the closer the error covariance comes to being spherical, the more likely it is that OLS estimates will be superior. Therefore, the overall similarity between the estimates of NLS and SUR is not surprising. It may result from the fact that residuals of the sectoral share equations do not have strong correlation. Thus, after considering the potential SUR problem, we still can confidently retain estimates of NLS.<sup>10</sup>

---

be much less efficient than those of NLS. (I discuss this empirical efficiency issue in the following subsection.) Thus, even if there were significant differences between N3SLS and NLS, we could not tell whether it is because of the SUR/endogeneity problem or just the inefficient estimation of N3SLS.

<sup>10</sup>Strictly speaking, a reliable test for the correlation of error terms should be done before we can retain estimate results of NLS. To test the correlation of error terms, many studies directly test for the zeros in off-diagonal error covariance matrix. (for example, Breusch and Pagan, 1980; Kariya, 1981; Shiba and Tsurumi, 1988). However, most of the tests are only justified under asymptotic arguments. Thus, standard multivariate likelihood-based asymptotic tests are unreliable in finite samples because test sizes deviate from the nominal significance levels for related simulation evidence (Dufour and Khalaf, 1998). Since the overall similarity between estimates of NLS and SUR, doing any (unreliable) test for SUR problem seems unnecessary in our study.

### 3.1.3 Two Stage Nonlinear Least Square Regressions

As for the endogeneity problem, the key task is to find enough valid IVs (at least as many as the endogenous variables to be replaced). Many economists (e.g. Eaton and Kortum, 2002; Melitz, 2003) have suggested various IVs such as tariffs, transport costs, and distance as these trade costs can only affect productivity through export variety. However, since all Chinese provinces, except the so called "Special Economic Zones", face the same tariffs against their exports, only IVs concerning transportation costs and distance will be useful here. Therefore, in order to find enough IVs, I also consider market demand and supply factors.

Besides all the included exogenous variables (except for the price index of nontraded goods), I find seven excluded IVs along three dimensions in addition to the land difference. They are weighted distance, road density, seaboard dummy, and border dummy for geography; lagged population density, and lagged CPI for market demand; and the policy dummy for market supply. All the relevant data except for the dummies and distance are available from provincial and national statistical yearbooks.

With respect to the distance IVs, since the export destinations are all over the world, I have to use a weighted distance to approximate the real trade distance facing each province. I measure the distances between the capital of each province to the capitals of the export destinations and weight them by the export shares of these destinations. Detailed distance data is obtained from CEPII<sup>11</sup> and export shares are calculated based on export

---

<sup>11</sup>I thank very much the anonymous referee who provided me with the CEPII data source.

data in Statistical Yearbook (1999–2006).

Road density is calculated by dividing total mileage of railway, national highway, and waterway by the provincial area. The lagged population density is measured by dividing the lagged effective provincial population by the provincial area. And effective population is the sum of the rural residents and the triple of urban residents. The idea of “effective population” comes from two reasons: firstly, the annual income of a representative urban resident is three times as high as that of a representative rural resident; secondly, an urban resident is more important for market demand than a rural counterpart because the former has to depend on trade for exchanging goods while the latter is more likely to produce most goods by himself.

The seaboard dummy indicates whether a province has seaboard and the border dummy indicates whether a province shares international borders with the other countries. Finally, the policy dummy is assigned to the three provinces having the Special Economic Zones<sup>12</sup> and four municipalities directly under central government. The remaining IVs are directly given.

The geographical IVs are the "classic" exogenous variables in explaining productivity since they may affect productivity via export variety but not the other way round. Likewise, the policy dummy is also a straightforward IV since these policies, which are based on political and geographical concerns, affect tariffs which may affect productivity via export variety. Since both CPI and effective population density are time series, they may cointe-

---

<sup>12</sup>The special economic zones are designed for the "experiments" of China's economic reform. In these zones, tariff and taxes are generally lower and legal procedures on business are simplified. By the year 2006, there were 6 special economic zones: Shenzhen, Zhuhai, Shantou, Xiamen, Pudong, and Hainan. The first three belong to Guangdong province, the rest belong to Fujian province, Shanghai, and Hainan province respectively.

grate with productivity which is also a time series. A treatment is to use the lags of CPI and effective population density. Intuitively, a current production plan may be based on the information in the last period such as last period's prices and number of target consumers. For example, a higher CPI or increase in population density may be treated as a positive signal to producers, and they may find it profitable to create new varieties to compete for the prospering market.

Theoretically, good (excluded) IVs should not only be exogenous with respect to the regressand(s), but also correlate non-trivially with the endogenous regressors (i.e. the excluded IVs should not be too "weak"). Following the method introduced in Staiger and Stock (1997), empirical studies often adopt the first-stage F-statistic to test for weak instruments to detect the "weak instruments" problem. The rule of thumb to reject the weak IVs hypothesis is that, in the first stage (when regressing the endogenous variables on all IVs), the F-statistics based on the Residual Sum of Squares of the OLS, with and without the excluded IVs, should be larger than 10.

Table 3 presents the results estimated by N2SLS. The results reveal that both the partial price effects on the share of industries and the Rybczynski effects of endowments are very similar to those estimated by NLS. The coefficient estimates on export varieties in (12b) change a little bit compared to NLS. The coefficients on wood & paper and Chemical & Plastics are above unity, but not significantly so. Furthermore, the fact that the first-stage F-statistics in column (8) are all much larger than 10 shows that the overall IVs are non-trivially correlated with the export variety variables ( $\lambda$ ).

However, we need to be cautious that the estimates of N2SLS may not

be more reliable than those of NLS. After all, if the endogeneity problem is not so serious, NLS is naturally superior to N2SLS since N2SLS actually replaces the most relevant explanatory variables with less explanatory ones (i.e. the IVs). The Hausman Test (Hausman,1978) is widely used to test for potential endogeneity. The null hypothesis is that the regressors in the structural equations are exogenous. Under the null hypothesis, both of the estimates (NLS and N2SLS) should be consistent, and the difference between these two vectors of parameter estimates should follow a Chi-square distribution with  $K$  degrees of freedom ( $K$  is number of unknown parameters). If the null hypothesis is not rejected, one can retain the estimates of NLS; otherwise, the estimates from N2SLS will have to be taken. Column (6) of table 4 reports that the p-value of the statistic is almost unity which means we cannot reject the null hypothesis of no endogeneity. This result is consistent with Kwan and Kwok (1995) that exports are exogenous in explaining China's economic growth (and thus productivity).

### 3.2 Specification Tests

As the NLS estimation results are still valid after checking the SUR and endogeneity problems, I proceed to test the overall validity of the restrictions imposed on the system based on the NLS estimation to check the overall robustness. Recall in section 2, I have imposed the following restrictions. For each of the share equations, homogeneity constraints on prices and endowments are imposed. The homogeneity constraint on endowments is imposed in the GDP function but not the homogeneity constraint on prices due to the possible measurement errors in nontraded good prices. The twenty-one sym-

metry constraints on the cross-price effects are also imposed on the whole system of equations.

I first test for all the homogeneity constraints one at a time. In each case, I constrain the variance-covariance matrix to be that of the unrestricted model. Conditional on all the accepted homogeneity constraints I further test for the symmetry constraints. This is done by comparing the value of the criterion function of the restricted model to a model with no symmetry constraints.

Table 4 presents the test statistics and the associated p-values for all the hypothesis tests. None of the homogeneity constraints for endowments are rejected, nor are the homogeneity constraints in prices on all sectoral share equations, except for the price constraint on agriculture industry, which can be rejected at the confidence level of 10%. The only violation of homogeneity constraint in prices is for the TFP equation, which we did not impose in the previous estimation.<sup>13</sup> Given that all the homogeneity constraints on endowments and prices are not rejected, Table 5 also reports that the null hypothesis on the 21 symmetry constraints on the seven sectors cannot be rejected either.

In summary, the results of hypothesis testing support my previous specification in terms of the imposed homogeneity constraints on endowment and prices as well as the symmetry constraints on cross-equation prices.

---

<sup>13</sup>Our estimation results should remain robust as long as the measurement error in non-traded goods is not systematically related to the province productivity or export variety.

### 3.3 Productivity Decomposition

To highlight the relationship between export variety and productivity, I carry out a panel of regressions of the estimated productivity on export variety ( $\sum_{n=1}^7 \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)}$ ) using the estimated parameters obtained in last section. Figure 2 plots the scatter graph. Both variables are averaged over time so actually we plot a “between” regression. From this graph, it is evident that the provincial export variety explains the productivity difference in a significant way.

To highlight the relationship between variety and productivity export sectors in China’s provinces, I perform a post-regression decomposition of estimated productivity based on the results in Table 1. Using (13), I compute the variance of estimated provincial TFP as:

$$\begin{aligned} var(Est.TFP_t^i) &= var(\hat{\alpha}_0^i) + var\left(\sum_{n=1}^7 \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)}\right) \\ &+ 2cov\left[\hat{\alpha}_0^i, \sum_{n=1}^7 \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)}\right] + var(\hat{\varepsilon}_t^i). \end{aligned} \quad (14)$$

The first term on the right hand side is the variance of provincial fixed effects, the second is the variance of export varieties, the third is the covariance between these two, and the fourth is the variance of error term. By removing variance of the fixed effects and the regression error, the “variety-induced” provincial TFP is defined as:

$$\text{Variety-induced } TFP_t^i \equiv \sum_{n=1}^N \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)}, \quad (15)$$

In addition, the first order difference of (15) within a province between years reveals the decomposition of the provincial productivity growth into two terms, which is the growth of variety induced provincial TFP and the change in regression errors:

$$\begin{aligned} \text{Growth of } TFP_t^i \equiv & \sum_{n=1}^7 \left( \frac{1}{2}(s_{nt}^i + s_{nt}^*) \frac{\ln \lambda_{nt}^*}{(1 - \hat{\sigma}_n)} - \frac{1}{2}(s_{nt-1}^i + s_{nt-1}^*) \frac{\ln \lambda_{nt-1}^*}{(1 - \hat{\sigma}_n)} \right) \\ & + (\hat{\varepsilon}_t^i - \hat{\varepsilon}_{t-1}^i). \end{aligned} \quad (16)$$

The variance in the growth rate of provincial TFP in export sectors is therefore the sum of the variance of the growth rate of variety-induced provincial TFP, and the variance of the difference in error terms, along with the covariance between the two terms. Table 5 shows the variance decomposition of provincial export TFP in levels and growth rates. Surprisingly, only 36.3% of the cross-province differences in the export TFP levels are explained by province fixed effects while variety-induced provincial TFP can account for about 44.1% of the provincial productivity levels. Furthermore, variety-induced TFP and province fixed effects are correlated, jointly contributing nearly 13.8% of the cross-province variation in export TFP levels. Compared to the results from the cross-country literature, our variety-induced TFP can account for a striking proportion of the total export TFP variation.<sup>14</sup> Such

<sup>14</sup>For instance, in FK(2008), the variety-induced TFP can account for only 3.3% of in OECD countries' TFP variations while ours account for 44.1% in China's provinces.

a big discrepancy may be attributed to the following two reasons: first, the difference between provinces within a country should be much smaller than that between countries. So the region fixed effect plays a less important role in this study. Secondly, my unbalanced data covers the period from 1998 to 2005 but with more observations in the most recent 4 years (2002–2005) than those prior to 2002. Taking into account that China entered WTO in 2001, freer trade significantly boosted China’s economic growth through fast growing exports.<sup>15</sup> Thus my data may contain a strong WTO effect which magnifies impact of export variety expansion on export TFP growth. The second column of Table 5 shows the decomposition of the provincial productivity growth. About 36.6% of the within-province growth in export TFP can be explained by the year-to-year growth in export variety, while the reminder is explained by the change in regression errors and the correlation between the two terms. Again, export variety growth in our estimation can explain much more TFP growth than in the cross-country literature. In any event, export variety nonetheless is important in explaining provincial productivity differences in both levels and growth rates.

To further illustrate the effects of export variety on productivity in export sectors, according to (13) a 1% increase in the export variety of each sector  $n$  would increase provincial productivity by  $\frac{1}{2}(s_{nt}^i + s_{nt}^*)\frac{1}{(1-\hat{\sigma}_n)}$  percent. Thus, at the sample mean, a 10% increase in export varieties of all industries could lead to a 1.4% increase in China’s productivity (as an average of provincial productivity). This effect is significant both statistically

---

<sup>15</sup> According to the national statistical yearbook (1995-2005), the average export growth rate is 29% during 2002-2005 compared to 10% during 1995-2001.

and economically.

## 4 Conclusions

Existing analyses of the export variety effects on TFP variation and growth have been restricted to cross country studies and mainly to OECD countries. In this paper I study the case for China by estimating the effects of export variety (via export revenue) on provincial productivity with multiple sectors and introducing export varieties into the provincial GDP function.

Estimating the seven share equations simultaneously with the GDP equation (transformed to become relative provincial productivity) allows me to identify and estimate the elasticities of substitution between export varieties in export sectors and then infer the contribution of export variety on provincial productivity in terms of level difference and growth. The resulting elasticity estimates (measuring the degree of contribution) range from a low of -0.235 in the wood & paper sector to a high of -8.025 in the food & beverage sector. The ranking I have obtained seems reasonable except for the wood & paper industry at first blush. Intuitively the contribution in the machinery industry should be the lowest since new variety in this industry generally can collect high sales revenue. However, considering the fact that China's forest-coverage rate was only 18.21% (only one fourth of the world average) by 2005, the export in wood & paper (for example, furniture in traditional Chinese style) sector is actually highly monopolistic, which may justify the lowest elasticity results. Additionally, the estimation also reveals that the land share in China's GDP is about 8.3%.

I consider the potential problems of correlated error terms in the seven sectoral share equations as well as endogeneity in the adjusted TFP equation. First of all, I conduct a SUR estimation and find that the estimates are very similar to those from NLS which implies that error term correlation is not a serious problem. Then I use a N2SLS estimation to address the potential endogeneity problem. A Hausman test is applied to test the endogeneity problem of the NLS estimation. That the p-value statistic is almost unity suggests the null hypothesis of no endogeneity in the system can not be rejected. This result reinforces the hypothesis suggested by many economists that China's (as well as many Southeast Asian countries') exports are exogenous in explaining productivity growth. Then, I check the overall robustness of the NLS estimation system by testing the specifications in my system. The results of these tests support my previous specifications on homogeneity constraints imposed on endowment and prices as well as the symmetry constraints imposed on cross-equation prices.

Finally, I also calculate the impact of export variety differences across provinces on their respective productivity. Surprisingly, export variety explains 44.1% of the total variation in provincial productivity while the results in previous cross-country studies typically report a much smaller export effect with a fairly big country-fixed effect. The reason may be that the fixed effect across provinces is much smaller than that across countries, so export variety plays a relatively much more important role in explaining TFP variation. Furthermore, export variety can explain 36.6% of the within-province productivity growth. At the sample mean, a 10% increase in export varieties of all industries leads to a 1.4% increase in productivity.

## References

- [1] Arkolakis, C., S. Demidova, P.J. Klenow, and A. Rodríguez-Clare (2008), "Endogenous Variety and the Gains from Trade", *NBER working paper* 13933.
- [2] Bernard, A.B., and J.B. Jensen (1999), "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47(1), 1-26.
- [3] Blonigen, B.A., and A. Soderbery (2009), "Measuring the Benefits of Product Variety with an Accurate Variety Set", *NBER working paper* 14956.
- [4] Breusch, T.S., and A.R.Pagan (1980), "The Lagrange multiplier test and its applications to model specification in econometrics," *Review of Economic Studies* 47, 239–254.
- [5] Broda, C., and D. Weinstein (2004), "Globalization and the Gain from Variety," *Quarterly Journal of Economics* 121, 541-585.
- [6] Broda, C., J. Greenfield, and D. Weinstein (2006), "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth", *NBER working paper* 12512.
- [7] Delgado, M.A., J.C.Farinas, and J.C.Ruano (2002), "Firm Productivity and Export Markets: a non-Parametric Approach", *Journal of International Economics* 57, 397-422.

- [8] Dufour, J.-M., J.F.Kiviet (1998), "Exact inference methods for 1st-order autoregressive distributed lag models," *Econometrica* 66, 79–104.
- [9] Eaton, J., and S. Kortum (2002), "Technology, Geography and Trade," *Econometrica*, 70(5), 1741-1780.
- [10] Feenstra, R.C., (1994) "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157-177.
- [11] Feenstra, R.C., D. Madani, and Yang, T.H. (1998), "Testing Endogenous Growth in South Korea and Taiwan," Unpublished, University of California, Davis.
- [12] Feenstra, R.C., and H.L. Kee (2004), "Export Variety And Country Productivity", NBER working paper 10830.
- [13] Feenstra, R.C., and H.L. Kee (2008), "Export Variety And Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity", *Journal of International Economics* 74(2), 500-518.
- [14] Funke, M., and R. Ruhwedel (2001), "Export Variety and Export Performance: Evidence from East Asia," *Journal of Asia Economics*, 12, 493-505.
- [15] Funke, M., and Ruhwedel, R. (2002), "Export Variety and Export Performance: Empirical Evidence for the OECD Countries", *Weltwirtschaftliches Archiv* 138, 97-114.

- [16] Funke, M., and Ruhwedel, R. (2005), "Export Variety and Economic Growth in Eastern European Transition Economies", *The Economics of Transition* 13, 25-50.
- [17] Ghironi, F., and M. J. Melitz (2005), "International Trade and Macroeconomic Dynamics with Heterogeneous Firms", *Quarterly Journal of Economics* 120, 865-915.
- [18] Grossman, G. M., and E. Helpman (1991), *Innovation and Growth in the Global Economy*. MIT press: Cambridge, MA.
- [19] Hausman, J.A. (1978), "Specification Tests in Econometrics," *Econometrica* 46, 1251-1271.
- [20] Heien, D. M., T. N. Chen, Y. L. Chien, and A. Garrido (1996), "Empirical Models of Meat Demand: How Do They Fit Out of Sample?", *Journal of Agribusiness* 12(1), 51-66.
- [21] Kariya, T. (1981) "Tests for the independence between two seemingly unrelated regression equations," *The Annals of Statistics* 9(2), 381–390.
- [22] A.C.C. Kwan, and B.Kwok (1995), "Exogeneity and the Export-led Growth Hypothesis: The Case of China", *Southern Economic Journal* 61(4): 1158-1166.
- [23] Maeshiro, A. (1980), "New Evidence on the Small Properties of Estimators", *Journal of Econometrics* (12), 177-187.

- [24] Melitz, M.J., (2003), "The Impact of the Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695-1725.
- [25] Romer, P., (1990), "Endogenous Technological Change," *Journal of Political Economy* 98(5), pt.2, S71-S102.
- [26] Sato, K., (1976), "The Ideal Log-Change Index Number," *Review of Economics and Statistics* 58, 223-228.
- [27] Shiba, T., and H.Tsurumi (1988), "Bayesian and non-Bayesian tests of independence in seemingly unrelated regressions," *International Economic Review* 29, 377-389.
- [28] Srivastava, V.K., and D.E.Giles (1987), *Seemingly Unrelated Regression Equations Models. Estimation and inference.* Marcel Dekker, New York.
- [29] Vartia, Y.O. (1976), "Ideal Log-Change Index Numbers," *Scandinavian Journal of Statistics* 3, 121-126.
- [30] Zellner, A. (1962), "An efficient method for estimating seemingly unrelated regressions and tests for aggregate bias," *Journal of the American Statistical Association* 57, 348-368.
- [31] Zhu, Z.Y., (2006), "China's Foreign Trade and Investment Policies for Promoting the International Balance of Payments", Speech on the 2007 China Industry Development Report Conference.

Figure 1. Provincial Export Share and TFP in China (2005)

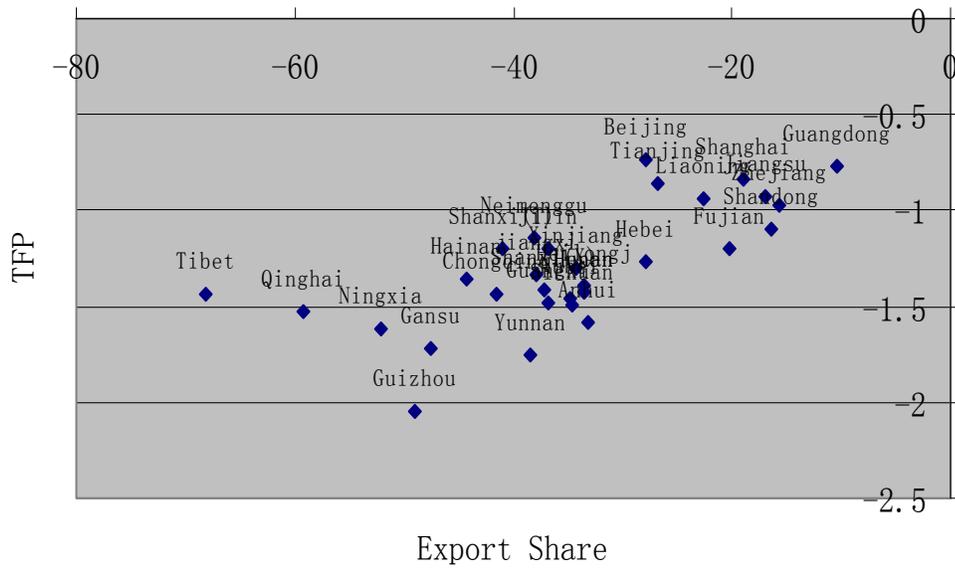
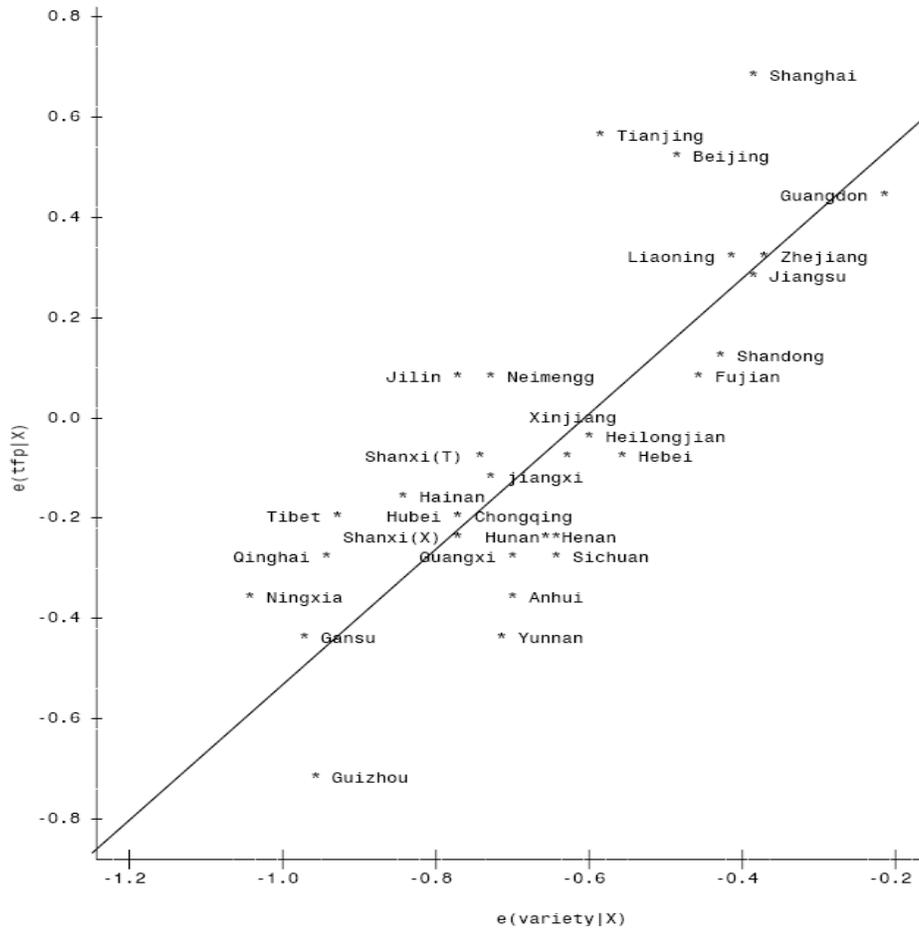


Figure 2. Provincial Productivity vs. Average Export Variety

coef=1.27499, se=0.18627, t=6.85



**Table 0: Main Variables and Parameters to be Estimated**

<b>Variables</b>	<b>Measure</b>
$T$	year-end arable land
$K$	capital stock <sup>a</sup>
$k(K/T)$	land adjusted capital stock
$L$	year-end labor input
$\ell(L/T)$	land adjusted labor input
$s_n$	sector share <sup>b</sup>
$s_L$	labor share <sup>c</sup>
$\lambda_n$	weighted variety ratio <sup>d</sup>
$P_S$	non-traded price index <sup>e</sup>
<b>Parameters</b>	<b>Intepretation</b>
$\sigma_n$	elasticity of substitution between varieties in export sector $n$
$\gamma$	price elasticity of supply
$\phi_L$	labor endowment effect on export (supply)
$\phi_K$	capital endowment effect on export (supply)
$\beta_k$	(negative) land share in GDP
$\beta_S$	non-traded price effect on TFP

a. Capital stocks are constructed by the perpetual inventory method. (see text for details)

b. Share of export sector  $n$  is calculated as ratio of value-added of sector  $n$  to GDP.

c. Labor share is calculated as the ratio of labor income to GDP.

d.  $\lambda_n$  is the ratio of provincial export revenue of sector  $n$  to its national counterpart.

e.  $P_S$ : equally weighted price indices of Education, Health & Child Care, and Rental for Housing.

Source: Author's calculations and various statistical yearbooks (national and provincial), 1999–2006.

Table 1: Dependent Variables-Industry shares in Column (1) to (7), and adjusted TFP in column (8)

Estimation Method: Non-linear Ordinary Least Squares Regressions

Total system observations: 1544

Observations per equation: 193

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Independent Variables	Agriculture	Textile& Garments	Wood & Paper	Chemical &Plastics	Mining& Metals	Machinery & Transportation	Food, Beverage & Tobacco	Adj.TFP
Agriculture	0.023*** (0.008)	-0.010*** (0.003)	0.005*** (0.002)	0.004 (0.003)	-0.022*** (0.008)	-0.016*** (0.004)	0.029*** (0.006)	0.382*** (0.044)
Textile & Garments	-0.010*** (0.003)	0.016*** (0.006)	0.003*** (0.001)	0.001 (0.002)	-0.004 (0.005)	0.002 (0.002)	-0.021*** (0.007)	0.431*** (0.140)
Wood & Paper	0.005*** (0.002)	0.003*** (0.001)	0.004*** (0.001)	-0.006*** (0.002)	-0.006** (0.002)	-0.002*** (0.001)	-0.006*** (0.002)	0.810*** (0.170)
Chemical &Plastics	0.004 (0.003)	0.001 (0.002)	-0.006*** (0.002)	0.013** (0.005)	-0.008 (0.006)	0.002 (0.002)	0.016*** (0.006)	0.418*** (0.101)
Mining& Metals	-0.022*** (0.008)	-0.004 (0.005)	-0.006** (0.002)	-0.008 (0.006)	0.125*** (0.027)	0.003 (0.005)	-0.021* (0.011)	0.170*** (0.029)
Machinery& Transportation	-0.016*** (0.004)	0.002 (0.002)	-0.002*** (0.001)	0.002 (0.002)	0.003 (0.005)	0.026*** (0.005)	-0.005 (0.003)	0.572*** (0.091)
Food, Beverage & Tobacco	0.029*** (0.006)	-0.021*** (0.007)	-0.006*** (0.002)	0.016*** (0.006)	-0.021* (0.011)	-0.005 (0.003)	0.060*** (0.017)	0.111*** (0.026)
Capital-Land Ratio	-0.059*** (0.005)	-0.002 (0.002)	-0.001 (0.001)	0.010*** (0.002)	0.030*** (0.006)	0.018*** (0.003)	-0.001 (0.004)	-0.083*** (0.018)
Labor-Land Ratio	0.057*** (0.007)	0.002 (0.003)	0.006*** (0.001)	-0.015*** (0.003)	-0.077*** (0.008)	-0.009** (0.004)	0.004 (0.006)	
Non-traded Goods Prices								-0.468*** (0.057)
Year Fixed Effects	YES	NO	YES	YES	YES	YES	NO	YES
Prov. Fixed Effects								YES
R-squared	0.7221	0.4855	0.2033	0.3325	0.5179	0.7638	0.3304	0.9761

Note: For columns (1) to (7), each log of relative export variety coefficient ( $\gamma_{nm}$ ) is the partial price effect of the industry in

that row on the share of the industry in the column. For column (8), each log of the relative export variety coefficient

( $1/(1 - \sigma_n)$ ) is the point estimate of the industry in that row.

\*, \*\*, and \*\*\* indicate significance at 90%, 95% and 99% confidence levels respectively, and White-robust standard errors are in parentheses.

Table 2: Dependent Variables-Industry shares in Column (1) to (7), and adjusted TFP in column (8)

Estimation Method: Iterated Non-linear SUR

Total system observations:1544

Observations per equation: 193

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Independent Variables</i>	Agriculture	Textile.& Garments	Wood & Paper	Chemical & Plastics	Mining& Metals	Machinery & Transportation	Food, Beverage & Tobacco	Adj.TFP
Agriculture	0.035*** (0.0106)	-0.005*** (0.002)	0.002 (0.0011)	-0.010** (0.004)	-0.030*** (0.010)	-0.017*** (0.005)	0.029*** (0.006)	0.297*** (0.043)
Textile & Garments	-0.005*** (0.002)	0.005*** (0.001)	0.002*** (0.000)	-0.000 (0.001)	0.000 (0.002)	-0.000 (0.001)	-0.008*** (0.002)	1.195*** (0.253)
Wood & Paper	0.002 (0.001)	0.002*** (0.000)	0.001*** (0.000)	-0.002** (0.001)	-0.004*** (0.001)	-0.001 (0.001)	-0.001 (0.001)	1.533*** (0.335)
Chemical & Plastics	-0.010** (0.004)	-0.000 (0.001)	-0.002** (0.001)	0.016** (0.007)	0.013** (0.005)	0.006** (0.003)	-0.000 (0.003)	0.286** (0.115)
Mining& Metals	-0.030*** (0.010)	0.000 (0.002)	-0.004*** (0.001)	0.013** (0.005)	0.087*** (0.020)	-0.003 (0.005)	-0.005 (0.007)	0.208*** (0.041)
Mach.& Transport.	-0.017*** (0.005)	-0.000 (0.001)	-0.001 (0.001)	0.006** (0.003)	-0.003 (0.005)	0.036*** (0.009)	-0.007** (0.004)	0.41*** (0.089)
Food, Beverage & Tobacco	0.029*** (0.006)	-0.008*** (0.002)	-0.001 (0.001)	-0.000 (0.003)	-0.005 (0.007)	-0.007** (0.004)	0.037*** (0.009)	0.235*** (0.045)
Capital-Land Ratio	-0.056*** (0.005)	-0.001 (0.002)	-0.000 (0.001)	0.009*** (0.003)	0.031*** (0.006)	0.018*** (0.004)	-0.002 (0.004)	-0.090*** (0.018)
Labor-Land Ratio	0.056*** (0.007)	0.002 (0.003)	0.003** (0.001)	-0.015*** (0.003)	-0.079*** (0.008)	-0.008* (0.005)	0.006 (0.005)	
Non-traded Goods Prices								-0.470*** (0.0567)
Year Fixed Effects	YES	NO	NO	YES	YES	YES	NO	YES
Prov. Fixed Effects								YES
R-squared	0.7077	0.4621	0.4916	0.4164	0.4659	0.7487	0.2633	0.9660

Note: For columns (1) to (7), each log of relative export variety coefficient ( $\gamma_{nm}$ ) is the partial price effect of the industry in

that row on the share of the industry in the column. For column (8), each log of the relative export variety coefficient ( $1/(1 - \sigma_n)$ ) is the point estimate of the industry in that row.

\*, \*\*, and \*\*\* indicate significance at 90%, 95% and 99% confidence levels respectively, and White-robust standard errors are in parentheses.

Table 3: Dependent Variables-Industry shares in Column (1) to (7), and adjusted TFP in column (8)

Estimation Method: Two Stage Nonlinear Least Squares Regressions

Total system observations: 1544

Observations per equation: 193

<i>Independent Variables</i>	(1) <i>Agriculture</i>	(2) <i>Textile &amp; Garments</i>	(3) <i>Wood &amp; Paper</i>	(4) <i>Chemical &amp; Plastics</i>	(5) <i>Mining &amp; Metals</i>	(6) <i>Machinery</i>	(7) <i>Food &amp; Beverage</i>	(8) <i>Adj.TFP</i>	<i>1<sup>st</sup>-Stg F-stat.</i>
Agriculture	0.0566*** (0.0172)	-0.0147*** (0.0046)	0.0037** (0.0019)	0.0065** (0.0029)	-0.0389** (0.0158)	-0.0342*** (0.0083)	0.0751*** (0.0182)	0.2506*** (0.0569)	33.15
Textile & Garments	-0.0147*** (0.0046)	0.0105** (0.0041)	-0.0002 (0.0005)	0.0006 (0.0009)	0.0022 (0.0042)	0.0007 (0.0018)	-0.0184*** (0.0064)	0.9475*** (0.1390)	17.22
Wood & Paper	0.0037** (0.0019)	-0.0002 (0.0005)	0.0030*** (0.0008)	-0.0023*** (0.0006)	-0.0054** (0.0022)	-0.0024** (0.0010)	-0.0031* (0.0018)	1.6886*** (0.4574)	22.07
Chemical & Plastics	0.0065** (0.0029)	0.0006 (0.0009)	-0.0023*** (0.0006)	0.0052*** (0.0016)	-0.0063* (0.0035)	-0.0023* (0.0015)	0.0130*** (0.00331)	1.564*** (0.3101)	14.00
Mining & Metals	-0.0389** (0.0158)	0.0022 (0.0042)	-0.0054** (0.0022)	-0.0063* (0.0035)	0.161*** (0.0409)	0.0105 (0.0072)	-0.0301* (0.0187)	0.1764*** (0.0403)	62.89
Machinery	-0.0342*** (0.0083)	0.0007 (0.0018)	-0.0024** (0.0010)	-0.0023* (0.0015)	0.0105 (0.0072)	0.0456*** (0.0117)	-0.0133** (0.0066)	0.4815*** (0.1094)	33.67
Food & Beverage	0.0751*** (0.0182)	-0.0184*** (0.0064)	-0.0031* (0.0018)	0.0130*** (0.00331)	-0.0301* (0.0187)	-0.0133** (0.0066)	0.0389 (0.0269)	0.0748*** (0.0299)	10.03
Capital-Land Ratio	-0.0457*** (0.0060)	-0.0033 (0.0022)	-0.0008 (0.0013)	0.0100*** (0.0024)	0.0286*** (0.0069)	0.0154*** (0.0037)	0.0054 (0.0055)	-0.14278*** (0.0237)	
Labor-Land Ratio	0.0502*** (0.0072)	0.0017 (0.0026)	0.0059*** (0.0017)	-0.0140*** (0.0031)	-0.0781*** (0.0089)	-0.0112** (0.0049)	-0.0004 (0.0068)		
Non-traded Goods Prices								-0.4275*** (0.0920)	
Year Fixed Effects	YES	YES	YES	YES	YES	YES	NO	YES	
Prov. Fixed Effects								YES	
R-squared	0.7045	0.4406	0.1061	0.2740	0.4459	0.7282	0.0313	0.9577	

Note: For columns (1) to (7), each log of relative export variety coefficient ( $\gamma_{mm}$ ) is the partial price effect of the industry in

that row on the share of the industry in the column. For column (8), each log of the relative export variety coefficient ( $1/(1 - \sigma_{\mu})$ ) is the point estimate of the industry in that row.

\*, \*\*, and \*\*\* indicate significance at 90%, 95% and 99% confidence levels respectively, and White-robust standard errors are in parentheses.

Table 4: Hypothesis Testing

Null Hypothesis	Homogeneity in		Symmetry in Cross Price Effects	Hausman Test OLS vs. N2SLS
	Endowments	Prices		
Degree of Freedom	1	1	21	139
Critical Value at 95%	3.84	3.84	32.67	167.5
Overall System			7.70 (0.9963)	94.33 (0.9986)
Agriculture	0.13 (0.7145)	3.51* (0.0610)		
Textile & Garments	0.01 (0.9080)	0.10 (0.7486)		
Wood & Paper	0.04 (0.8355)	0.05 (0.8216)		
Chemical & Plastics	0.35 (0.5556)	0.91 (0.3392)		
Mining & Metals	1.68 (0.1947)	0.70 (0.4023)		
Machinery & Transportation	0.95 (0.3297)	0.03 (0.8690)		
Food, Beverage & Tobacco	0.05 (0.8287)	0.84 (0.3587)		
Adj. TFP	0.92 (0.3387)	121.66*** (<.0001)		

Notes: All test statistics are asymptotically Chi-squared distributed with degree of freedom equals number of restrictions. Numbers in parentheses denote p-value of the test statistics.

Table 5: Productivity Decomposition

	Level Decomposition (in % of TFP)	Growth Decomposition (in % of TFP)
Variance of Estimated Province TFP	0.0954 (100)	0.0050 (100)
Variance of Province Fixed Effects	0.0347 (36.3%)	-
Variance of Variety Induced TFP	0.0421(44.1%)	0.0018 (36.6%)
2*Covariance between Province Fixed Effects and Variety Induced TFP	0.0138 (14.4%)	-

Source: Author's calculation based on regression results of Table 1.