Long-Run Estimates of Interfuel and Interfactor Elasticities

Chunbo Ma
School of Agricultural and Resource Economics, University of Western Australia, Crawley, WA 6009, AUSTRALIA. Email: chunbo.ma@uwa.edu.au.

David I. Stern
Crawford School of Public Policy, The Australian National University, 132 Lennox Crossing, Acton, ACT 2601, AUSTRALIA. E-mail: david.stern@anu.edu.au. Phone: +61-2-6125-0176

18 December 2014

Abstract
Econometric theory suggests that the between estimator might generate the best estimates of long-run elasticities but no existing estimates of the elasticities of substitution have used it. Chirinko *et al*. argued instead for using an interval difference estimator. We provide estimates of interfuel and interfactor elasticities of substitution using the between and long-run difference estimators using and a panel of Chinese provincial data. We also improve on previous studies by adding total factor productivity terms to our regressions. Our results are quite different and more plausible than previous research and, as expected, our estimates of elasticities are larger than traditional fixed effects estimates.

Key Words: panel data, elasticity of substitution, demand, China, energy

JEL Codes: D24, Q40

Acknowledgements: We thank the Australian Research Council for funding through grant DP120101088: "Energy transitions: Past, present and future." We thank participants at the ANU-Tsinghua Forum in Beijing on 15th September 2014 for useful comments and Shuang Liu for producing the maps.

1. Introduction

In this paper, we use the between estimator and a long-run difference estimator to estimate interfuel and interfactor elasticities of substitution using the translog cost function and a panel of Chinese provincial data. We also improve on previous studies by adding total factor productivity terms to our regressions, which should eliminate the main source of correlation between the regressors and the error terms. Our results are quite different to some previous estimates for China but are close to what would be expected for long-run estimates of elasticities of substitution from meta-analyses.

Meta-analysis of interfuel shadow elasticities of substitution for coal, oil, gas, and electricity (Stern, 2012) shows that the type of data – time series, panel, or cross-section – and the estimator used in the primary studies strongly affect their econometric results. Stern (2012) found that interfuel shadow elasticities of substitution from cross-section studies are greater than unity for all combinations of fuels apart from coal and electricity. Elasticities of substitution are generally smaller for OLS and fixed-effects panel estimates and time series estimates. Koetse *et al.* (2008) found similar results for capital-energy elasticities of substitution. But only two of the studies in the database analyzed by Stern (2012) used cross-section estimates and so more cross-section estimates would be desirable. However, cross-section estimates may be biased as they only utilize a single time series observation (Pesaran and Smith, 1995). Econometric theory (Pesaran and Smith, 1995; Griliches and Mairesse, 1984; Mairesse, 1990; Hauk and Wacziarg, 2009) suggests that the between estimator – a cross-sectional regression on the mean values over time for each individual - will produce consistent estimates of long-run elasticities under ideal assumptions and produce less biased estimates than traditional panel data estimators in the presence of misspecifications. But this estimator has not been used in the context of interfuel and energy-capital substitution. Chirinko *et al.* (2011) suggested instead that the interval difference estimator, which uses the differences between two time series averages, might provide superior estimates of long-run substitution possibilities. They used this estimator to estimate the elasticity of substitution between capital and labor using a CES production function but this estimator has also not been applied to interfuel or capital-energy substitution possibilities.

Estimates of interfuel and energy-capital elasticities of substitution are particularly relevant to the estimation of the costs of climate mitigation. There are widely divergent opinions on the costs of climate mitigation policies and their impact on economic growth. There has been extensive work on modeling the costs of climate change mitigation and adaptation using the tools of computable general equilibrium (CGE) models. Such models critically depend on research on the possibilities for technological change and substitution between energy and other inputs and among fuels. The parameters that govern these possibilities – the elasticities of substitution - "are the single most important parameters that affect the[ir] results." (Bhattacharya, 1996, 159). Furthermore, "in the economic literature, there is little consensus about different elasticities for energy products" (Bhattacharya, 1996, 159). Stern (2012) found a large dispersion in the estimated elasticities of substitution between fuels and that estimates based on time-series such as those used in the G-Cubed (McKibbin and Wilcoxen, 1999) or IGEM (Goettle *et al*., 2007) models tend to underestimate the long-run possibilities of substitution between inputs. Similar results were found by a meta-analysis of substitution possibilities between energy and capital (Koetse *et al*., 2008). As China is the largest emitter of greenhouse gases, better estimates of Chinese elasticities are important for estimates of mitigation possibilities and costs.

A second methodological contribution of this paper is that we add province level inefficiency and technical change terms to our interfactor cost function equation. We compute these using index number methods and then add them to the regression models. There are very substantial productivity differences between provinces. Taking these into account means that our results are quite different from previous estimates of substitution elasticities for China.

We use Chinese data that is similar to that previously used by H. Ma *et al*. (2008, 2009).[1] This Chinese data has a good balance of time and cross-section dimensions compared to other datasets used in the literature (Stern, 2012) and so is suitable for evaluating different panel data estimators. Our dataset consists of province level data on quantities of the final use of individual fuels and electricity, capital, and labor, and the price series from the provincial capitals, which we use to proxy provincial prices. This data has a cross-section dimension of 30 and a time series dimension of eleven years. We aggregate the various types of coal into a single coal input and estimate the energy cost function and cost share equations for coal, gasoline, diesel, and electricity. We also aggregate all energy types into a single energy input and carry out a similar analysis for capital, labor, and energy. The disadvantage of this

[1] Our data covers the period 2000-2010 whereas H. Ma *et al*. (2008) used the period 1995-2004.

dataset is that there are no natural gas variables in the data and so we can only provide information on three of the elasticities of substitution analyzed by Stern (2012).

The paper follows the usual layout with the methods section following this introduction, a section describing the data and then the econometric results are presented. The final two sections discuss the results in the context of previous estimates and provide conclusions.

2. Methods

a. Background

Differences between time-series and cross-section estimates have long been discussed in the econometric literature (Baltagi and Griffin, 1984). In recent decades, this interest has been transferred to panel data, as time-series and cross-sections can be seen as special cases of panels with a cross-section or time dimension of one, respectively.

Apostolakis (1990) and Bacon (1992) surveyed some of the early studies of interfuel substitution elasticities in the OECD countries. Bacon found that panel data studies tended to find more substitutability between fuels as measured by their cross-price elasticities than did time-series studies. He suggested that this was because this data represented long-run elasticities, while time-series data generated short-run elasticities. Apostolakis (1990) came to similar conclusions regarding substitution between aggregate energy and capital. Though short-run elasticities of substitution can be defined and estimated (Mundlak, 1968; Sharma, 2002), the usual definitions of elasticities of substitution are based on long-run responses and, therefore, long-run estimates are desirable. As mentioned in the introduction, recent meta-analyses of the interfuel elasticities of substitution (Stern, 2012) and the capital-energy elasticity of substitution (Koetse *et al*., 2008) literatures find that the largest elasticities of substitution are produced by cross-section estimates and the smallest by time series estimates with fixed effects estimates somewhere in between. Koetse *et al*. (2008) find a mean Morishima elasticity of substitution in time-series data of 0.22, in panel data of 0.59, and in cross-section data of 0.85. Stern (2012) finds an average shadow elasticity of interfuel substitution of 0.49 for time series estimates, 1.05 for OLS panel estimates, 1.06 for fixed effects estimates, and 1.60 for cross-section data.

Pesaran and Smith (1995) point out that, if the true data generating process is static, the explanatory variables are uncorrelated with the error term, and any parameter heterogeneity across individuals is random and distributed independently of the regressors, all the usual estimators – time-series, cross-section, and panel OLS, fixed and random effects, and between estimates - should be consistent estimators of the coefficient means. It is the presence of dynamics and/or correlation between the regressors and the error term that results in differences between the estimators. There is no essential difference between time-series and panel estimates, only differences in the likely importance and impact of misspecification. They argue further that, in the absence of correlation between the regressors and the error components, the cross-sectional average of dynamic time-series models for each individual and BE are consistent. But a traditional cross-section estimate – BE for a single period - may suffer from a high level of bias. In the presence of coefficient heterogeneity, FE and RE estimators for dynamic models will be inconsistent, as forcing the coefficients to be equal induces serial correlation in the disturbance, which results in inconsistency when there are lagged dependent variables. If the true model is static, static FE and RE should be consistent in the absence of other misspecifications. When the true model is dynamic, the higher the level of correlation between the dependent variable and the lagged variables omitted by a static estimator, the closer static estimates will be to the long-run coefficients (Baltagi and Griffin, 1984). In the non-stationary case, static time-series estimates are superconsistent when the variables are I(1) and cointegrate. But, if the parameters vary across groups, the pooled estimates need not cointegrate. BE also consistently estimates the long-run coefficients when the explanatory variables are non-stationary but strictly exogenous even if there is no cointegration (Pesaran and Smith, 1995).

However, the assumption that the regressors and errors are uncorrelated does not necessarily hold. The one-way error components model assumes that the error term in a panel model is composed of an individual effect, which varies across individuals but is constant over time, and a remainder disturbance that varies over both time and individuals (Baltagi, 2008). If omitted explanatory variables are correlated with the included regressors, the regressors will be correlated with the individual effects and/or the remainder disturbance (Griliches and Mairesse, 1987). The fixed effects estimator eliminates the individual effects prior to estimation while the between estimator averages over the remainder disturbances of each individual. Therefore, OLS panel, RE, BE, and cross-section estimators will be biased if the regressors are correlated with the individual effects and FE and time-series estimators will be unbiased. But if the correlation is with the remainder disturbance instead, BE will be consistent and all the other estimators will be inconsistent (Griliches and Mairesse, 1987).

Measurement error in the explanatory variables is also problematic in this context as it induces a correlation between the error term and the regressors and biases the estimates towards zero (Hausman, 2001). If measurement errors are non-systematic, BE will average them out over time and will be consistent but biased when the time-series dimension is small, while FE amplifies the noise to signal ratio by subtracting individual means from each time-series (Mairesse, 1990). Hauk and Wacziarg (2009) conducted a Monte Carlo analysis of an economic growth equation to examine the effects of combined measurement error and omitted variables on alternative panel estimators. The former would be expected to affect FE more and the latter to affect BE more. They found BE to have the minimum bias relative to FE, RE, and some GMM estimators commonly used in the growth literature. Other papers that find superior performance for BE compared to other potential estimators are Pirotte (1999) and Egger and Pfaffermayr (2004).

We can reduce the potential correlation between the regressors and the individual effects by including additional variables that vary across individuals and are usually omitted from regression analyses. In the case of cost share and cost function equations, the most important omitted variable is likely to be the state of technology. Total factor productivity (TFP) varies across Chinese provinces and it is likely that TFP is correlated with input prices. For example, coal is cheaper and TFP lower in the poorer inland provinces. Of course, wage rates will be highly correlated with TFP. Therefore, we may obtain more consistent results by including TFP as an additional variable in the cost function equation for the interfactor estimates. In this study, we compute these TFP indices both across provinces and over time and include these in the regressions as appropriate.

Chirinko *et al*. (2011) also propose a method intended to capture long run rather than short run variation – the interval difference estimator (IDE). They compute the average of each variable over two periods of seven years and then compute the difference between the two periods. The estimator uses the cross-section of these interval differences. They interpret this estimator "in terms of a low-pass filter placing relatively more weight on low-frequency movements than the traditional approach of first-differencing" (588). They argue that IDE is robust to several potential issues including unit roots, omitted variables bias, misspecified dynamics, and measurement error. With the exception of omitted variables bias, these are the same problems that motivate adoption of BE. In a departure from Chirinko *et al*. (2011), we compute the differences from the first to the last time series observations in our sample. This allows us to exploit more of the variation in our short sample of 11 years. Differences computed in the two different ways are highly correlated. This estimator then uses the cross-section of differences over time whereas the between estimator uses the cross-section of averages over time.

Chirinko *et al*. (2011) argue that the differenced regressors will likely not be correlated with the productivity shocks but they also provide constant returns to scale and instrumental variable estimates in case they are. Instead, we include estimates of the productivity shocks by including the long-run differences of the provincial TFP variables mentioned above.

The relatively regulated energy price regime in China means that the assumption that energy prices at the provincial level are exogenous and driven mostly by differences in transportation costs is not unreasonable. However, following Pindyck (1979), we also use an IV method to take into account the possible endogeneity of price indices for coal and energy in our interfuel and interfactor analyses respectively in both our IDE and BE models.

b. Model

Assuming constant returns to scale, the translog cost function for a panel of provinces is given by:

$$\ln C_{it} = \beta_0 + \ln D_{it} + f_t + \sum_{j=1}^{J} \beta_j \ln P_{jit} + 0.5 \sum_{j=1}^{J} \sum_{k=1}^{J} \beta_{jk} \ln P_{jit} \ln P_{kit} + \sum_{j=1}^{J} \gamma_j t \ln P_{jit} + \varepsilon_{it} \tag{1}$$

where C is unit output cost, P are the prices of the J inputs indexed by j and k, i indexes provinces, and t years. All log prices and the linear time trend are normalized at the sample mean. The first parameter on the RHS is a national mean effect, the second a provincial efficiency effect, which varies over time and is zero in the most efficient province,[2] and the third is a national time effect with mean zero. The final term in the equation is a stationary random error term with a mean of zero.

[2] The inefficiency term is the log of the distance of the province from the efficient frontier. A distance of unity places a province on the frontier and hence the log of distance in this province – in our sample Shanghai – is zero. See the next subsection for details of computation of distance and TFP.

As we assume constant returns to scale, in our interfactor analysis C is total cost per unit of real gross output.[3] In our interfuel analysis, C is the cost per unit of energy calculated as cost of energy divided by the aggregate energy input index (Pindyck, 1979). We do not include provincial TFP terms in the interfuel function. However, we do include the technical change bias terms. The IDE model allows implicitly for national level technological change if it includes a constant term in the cost function equation, as we will see below.

We impose the standard homogeneity and symmetry conditions on the parameters in all our estimates. Homogeneity of degree one in prices is imposed by deducting the log price of the Jth input from the prices of the first J-1 inputs. For the interfuel analysis we use the price of diesel as the Jth or numeraire and for the interfactor analysis we use the price of labor as the Jth or numeraire. The cost function should also be concave in input prices. We test the concavity of the cost function at the reference point. We found that the concavity assumption was violated for the interfuel substitution model but not for the interfactor substitution model. Therefore, we imposed concavity on the interfuel model using the method of Ryan and Wales (2000). However, we report the original coefficients of the cost function as in equation (1) rather than the coefficients estimated in the Ryan and Wales method and we compute their standard errors using the delta method.

The standard cost share equations based on Shephard's Lemma (Shephard, 1953) are given by:

$$S_{jit} = \beta_j + \sum_{k=1}^{J} \beta_{jk} \ln P_{kit} + \gamma_j t + \varepsilon_{jit} \,,\ \forall j = 1, \ldots, J-1 \tag{2}$$

Only the first J-1 equations need to be estimated as the shares sum to unity. Using the Jarque-Bera test we could not reject the hypotheses that both the share data across provinces and the differenced share data are normally distributed. Hence the simple cost share functional form is appropriate. Applying the between estimator to (2) implies estimating the cross-sectional regressions:

$$M\left(S_{jit}\right) = \beta_j + \sum_{k=1}^{J} \beta_{jk} M\left(\ln P_{kit}\right) + M\left(\varepsilon_{jit}\right),\ \forall j = 1, \ldots, J-1 \tag{3}$$

[3] See the following subsection for details of the computation of gross output.

where $M()$ is the mean over time operator. As explained above, we deduct the sample mean of the logs of the price variables prior to averaging. β_j is then an estimate of the cost share when all prices are at their means in all provinces and can be used in elasticity formulae (Stern, 2011). Because of its zero mean, the technical change bias has been averaged away. The simplest approach is to estimate only the *J*-1 equations (3), imposing the cross-equation symmetry restrictions. However, better estimates might be obtained and degrees of freedom increased by jointly estimating (3) and the cost function itself (Leon-Ledesma *et al*., 2010). Averaging (1) over time yields:

$$M\left(\ln C_{it}\right)=\beta_0+M\left(\ln D_{it}\right)+\sum_{j=1}^{J}\beta_j M\left(\ln P_{jit}\right)+0.5\sum_{j=1}^{J}\sum_{k=1}^{J}\beta_{jk}M\left(\ln P_{jit}\ln P_{kit}\right)+\sum_{j=1}^{J}\gamma_j M\left(t\ln P_{jit}\right)+M\left(\varepsilon_{it}\right) \tag{4}$$

The sample means of the log prices and the time trend are first deducted before any other operations – creation of the various interactions and averaging – are carried out. We also deduct the mean of the log of cost, though this only affects the estimate of β_0, but we do not demean the distance variable. Also, the national time effect, but not the technical change bias terms, has been averaged away.[4] We impose a coefficient of unity on the distance variable by subtracting it from both sides of (4). Estimating (3) and (4) jointly takes advantage of cross-equation restrictions but no cross-equation restrictions can be imposed on the γ_j coefficients.

As IDE involves differencing rather than averaging, the estimation equations are different. Differencing (2) yields:

$$D\left(S_{jit}\right)=\sum_{j=1}^{J}\beta_{jk}D\left(\ln P_{kit}\right)+\gamma_j D\left(t\right)+D\left(\varepsilon_{jit}\right),\forall j=1,\ldots,J-1 \tag{5}$$

where *D()* is the interval differencing operator and the constant term has been differenced away. The estimation equation for the cost function is:

[4] As the translog function is non-linear the mean value of cost may not coincide with the sample mean of the prices. Therefore, it is not appropriate to instead use (1) with the time averaged means of the variables substituted in places of the time series of the variables. Instead the interaction terms should be computed first and then averaged. Also the biased technical change component of the time effect is not averaged away because, in general, $M\left(t\right)M\left(\ln P_{jit}\right)=0$ but $M\left(t\ln P_{jit}\right)\neq 0$.

$$D\left(\ln C_{it}\right)=\sum_{j=1}^{J}\beta_j D\left(\ln P_{jit}\right)+0.5\sum_{j=1}^{J}\sum_{k=1}^{J}\beta_{jk} D\left(\ln P_{jit}\ln P_{kit}\right)+\sum_{j=1}^{J}\gamma_j D\left(t\ln P_{jit}\right)-D\left(\ln A_{it}\right)+D\left(\varepsilon_{it}\right) \tag{6}$$

where A_i is an index of TFP in province i in year t and combines the provincial inefficiency and national time effects in (1). We impose a coefficient of minus one on this variable by adding it to both sides of (6). As we did for BE, the sample means of each price and time variable are first deducted before any other operations – creation of the various interactions and differencing – are carried out. Estimating (5) and (6) jointly takes advantage of cross-equation restrictions but no cross-equation restrictions can be imposed on the β_j coefficients. For the interfuel model we do not include the TFP term but we do include the constant $D(t)$ while in the interfactor model there is no constant in the cost function equation.

In order to compare our proposed estimators to more conventional approaches we also estimate a fixed effects model. The cost function is estimated using:

$$F\left(\ln C_{it}\right)=F\left(\ln D_{it}\right)\sum_{j=1}^{J}\beta_j F\left(\ln P_{jit}\right)+0.5\sum_{j=1}^{J}\sum_{k=1}^{J}\beta_{jk} F\left(\ln P_{jit}\ln P_{kit}\right)+\sum_{j=1}^{J}\gamma_j F\left(t\ln P_{jit}\right)+\varepsilon_{it} \tag{7}$$

where $F()$ is the fixed effects operator that first subtracts provincial means from each of the time series and then subtracts national time means from each time period. Therefore, we use both provincial and annual effects. We estimate the cost share equations using the following equations:

$$F\left(S_{jit}\right)+S_j=\beta_j+\sum_{k=1}^{J}\beta_{jk} F\left(\ln P_{kit}\right)+\varepsilon_{jit}\ ,\ \forall j=1,\ldots,J-1 \tag{8}$$

where S_j is the national mean cost share, which we add in order to get better estimates of the β_j parameters by equating their value in (7) and (8).

As we aggregate various types of coal into a single coal input using the Divisia index in the interfuel model and similarly aggregate the various types of energy into a single index in the interfactor model, the aggregated coal and energy prices is quantity weighted and endogenously chosen. We follow Pindyck (1979) and use instrumental variables to estimate the models. We represent the price of aggregated coal (i.e. the average cost of aggregated coal for a producer choosing different coal products) by a homothetic translog cost function with constant returns to scale. Estimation of the share equations implied by this cost function allows us to generate the fitted cost function, which provides an instrumental variable for the price of the aggregated coal input. We also create additional instrumental variables for the interactions between the coal price and other variables by interacting the fitted price of the aggregated coal with the prices of other inputs and time. We follow the same procedure for the interfactor substitution estimation. We estimate the share equations implied by a homothetic translog cost function for the price of the aggregated energy input, and then generate instruments using the fitted energy cost and its associated interactions.

c. Computing Gross Output and Distance and Total Factor Productivity

Gross output and value added are related as follows:

$$P_G G = P_Q Q - P_E E \tag{9}$$

where Q is gross output, G is GDP , and E is energy and the P_i are their prices as indicated by the subscripts.[5] As the price of gross output is unobserved, assume that the price of GDP and gross output are equal and then compute gross output as follows:

$$Q = G + \frac{P_E}{P_G} E \tag{10}$$

Based on Hsieh (2002), we compute total factor productivity starting from the assumption that the value of output must equal the value of input and so the ratio of the value of output in two different provinces or years *i* and *j* must equal the ratio of the value of their inputs:

$$\frac{P_{Qj} Q_j}{P_{Qi} Q_i} = \frac{\sum_k P_k X_{kj}}{\sum_k P_{ki} X_{ki}} = \bar{P}_{ji} \bar{X}_{ji} \tag{11}$$

where $\bar{P}_{ji}$ and $\bar{X}_{ji}$ are indices for the difference in prices and difference in quantities of inputs across the two provinces or years. Then rearranging (9) we have:

[5] Obviously, gross output should really also add back in the value of all other intermediate inputs besides energy, but then we would need to also include these inputs in our cost function. Given data limitations, the only intermediate input we consider is energy.

$$\frac{Q_j}{Q_i} / \overline{X}_{ji} = \frac{P_{Qi}}{P_{Qj}} \overline{P}_{ji} \tag{12}$$

Given that the LHS is the primal index of the TFP difference between two provinces or years as given by Feenstra *et al*. (2013), the RHS is the dual index of the TFP difference between the two provinces. The logic of the formula on the RHS is that if input prices are relatively high in a province or year compared to output prices then that province or year must be relatively productive (e.g. high wage). Following Feenstra *et al*. (2013), who use the Divisia index to approximate the quantity index, we approximate the price index using the Divisia index.

For IDE we compute a TFP time series for each province using the RHS of (12) and then take the difference in log TFP between the last and first years for each province. For BE we compute the relative TFP of each province *i* to the TFP in the most productive province, *j* (Shanghai). This is, therefore, the distance of each province from the efficient frontier. We compute the distance of all provinces in each year; take logs, and then average across years in each province.

d. Elasticities

We compute own and cross-price elasticities and Morishima and shadow elasticities of substitution. For the normalized translog cost function, the own and cross-price net elasticities at the sample mean are given by:

$$\eta_{ii} = \frac{\partial \ln X_i}{\partial \ln p_i} = \frac{\beta_{ii} + \beta_i^2 - \beta_i}{\beta_i} \tag{13}$$

$$\eta_{ij} = \frac{\partial \ln X_i}{\partial \ln p_j} = \frac{\beta_{ij} + \beta_i \beta_j}{\beta_j} \tag{14}$$

We compute standard errors for these mean elasticities using the delta method. Positive cross-price elasticities indicate p-substitutes and negative cross-price elasticities p-complements.[6] Morishima and shadow elasticities of substitution measure the difficulty of

[6] p-substitutes and complements are the standard definitions of substitutes and complements measuring the response of factor quantities to changes in factor prices. By contrast q-substitutes and q-complements are defined by the reaction of factor prices to factor quantities. Inputs are usually q-complements – an increase in the level of other inputs increases their

substitution by measuring the response of the factor quantity ratio to a change in the factor price ratio holding the prices of other inputs and output constant. Values between zero and unity indicate poor substitutability and values above unity high substitutability.[7] Unless cost is also held constant, the response of the factor quantity ratio depends on which price in the price ratio changes. The Morishima elasticities are asymmetric because they do not hold cost constant while the shadow elasticities are symmetric because they hold costs constant. The symmetric shadow elasticities are good summary statistics of the overall degree of substitutability between inputs.

At the sample mean, the Morishima elasticities of substitution for a change in the price of input *i* is given by:

$$\mu_{ij} = \frac{\partial \ln X_j}{\partial \ln p_i} - \frac{\partial \ln X_i}{\partial \ln p_i} = \frac{\beta_{ij} + \beta_i \beta_j}{\beta_j} - \frac{\beta_{ii} + \beta_i^2 - \beta_i}{\beta_i} \tag{15}$$

The shadow elasticity of substitution at the sample mean can be expressed as the share weighted mean of the Morishima elasticities:

$$\sigma_{ij} = \frac{S_i}{S_i + S_j}\mu_{ij} + \frac{S_j}{S_i + S_j}\mu_{ji} = \frac{\beta_i}{\beta_i + \beta_j}\left(\frac{\beta_{ij} + \beta_i \beta_j}{\beta_j} - \frac{\beta_{ii} + \beta_i^2 - \beta_i}{\beta_i}\right) + \frac{\beta_j}{\beta_i + \beta_j}\left(\frac{\beta_{ij} + \beta_i \beta_j}{\beta_i} - \frac{\beta_{jj} + \beta_j^2 - \beta_j}{\beta_j}\right) \tag{16}$$

Allen-Uzawa elasticities of substitution are frequently reported in studies of substitution possibilities but we do not report them.[8] When there are only two inputs and constant returns to scale then the elasticity of substitution is unambiguously defined – the Morishima and Shadow elasticity formulae reduce to the Allen-Uzawa elasticity of substitution. But when there are more than two inputs this is not the case – the Allen Uzawa elasticities have the same sign as the cross-price elasticities and no longer measure the difficulty of substitution on a zero to infinity scale. On the other hand, they do not add any information beyond that

---

marginal product but could be p-complements or p-substitutes. Simply referring to inputs as complements or substitutes is, therefore, confusing (Stern, 2011).

[7] The former are often referred to as complements and the latter substitutes but this terminology is again confusing. When there are only two inputs they must be net p-substitutes irrespective of the value of the elasticity of substitution.

[8] The Allen-Uzawa elasticities of substitution are equal to the cross-price elasticities of substitution divided by the relevant cost share: $\alpha_{ij} = \left(1/S_j\right)\eta_{ij}$.

contained in the cross-price elasticities and in fact obscure that information by dividing the cross-price elasticity by a cost share.

Following Pindyck (1979), the total own and cross-price elasticities for fuel *i* with respect to the price of fuel *j* are given by:

$$\eta_{ij}^{T} = \eta_{ij} + \eta_{EE}\hat{S}_j \tag{17}$$

where η_{EE} is the own price elasticity of aggregate energy and $\hat{S}_j$ is the fitted cost share for fuel *j*. To compute the elasticity at the reference point we again substitute the estimate of β_j for $\hat{S}_j$. This only makes a difference for the own and cross-price elasticities and not the Morishima and shadow elasticities of substitution. To compute the total elasticities we will assume that the energy own price elasticity is a fixed known parameter.

3. Data

Our dataset consists of provincial level data for China for real provincial GDP, the quantities of the final use of seven individual fuels and electricity, capital, and labor, price series for capital and labor, and energy price series from the provincial capitals, which we use to proxy provincial energy prices. The data cover all provinces, province level municipalities, and autonomous regions of the People's Republic of China except Tibet for the years 2000 to 2010. Therefore, the panel has a cross-section dimension of 30 provinces and a time series dimension of eleven years. Table 1 provides summary statistics for these raw variables. This shows the extraordinary growth rates of many variables in China over this period. The exceptionally high growth rate for coal briquettes reflects growth from a very small base in 2000 when several provinces did not report any use of this fuel. The growth rate for China as a whole was 32% per annum over the 11 years.

We use three factor inputs: aggregate energy use, capital services, and labor use for the interfactor substitution analysis and four energy inputs: aggregated coal, gasoline, diesels and electricity for the interfuel substitution analysis. The energy data include five types of coal products, diesel, gasoline, and electricity. For the interfuel substitution analysis we aggregated the five types of coal into a single coal input. For the interfactor analysis we aggregated the eight types of fuel into a single energy input. The between estimator depends on variation in relative prices across the provinces. Therefore, we cannot set the base year price or quantity index in each province to an arbitrary level as is often done in cost function analyses. Also, if we simply compute the average cost of energy per Joule in each province then the price will depend on the mix of fuels. High quality fuels such as electricity have much higher prices than low quality fuels such as coal. Therefore, we need to use index number methods to compute proper price indices for coal and energy for each province in each year.

To do this, we first computed Laspeyres price indices for each province in each year, using the national average price of coal or energy in RMB per Joule as the base price. The Laspeyres index evaluates the price level in each province using the national average quantities as weights. The energy or coal price index in province i in year t is given by:

$$P_{it} = \frac{\sum_j p_{ijt} q_{Njt}}{\sum_j p_{Njt} q_{Njt}} \tag{18}$$

where j indexes the different fuels or different forms of coal, p is the price of each fuel and q its quantity, and the N subscript indicates the national average price for that fuel. This gives a relative energy or coal price index for each province in each year but these prices cannot be compared across years. This can be addressed by multiplying these relative price indices by a national price index time series to obtain a time series in each province in each year. We use the Divisia index to compute the national level price index.

The individual fuel consumption data are obtained from the China Energy Yearbooks (CEY). The CEY provides detailed data on final consumption of different fuel types by sector, province and year. Energy used as intermediate inputs, such as coal used to generate electricity and heat and to produce coke, is excluded. This study covers eight fuel types: steam coal, coking coal, coke, briquettes, coal gas, gasoline, diesel and electricity. We aggregate final consumption of each fuel type from the five sectors: “Farming, Forestry, Animal Husbandry, Fishery and Water Conservancy”, “Industry”, “Construction”, “Transport, Storage and Post” and “Wholesale, Retail Trade, Hotel and Catering”. Rural and urban residential consumption and fuel use for non-energy purposes are excluded.

We collect fuel price data from two sources. China’s National Development and Reform Commission (NDRC) collects perhaps the most authoritative commodity price data from 36 large cities including all provincial capitals. We use fuel prices collected in the provincial capital cities as an indicator of provincial prices. As price collection is conducted by the NDRC three times a month, we have constructed an annual price series by taking the average of these price data for each province. *China Price Statistical Yearbooks* (*CPSY*s) provide the second source of data. The *CPSY*s reported provincial prices on twelve fuels including seven used in this study. Both the NDRC data and the *CPSY* data have missing values. Most missing values in the NDRC data appeared in early years while the *CPSY* ceased to report fuel prices after 2006. Where the value is missing in the NDRC data but available in the *CPSY* data, we use the value reported in the *CPSY* data to replace the missing value in the NDRC data. Remaining missing values in the NDRC data are then linearly interpolated.

Total employment data, which include all employed persons in urban and rural sectors are obtained from the *China Statistical Yearbooks* (*CSY*). The *CSY* also provides an income approach decomposition of GDP as the sum of compensation of employees, net taxes of production, depreciation of fixed assets, and operating surplus. Compensation of employees includes the total payment of various forms to employees for the productive activities they are engaged in. It includes wages, bonuses and allowances, which the employees earn in cash or in kind. It also includes the free medical services provided to the employees and medicine expenses, transport subsidies and social insurance, and housing funding paid by the employers. Compared to wages, which were used in H. Ma *et al*. (2008), this is clearly a superior indicator for actual compensation of labor. Using total employment and total compensation, we can construct the price index for labor.

The sum of the other three components of the decomposed GDP - net taxes of production, depreciation of fixed assets, and operating surplus - effectively gives the compensation of capital. Unfortunately, the *CSY* does not provide capital stock statistics. The capital stock series are taken from Wu (2009) with updated statistics obtained from the author. Using total compensation of capital and the capital stock series, we have constructed the price index for capital. Our measure of total compensation of capital is also superior to that used by H. Ma *et al*. (2008), who simply used the product of the capital stock and the price index of fixed assets reported in the *CSY*.

Figures 1 to 3 show the variation in some of our constructed variables across China. In 2010 the coal price index varies by roughly a factor of two across provinces (Figure 1). In general, prices are higher in the coastal provinces and cheapest in the mid-western and northwestern provinces – Chongqing, Shaanxi, Ningxia, Nei Mongol, and Xinjiang. An anomaly is Shanxi, which has relatively high prices despite being a major coal-producing province, though this wasn't the case in earlier years.

Distance from the technological frontier varies by a factor of roughly three. Shanghai is the most productive province and Guangdong and Tianjin and other coastal provinces are relatively close to the frontier. The further from the frontier are the western provinces of Qinghai, Ningxia, and Guizhou. Figure 3 shows that there has been some convergence in TFP over time. The slowest TFP growth was in Beijing and the highest in Sichuan and Hainan.

4. Econometric Results

We use iterative seemingly unrelated regressions to estimate the fitted prices of the aggregated coal input for use as instruments in the interfuel substitution model and the fitted prices of the aggregated energy input for the interfactor substitution model.[9] Then we estimate the BE model (3) and (4) simultaneously using iterative 3-stage least squares regression imposing symmetry and homogeneity restrictions for both the interfuel and interfactor cost function systems.[10] We estimate the IDE model (5) and (6) in the same way. The diesel and labor prices are treated as the numeraire in the interfuel and interfactor estimations and we drop the corresponding cost share equations. We retrieve the parameters for these variables using the homogeneity restrictions and compute their standard errors using the delta method.[11] We also test the concavity of the cost function at every data point and impose concavity using the method of Ryan and Wales (2000) if concavity is violated.

4.1 Interfuel Substitution

We first estimated the unconstrained interfuel substitution model using BE and found that the estimated energy cost function violates the concavity condition at all sample observations. The energy cost function estimated using unconstrained IDE is only concave at 58 out of the total 330 sample observations. Therefore, we impose concavity. Table 2 presents the results from the constrained BE and IDE estimation. Table 3 provides the implied own and cross price elasticities and Table 4 the Morishima and shadow elasticities of substitution. Figures 4 and 5 illustrate the results graphically.

---

[9] To save space, these first-stage results are not reported but available upon request.
[10] We use the procedure NLSYSTEM in RATS. This estimates the model using the generalized method of moments and an optimal weighting matrix (Estima, 2010). Hence there is no need to request the program to compute robust standard errors.
[11] We use the procedure SUMMARIZE in RATS.

The parameters estimated using the two different methods are remarkably similar – the main differences are in the technical change biases, which are not subject to cross-equation restrictions in the BE estimates (Table 2). Both methods find coal-using technical change, but IDE finds gasoline saving and BE electricity saving technical change, with the other biases statistically insignificant. It is particularly remarkable that the estimated β_i's are very similar for the two estimators when IDE does not explicitly utilize cross-equation restrictions to estimate them. However, the concavity restrictions involve these parameters. The fixed effects estimates differ substantially in several cases from the BE and IDE estimates but are similar in most cases.

A seemingly surprising result is that the average rate of cost reduction, *ceteris paribus*, given by the IDE estimates is 1% per annum. We would have assumed that it is hard to make technological progress in composing an energy aggregate from different fuels but this is apparently the case.

The BE and IDE estimates of own- and cross-price elasticities are similar, but the IDE estimates are consistently greater in absolute value (Table 3, Figure 4). For IDE, all the own-price elasticities are negative and statistically significant - gasoline demand is elastic and the demand for the other fuels inelastic especially that of coal. Among the significant cross-price elasticities, coal and electricity and diesel and electricity are complements and gasoline and diesel substitutes as would be expected. By contrast, the fixed effects estimates of elasticities are all very small. The reason for this is that the eigenvalues of the Hessian of the cost function are all positive for the unconstrained model. When we impose concavity on the model the elasticities are all forced towards zero in order to meet the restrictions.

Moving on to the elasticities of substitution (Table 4, Figure 5), we make the following observations:

- As expected given the parameter estimates, the elasticities estimated by the two estimators are similar but the BE elasticities are mostly smaller in absolute value and none of them is statistically significant.

- Shadow elasticities of substitution between fuels are less than unity (indicating poor substitutability) except for gasoline and diesel, which are good substitutes. The

Morishima elasticities show that gasoline is a good substitute for coal and electricity if the gasoline price changes but not otherwise.

- The substitution elasticity between coal and electricity is very close to zero.

- In general, elasticities involving electricity appear to be smaller than other elasticities.

The elasticities of substitution estimated using the fixed effects estimator are all indistinguishable from zero.

4.2 Interfactor Substitution

Again, we first estimated the cost function systems without imposing the concavity constraints. Using BE, the estimated cost function satisfies the concavity condition for most sample observations and is concave at the reference point. Using IDE, the estimated cost function is concave for all sample observations. Therefore, we did not impose concavity on the interfactor substitution models. Table 5 presents the results from the unconstrained BE and IDE estimation. Table 6 provides the implied own and cross price elasticities and Morishima and shadow elasticities of substitution.

The BE and IDE parameter estimates differ more than the interfuel parameter estimates do (Table 5).[12] BE finds no statistically significant technical change biases but some indication of energy-using and labor-saving biases, while IDE finds capital-using and labor-saving technical change. Fixed effects estimates also differ by a similar degree. Because of the much larger number of degrees of freedom all the FE parameter estimates are highly statistically significant.

The BE and IDE estimates of the own- and cross-price elasticities (Table 6, Figure 4) are fairly different with no apparent pattern to the differences. All the own price elasticities are inelastic. The BE estimates finds that energy and capital are both substitutes for labor. The IDE estimates find that energy and capital are also substitutes. The FE estimates are mostly smaller or much smaller in absolute value.

[12] Wondering whether this is because there are no cross-equation restrictions on the β_i's and no concavity restrictions to impose them implicitly we restricted the β_i's to be equal to the mean cost shares. This did not change the other parameters by very much.

Moving onto the elasticities of substitution (Table 6, Figure 5), we have four key observations:

- All Morishima and shadow elasticities of substitution between factors are significant and less than unity for the IDE estimator.

- BE finds that capital and energy are poor substitutes with an elasticity of substitution that is insignificantly different from zero and finds that energy and labor and capital and labor are good substitutes, though only when it is the price of labor which changes.

- For both estimators substitution elasticities are higher between energy and labor than between energy and capital.

All the FE estimates of elasticities of substitution are smaller than their BE and IDE counterparts.

5. Comparison with Existing Literature

5.1 Interfuel Substitution

The elasticities estimated in this study are mostly smaller than the elasticities for the United States estimated using meta-analysis (Stern, 2012). The latter are all greater than unity with the exception of the coal-electricity elasticity (0.176).

But how do our estimates compare to previous estimates for China? Stern (2012)'s meta-analysis included four studies: Fisher-Vanden *et al*. (2004), Hang and Tu (2007), H. Ma *et al*. (2008), and H. Ma *et al*. (2009). Two more recent studies are Serletis *et al*. (2011) who investigate interfuel substitution for a number of countries including China and Smyth *et al*. (2012) who examine interfuel substitution in the Chinese iron and steel sector. We summarize the results of these studies in Table 7 as shadow elasticities. We computed the shadow elasticities as described in Stern (2012). We use long-run Morishima elasticities from Serletis *et al*. (2011) and weight them using the average cost shares from the current study. The averages of the seven regions in H. Ma *et al*. (2009) are very close to the national estimates and so we do not report results from that study. We used the most recent estimates of Hang and Tu (2007) and Smyth *et al*. (2012). As they estimated a positive own price elasticity for electricity, their oil-electricity elasticity is negative in violation of theory. Data types and methods of estimation differ across the studies. Fisher-Vanden *et al*. (2004) is a firm-level cross-sectional analysis. Hang and Tu (2007) use a national level time series and unconstrained logarithmic demand curves. H. Ma *et al*. (2008) also use a panel data set of seven Chinese regions and fixed effects estimation using the translog cost function. Smyth *et al*. (2012) estimate a translog production function using ridge regression for a time series from the iron and steel industry. Finally, Serletis *et al*. (2011) use a normalized quadratic cost function system with concavity imposed and fixed effects for a panel of data for China, India, South Africa, and Thailand.

The most directly comparable study to ours should be that of H. Ma *et al*. (2008, 2009) who use an earlier version of the same dataset as us. However, the shadow elasticities from that study are almost inversely related to those from our study. In particular, their highest elasticity is for coal-electricity substitution which is our lowest elasticity and our highest elasticity which is for gasoline-diesel substitution is their second lowest. Our results seem more plausible as one would expect that it is easier to substitute between gasoline and diesel than between coal and electricity. The coal-electricity elasticity is the smallest in Stern's (2012) meta-analysis. Apart from the difference in estimators – fixed effects vs. specifically long-run estimators – we impose concavity, which is one reason why even our fixed effects estimates are very different - and our data covers the period 2000-2010 whereas they used the period 1995-2004. We also include eight different fuel inputs while they only included four and there are other improvements in our data as mentioned above.

Smyth *et al*.'s (2012) results are likely driven by using ridge regression with a single production function equation. Our results are more similar to those of Serletis *et al*. (2011) Fisher-Vanden *et al*. (2004), and Hang and Tu (2007). All these studies find a larger value for the oil-coal elasticity than for the coal-electricity elasticity. The elasticities found by both Serletis *et al*. (2011) and this study are smaller than those found by Fisher-Vanden *et al*. (2004). This is likely because the latter uses firm level intra-industry data and the former use macro-level data. Stern (2012) finds that macro-level elasticities are smaller than elasticities for disaggregated data.

We also compare our results to a few single equation estimates of demand elasticities that use panels of Chinese provincial data. Burke and Liao (2014) find that the coal own-price elasticity in China has increased over time. Using a panel of provincial data from 1998 to 2012 they obtain a "two-year elasticity" of -0.14 for a model that does not allow the elasticity to evolve over time. This is very close to our estimates. Allowing evolution over time, they estimate an elasticity of -0.3 for 2005 (personal communication). Cao and Xie (2011) estimate that the long-run elasticity of demand for diesel is -0.86 between 1999 and 2007. This is very close to our IDE estimate of -0.88.

5.2 Interfactor Substitution

Koetse *et al*. (2008) conducted a meta-analysis of 34 studies of capital-energy substitution. Most studies included in this meta-analysis used data from before 1990 and there were no studies that included China. In fact, there are very few studies on interfactor substitution in China. Koetse *et al*. (2008) found that the short-run (time series) elasticities are around 0.4 for North America and 0.15 for Europe. The long-run (cross-section) elasticities are slightly greater than unity in North America and around 0.8 for Europe. Our BE elasticity is obviously much smaller than these long-run elasticities (and insignificantly different from zero), while our IDE estimate is close to estimates for Europe and to Stern and Kander's (2012) estimate for Sweden. The consensus is that the capital-labor elasticity is less than unity for the United States (Acemoglu, 2003; Klump *et al*., 2007). Here our IDE estimate is again closest to expectations for developed countries.

Table 8 reports the available estimates for China. We use the cost shares from our study to weight the Morishima elasticities from other studies. We use Fan *et al*.'s (2007) estimate for 1993-2003 and Smyth *et al*.'s (2011) estimates for 2007. Again a variety of methods and data were used to produce these results. Ma *et al*. (2008) estimate a translog cost function and share equations for a panel of seven Chinese regions, Fan *et al*. (2007) used only the cost share equations and a time series of national data, Mallick estimates a two input CES production function for a national time series, and Smyth *et al*. (2011) estimate a translog production function using ridge regression for a time series from the iron and steel industry.

Our results are very different to Fan *et al*. (2007) and Smyth *et al*. (2011) who use less sophisticated methods. Our IDE capital-labor elasticity of substitution is almost identical to Mallick's (2012). Compared to H. Ma *et al*. (2008), our energy-capital elasticity of substitution is smaller and our energy-labor and capital-labor elasticities of substitution are higher. The IDE estimates are closer to their estimates than the BE elasticities. There are a number of differences between our study and H. Ma *et al*.'s. We use a broader coverage of fuel inputs and a dataset covering a different period. More importantly, using wages will underestimate the actual compensation for labor and their studies also seemed to have used incorrect measures of capital services, which will undoubtedly overestimate the actual compensation for capital. We employ the income approach to GDP, which provides more appropriate measures of labor and capital compensation.

6. Conclusions

In this paper, we use between and long-run difference estimators to provide estimates of interfuel (coal, gasoline, diesel, and electricity) and interfactor (capital, labor, and energy) elasticities of substitution for China. We also improve on previous studies by adding total factor productivity terms to our regressions. The results show that demand for coal and electricity in China is very inelastic, while demand for diesel and gasoline is elastic. There are limited substitution possibilities among the fuels with the exception of gasoline and diesel. The elasticity of substitution between electricity and coal is particularly low. Substitution possibilities are greater between energy and labor than between capital and labor. The results are quite different to some previous studies for China but coincide well with the patterns in meta-analyses for long-run estimates of elasticities of substitution. Interfuel and capital-energy elasticities of substitution are generally lower than estimates for the United States derived from meta-analyses (Koetse *et al*., 2007; Stern, 2012). These results need to be taken into account in CGE models used for assessing climate policy options in China. The marginal cost of abatement for a given reduction in emissions will be higher than would be predicted assuming that Chinese elasticities are the same as US elasticities as some models assume (Lu and Stern, 2014).

Our findings potentially have significant policy implications at a time when the Chinese government is under mounting domestic and international pressure to reduce greenhouse gases emissions from burning fossil fuels and particularly coal. Some previous studies (e.g. H. Ma *et al*., 2008) are generally optimistic about the degree of difficulty in substituting cleaner energy for dirty coal. Our very low estimate of the elasticity of substitution between coal and electricity in final energy consumption suggests that replacing coal with renewably generated electricity in end-use applications will be costly. However, because our data are for final energy consumption, no inferences can be drawn from our study with regard to substitution between coal and cleaner sources including natural gas, wind and solar in the generation of electricity. Some may be optimistic about this substitution given the measures taken by the Chinese government in recent years to promote the penetration of cleaner energy sources and renewable energy in particular. However, this substitution potential may be limited by the fact that China still accounts for half of the annual global coal consumption and half of that is consumed in the electricity sector. The government has implemented a number of retrofitting mandates to close down small and old generation units and replace them with new and large coal-fired units. This massive fleet of newly installed coal-fired units will lock the Chinese economy into a coal-dominated energy supply for the next two to three decades.

The BE and IDE results are much more similar for the interfuel substitution system where concavity was imposed than for the interfactor substitution system where we did not need to impose concavity. It is unclear which of the two estimators should be preferred when they do diverge. As IDE imposes cross-equation restrictions on the technical change biases it probably produces more consistent estimates of these. Further research is needed on the performance of the estimators under alternative restrictions and conditions. However, we think that both estimators show potential for estimation of long-run elasticities of substitution. Our FE estimates of the interfuel cost share system had very poor curvature properties and when we imposed concavity all the interfuel elasticities of substitution effectively became zero. FE estimates of the interfactor system had adequate curvature properties but all the elasticities of substitution are smaller than their BE and IDE counterparts. This suggests that IDE and BE do in fact capture long-run substitution possibilities and FE short-run substitution possibilities as we predicted.

References

Acemoglu, Daron (2003) Labor- and capital-augmenting technical change, *Journal of the European Economic Association* 1(1): 1-37.

Apostolakis, Bobby E. (1990) Interfuel and energy-capital complementarity in manufacturing-industries, *Applied Energy* 35(2): 83-107.

Bacon, Robert (1992) Measuring the possibilities of interfuel substitution, *Policy Research Working Paper Series* 1031. The World Bank: Washington, DC.

Baltagi, Badi H. (2008) *Econometric Analysis of Panel Data*, 4th edition. Chichester: John Wiley & Sons.

Baltagi, Badi H. and James M. Griffin (1984) Short and long run effects in pooled models, *International Economic Review* 25(3): 631-645.

Bhattacharya, Subhes C. (1996) Applied general equilibrium models for energy studies, *Energy Economics* 18: 145-164.

Burke, Paul J. and Hua Liao (2014) *The Price Elasticity of Coal Demand in China is Increasing*. Manuscript, Arndt-Corden Department of Economics, Crawford School of Public Policy, The Australian National University, Canberra.

Cao, Jing and Yang Xie (2011) Diesel demand per capita elasticities in China, 1999-2007, *Qinghua Daxue Xuebao/Journal of Tsinghua University* 51(2): 241-244.

Egger, Peter and Michael Pfaffermayr (2004) Estimating long and short run effects in static panel models, *Econometric Reviews* 23(3): 199–214.

Estima (2010) *RATS Version 8: User's Guide*. Evanston IL: Estima.

Fan, Ying, Hua Liao, and Yi-Ming Wei (2007) Can market oriented economic reforms contribute to energy efficiency improvement? Evidence from China, *Energy Policy* 35: 2287-2295.

Feenstra, Robert C., Inklaar, Robert, Timmer, Marcel P., 2013. The next generation of the Penn World Table. *NBER Working Papers* 19255.

Fisher-Vanden, Karen, Gary H. Jefferson, Hongmei Liu, and Quan Tao (2004) What is driving China's decline in energy intensity? *Resource and Energy Economics* 26: 77–97.

Goettle, Richard J., Mun S. Ho, Dale W. Jorgenson, Daniel T. Slesnick, and Peter J. Wilcoxen (2007) *IGEM, an Inter-temporal General Equilibrium Model of the U.S. Economy with Emphasis on Growth, Energy and the Environment*, Prepared for the U.S. Environmental Protection Agency (EPA), Office of Atmospheric Programs, Climate Change Division.

Griliches, Zvi and Jacques Mairesse (1987) Productivity and R&D at the firm level. In: Griliches, Z. (ed.) *R & D, Patents, and Productivity*. University of Chicago Press.

Hang, Leiming and Meizeng Tu, (2007) The impacts of energy prices on energy intensity: evidence from China, *Energy Policy* 35(5): 2978–2988.

Hausman, Jerry A. (2001) Mismeasured variables in econometric analysis: problems from the right and problems from the left, *Journal of Economic Perspectives* 15(4): 57–67.

Hauk, William R. and Romain Wacziarg (2009) A Monte Carlo study of growth regressions, *Journal of Economic Growth* 14: 103-147.

Hsieh, Chang-Tai (2002) What explains the industrial revolution in East Asia? Evidence from the factor markets, *American Economic Review* 92(3): 502-526.

Klump, Rainier, Peter McAdam, and Alpo Willman (2007) Factor substitution and factor-augmenting technical progress in the united states: a normalized supply-side system approach, *Review of Economics and Statistics* 89(1): 183-192.

Koetse, Mark J., Henri L. F. de Groot, and Raymond J. G. M. Florax (2008) Capital-energy substitution and shifts in factor demand: A meta-analysis, *Energy Economics* 30: 2236–2251.

Leon-Ledesma, Miguel A., Peter McAdam, and Alpo Willman (2010) Identifying the elasticity of substitution with biased technical change, *American Economic Review* 100: 1330-1357.

Lu Yingying and David I. Stern (2014) Substitutability and the cost of climate mitigation policy, *CAMA Working Papers* 28/2014.

Ma, Chunbo and David I. Stern (2008) China's changing energy intensity trend: a decomposition analysis, *Energy Economics* 30(3): 1037-1053.

Ma, Hengyun, Les Oxley, and John Gibson (2009) Substitution possibilities and determinants of energy intensity for China, *Energy Policy* 37, 1793–1804.

Ma, Hengyun, L. Oxley, John Gibson, and Bongguen Kim (2008) China's energy economy: Technical change, factor demand and interfactor/interfuel substitution, *Energy Economics* 30: 2167–2183.

Mairesse, Jacques (1990) Time-series and cross-sectional estimates on panel data: Why are they different and why should they be equal? In J. Hartog, G. Ridder, J. Theeuwes (eds.) *Panel Data and Labor Market Studies*. Contributions to Economic Analysis No. 192, North-Holland, Amsterdam.

Mallick, Debdulai (2012) The role of the elasticity of substitution in economic growth: A cross-country investigation, *Labour Economics* 19: 682-694.

McKibbin, Warwick J. and Peter J. Wilcoxen (1999) The theoretical and empirical structure of the G-Cubed model, *Economic Modelling* 16: 123-148.

Mundlak, Yair (1968) Elasticities of substitution and the theory of derived demand, *Review of Economic Studies* 35: 225-236.

Pesaran, M. Hashem. and Ron Smith (1995) Estimating long-run relationships from dynamic heterogeneous panels, *Journal of Econometrics* 68(1): 79-113.

Pindyck, Robert S. (1979) Interfuel Substitution and the industrial demand for energy: an international comparison, *Review of Economics and Statistics* 61(2): 169-179.

Pirotte, Alain (1999) Convergence of the static estimation toward the long run effects of dynamic panel data models, *Economics Letters* 63(2): 151-158.

Ryan, David L. and Terrence J. Wales (2000) Imposing local concavity in the translog and generalized Leontief cost functions, *Economics Letters* 67: 253–260.

Serletis, Apotolos, Govinda Timilsina, and Olexandr Vasetsky (2010) Interfuel substitution in the United States, *Energy Economics* 32: 737-745.

Serletis, Apotolos, Govinda Timilsina, and Olexandr Vasetsky (2011) International evidence on aggregate short-run and long-run interfuel substitution, *Energy Economics* 33: 209-216.

Sharma, Subhash C. (2002) The Morishima elasticity of substitution for the variable profit function and the demand for imports in the United States, *International Economic Review* 43(1): 115-135.

Shephard, Ronald (1953) *Theory of Cost and Production Functions*, Princeton University Press.

Smyth, Russell, Paresh K. Narayan, and Hongliang Shi (2012) Inter-fuel substitution in the Chinese iron and steel sector, *International Journal of Production Economics* 139: 525-532.

Smyth, Russell, Paresh K. Narayan, and Hongliang Shi (2011) Substitution between energy and classical factor inputs in the Chinese steel sector, *Applied Energy*, 88(1): 361-367.

Stern David I. (2011) Elasticities of substitution and complementarity, *Journal of Productivity Analysis* 36(1): 79-89.

Stern, David I. (2012) Interfuel substitution: A meta-analysis, *Journal of Economic Surveys* 26(2): 307-331.

Stern, David I. and Astrid Kander (2012) The role of energy in the industrial revolution and modern economic growth, *Energy Journal* 33(3): 125-152.

Wu, Yanrui R. (2009) China’s capital stock series by region and sector, *Discussion Paper* 09.02, Business School, University of Western Australia.

Table 1. Summary Statistics 2000-2010

Variables	Unit	Provincial Mean		Annual Provincial Growth Rate	
		Mean†	S.D. †	Mean	S.D.
Steam Coal Price	*yuan/tonne*	333	111	9%	3%
Coking Coal Price	*yuan/tonne*	547	130	13%	3%
Briquette Price	*yuan/100 kg*	32.9	9.62	8%	4%
Coke Price	*yuan/tonne*	937	182	13%	2%
Coal Gas Price	*yuan/m*^{3}	1.27	0.29	2%	3%
Gasoline Price	*yuan/tonne*	5518	182	10%	1%
Diesel Price	*yuan/tonne*	4730	89.8	9%	1%
Electricity Price	*yuan/kwh*	0.60	0.09	6%	4%
Steam Coal Quantity	*mil. tonnes*	23.8	16.0	6%	4%
Coking Coal Quantity	*1000 tonnes*	952	1108	18%	23%
Briquette Quantity	*1000 tonnes*	226	309	42%	23%
Coke Quantity	*mil. tonnes*	6.9	8.0	13%	12%
Coal Gas Quantity	*bil. m*^{3}	5.41	6.78	1%	23%
Gasoline Quantity	*mil. tonnes*	1.73	1.19	7%	5%
Diesel Quantity	*mil. tonnes*	3.31	2.33	14%	6%
Electricity Quantity	*bil. kwh*	72.5	52.2	12%	3%
Capital††	*bil. yuan*	2425	1698	18%	3%
Labor	*mil. persons*	22.8	15.2	2%	1%
Capital Compensation†††	*bil. yuan*	410	341	17%	3%
Labor Compensation†††	*bil. yuan*	341	263	15%	2%

Notes: †The provincial mean is the mean over time of the variable in a province. Then the mean and S.D. reported here are the mean and standard deviation of those values ††in constant 2000 yuan; †††the sum of the two compensation items is equal to GDP; *mil*. and *bil*. indicate million and billion.

Table 2. Fuel Cost Function Parameter Estimates

Parameter	Between Estimator		Long-Run Difference Estimator		Fixed Effects	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
β_0 or $D(f_t)$	-0.0090	0.0001***	-0.0099	0.0001***		
β_C	0.2397	0.0004***	0.2272	0.0027***	0.2358	0.0015***
β_G	0.1081	0.0029***	0.1038	0.0047***	0.1080	0.0010***
β_X	0.4806	0.0005***	0.4877	0.0023***	0.4811	0.0011***
β_D	0.1715	0.0029***	0.1813	0.0050***	0.1751	0.0013***
β_{CC}	0.1810	0.0038***	0.1667	0.0143***	0.1802	0.0009***
β_{GC}	-0.0325	0.0199	-0.0280	0.0297	-0.0255	0.0049***
β_{GG}	-0.0237	0.1015	-0.0763	0.0751	0.0956	00127***
β_{XC}	-0.1147	0.0029***	-0.1007	0.0116***	-0.1134	0.0045***
β_{XG}	-0.0430	0.0120***	-0.0298	0.0088***	-0.0513	0.0058***
β_{XX}	0.2490	0.0020***	0.2369	0.0062	0.2491	0.0041***
β_{DC}	-0.0337	0.0190*	-0.0380	0.0302	-0.0413	0.0011***
β_{DG}	0.0992	0.0988	0.1340	0.0705*	-0.0188	0.0076***
β_{DX}	-0.0912	0.0110*	-0.1064	0.0127***	-0.0843	0.0061***
β_{DD}	0.0258	0.0983	0.0105	0.0740	0.1444	0.024***
γ_C	0.0031	0.0003***	0.0022	0.0006***	0.0036	0.0003***
γ_G	0.0011	0.0029	-0.0032	0.0014**	-0.0037	0.0012***
γ_X	-0.0021	0.0003***	0.0005	0.0004	-0.0015	0.0003***
γ_D	-0.0021	0.0029	0.0004	0.3806	0.0001	0.0011

Note: $C, G, X,$ *and* D denote coal, gasoline, electricity, and diesel respectively. For BE first parameter is β_0, while for IDE it is $D(f_t)$ - the average rate of cost reduction. Significance levels: * 10%, ** 5%, *** 1%.

Table 3: Fuel Total Own and Cross-Price Elasticities

Elasticity	Between Estimator		Long-Run Difference Estimator		Fixed Effects	
	Estimate	Standard Errors	Estimate	Standard Errors	Estimate	Standard Errors
η^T_{CC}	-0.076	0.015***	-0.193	0.068***	-0.014	0.003***
η^T_{CG}	-0.060	0.083	-0.090	0.131	-0.007	0.021
η^T_{CX}	-0.141	0.012***	-0.285	0.054***	-0.029	0.020
η^T_{CD}	-0.020	0.079	-0.109	0.135	-0.011	0.004**
η^T_{GC}	-0.132	0.184	-0.196	0.291	-0.015	0.045
η^T_{GG}	-1.143	0.938	-1.701	0.727**	-0.013	0.118
η^T_{GX}	-0.060	0.115	-0.129	0.088	-0.023	0.055
η^T_{GD}	1.038	0.907	1.350	0.699*	-0.010	0.070
η^T_{XC}	-0.070	0.006***	-0.133	0.024***	-0.014	0.010
η^T_{XG}	-0.013	0.025	-0.027	0.018	-0.005	0.012
η^T_{XX}	-0.144	0.004***	-0.357	0.012***	-0.030	0.009***
η^T_{XD}	-0.069	0.024***	-0.160	0.027***	-0.011	0.013
η^T_{DC}	-0.028	0.111	-0.136	0.168	-0.014	0.006**
η^T_{DG}	0.654	0.578	0.772	0.383**	-0.006	0.043
η^T_{DX}	-0.194	0.068***	-0.429	0.072***	-0.030	0.035
η^T_{DD}	-0.729	0.573	-0.884	0.408**	-0.011	0.012

Note: C, G, X, and D denote coal, gasoline, electricity and diesel respectively. First subscript is the quantity and second subscript is the price. Significance levels: * 10%, ** 5%, *** 1%.

Table 4: Interfuel Morishima and Shadow Elasticities of Interfuel Substitution

Elasticity	Between Estimator		Long-Run Difference Estimator		Fixed Effects	
	Estimate	Standard Errors	Estimate	Standard Errors	Estimate	Standard Errors
Morishima						
μ_{CG}	-0.056	0.186	-0.003	0.250	-0.0004	0.0421
μ_{GC}	1.083	0.965	1.612	0.765**	0.0060	0.1258
μ_{CX}	0.006	0.017	0.060	0.091	0.0001	0.0125
μ_{XC}	0.003	0.012	0.071	0.062	0.0013	0.0235
μ_{CD}	0.048	0.112	0.057	0.224	0.0001	0.0084
μ_{DC}	0.709	0.573	0.775	0.489	0.0001	0.0153
μ_{GX}	1.129	0.934	1.674	0.733**	0.0075	0.1299
μ_{XG}	0.084	0.118	0.228	0.089**	0.0068	0.0571
μ_{GD}	1.797	1.511	2.474	1.093**	0.0068	0.1546
μ_{DG}	1.767	1.476	2.233	1.096**	0.0010	0.0820
μ_{XD}	-0.05	0.064	-0.073	0.072	0.0006	0.0380
μ_{DX}	0.66	0.576	0.724	0.393*	-0.0000	0.0001
Shadow						
σ_{CG}	0.298	0.368	0.503	0.356	0.0016	0.0580
σ_{CX}	0.004	0.013	0.068	0.071	0.0009	0.0197
σ_{CD}	0.324	0.250	0.376	0.319	0.0001	0.0106
σ_{GX}	0.276	0.190	0.481	0.167***	0.0069	0.0686
σ_{GD}	1.779	1.485	2.321	1.074**	0.0032	0.1062
σ_{XD}	0.137	0.166	0.143	0.097	0.0004	0.0280

Note: C, G, X, and D denote coal, gasoline, electricity and diesel respectively. For cross-price elasticities, the first subscript is the quantity and second subscript is the price. For Morishima elasticities the first subscript is the price that changes. Significance levels: * 10%, ** 5%, *** 1%.

Table 5: Factor Cost Function Parameter Estimates

Parameter	Between Estimator		Long-Run Difference Estimator		Fixed Effects	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
β_0	-0.5661	0.0053***				
β_E	0.1268	0.0054***	0.2179	0.0750***	0.1235	0.0006***
β_K	0.4763	0.0053***	0.4171	0.0401***	0.4686	0.0011***
β_L	0.3968	0.0039***	0.3650	0.0363***	0.4080	0.0012***
β_{EE}	0.0687	0.0340**	0.0174	0.0494	0.1008	0.0090***
β_{KE}	-0.0863	0.0295***	-0.0388	0.0218*	-0.0608	0.0053***
β_{KK}	0.0569	0.0309*	0.1140	0.0298***	0.1480	0.0063***
β_{LE}	0.0175	0.0126	0.0214	0.0358	-0.0399	0.0063***
β_{LK}	0.0294	0.0133**	-0.0752	0.0285***	-0.0871	0.0053***
β_{LL}	-0.0469	0.0101***	0.0538	0.0367	0.1271	0.0067***
γ_E	0.0144	0.0093	-0.0012	0.0022	0.0082	0.0010***
γ_K	-0.0039	0.0058	0.0138	0.0034***	0.0092	0.0010***
γ_L	-0.0105	0.0066	-0.0126	0.0035***	-0.0174	0.0007***

Note: *E*, *K, and L* denote capital, energy, and labor respectively. Significance levels: * 10%, ** 5%, *** 1%.

Table 6: Factor Own and Cross-Price Elasticities, Morishima, and Shadow Elasticities of Substitution

Elasticity	Between Estimator		Long-Run Difference Estimator		Fixed Effects	
	Estimate	Standard Errors	Estimate	Standard Errors	Estimate	Standard Errors
Own & Cross-price						
η_{EE}	-0.331	0.266	-0.702	0.235***	-0.060	0.073
η_{EK}	-0.204	0.232	0.239	0.100**	-0.024	0.043
η_{EL}	0.535	0.102***	0.463	0.180***	0.084	0.051*
η_{KE}	-0.054	0.062	0.125	0.079	-0.006	0.011
η_{KK}	-0.404	0.065***	-0.310	0.070***	-0.216	0.013***
η_{KL}	0.629	0.105***	0.020	0.147	-0.298	0.043***
η_{LE}	0.171	0.031***	0.277	0.126**	0.026	0.015*
η_{LK}	0.550	0.032***	0.211	0.104**	0.255	0.013***
η_{LL}	-0.721	0.023***	-0.488	0.100***	-0.281	0.016***
Morishima						
μ_{EK}	0.277	0.324	0.827	0.268***	0.054	0.081
μ_{KE}	0.200	0.292	0.548	0.144***	0.192	0.052***
μ_{EL}	0.502	0.282*	0.979	0.319***	0.086	0.086
μ_{LE}	1.256	0.112***	0.951	0.252***	0.365	0.063***
μ_{KL}	0.955	0.082***	0.521	0.153***	0.471	0.024***
μ_{LK}	1.350	0.118***	0.508	0.209**	-0.017	0.053
Shadow						
σ_{EK}	0.217	0.295	0.644	0.174***	0.163	0.055***
σ_{EL}	1.074	0.138***	0.961	0.267***	0.300	0.066***
σ_{KL}	1.134	0.090***	0.515	0.159***	0.244	0.036***

Note: *E*, *K* and *L* denote capital, energy and labor respectively. For Morishima elasticities the first subscript is the price that changes. Significance levels: * 10%, ** 5%, *** 1%.Significance levels: * 10%, ** 5%, *** 1%.

Table 7: Interfuel Elasticities of Substitution for China from the Existing Literature

	Coal-Gasoline	Coal-Elec.	Coal-Diesel	Gasoline-Elec.	Gasoline-Diesel	Elec.-Diesel	Oil-Elec.	Oil-Coal
Fisher-Vanden *et al*. (2004)		0.33					0.65	1.06
Hang & Tu (2007)		0.20					-0.33	0.53
H. Ma *et al*. (2008)	0.28	1.17	0.06	0.67	0.16	0.69	*0.68*	*0.17*
Serletis *et al*. (2011)		0.11					0.05	0.46
Smyth *et al*. (2012)		1.01					1.09	0.90
This study BE	0.30	0.00	0.32	0.28	1.78	0.14	*0.21*	*0.31*
This study IDE	0.50	0.07	0.38	0.48	2.32	0.14	*0.31*	*0.44*
This study FE	0.00	0.00	0.00	0.01	0.00	0.00		

Note: Figures in italics are averages for gasoline and diesel elasticities. Data for Smyth *et al*. are the Hicks elasticity of substitution.

Table 8: Interfactor Elasticities of Substitution for China from the Existing Literature

	Energy-Capital	Energy-Labor	Capital-Labor
Fan *et al*. (2007)	1.44	-0.07	0.77
Mallick (2007)			0.55
H. Ma *et al*. (2008)	0.80	0.61	0.34
Smyth *et al*. (2011)	1.01	0.68	0.98
This study BE	0.22	1.07	1.13
This study IDE	0.64	0.96	0.52
This study FE	0.16	0.30	0.24

Figure 1. Laspeyres Coal Price Index 2010

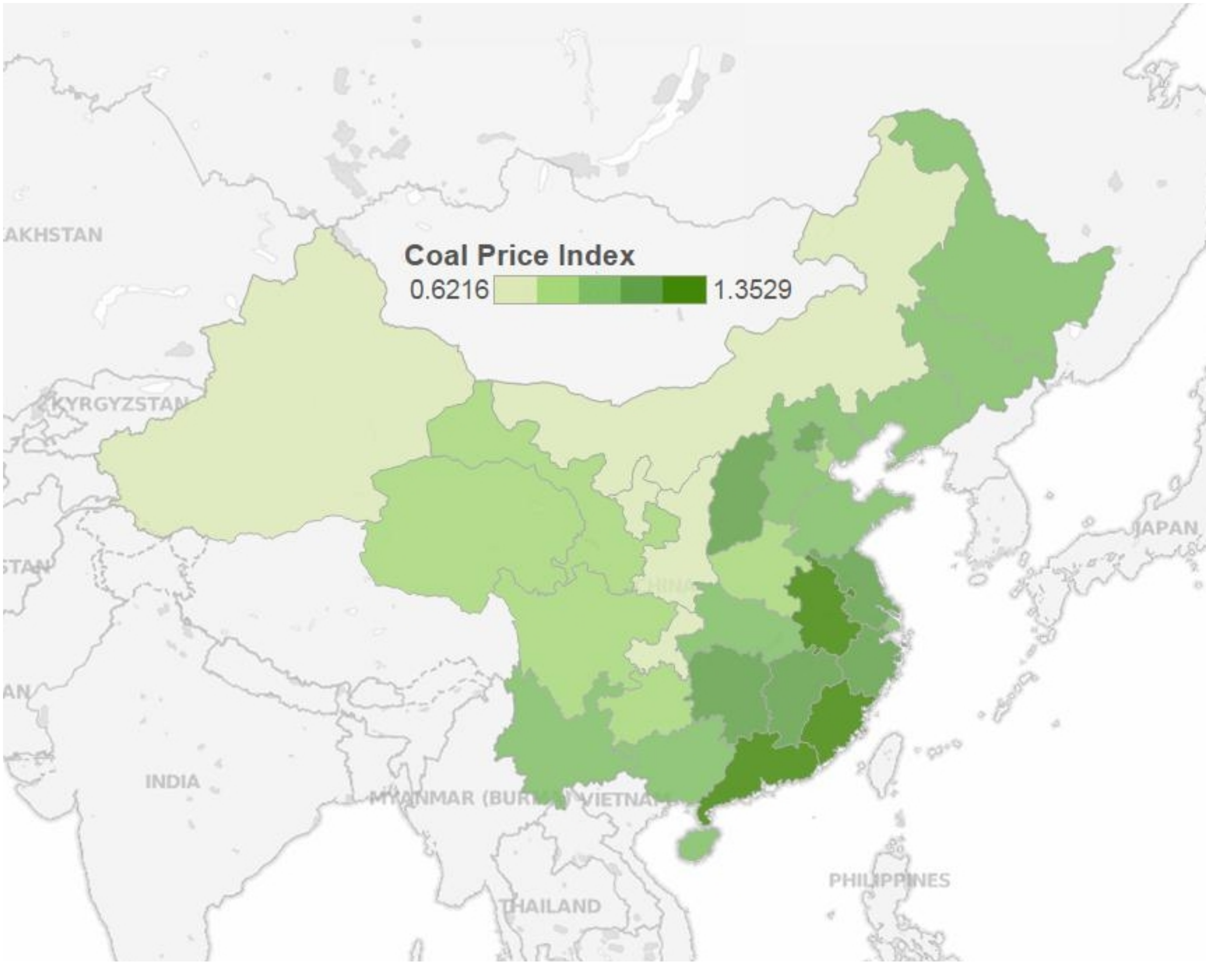

Figure 2. Log Distance from the Frontier 2010

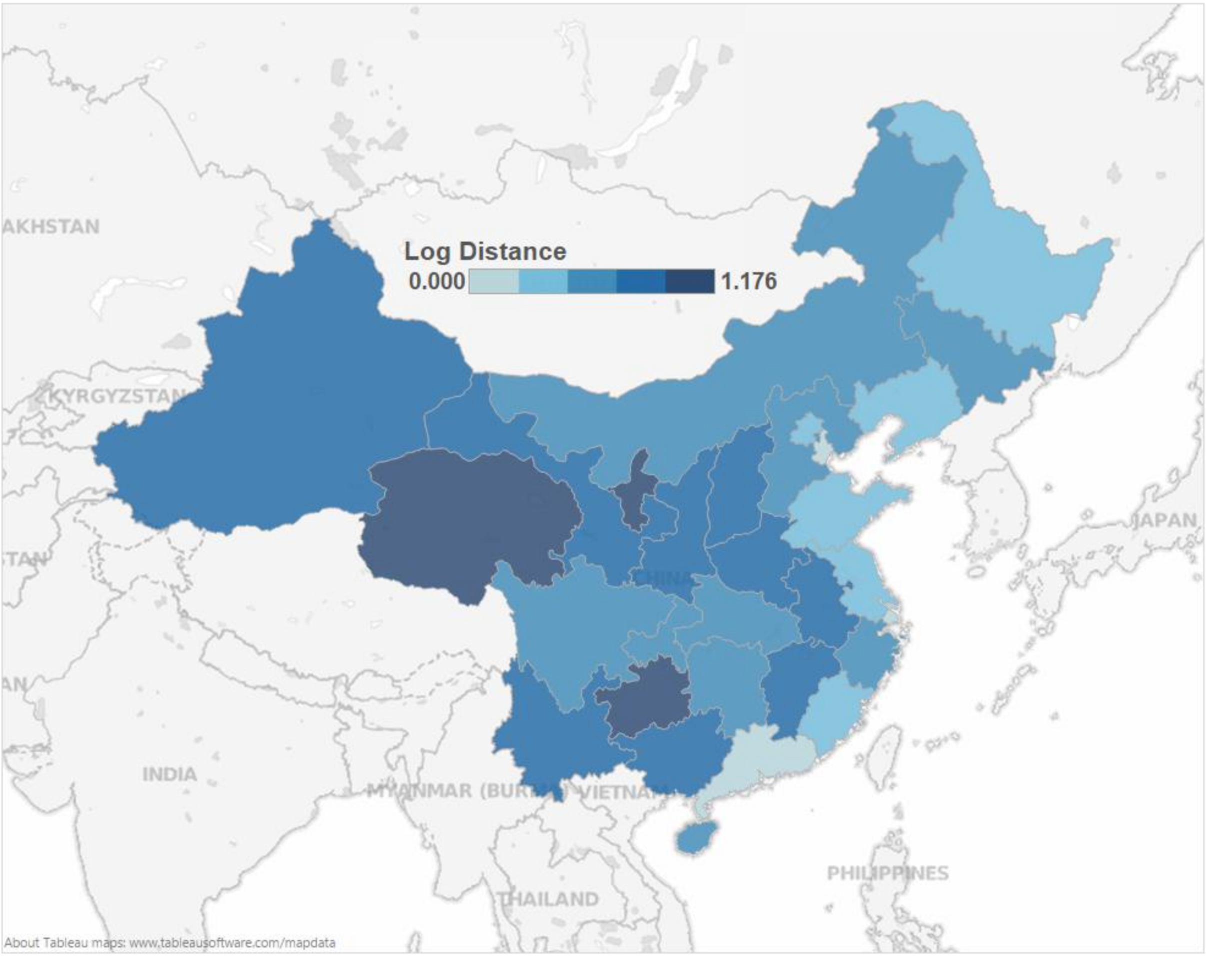

Figure 3. Growth Rate of TFP 2000-2010

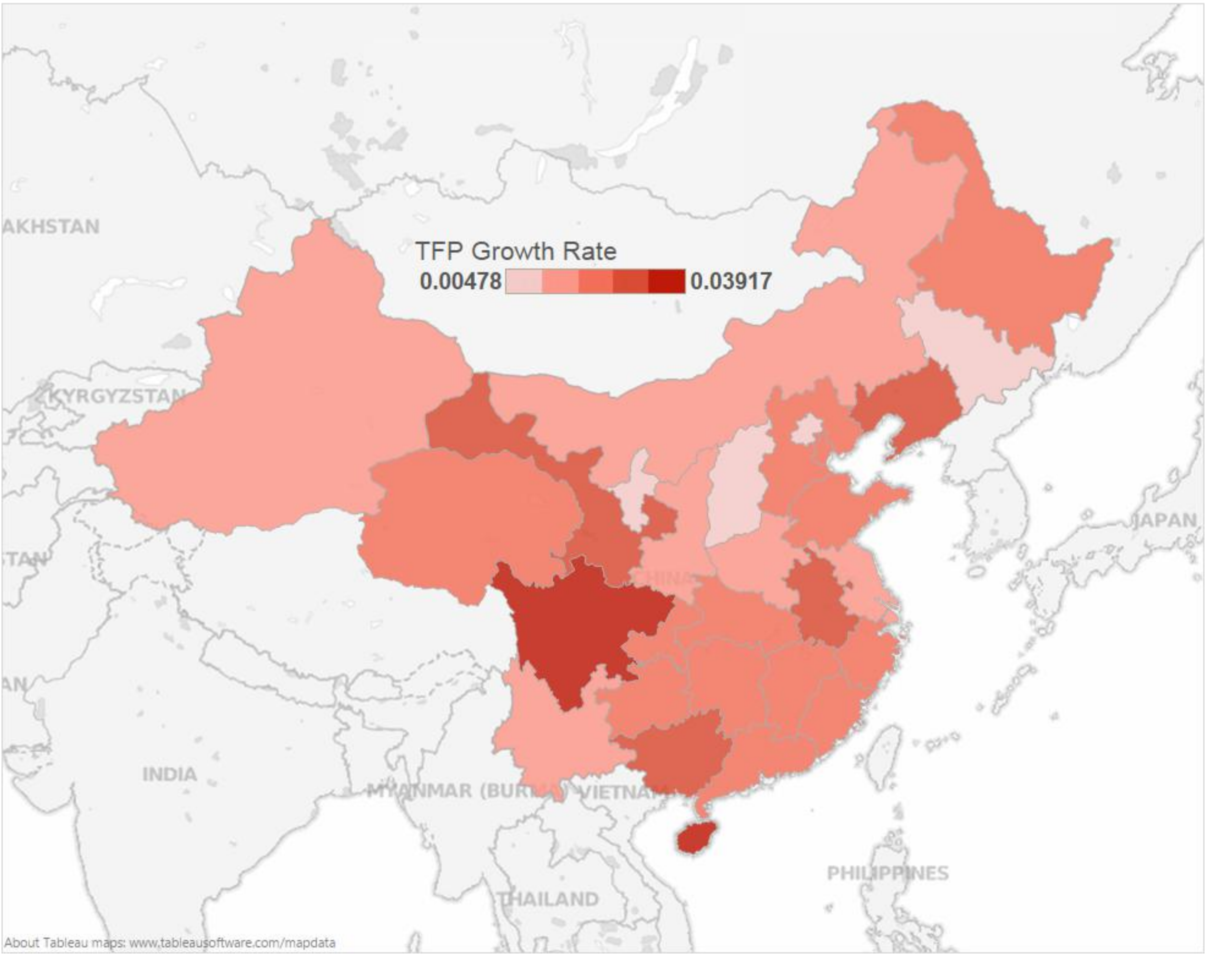

Figure 4. Own and Cross Price Elasticities

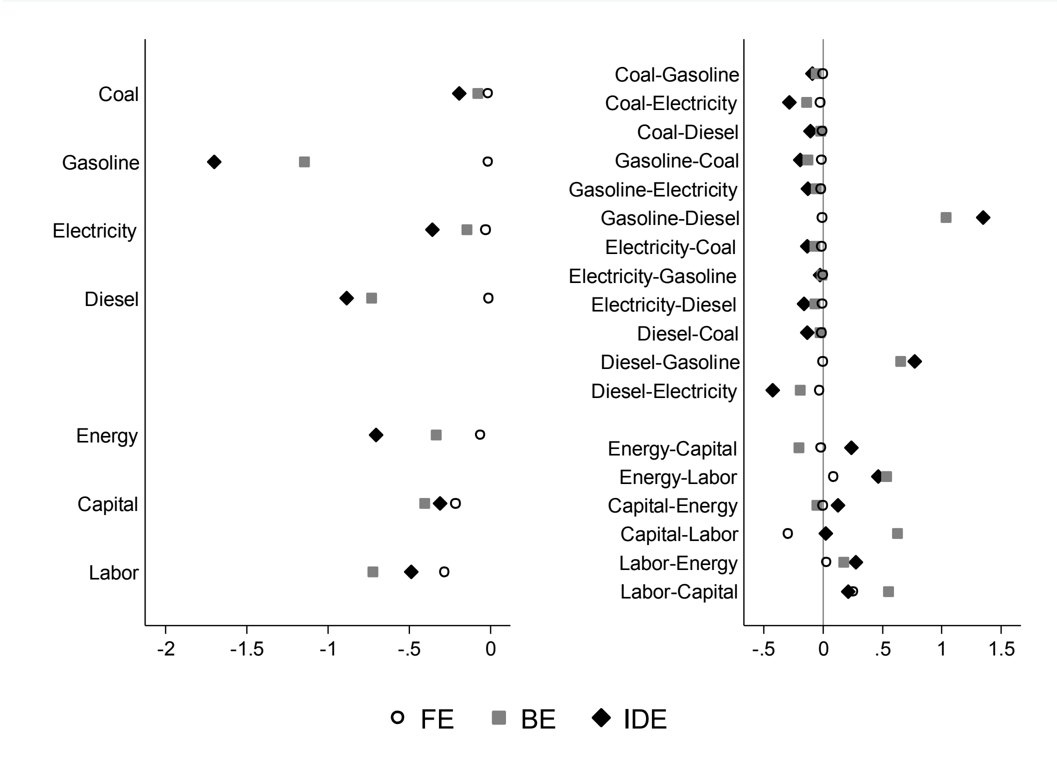

Figure 5. Shadow Elasticities of Substitution