Estimating the local effectiveness of institutions: a latent-variable approach

Antonia Reinecke* Hans-Jörg Schmerer†

Abstract

There is some consensus in the academic debate that strong institutions promote exports. Most existing studies in this field build on country- or province-level estimates of institutional quality without taking regional heterogeneity into consideration. Our new identification strategy allows constructing more detailed proxies on the effectiveness of institutions at the firm-level. An otherwise identical firm located in an environment with good institutions and better government effectiveness should be able to export more. Put differently, the export performance comprises information on the institutional quality the firm is faced with. We argue that this interdependence can be uncovered as government effectiveness is systematically linked to the firm's distance from its political hub. Development in earlier stages is likely more focused on provincial capitals. Development in the hinterland takes place with a certain lag. This second hypothesis enables us to use distance as a proxy for the effectiveness of institutions. Using methods of structural equation modeling we predict unobserved institutional quality at the zip-code level. We test the impact of this novel identification strategy in a treatment effect analysis that compares the coefficients obtained from provincial measure of institutions with our new index. We can show that the magnitude of the effect of institutional quality on the propensity to export is biased in estimates that neglect regional differences in institutional effectiveness.

*FernUniversität in Hagen. E-mail: antonia.reinecke@fernuni-hagen.de

†FernUniversität in Hagen, CESifo and IAB. E-Mail: hans-joerg.schmerer@fernuni-hagen.de

1. Introduction

There is some consensus in the academic debate that weak institutions hinder firm performance through higher transaction costs. Moreover, a recent strand of literature shows that more productive firms are also more likely to export. Taking theses findings together brings us to the conclusion that institutions likely affect firms' export performance. There is a large and emerging literature that tries to identify this hypothesized link. Anderson and Marcouiller (2002) estimate the effect of institutional quality on trade by treating instability due to weak institutions as hidden tax. They identify a strong positive effect of strong institutions and trade. Reinecke and Schmerer (2017) find a negative effect of low government efficiency, as a determinant of institutional quality, on the positive relation of firm size and export volume. De Groot et al. (2004) provide empirical evidence that the degree of homogeneity in institutions gives some security to all trading partners associated with soaring export volumes among them. More similar business proceedings can reduce transaction cost, which has a stimulating effect on trade. Ranjan and Young Lee (2007) show that stronger contract enforcement has a positive effect on trade volumes. Gani (2017) identifies a significant and positive relation between improvements in contract enforcement and trade volumes by analyzing time and financial expenditures to enforce contracts.

Other than the mentioned transaction cost channels, firm selection into the export market should be affected by institutional quality as well. According to Melitz (2003) only the more productive firms are able to cover exporter fixed cost but productivity in the Melitz model is constant over time. Existing studies by Hall and Jones (1999), Tanzi and Davoodi (2000), Rodrik et al. (2004) suggest a link between institutions and firm productivity. Moreover, Bryant and Javalgi (2016) find that investment into human capital is associated with soaring exports but the effect is mitigated by corruption.

All studies mentioned above have in common that institutional quality[1] is measured at the highly aggregated country level. These measures are appropriate for cross-country comparisons but less suitable for firm-level studies that allow for a more

[1] Including government efficiency, corruption, contract enforcement, etc.

detailed analysis of the effects of institutions on various outcome variables.

On a more dis-aggregated level, local governments likely support particular branches or regions through special policy actions on a more detailed regional level. Konstantynova and Wilson (2017) argue that the local community and informal institutions within a region determine the region's effectiveness in executing certain policy measures implemented on the country or the provincial level. Furthermore, governmental departments differ at the regional level as well. Institutional heterogeneity therefore affects the effectiveness of how policies on the upper level are executed in the different regions. Cluster policies affect the heterogeneity of institutions and competitiveness in particular regions (Konstantynova and Wilson (2017)). Dougherty (2013) analyzes the impact of the degree of legal enforcement quality on firm performance using Mexican firm-level data and state-level data for institutional quality. The distance between the cities in which firms are located in and the American border is considered as proxy for institutional quality. States closer to the American border are supposed to benefit from better institutions. Second, firms located in states associated with better institutions tend to be larger and more productive.

Another strand of literature focuses on regional planning and the role of institutions in regional (re-)structuring. Hahn (1970) emphasizes that rural governments are different from governments in urban areas: they are characterized by higher level of informality, slower rates of adjustment and less experienced in handling crises. Additionally, infrastructure is less developed and the concentration of institutional locations is lower than in urbanized regions (Frank and Reiss (2014)). Moreover, the transition from rural to more urban economies can be associated with a change in government efficiency. Tailoring the right policies becomes more difficult when societies are more diverse across or within regions.[2] Pemperton and Goodwin (2010) highlight the role of local governments, which are spatially delimited implementing political strategies: The regional diversity regarding political, social and cultural nature is decisive for institutional development.

The discussion above highlights that regional disparities in institutional quality must

[2]See Hahn (1970) for more details on this issue.

be taken into consideration when estimating institutional impact on firms' performance. However, there is no good proxy or instrument waiting in wings. We suggest using an identification strategy that treats regional effectiveness as latent variable that is determined by the institutional setting on the province level and the distance between the firm's location and the political hub of the province. The latter is defined as the capital of the respective province. We are arguing that the level of institutions measured in the data reflects the situation in the core area of a province. Higher remoteness is associated with lower enforcement power, mainly due to a sequential development strategy that focuses on the core areas first. The firm's effectiveness in dealing with legal issues is unobserved but we are able to uncover it as latent variable estimated in a structural equation modeling approach. The identification is based upon information on provincial government efficiency and a regional discount factor that depends on the respective firm's remoteness from the provincial policy hub.

We apply this new method in a treatment effect analysis. The effect of institutional quality is four times larger in regressions that do not account for the regional discount factor on provincial institutional quality.

First Glimpse on the data. Despite the political efforts to reduce disparities across Chinese provinces since the 1990's, the differences in regional development prevail across and within regions. The political landscape of China is characterized by policy clusters across narrowly defined areas. Hence, political measures were often implemented at the local level of provinces concerning particular industries. In Guangdong the *Zujiang Delta* (Pearl River Delta) describes a spatially delimited area where export-processing and equipment supply were favored by the government. The economic performance in those regions is mainly driven by foreign investors. In contrast, the *Wenzhou Model* in Zhejiang promoted light industries dominated by local, private owned enterprises. In opposition to those contrasting development strategies, the *Sunan Model* in Jiangsu emphasized the role of local governments investing into state owned enterprises located in Jiangsu. The diverging development strategies in different areas led to high economic and institutional disparities across and within provinces (e.

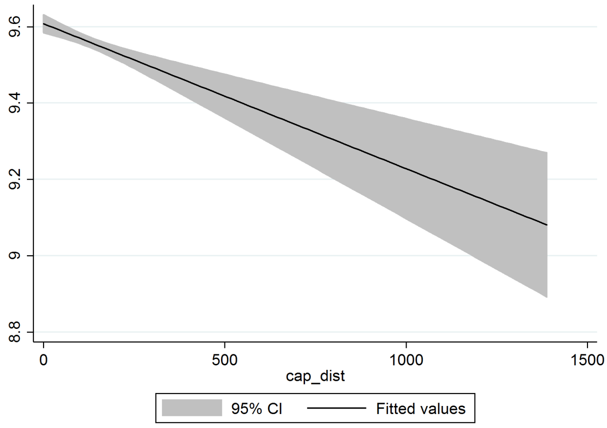
g. Xinyue and Wei (2005), Wei and Kim (2002), Gu et al. (2001)). The choice of certain regions was also driven by comparative advantage determined by its proximity to the economic hubs (e. g. Hong Kong, Shanghai, or the respective provincial capital). Thus, distance to those economic hubs comprises some information on the development strategy that can be used to estimate a more detailed proxy of government efficiency.

We build our analysis on a province-level index of government efficiency, which is one of the main determinants regarding the allocation of legal systems, public services, public infrastructure and reasonable decision making, and thereby covers important aspects of institutional quality in China.

Based on this information, we are using the proximity to the provincial capital as discount factor for local government efficiency. Decisions about certain policy measures are made in the capital of a province and we argue that areas closer to the capital tend to be characterized by higher government efficiency than areas located in rural areas, far away from the provincial capital. With increasing distance from the area the firm is located to the provincial capital city efficiency of local governments decreases. Secondly, we argue that firms' export behavior is determined by the government's efficiency and therefore comprises some information on institutions at the firm's location. The firm-level has the advantage that more detailed information is available that allows us to reveal the discount factor on a narrowly defined area. The higher the proximity between a firm's location and the provincial capital, the higher the expected government efficiency associated with this area and, ceteris paribus, the higher is the expected export share of the observed firm. In figure 1 we plot the export share of a firm against the distance to the capital.

The stylized fact presented in figure 1 supports the hypothesized link described above: firms located closer to the provincial capital tend to export more than firms located in a higher distance to the capital city. However, this result may be driven by sorting of firms into different regions. The more productive firms may sort into areas closer to the economic hub and those firms likely export more. Thus, we try to uncover the regional discount factor of distance to economic hubs using more sophisticated empirical methods that uses both, the information on the provincial level and the

Figure 1: *Correlation between export volume and distance to provincial capital*



detailed information on the firm level. The latter allows us to isolate the effect of distance on governmental efficiency by including firm-level controls in the regression. The analysis is structured as follows: In section two the empirical strategy is presented, section three presents the data used for the empirical analysis. In section four results are presented and discussed. Section five concludes.

2. Empirical Strategy

Our paper's contribution is a new strategy that allows to translate more aggregate measures of institutional measures to a more dis-aggregated level using firm-level data. We use this index to demonstrate the power of this new identification strategy based on a treatment effect analysis.

2.1. Estimating government efficiency across regions

Our identification strategy builds on a broad measure of government efficiency on the province level that can be merged to our enterprise data. This approach generates a multi-level setup with information on both the firm and the province level. However, the address of the firm contains information on the zip-code of firms' location, which can be used to identify the distance between the firm's location and the important political and economic hubs located at arms length to the firm. As argued in the introduction, the distance to the economic hubs in China (provincial capital, proximity to Hong

Kong, Shanghai and the coast) can be used to estimate institutional discount factors on the zip-code level. These discount factors allow us to disentangle the provincial information on government efficiency into a more detailed measure on the zip-code level.

Generalized Structural Equation Modeling

Suppose the following model, which explains a dependent variable on the firm-level using a constant, α, government-efficiency, GE, on the zip-code level denoted by index z, and additional control variables on the firm-level:

$$y_i = \alpha + \gamma_1 GE_z + \gamma_2 CV_i + \epsilon_{i1} \quad (1)$$

All variables other than GE_z are observable and information on the control variables is available in the data set at hand. The variable GE_z is treated as latent variable in our setting.

We are assuming that local government efficiency is a function of provincial government efficiency, GE_p, and the distance between the zip-code area the firm is located and the provincial capital, $ln(dist)$:

$$GE_z = \beta_1 + \lambda_1 GE_p + \lambda_2 ln(dist) + \epsilon_{i2} \quad (2)$$

Substituting GE_z in equation (1) by the relation derived in equation (2) leads to:

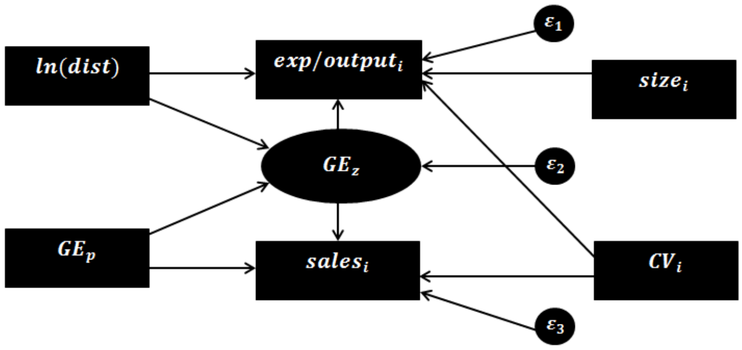
$$y_i = \alpha + \gamma_1(\beta_1 + \lambda_1 GE_p + \lambda_2 ln(dist) + \epsilon_{i2}) + \gamma_2 CV_i + \epsilon_{i1} \quad (3)$$

As dependent variable we are using firms' export share, which contains useful information about the institutional environment at the firm's location. The effect of government efficiency on a firm's export share in equation (3) depends on two coefficients, γ_1 and λ_1. Standard linear models cannot identify the two coefficients isolated from each other. Thus, we are using a structural equation approach.

To sum up, we suppose a relation between government efficiency and firm export.

Government efficiency can be observed at the province level but not at the zip-code level. However, we hypothesize a relationship between the distance from the zip-code area of a firm's is location to provincial capital and its efficiency in implementing those policies. The identification strategy can be summarized by the following path diagram:

Figure 2: *Path diagram - structural equation modeling*



Due to the fact we control for firm ownership by including ownership dummies and for firm size by defining five size classes based on the observed output we estimate the given relations by *Generalized Structural Equation Modeling - GSEM*. Further advantage using *GSEM* is that joint normality of exogenous and endogenous observed as well as latent variables in mean and variance matrix is not necessary. Observed, exogenous variables[3] are treated as given. The predictions for the variable GE_z are computed conditional on the values of those exogenous variables. Hence, the assumption of joint normality conditional on exogenous, observed variables is sufficient. To handle non-normality of estimated standard errors those will be adjusted by using quasi-maximum likelihood.

Based on this structural equation model we predict the latent variable government efficiency at the zip-code level, GE_z. This index can be used to separate firms into a treated and a non-treated group. For this purpose we have to standardize the magnitude of our latent variable to values ranging from -1 and $+1$ to make it comparable to the efficiency index at the provincial level.

[3]Observed, exogenous variables are: Government Efficiency at province-level, GE_p, distance between zip code area firm is located and provincial capital city, $ln(dist)$, and our control variables size, ownership, date of establishment as well as level of employment and capital.

2.2. Application of the new index

The impact of this new identification strategy is demonstrated in a treatment analysis. We are testing if government efficiency has a causal effect on its export performance. We run this analysis using the standard measure on the province level and compare those results to the results obtained using the more detailed measure on the zip-code level. We are able to show that the matching of firms is much better in application with more detailed proxies of government efficiency.

Treatment effect analysis

We want to test if the propensity of exporting is higher for firms treated with higher government efficiency. For this purpose we implement a treatment effect analysis with nearest neighbor matching (NNM)[4]. Our potential outcome model reads:

$$y_i = (1 - g)y_0 + gy_1 \tag{4}$$

where y_0 and y_1 describe the potential outcome under consideration of the treatment that is given by g. If $g = 0$ the observed outcome equals y_0, while the observed outcome equals y_1 if $g = 1$. The functional form of the outcome in our analysis is given by:

$$y_1 = \gamma_1(g) + \gamma_2(size_i) + \gamma_3(CV_i) + \varepsilon_i \tag{5}$$

The treatment model can be specified as:

$$g = \begin{cases} 1 & \text{if} \quad w'\rho + \eta > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where w specifies a treatment indicator with $w = 1$ in case of treatment, and $w = 0$ in case of non-treatment. The parameter ρ defines the unconditional probability of treatment, $P \equiv \rho(w = 1)$, and η denotes a error term.

[4]To determine the nearest neighbor the Mahalanobis distance is employed that use weights based on the inverse variance-covariance matrix of the covariates.

Effects of government efficiency cannot be observed immediately but it needs some periods to establish new infrastructure projects or changes in public spending. Simultaneously, the impact of a high efficient government does not vanish without delay. Due to this long term effect of government efficiency, a firm is defined as treated if the *Chinese provincial government efficiency index* of the province where the firm is located is $GE > 0$ for the years 2001, 2002 and 2003 in succession. Otherwise the firm is defined as member of the control group. As dependent variable, y_i, we are using the firm's export share. Firm size is controlled for by including sales, employment level and capital stock. Additionally, we include ownership dummies and the year of establishment as control variables.

Abadie and Imbens (2006) illustrate NNM estimations being biased if matching is conducted for more than two continuous covariates. Therefore, we apply a correction term provided by Abadie and Imbens (2011) that leads to consistent estimators. The firm's ownership is matched exactly and robust standard errors are implemented. The average treatment effect (ATE) and the average treatment effect on the treated (ATET) are estimated individually for every year in the time span between 2001 and 2006:

$$ATE = E(y_1 - y_0) \tag{7}$$

$$ATET = E(y_1 - y_0 | g = 1) \tag{8}$$

whereby the calculation of the $ATET$ seems to be more reliable because it allows to relax the conditional independence assumption (CI)[5] as well as the overlap assumption[6].

3. Data

Our analysis based on a panel covering the time span 2001 to 2006. To identify a firm's export volume NBS firm level data find application. It surveys firms from a turnover of

[5] The CI assumption requires that only observable covariates x affect potential and treatment outcome. Unobservable variables affecting the treatment must be independent of potential outcome. Calculating the $ATET$ instead the ATE requires only conditional independence of the potential outcome of the control group.

[6] The overlap assumption demand for a positive probability of obtaining each treatment level for each firm. Calculating the $ATET$ this assumption is relaxed by providing consistent estimates, if the probability of receiving each treatment is positive given the structure of covariates for which a firm is treated.

5 Million RMB between 1998 and 2006. Overall we have 1,728,740 observations during the whole period, where the number of firms covered varies between 146,106 in 1999 and 278,739 in 2006. The established firm level controls similarly origin from the NBS firm level data set. Therefore, we are able to calculate a firm's export share by dividing export volume by total output of the respective firm. Due to some inconsistency in the data, we drop duplicate firms and firms with a year of establishment earlier than 1850 and later than 2006.

The *Chinese provincial government efficiency index* is derived from Tang et al. (2014). The data cover 31 Chinese provinces[7] in a time span between 2001 and 2010.[8] The raw data to calculate the index stems from diverse Chinese statistical Year Books. Tang et al. (2014) adopt the calculation methods of the International Institution for Management (IMD) by computing averages and standard deviations of the respective indexes. The results are weighted and standardized. Therefore, the index ranges from -1 to $+1$, where a value of -1 is associated with the lowest level of government efficiency, while $+1$ describes the highest level.

To calculate the distances between zip-code area the firm is located and the provincial capital city we first geocode the zip code given in the NBS firm-level dataset to obtain coordinates of the respective area. We are not able to identify the exact coordinates of a business address because the NBS data only provide information about the zip code the firm is located. The coordinates always represent the hub of the zip-code area. Second, we use a Geographic Information System (GIS) to calculate the distances between the coordinates of the corresponding zip-code areas and the provincial capital city. A descriptive statistic can be found in Appendix II Table 10.

[7]Anhui, Beijing, Chongqing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Inner Mongolia, Jiangsu, Jiangxi, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shandong, Shanghai, Shanxi, Sichuan, Tianjin, Tibet, Xinjiang, Yunnan, Zhejiang

[8]The *Chinese provincial government efficiency index* contains government public services (24 indexes), government public infrastructure (11 indexes) and residents' economic welfare (7 Indexes) as well as size of the government. Hence, the index covers a wide range of characteristics determining government efficiency. For more detailed information see Appendix I.

4. Results

The structural equation modeling outcomes are discussed in the first part of the results section. The predicted government efficiency index is illustrated across provinces and zip-code regions. This comparison gives a first intuition for the impact of this new measure. The second part of this section presents the application results obtained from the treatment effect analysis. Those results underscore the first glance at the stylized facts. More detailed information on regional institutions allow for much more accurate matching of statistical twins.

4.1. Generalized SEM outcomes

In the first step the latent variable *Government Efficiency* at the *zip-code level* is predicted. Table (1) presents the estimates of the *gsem* using three different equations:

$$Sales_{it} = Constant_1 + \alpha_1 \times Controls + \alpha_2 \times GE_z + a_{it} \tag{9}$$

$$GE_{zt} = Constant_2 + \beta_1 \times Distance + \beta_2 \times GE_p + b_{it} \tag{10}$$

$$(exp/output)_{it} = Constant_3 + \gamma_1 \times Controls + \gamma_2 \times GE_z + c_{it} \tag{11}$$

The first equation captures the link between government efficiency and sales. We argue that firms located in areas with higher efficiency have better access to national markets and therefore higher sales. It is necessary to define starting values for the latent variable for identification of the parameters, which can be achieved through normalizing the coefficient α_2 to unity. All other coefficients must be interpreted relative to α_2. The second equation specifies the relationship between local government efficiency, provincial government efficiency and distance. The third equation specifies the relationship between exports, government efficiency and other control variables. Each equation has its own error term: a_{it}, b_{it} and c_{it}. Equation specific constants are included as well.

The main analysis is the effect of government efficiency on the export share of a firm. We argue that distance to the provincial capital has an effect on efficiency and that this effect is revealed by export dependency. Consistently, we aim at controlling for the major firm characteristics determining the export decision. On the one hand, these include firm characteristics as size and ownership. Additionally the choice of location may affect a firm's export share via for example advantages concerning infrastructure. Hence, results in the export equation will change depending on the controls included. In contrast, controls in the equations including *sales* and GE_Z as dependent variables will remain unchanged during the whole analysis, and thus the estimated coefficients, too. This approach helps us to predict reasonable values of the latent variable government efficiency at the zip code level.

The observed dependent variables are firm *sales* and export share, which are used to identify the latent variable government efficiency at the zip-code level, GE_z. In each specification the dependent variable *sales* is regressed on government efficiency, GE_z, and the capital-labor ratio to take production technology into account. The endogenous latent variable GE_z, government efficiency at the zip-code level, depends on the observed government efficiency at the province level, GE_p, and on the distance between the zip-code area the firm is located and the provincial capital city, $ln(dist)$. A firm's export share, $exp/output$, depends on capital distance, $ln(dist)$, government efficiency on zip-code level, GE_z, year of establishment, *birth date*, as well as on the capital-labor ratio, $capital/labor$. These benchmark results can be found in column (1) of table (1). In column (2) we additionally control for firm ownership including ownership dummies[9] with SOE being the base category in the export equation. In the third specification (column (3) of table (1)) we consider firm size in the export decision by including five size classes[10]. In column (4) we control for both, size and ownership estimating the specification with size groups and ownership dummies. In all four specifications the time trend is controlled for by including year-factor variables with 2001 serving as basis year.

In compliance with related academic literature, larger firms tend to export more,

[9]Ownership dummies are defined for state owned enterprises (SOE) including also collectively owned enterprises, private owned enterprises and foreign owned enterprises (also including firms from Hong Kong and Taiwan).

[10]Size classes are defined in quintiles based on the observed firms' output in the underlying data set.

Table 1: *Benchmark Regression*

	(1)		(2)		(3)		(4)	
	b	se	b	se	b	se	b	se
Sales								
GE_z	normalized to 1		normalized to 1		normalized to 1		normalized to 1	
Capital/Labour	0.048^{**}	(0.02)	0.048^{**}	(0.02)	0.048^{**}	(0.02)	0.048^{**}	(0.02)
Constant	10.581^{***}	(0.03)	10.581^{***}	(0.03)	10.581^{***}	(0.03)	10.581^{***}	(0.03)
Export/Output								
Distance (ln)	0.025^{***}	(0.00)	0.032^{***}	(0.00)	0.025^{***}	(0.00)	0.032^{***}	(0.00)
Capital/Labour	-0.005^{***}	(0.00)	-0.008^{***}	(0.00)	-0.005^{***}	(0.00)	-0.008^{***}	(0.00)
Birth	0.004^{***}	(0.00)	0.000^{***}	(0.00)	0.004^{***}	(0.00)	0.000^{***}	(0.00)
GE_z	0.524^{***}	(0.02)	0.255^{***}	(0.02)	0.508^{***}	(0.02)	0.245^{***}	(0.02)
Private (dummy)			0.044^{***}	(0.00)			0.043^{***}	(0.00)
Foreign (dummy)			0.375^{***}	(0.00)			0.374^{***}	(0.00)
Size group 2					0.034^{***}	(0.00)	0.029^{***}	(0.00)
Size group 3					0.042^{***}	(0.00)	0.034^{***}	(0.00)
Size group 4					0.053^{***}	(0.00)	0.038^{***}	(0.00)
Size group 5					0.051^{***}	(0.01)	0.027^{***}	(0.01)
Constant	-7.045^{***}	(0.08)	-0.692^{***}	(0.07)	-6.931^{***}	(0.08)	-0.595^{***}	(0.08)
Time trend	yes		yes		yes		yes	
GE$_z$								
Distance (ln)	-0.014^{***}	(0.00)	-0.014^{***}	(0.00)	-0.014^{***}	(0.00)	-0.014^{***}	(0.00)
GE_p	0.280^{***}	(0.01)	0.280^{***}	(0.01)	0.280^{***}	(0.01)	0.280^{***}	(0.01)
Var(e.GE_z)	0.086^{***}	(0.00)	0.026^{***}	(0.01)	0.062^{***}	(0.01)	0.011	(0.01)
Var(e.sales)	1.769^{***}	(0.01)	1.829^{***}	(0.02)	1.793^{***}	(0.02)	1.844^{***}	(0.02)
Var(e.exp/output)	0.111^{***}	(0.00)	0.108^{***}	(0.00)	0.118^{***}	(0.00)	0.109^{***}	(0.00)
Number of obs.	324,300		324,300		324,300		324,300	

Standard errors are reported in parentheses. Coefficients are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. Specification is estimated by generalized simulation equation modeling. Dependent variables are **sales**, export share, **exp/output** and government efficiency at zip-code level, **GE_z**, which is simultaneously a latent variable. GE_p is observed government efficiency at province level. $capital/labor$ specifies the capital-labor ratio. $ln(dist)$ measures the distance between the zip-code area the firm is located and the provincial capital city. *birth date* identifies the year of establishment. Firms are classified into $size\ groups$ defined by output quintiles. *SOE*, *Private* and *Foreign* specify firm ownership and stand for state owned enterprise (including collective owned enterprises), private owned firms as well as foreign owned firms (including firms of Hong Kong and Taiwan). *SOE* serves as base category. All specifications include yearly categories from 2001 to 2006 with 2001 being the base year. AIC specifies Akaike's information criterion, while BIC represents Bayesian information criterion.

represented by the highly significant and positive estimates of *size groups*.[11] The impact of a relative capital intensive production technology on *sales* is significantly positive and in line with expectations. However, a firm's export share is significantly negative related with a production technology being relatively capital intensive, indicated by the highly significant and negative coefficient of the capital-labor ratio, $capital/labor$. This result is counter intuitive: A well capitalized firm would be expected to be high productive and characterized by higher sales and export volume than firms less capitalized. This result can be explained by frictions in the Chinese financial market: China's company landscape is still characterized by a high level of SOEs and those have an easier access to the credit market. Consistently, SOEs are able to invest more in physical capital. Moreover, SOEs tend to export at a lower extend compared to firms owned by private or foreign investors: The coefficients of private and foreign ownership in column (2) and (4) of table (1) are highly significant and positive. Another explanation could be foreign firms well established in the global supply chain. Focusing on the coefficients of the ownership dummies again, it is conspicuous that foreign owned firms export at a higher extend than private owned firms. China is characterized by labor abundance, hence we would expect that foreign producers use a labor intensive production technology.[12] The estimates of *birth date* is significantly positive implying young firms tend to export more.[13]

The coefficient of the latent variable government efficiency at zip-code level, GE_z, is highly significant and positively associated with a firm's export share, and thereby underline the treatment effect results in tables (3) to (**??**). The estimates of $ln(dist)$ is positively related to exports, which is against our expectations. However, the endogenous latent variable GE_z is significantly negative affected by capital distance, $ln(dist)$, and thereby coincide with our hypothesis in the beginning: the proximity

[11] E. g. Egger and Kreickemeier (2009), Melitz (2003), Baumgarten (2013) for empirical evidence.

[12] The Heckscher Ohlin type specialization find only little empirical evidence in the academic literature. However, Ito et al. (2016) find empirical evidence for specialization according to respective sector endowments analyzing trade value added instead of trade volumes.

[13] According to Reinecke and Schmerer (2017) especially shortly before China's WTO entry in 2001 a massive increase of private and foreign owned firms establishments is observable. As seen in the coefficients of the ownership dummies *private* and $foreign$, these young companies are firms characterized by a higher export share than SOEs, which are older by trend.

of a firm location to the provincial capital city is an important determinant of the government efficiency level. According to our results, government efficiency declines with the distance between the zip code area the firm is located and the provincial capital city. In line with intuition government efficiency at zip-code level is positively affected by the initial level of government efficiency at the province level, indicated by the highly significant and positive coefficient of GE_p.

To decide which is the most appropriate specification to predict the latent variable GE_z we revert to Akaike's and Bayesian information criterion. Lower values mean one specification is preferred over another. Therefore, we will predict the latent variable based on the specification found in column (4) in table (1) including size groups as well as ownership dummies. To make the latent variable predictions comparable to the government efficiency measure at province level, we standardize them to $[-1, 1]$. The following table (2) shows a descriptive statistics of both, predicted and observed government efficiency:

Table 2: *Descriptive statistics for predicted and observed government efficiency*

	Obs.	Mean	Stand. Dev.	Min	Max
GE_z	326,751	-0.201	0.218	-1	1
GE_p	389,598	.157	.230	-.54	.77

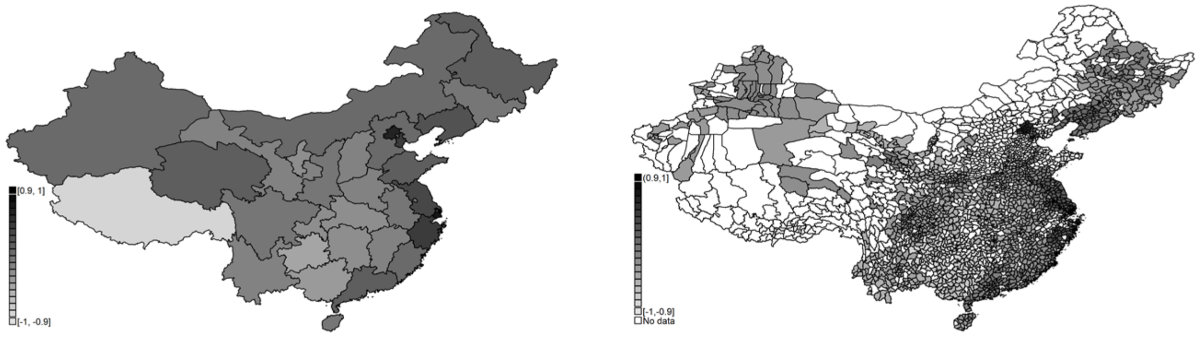
The prediction procedure produces 326,751 observations for the unobserved government efficiency at zip code level. Furthermore, we observe a higher standard deviation and a lower mean of GE_z. These results indicate higher diversity as well as an overestimation of government efficiency.

The following maps in figure (3) compare the observed and the predicted government effciency graphically:

The left panel visualizes government efficiency by province, while the right panel shows efficiency at a more dis-aggregated level.[14] The index is classified into increments

[14]The predicted Government Efficiency Index is even more dis-aggregated. We built mean values for specific areas to roughly illustrate the differences between the provincial and more dis-aggregated efficiency index. Firms in the underlying data set do not cover all areas represented by the white areas ("no data")

Figure 3: *Comparison of observed and predicted government efficiency*



of 0.1. Darker areas are associated with higher efficiency of the governments. The comparison of the two maps in figure (3) illustrates the provincial governmental heterogeneity: within a single province diverse levels of government efficiency are observable.

4.2. Treatment effect analysis outcomes

The following tables present the outcomes of the treatment analysis. Results in the first paragraph are based on the provincial measure on government efficiency. Paragraph 2 reports the results obtained from an analysis that builds on the more detailed measure of government efficiency constructed using the structural equation modeling approach introduced in this paper.

Province-level analysis. Table 3 represents the benchmark results of our treatment effect analysis estimating the effect of provincial government efficiency on exports using nearest neighbor matching. The regression is conducted for each year separately, robust standard errors and the correction term by Abadie and Imbens (2011) find application in each regression. We build a balanced panel starting in 2001 to avoid biased estimates due to a massive entry of firms. Balancing conditions can be found in tables (13) to (18).[15]

The results suggest a positive effect of government efficiency on a firm's export

[15] A good fit of the model is given, if after matching the standardized differences equal zero, while the variance ratios are close to one. The results in table (13) to (18) suggest a good fit of our model for the observed time period.

Table 3: *Treatment Effect regression results - 1*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.105***	0.107***	0.112***	0.109***	0.103***	0.105***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.103***	0.107***	0.111***	0.109***	0.102***	0.105***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	42,579	42,636	42,694	42,700	42,681	42,472
Control obs.	21,763	21,879	21,891	21,894	21,873	21,717
Number of obs.	64,342	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1 \; if \; GE_p > 0$ and $g = 0 \; if \; GE_p \leq 0$. Age, sales, employment level and capital stock of a firm are included as controls.

share. Both the $ATET$ and the ATE are highly significant for all years and the estimated magnitude of both effects is very similar. Consistently, we conclude: A firm located in a province characterized by high government efficiency tend to export between 0.104 and 0.112 percentage points more than a firm not treated by government efficiency. This result is very robust. To take ownership of a firm into consideration we apply ownership dummies in the next regression. Ownership is matched exactly. Results are shown in the Appendix. See Table 11 for more details.

Our estimates are robust against the inclusion of ownership dummies. The magnitude of the highly significant and positive coefficients slightly decreases. According to these results, firms located in a province associated with high government efficiency export between 0.076 and 0.084 percentage points more than firms being located in a less efficient province, which is consistent with our benchmark results.

More detailed analysis. We repeat the treatment effect analysis using the predicted values of government efficiency on zip-code level. A firm is again defined as treated if $GE_z > 0$ for the years 2001, 2002 and 2003 in succession. Table (4) represents the results of the treatment effect analysis using the predicted government efficiency at zip-code level.[16]

[16] Balancing Condition can be found in Appendix III, tables (19) to (24).The results in table (19) to (24) suggest a good model fit: standardized differences are close to zero, whereas variance ratios are close to

Table 4: *Treatment results based on Benchmark results (table (1))*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.051***	0.055***	0.058***	0.050***	0.050***	0.057***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.055***	0.059***	0.057***	0.053***	0.060***	0.059***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	11,642	11,659	11,662	11,660	11,661	11,639
Control obs.	52,701	52,856	52,923	52,934	52,893	52,550
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, employment level and capital stock of a firm are included as controls.

Observing treated vs. control observation it is conspicuous that the distribution of treated and non-treated firms changes massively in the regression analysis including the predicted government efficiency index at the zip-code level.[17] The estimated treatment effect is highly significant and positive, indicating that firms being located in an area characterized by higher efficiency tend to export more. This result goes hand in hand with the estimates using provincial government efficiency. However, examining the magnitude of the estimated treatment effect, it is obviously less pronounced than in the benchmark regression with provincial efficiency. The estimated $ATET$ ranges between 0.050 and 0.058, hence a firm being located in an area with efficient government is associated with a 0.050 to 0.058 percentage points higher export share than a firm being located in a less efficient area. Consequently, the effect of government efficiency is strongly overestimated neglecting a more detailed consideration of government efficiency at the zip-code instead of the provincial level.

To check the robustness of our results, we conduct the treatment regression in the next step controlling for firm ownership. Ownership is matched exactly again. Results are represented in table (12).

one.

[17]At the provincial level treated observations range between 42,523 and 42,750 (in average $\sim$ 66%), while control observations vary between 21,909 and 22,159. Including the predicted values of government efficiency in the regression 11,639 to 11,662 (in average $\sim$ 18%) observations are treated and 52,550 to 52,934 serves as control observations.

As the estimates before, the results in table (12) indicate a statistically significant and positive effect of government efficiency on a firm's export share. Similarly, the composition of treated and control observations differs from the estimation results shown in table (11), without consideration of government efficiency at the zip-code level. By the same token, the magnitude of the estimated effect is lower than in table (11) and thereby support the results in table (4): The effect of government efficiency is overestimated using provincial instead of zip-code government efficiency. However, the impact of government efficiency on a firm's export share is higher including firm ownership dummies. A firm located in an area associated with a high level of government efficiency tend to be characterized by a export share 0.070 to 0.076 percentage points higher than a firm located in a less efficient area.

4.3. *The role of distance to economic hubs in China*

So far, we assumed that a firm's export share exclusively depends on specific firm characteristics and the distance between firm location and provincial capital. But location choice as well as export decision of a firm may depend on other advantages concerning for example infrastructure as well. The distance to major harbors or the access to important transportation structure are expected to have an effect on the export decision, and thereby on the export share of a firm. East China is characterized by a higher development level than Central or Western China. The proximity to the sea represents a huge advantage for exporting firms because in the observed time frame from 2001 to 2006 shipping is the major means of transportation for exporting goods into the world. To take the importance of the sea route regarding exports into account, we include the distance between a firm location and the five biggest harbors of China.[18] Due to the fact, that these five biggest harbors are distributed along the east coast and we assume that these harbors are substitutes we choose to define catchment areas to avoid biased estimates. Doing so, we built for each harbor the average distance between the five harbors.

[18] These harbors are: Port of Shanghai, port of Shenzen, port of Ningbo, port of Guangzhou and port of Quingdao.

Table 5: *Regression results - including coastline and harbor distances*

	(1)		(2)		(3)		(4)	
	b	se	b	se	b	se	b	se
Sales								
GE_z	normalized to 1		normalized to 1		normalized to 1		normalized to 1	
Capital/Labor	0.048**	(0.02)	0.048**	(0.02)	0.048**	(0.02)	0.048**	(0.02)
constant	10.581***	(0.03)	10.581***	(0.03)	10.581***	(0.03)	10.581***	(0.03)
Export/Output								
ln(dist)	0.019***	(0.00)	0.021***	(0.00)	0.019***	(0.00)	0.021***	(0.00)
ln(dist_Taiwan)	−0.100***	(0.00)	−0.073***	(0.00)	−0.100***	(0.00)	−0.073***	(0.00)
ln(dist_Japan)	−0.095***	(0.00)	−0.060***	(0.00)	−0.095***	(0.00)	−0.060***	(0.00)
ln(dist_Hong Kong)	−0.105***	(0.00)	−0.068***	(0.00)	−0.104***	(0.00)	−0.067***	(0.00)
D × ln(dist_QD)	0.004***	(0.00)	0.002***	(0.00)	0.004***	(0.00)	0.002***	(0.00)
D × ln(dist_SH)	0.000	(0.00)	0.012***	(0.00)	0.000	(0.00)	0.012***	(0.00)
D × ln(dist_SZ)	−0.008***	(0.00)	−0.010***	(0.00)	−0.008***	(0.00)	−0.010***	(0.00)
D × ln(dist_NB)	−0.011***	(0.00)	−0.012***	(0.00)	−0.011***	(0.00)	−0.013***	(0.00)
D × ln(dist_GZ)	0.009***	(0.00)	0.011***	(0.00)	0.009***	(0.00)	0.011***	(0.00)
D × ln(dist_coastline)	−0.004***	(0.00)	−0.004***	(0.00)	−0.004***	(0.00)	−0.004***	(0.00)
capital/labor	−0.005***	(0.00)	−0.008***	(0.00)	−0.005***	(0.00)	−0.008***	(0.00)
birth date	0.002***	(0.00)	0.000	(0.00)	0.002***	(0.00)	−0.000	(0.00)
GE_z	0.654***	(0.00)	0.266***	(0.00)	0.647***	(0.00)	0.262***	(0.00)
Private			0.028***	(0.00)			0.027***	(0.00)
Foreign			0.313***	(0.00)			0.313***	(0.00)
size group 2					0.018***	(0.00)	0.018***	(0.00)
size group 3					0.027***	(0.00)	0.023***	(0.00)
size group 4					0.038***	(0.00)	0.029***	(0.00)
size group 5					0.039***	(0.01)	0.020***	(0.00)
constant	−2.105***	(0.08)	1.439***	(0.08)	−2.049***	(0.08)	1.497***	(0.08)
time trend	yes		yes		yes		yes	
GE_z								
ln(dist)	−0.014***	(0.00)	−0.014***	(0.00)	−0.014***	(0.00)	−0.014***	(0.00)
GE_p	0.280***	(0.01)	0.280***	(0.01)	0.280***	(0.01)	0.280***	(0.01)
var(e.GE)	0.049***	(0.00)	0.014**	(0.01)	0.031***	(0.00)	0.000	(0.00)
var(e.sales)	1.806***	(0.01)	1.841***	(0.02)	1.824***	(0.02)	1.855***	(0.01)
var(e.expv)	0.098***	(0.00)	0.103***	(0.00)	0.106***	(0.00)	0.103***	(0.00)
Number of obs.	324,300		324,300		324,300		324,300	

Standard errors are reported in parentheses. Coefficients are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. Specification is estimated by generalized simulation equation modeling. Dependent variables are **sales**, export share, **exp/output** and government efficiency at zip-code level, **GE_z**, which is simultaneously a latent variable. GE_p is observed government efficiency at province level. *capital/labor* specifies the capital-labor ratio. *ln(dist)* measures the logarithmized distance between the zip-code area the firm is located and the provincial capital city. *ln(dist_Quingdao)*, *ln(dist_Shanghai)*, *ln(dist_Shenzen)*, *ln(dist_Ningbo)*, *ln(dist_Guangzhou)* are the distances between firm location and the respective harbor. *ln(dist_coastline)* represents the shortest distence between coastline and firm location. *birth date* identifies the year of establishment. Firms are classified into *size groups* defined by output quintiles. *SOE*, *Private* and *Foreign* specify firm ownership and stand for state owned enterprise (including collective owned enterprises), private owned firms as well as foreign owned firms (including firms of Hong Kong and Taiwan). *SOE* serves as base category. All specifications include yearly categories from 2001 to 2006 with 2001 being the base year. *AIC* specifies Akaike's information criterion, while *BIC* represents Bayesian information criterion.

Afterwards, we define different dummy variables that take the value 1 if the harbor lies in a specific radius of a firm, defined as the average distance between the respective harbors, the dummy variable takes a value of one. This dummy is than interacted with the actual distance between firm location and harbor. Due to the fact, China has approximately 2000 harbors, we additionally include the shortest distance between firm location and coastline. Just as in the case before, we built a dummy that takes a value of one if the coast is in the catchment area[19], and is then interacted with the actual distance between coastline and firm location.

Moreover, the proximity to important neighboring trading partners may be a further factor affecting firm location choice. Therefore, we include the distance between firm location and Hong Kong, Taiwan and Japan. Hong Kong additionally is an important harbor and at the same time, a financial hot spot of China. Hence, the distance to Hong Kong seems also be important to consider.

Sales again depend on government efficiency at the zip-code level and the capital-labor ratio. The latent dependent variable government efficiency at the zip code level is a function of the provincial level of efficiency and the distance between firm location and provincial capital city. The results including additional distances to major harbors and the coastline can be found in table (5)

As discussed in the previous section: The estimation equations including *sales* and GE_Z as endogenous variables remain unchanged. Hence, estimated coefficients of these two equation are not affected by the changes in the export equation.

The variables included in the specification before (table (1)) remain unchanged: firm size approximated by output quintiles is positively associated with a firm's export share, indicated by the highly significant coefficients of size group 2 to 5. While sales are significantly positive affected by the capital-labor ratio, the capital-labor ratio has a negative effect on exports. Private as well as foreign owned firms tend to export more relative to SOE. Similarly, younger firms tend to be characterized by higher export shares, indicated by the highly significant coefficient of *birth date*. The latent dependent variable GE_z, government efficiency at the zip-code level, is positively affected by provincial government efficiency and negatively by the distance between firm location and provincial capital.

In line with our expectation, the export share of a firm is negatively affected by

[19]The catchment area is defined as 250km radius (around half of the mean value).

the distance between location of a firm and important trading partners; Taiwan, Japan and Hong Kong. Similarly, the export share of a firm is positively affected by higher proximity to the coastline, illustrated by the highly significant and negative coefficient of $D \times ln(dist\_coastline)$. The estimation results concerning the distances between firm location and harbors are not that clear. The distance of a firm to the port of Ningbo and Shenzen is positively associated with a firms export share. In contrast, the estimated coefficients of $D \times ln(dist\_Quingdao)$ and $D \times ln(dist\_Guangzhou)$ are statistically significant and positive, suggesting a higher distance is associated with a higher export share of a firm. The estimates concerning the distance to the port of Shanghai are ambiguous: In column (1) and (3) the effect is insignificant, whereas the effect is significantly positive in column (2) and (4). An explanation for these results is that firms may be not located directly in the harbor area but in a more indirect proximity because other transportation channels play a role.

According to the AIC and BIC the last specification is preferable over the other ones. Therefore, we again predict our latent variable based on the last specification. A descriptive summary can be found in table (6).

Table 6: *Descriptive statistics for predicted and observed government efficiency - 2*

	Obs.	Mean	Stand. Dev.	Min	Max
GE_z	326,751	-.072	.304	-1	1
GE_p	389,598	.157	.230	-.54	.77

Again, the mean of the predicted government efficiency at the zip-code level is lower than those of the observed provincial government efficiency. Similarly, the standard deviation is higher in case of predicted variables. Hence, we can conclude that government efficiency is overestimated at the provincial level and that deviation is higher at the zip-code level.

To test how these results affect the impact of government efficiency on a firm's export share, we conduct our treatment effect analysis again. Results can be found in table (7) to (9).[20]

[20]Balancing Condition can be found in Appendix III, tables (25) to (30).The results in table (25) to (30)

Table 7: *Treatment results based on based on coastline and harbor regression (table (5))*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.018***	0.009***	0.010***	0.018***	0.016***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.018***	0.012***	0.011***	0.020***	0.018***	0.021***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	27,259	27,224	27,303	27,308	27,305	27,240
Control obs.	37,084	37,291	37,282	37,286	37,249	36,949
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, labor productivity and total factor productivity of a firm are included as controls.

Firstly, it is conspicuous the number of treated and control observations changes massively. Based on the prediction of the latent variable GE_z including additionally the distance to the coastline as well as to important harbors leads to a distribution of treated vs. control observation of approximately 42.4% to 57.6%. Hence, treated and control observation are distributed more equally compared to the prediction exclusively including the distance to the capital city and the estimation based on the provincial level of government efficiency. Secondly, the effect of government efficiency on a firm's export share is highly significant and positive but further declines: Based on this estimation results, we conclude that firms being treated by high government efficiency tend to export 0.009 to 0.02 percentage points more than firms not treated by high government efficiency. This estimation support the result that the effect of efficiency on a firm's export share is highly overestimated neglecting the within provincial heterogeneity of government efficiency by exclusively including government efficiency at the provincial level.

As a robustness check we again conduct the treatment effect analysis including ownership dummies in table (8).

Including ownership dummies in the treatment effect analysis leads to an increase

suggest a good model fit: standardized differences are close to zero, whereas variance ratios are close to one.

Table 8: *Treatment results based on based on coastline and harbor regression (table (5)) - including ownership dummies*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.024***	0.016***	0.014***	0.023***	0.017***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.018***	0.014***	0.013***	0.019***	0.017***	0.017***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	27,259	27,224	27,303	27,308	27,305	27,240
Control obs.	37,084	37,291	37,282	37,286	37,249	36,949
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, labor and capital stock of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

of the significantly positive effect of government efficiency on the export share of a firm, which supports the results of the estimation conducted before. A firm located in an area characterized by high government efficiency tend to export 0.019 percentage points more in average than a firm not located in an area associated with an efficient government. To show the robustness of the estimation results, the next specification includes labor and total factor productivity instead of labor and capital stock. Table (9) represents the results.

The results shown in table (9) support the results of the estimations made before: The effect of government efficiency on a firm's export share is positive and highly significant suggesting firms being located in an efficient province tend to export 0.019 percentage points more of their total output than firms being located in an area characterized by low government efficiency. Moreover, the estimated effect is lower than those estimated based on provincial government efficiency data.

5. Conclusion

In this paper we investigated in how far measures of institutional quality at the national level are appropriated measures to analyze the relation between trade and institutional quality.

Table 9: *Treatment results based on based on coastline and harbor regression (table (5)) - including productivity measures*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.024***	0.017***	0.018***	0.021***	0.021***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.021***	0.016***	0.016***	0.018***	0.019***	0.019***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	26,946	26,880	26,881	26,018	25,216	24,579
Control obs.	36,302	36,488	36,456	35,252	34,703	33,634
Number of obs.	63,248	63,368	63,337	61,270	59,919	58,213

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, labor productivity and total factor productivity of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

Our analysis builds on the hypothesis that the quality of institutions decreases with increasing distance to the regional capital city. We argue that the proximity of a firm location to the regional capital city is an important factor determining local institutional quality. Political decisions are made in capital cities, hence we expected areas close to those political hubs being characterized by a higher level of institutional quality. Based on this hypothesis we argue that the eased quality is associated with a negative effect on a firm's export share. Hence, the closer a firm is located to the regional capital city the higher is the institutional quality and the higher the expected export share of the observed firm.

Based on Chinese firm-level data we verify the hypothesized relation. Furthermore, institutional quality was approximated by government efficiency at provincial level. Firstly, we estimate a baseline regression conducting a treatment effect analysis. The results suggest a positive effect of government efficiency on a firm's export share. In the next step, the latent variable government efficiency at the zip-code level is predicted using *Generalized Simulation Equation Modeling - gsem*, where government efficiency at zip-code level is a function of provincial government efficiency and the distance between firm location and capital city. The analysis support the hypothesized negative relation between distance to the capital city and government efficiency. Based on these results we predict values for our unobservable efficiency at zip-code level. In the last step, we again estimate the treatment effect using the predicted, more detailed level of government efficiency.

Both treatment effects analyses suggest a highly significant and positive effect of government efficiency on a firm's export share. However, compared to the treatment effect analysis using the more detailed, predicted values of government efficiency, the baseline specification overestimate the effect of government efficiency. The estimated magnitude is less pronounced and this result is robust against the inclusion and substitution of diverse variables.

Hence, we conclude that measuring institutional quality at a national or county level yields to upward biased results estimating the relation between the quality of institutions and international trade.

References

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics*, 29(1):1–11.

Anderson, J. E. and Marcouiller, D. (2002). Insecurity and the pattern of trade: An empirical investigation. *The Review of Economics and Statistics*, 84(2):342–352.

Baumgarten, D. (2013). Exporters and the rise in wage inequality: Evidence from german linked employerâĂŞemployee data. *Journal of International Economics*, 90:201 – 217.

Bryant, C. E. and Javalgi, R. G. (2016). Global economic integration in developing countries: The role of corruption and human capital investment. *Journal of Business Ethics*, 136:437–450.

De Groot, H. L. F., Linders, G.-J., Rietvield, P., and Sumbramanian, U. (2004). The institutional determinants of bilateral trade patterns. *KYKLOS*, 57(1):103–124.

Dougherty, S. (2013). Legal reform, contract enforcement and firm size in mexico. OECD Economics Department Working Papers 1042, OECD.

Egger, H. and Kreickemeier, U. (2009). Firm heterogeneity and the labor market effects of trade liberalization. *International Economic Review*, 50(1):187–216.

Frank, K. I. and Reiss, S. A. (2014). The rural planning perspective at an opportune time. *Journal of Planning Literature*, 29(4):386 – 402.

Gani, A. (2017). Contract enforcement and trade. *Journal of Industry, Competition and Trade*, pages 1–13.

Gu, C., Shen, J., Wong, K.-y., and Zhen, F. (2001). Regional polarization under the socialist-market system since 1978: a case study of guangdong province in south china. *American Economic Review*, 98(5):1978–97.

Hahn, A. J. (1970). Planning in rural areas. *Journal of the American Institute for Planners*, 36(1):44–49.

Hall, Robert, E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1):83–116.

Ito, T., Rotunno, L., and Vezina, P.-L. (2016). Heckscher-ohlin: Evidence from virtual trade in value added. *Review of International Economics*.

Konstantynova, A. and Wilson, J. R. (2017). Cluster policies and cluster institutions: an oppportunity to bind economic and social dimensions. *Economia e Politica Industriale*, 44(4):457–472.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Pemperton, S. and Goodwin, M. (2010). Rethinking the changing structures of rural local government - state power, rural politics and local political strategies. *Journal of Rural Studies*, 3(3):272.283.

Ranjan, P. and Young Lee, J. (2007). Contract enforcement and international trade. *ECONOMICS & POLITICS*, 19(2):191–218.

Reinecke, A. and Schmerer, H.-J. (2017). Government efficiency and exports in china. *Journal of Chinese Economic and Business Studies*, 15(3):249–268.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9.

Tang, R., Tang, T., and Lee, Z. (2014). The efficiency of provincial governments in china from 2001 to 2010: measurement and analysis. *Journal of Public Affairs*, 14(2):142–153.

Tanzi, V. and Davoodi, H. R. (2000). Corruption, growth, and public finances. IMF Working Paper 00/182, International Monetary Fund.

Wei, Y. D. and Kim, S. (2002). Widening inter-county inequality in jiangsu province, china, 1950âĂŤ95. *Journal of Development Studies*, 38:142–164.

Xinyue, Y. and Wei, Y. D. (2005). Geospatial analysis of regional development in china: The case of zhejiang province and the wenzhou model. *Eurasian Geography and Economics*, 46:895–921.

6. Appendix

6.1. Chinese Provincial Government Efficiency Index composition

Indicator	Sub-indicator	Index
Government public service	Service of science, education, culture and health	Expenditure on science per capita (new product trial fees, testing fees, and subsidies for major research projects, in yuan), qualified product rate (in %), patent application granted (invention patents, utility model patents, and design patents (per 100 000 people)), turnover per capita in technology market (in yuan), primary school students teachers ratio (reverse index), secondary school students teachers ratio (reverse index), functional illiteracy rate (in %, reverse index), state education budget (in % of GDP), literary publishing units (per 100 000 people), health beds (100 000 people), health personnel (100 000 people)
	Public security service	accident rate (traffic accidents, fires, and environmental pollution (one per 100 000 people, reverse index), loss of three accidents per capita (in yuan, reverse index), legislation (new legislation, amendments including laws, bills, and regulations), accepted case in court of first instance, concluded case in court of first instance, arrested suspect by the prosecutor's office, criminal case cracked or registered by the Public Security Bureau, criminal case rate (per thousand people, reverse index)
	Meterological service	Site for agricultural meteorology service (per 100 000 people), seismic monitoring stations (per 100 000 people)
	Social security Service	Employment agencies (per 100 000 people), urban community service facilities (per 100 000 people), rural social security network (per 100 000 people)
Government public infrastructure	Basic social infrastructure	basic infrastructure and renovation investment within state budget (RMB 100m), ratio of local projects and central government projects in basic infrastructure and renovation investment, basic infrastructure and renovation project completed rate (in %), industrial 'three wastes' treatment efficiency (waster, residue, and gas waste, in %), reservoir capacity per 10 000 people (100 million cubic meters per 10 000 people), ratio of nature reserve area (in %)
	Basic urban	city gas penetration rate (in %), urban public transport vehicles per 10 000 people, urban road area per capita (square meters), urban public green area per capita (square meters), urban public toilets per 10 000 people
Size of Government		proportion of administrative staff in the total population (per 10 000 people, reverse index), proportion of administrative employment in total employment (in %, reverse index), ratio of government consumption and final consumption (in %, reverse index), ratio of government consumption and GDP (in %, reverse index), ratio of income from confiscation and administration on total revenue (in %, reverse index)
Residents' economic welfare		per capita net income of rural households (in yuan), per capita urban disposable income of households (in yuan), Engel coefficient of rural residents (in %, reverse index), Engel coefficient of urban residents (in %, reverse index), consumer price index (previous year = 100, reverse index), per capita GDP (in yuan), ratio of subsidy expenditure and financial expenditure (in %)

Source: Tang, R., Tang, T. and Lee, Z. (2014): The efficiency of procincial governments in China from 2001 to 2010: measurement and analysis, p. 147/148

6.2. Summary statistics

Table 10: *Descriptive statistics*

	Obs.	Mean	Stand. Dev.	Min	Max
$export/output$	1,689,972	.169	.343	0	26.0212
GE_p	1,284,366	.147	.237	-.88	.77
$sales$	1,689,972	9.819	1.418	0	18.878
age	1,723,523	10.851	12.270	0	155
$ln(labor)$	1,696,413	4.753	1.184	0	12.053
$ln(capital)$	1,700,289	3.846	1.671	-5.478	13.789
TFP	1,577,162	-.399	1.357	-15.907	9.522
LP	1,673,894	5.054	1.189	-8.120	14.458
SOE	1,723,523	.463	.497	0	1
$Private$	1,723,523	.331	.471	0	1
$Foreign$	1,723,523	.205	.404	0	1
N	1,723,523				

6.3. Robustnes checks I

Table 11: *Treatment analysis with provincial government efficiency including ownership dummies*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.078***	0.083***	0.083***	0.080***	0.076***	0.080***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.074***	0.079***	0.082***	0.078***	0.074***	0.080***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	42,579	42,636	42,694	42,700	42,681	42,472
Control obs.	21,763	21,879	21,891	21,894	21,873	21,717
Number of obs.	64,342	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_p > 0$ and $g = 0\ if\ GE_p \leq 0$. Age, sales, employment level and capital stock of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

Table 12: *Treatment analysis with more detailed government efficiency including ownership dummies*

Dependent Variable: export share

	(1) 2001	(2) 2002	(3) 2003	(4) 2004	(5) 2005	(6) 2006
ATET	0.072***	0.072***	0.076***	0.070***	0.071***	0.072***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.068***	0.072***	0.072***	0.065***	0.070***	0.069***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Treated obs.	11,642	11,659	11,662	11,660	11,661	11,639
Control Obs.	52,701	52,856	52,923	52,934	52,893	52,550
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, employment level and capital stock of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

6.4. Test Statistics - Balancing Condition

Balancing conditions - Benchmark regression

Table 13: *Balancing condition - Benchmark 2001*

	Stand. diff.		Variance ratio	
	raw	matched	raw	matched
age	-0.244	0.002	0.706	1.010
age^2	-0.179	0.003	0.946	1.030
sales	0.088	0.003	0.833	1.042
lnl	-0.212	-0.003	0.869	1.033
lnk	-0.124	0.001	0.997	1.043
Number of obs.	64,342	85,158		
treated obs.	42,579	42,579		
control obs	21,763	42,579		

Table 14: *Balancing condition - Benchmark 2002*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.234	0.002	0.706	1.011
age^2	-0.181	0.003	0.974	1.038
sales	0.058	0.003	0.852	1.049
lnl	-0.186	-0.002	0.882	1.034
lnk	-0.106	0.002	0.981	1.043
Number of obs.	64,515	85,272		
treated obs.	42,636	42,636		
control obs	21,879	42,636		

Table 15: *Balancing condition - Benchmark 2003*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.218	0.003	0.703	1.025
age^2	-0.177	0.006	0.999	1.163
sales	0.061	0.000	0.833	1.040
lnl	-0.175	-0.004	0.894	1.032
lnk	-0.096	-0.001	0.967	1.040
Number of obs.	64,585	85,388		
treated obs.	42,694	42,694		
control obs	21,891	42,694		

Table 16: *Balancing condition - Benchmark 2004*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.200	0.002	0.659	1.009
age^2	-0.175	0.003	0.855	1.016
sales	0.031	0.001	0.830	1.049
lnl	-0.155	-0.003	0.896	1.035
lnk	-0.099	0.000	0.958	1.044
Number of obs.	64,594	85,400		
treated obs.	42,700	42,700		
control obs	21,894	42,750		

Table 17: *Balancing condition - Benchmark 2005*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.186	0.002	0.661	1.010
age^2	-0.172	0.003	0.860	1.026
sales	0.008	-0.001	0.836	1.048
lnl	-0.149	-0.004	0.913	1.038
lnk	-0.107	-0.001	0.960	1.046
Number of obs.	64,554	85,362		
treated obs.	42,681	42,681		
control obs	21,873	42,681		

Table 18: *Balancing condition - Benchmark 2006*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.177	0.003	0.654	1.011
age^2	-0.170	0.003	0.852	1.018
sales	0.028	-0.002	0.836	1.037
lnl	-0.139	-0.005	0.935	1.038
lnk	-0.110	-0.002	0.959	1.044
Number of obs.	64,189	84,944		
treated obs.	42,472	42,472		
control obs	21,717	42,472		

Balancing conditions - latent variable estimation

Table 19: *Balancing condition - Benchmark latent 2001*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.049	0.004	1.076	1.011
age^2	-0.002	0.003	1.544	1.017
sales	0.259	0.008	1.043	1.033
lnl	0.138	0.006	1.055	1.020
lnk	0.079	0.005	1.037	1.025
Number of obs.	64,343	23,284		
treated obs.	11,642	11,642		
control obs	52,701	11,642		

Table 20: *Balancing condition - Benchmark latent 2002*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.046	0.003	1.087	1.008
age^2	-0.001	0.003	1.556	1.018
sales	0.260	0.009	1.025	1.038
lnl	0.137	0.004	1.041	1.025
lnk	0.098	0.003	1.037	1.029
Number of obs.	64,515	23,318		
treated obs.	11,659	11,659		
control obs	52,856	11,659		

Table 21: *Balancing condition - Benchmark latent 2003*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.034	0.002	1.103	1.006
age^2	-0.005	0.002	1.575	1.006
sales	0.259	0.008	1.061	1.032
lnl	0.142	0.003	1.035	1.024
lnk	0.110	0.004	1.042	1.027
Number of obs.	64,585	23,324		
treated obs.	11,662	11,662		
control obs	52,923	11,662		

Table 22: *Balancing condition - Benchmark latent 2004*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.017	0.001	1.163	1.010
age^2	-0.021	0.002	1.851	1.034
sales	0.255	0.006	1.042	1.033
lnl	0.133	0.002	1.033	1.026
lnk	0.126	0.002	1.048	1.024
Number of obs.	64,594	23,320		
treated obs.	11,660	11,660		
control obs	52,934	11,660		

Table 23: *Balancing condition - Benchmark latent 2005*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.016	0.003	1.160	1.012
age^2	-0.021	0.003	1.783	1.026
sales	0.264	0.007	1.037	1.034
lnl	0.137	0.003	1.041	1.020
lnk	0.140	0.005	1.054	1.025
Number of obs.	64,554	23,322		
treated obs.	11,661	11,661		
control obs	52,893	11,661		

Table 24: *Balancing condition - Benchmark latent 2006*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.014	0.002	1.185	1.013
age^2	-0.024	0.003	1.826	1.028
sales	0.254	0.009	1.017	1.035
lnl	0.128	0.005	1.031	1.025
lnk	0.149	0.008	1.064	1.033
Number of obs.	64,554	23,320		
treated obs.	11,639	11,639		
control obs	52,55	11,639		

Table 25: *Balancing condition - including harbors and coastline 2001*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.067	0.002	0.914	1.007
age^2	-0.046	0.002	1.107	1.018
sales	0.161	0.006	0.840	1.032
lnl	-0.097	0.002	0.900	1.022
lnk	-0.047	0.001	0.941	1.024
Number of obs.	64,343	54,518		
treated obs.	27,259	27,259		
control obs	37,084	27,259		

Table 26: *Balancing condition - including harbors and coastline latent 2002*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.066	0.002	0.912	1.006
age^2	-0.046	0.002	1.108	1.016
sales	0.162	0.005	0.827	1.034
lnl	-0.082	0.000	0.885	1.022
lnk	-0.031	-0.000	0.937	1.026
Number of obs.	64,515	54,448		
treated obs.	27,224	27,224		
control obs	37,291	27,224		

Table 27: *Balancing condition - including harbors and coastline 2003*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.051	0.002	0.926	1.001
age^2	-0.038	0.001	1.156	0.966
sales	0.152	0.004	0.862	1.031
lnl	-0.081	-0.001	0.881	1.024
lnk	-0.024	0.000	0.944	1.026
Number of obs.	64,585	54,606		
treated obs.	27,303	27,303		
control obs	37,282	27,303		

Table 28: *Balancing condition - including harbors and coastline 2004*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.039	0.003	0.914	1.010
age^2	-0.034	0.003	1.282	1.036
sales	0.145	0.004	0.848	1.032
lnl	-0.084	-0.001	0.869	1.023
lnk	-0.018	-0.000	0.949	1.026
Number of obs.	64,594	54,616		
treated obs.	27,308	27,308		
control obs	37,286	27,308		

Table 29: *Balancing condition - including harbors and coastline latent 2005*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.029	0.002	0.920	1.010
age^2	-0.029	0.003	1.241	1.025
sales	0.136	0.004	0.861	1.036
lnl	-0.082	-0.000	0.879	1.024
lnk	-0.015	0.001	0.958	1.028
Number of obs.	64,554	54,610		
treated obs.	27,305	27,305		
control obs	37,249	27,305		

Table 30: *Balancing condition - including harbors and coastline latent 2006*

	Standardized differences		Variance ratio	
	raw	matched	raw	matched
age	-0.022	0.002	0.929	1.011
age^2	-0.024	0.003	1.220	1.028
sales	0.123	0.004	0.874	1.032
lnl	-0.088	-0.001	0.887	1.025
lnk	-0.012	0.001	0.968	1.031
Number of obs.	64,189	54,480		
treated obs.	27,240	27,240		
control obs	36,949	27,240		

6.5. Robustness check - including East Dummy

The east coast of China has some obvious advantages over central and western China: The proximity to the sea facilitate trade by providing short ways to ship goods to the rest of the world. This geographic advantage leads to special treatment of certain areas, which in particular includes measures supporting the development of international relations. To take these regional distinctions into consideration an *East Dummy*[21] is included in the export equation of the structural equation model. The other two equations, those determining sales and government efficiency at the zip code level, remain unchanged. Table (31) presents the results:

The sales as well as the latent variable equation remain unchanged, therefore the coefficients in these to equations are similar to these in table (1) and (5). *Sales* are positively affected by the capital-labor ratio, $capital/labor$. The latent variable, government efficiency at the zip-code level, GE_z, is positively related with provincial government efficiency, GE_p, and decreases with increasing distance of the firm location to the provincial capital city, $ln(dist)$.

[21] Eastern provinces are: Heilongjiang, Jiangsu, Shanghai, Hebei, Beijing, Shandong, Jilin, Liaoning, Zhejiang, Fujian, Guangdong, Hainan.

Table 31: *Including coastline, harbor distances and east Dummy*

	(1) b/se	(2) b/se	(3) b/se	(4) b/se
sales				
capital/labor	0.048**	0.048**	0.048**	0.048**
	(0.02)	(0.02)	(0.02)	(0.02)
GE_z	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)
constant	10.581***	10.581***	10.581***	10.581***
	(0.03)	(0.03)	(0.03)	(0.03)
exp/output				
ln(dist)	0.018***	0.020***	0.018***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)
ln(dist_Taiwan)	-0.098***	-0.071***	-0.098***	-0.071***
	(0.00)	(0.00)	(0.00)	(0.00)
ln(dist_Japan)	-0.073***	-0.043***	-0.073***	-0.043***
	(0.00)	(0.00)	(0.00)	(0.00)
ln(dist_Hong Kong)	-0.101***	-0.065***	-0.101***	-0.065***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_Quingdao)	0.004***	0.003***	0.004***	0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_Shanghai)	0.000	0.012***	0.000	0.012***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_Shenzen)	-0.009***	-0.011***	-0.009***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_Ningbo)	-0.011***	-0.012***	-0.011***	-0.012***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_Guangzouh)	0.009***	0.011***	0.009***	0.011***
	(0.00)	(0.00)	(0.00)	(0.00)
D × ln(dist_coastline)	-0.005***	-0.004***	-0.005***	-0.004***
	(0.00)	(0.00)	(0.00)	(0.00)
East Dummy	0.025***	0.019***	0.024***	0.018***
	(0.00)	(0.00)	(0.00)	(0.00)
capital/labor	-0.005***	-0.008***	-0.005***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)
birth date	0.002***	0.000	0.002***	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
GE_z	0.598***	0.224***	0.592***	0.221***
	(0.03)	(0.02)	(0.03)	(0.02)
Private		0.028***		0.027***
		(0.00)		(0.00)
Foreign		0.313***		0.313***
		(0.00)		(0.00)
size group 2			0.017***	0.018***
			(0.00)	(0.00)
size group 3			0.026***	0.023***
			(0.00)	(0.00)
size group 4			0.038***	0.028***
			(0.00)	(0.00)
size group 5			0.038***	0.019***
			(0.01)	(0.00)
constant	-2.309***	1.278***	-2.250***	1.338***
	(0.08)	(0.08)	(0.08)	(0.08)
time trend	yes	yes	yes	yes
GE_z				
ln(dist)	-0.014***	-0.014***	-0.014***	-0.014***
	(0.00)	(0.00)	(0.00)	(0.00)
GE_p	0.280***	0.280***	0.280***	0.280***
	(0.01)	(0.01)	(0.01)	(0.01)
var(e.GE_z)	0.053***	0.015**	0.034***	0.000*
	(0.00)	(0.01)	(0.00)	(0.00)
var(e.sales)	1.801***	1.840***	1.821***	1.855***
	(0.01)	(0.02)	(0.02)	(0.01)
var(e.exp/output)	0.100***	0.103***	0.107***	0.103***
	(0.00)	(0.00)	(0.00)	(0.00)
Number of obs.	324,300	324,300	324,300	324,300
AIC	1349534.081	1305481.504	1349398.309	1305348.391
BIC	1349854.764	1305823.566	1349761.749	1305722.521

Standard errors are reported in parentheses. Coefficients are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. Specification is estimated by generalized simulation equation modeling. Dependent variables are **sales**, export share, **exp/output** and government efficiency at zip-code level, **GE_z**, which is simultaneously a latent variable. GE_p is observed government efficiency at province level. *capital/labor* specifies the capital-labor ratio. *ln(dist)* measures the logarithmized distance between the zip-code area the firm is located and the provincial capital city. *ln(dist_Quingdao)*, *ln(dist_Shanghai)*, *ln(dist_Shenzen)*, *ln(dist_Ningbo)*, *ln(dist_Guangzhou)* are the distances between firm location and the respective harbor. *ln(dist_coastline)* represents the shortest distence between coastline and firm location. *East Dummy* takes a value of 1, if the firm is located in a east coast province. *birth date* identifies the year of establishment. Firms are classified into *size groups* defined by output quintiles. *SOE*, *Private* and *Foreign* specify firm ownership and stand for state owned enterprise (including collective owned enterprises), private owned firms as well as foreign owned firms (including firms of Hong Kong and Taiwan). *SOE* serves as base category. All specifications include yearly categories from 2001 to 2006 with 2001 being the base year. *AIC* specifies Akaike's information criterion, while *BIC* represents Bayesian information criterion.

Coefficients in the export equation change slightly, significance level and sign are robust against the inclusion of the *East Dummy*. Distances to important neighboring trading partners, *ln(dist_Taiwan)*, *ln(dist_Japan)*, *ln(dist_HongKong)*, are again negatively related with a firm's export share. The same holds for the distance to the coastline, $D \times ln(dist\_coastline)$. As in the regression before (table (5)), the impact of the distance to the five biggest harbors is ambiguous. The proximity to the harbors of Shenzen and Ningbo is significantly positive associated with the export share of a firm, while the proximity to the ports of Quingdao and Guangzouh is negatively associated with exports. The effect of the distance between the port of Shanghai and a firm location is again insignificant (column (1) and (3)) or positive (column (2) and (4)). This result support the hypothesis that firms are not directly located in the harbor area but other infrastructure is more important. The new implemented *East Dummy* exhibit the expected sign: Firms located in the eastern provinces are associated with an higher export share, indicated by the highly significant and positive coefficient of *East Dummy* in all specifications.

The estimates of the firm controls are robust against the changes implemented in this regression: A higher capital labor ratio, $capital/labor$, is associated with a lower export share of a firm. In contrast, younger firms tend to export more, indicated by the positive and highly significant coefficient of *birth date*. The coefficients of private and foreign firm ownership are again significantly positive, suggesting that these firms are characterized by a higher export share relative to state owned enterprises. Similarly, size is positively associated with a firm's export share, as discussed in relevant trade literature.

In compliance with the AIC and BIC, the last specification including both ownership dummies and size groups, is preferable over the other regressions indicated by the lowest values of AIC and BIC. Consequently, the latent variable GE_z is predicted based on the specification in column (4) of table (31). Comparable summary statistics can be found in table (32).

Table 32: *Descriptive statistics for predicted and observed government efficiency - 3*

	Obs.	Mean	Stand. Dev.	Min	Max
GE_z	326,751	-.072	.304	-1	1
GE_p	389,598	.157	.230	-.54	.77

Mean value and standard deviation do only change in a minor way. As in the predictions before, the mean value of the latent variable GE_z is lower than the mean value of the observed government efficiency at the provincial level, suggesting an overestimation of government efficiency at the highly aggregated level. The higher standard deviation for the predicted variable suggests a higher heterogeneity of government efficiency considering the exact zip-code area the firm is located. Based on this prediction we run the treatment effect analysis with nearest neighbor matching a third time. As before, a firm is defined as treated if government efficiency at the zip-code level is greater than zero for the years 2001, 2002 and 2003 in succession. Estimation results exclusively controlling for age, labor and capital stock as well as firm size by including logarithmic sales can be found in table (33).

Table 33: *Treatment results based on based on coastline, harbor and East Dummy regression (table (31))*

	2001	2002	2003	2004	2005	2006
	b/se	b/se	b/se	b/se	b/se	b/se
ATET	0.018***	0.009***	0.010***	0.018***	0.016***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.018***	0.012***	0.011***	0.020***	0.018***	0.021***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189
Treated obs.	27,259	27,224	27,303	27,308	27,305	27,240
Control obs.	37,084	37,291	37,282	37,286	37,249	36,949

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1\ if\ GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0\ if\ GE_z \leq 0$. Age, sales, labor productivity and total factor productivity of a firm are included as controls.

The results of the treatment effect analysis do not change, neither in significance level nor in coefficients' magnitude. Furthermore, the distribution between treated and control group remain unchanged. Including ownership dummies or substituting labor and capital stock by productivity measures[22] do not lead to changes in the results compared to the specification in tables (8) and (9): A firm's export share is positively affected by good institutions, in particular efficient governments. Moreover, the estimated magnitude of the effect is lower considering government efficiency at a highly dis-aggregated level, namely the zip-code level, than estimating the effect of

[22] Results can be found in table (34) and (35).

efficiency at the provincial level. This result is highly significant and robust.

Table 34: *Treatment results based on based on coastline, harbor and East Dummy regression (table (31)) - including ownership*

	2001	2002	2003	2004	2005	2006
	b/se	b/se	b/se	b/se	b/se	b/se
ATET	0.024***	0.016***	0.014***	0.023***	0.017***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.018***	0.014***	0.013***	0.019***	0.017***	0.017***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of obs.	64,343	64,515	64,585	64,594	64,554	64,189
Treated obs.	27,259	27,224	27,303	27,308	27,305	27,240
Control obs.	37,084	37,291	37,282	37,286	37,249	36,949

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1 \; if \; GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0 \; if \; GE_z \leq 0$. Age, sales, labor and capital stock of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

Table 35: *Treatment results based on based on coastline, harbor and East Dummy regression (table (31)) - including productivity measures*

	2001	2002	2003	2004	2005	2006
	b/se	b/se	b/se	b/se	b/se	b/se
ATET	0.024***	0.017***	0.018***	0.021***	0.021***	0.020***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ATE	0.021***	0.016***	0.016***	0.018***	0.019***	0.019***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of obs.	63,248	63,368	63,337	61,270	59,919	58,213
Treated obs.	26,946	26,880	26,881	26,018	25,216	24,579
Control obs.	36,302	36,488	36,456	35,252	34,703	33,634

$ATET$ (average treatment effect on the treated) and ATE (average treatment effect) build on a treatment effect analysis with nearest neighbor matching. Standard errors are reported in parentheses. Marginal effects are significant at the 10 percent (* p<0.10), 5 percent (** p<0.05) or 1 percent (*** p<0.01) level. The dependent variable is the export share on total output. The treatment is defined as $g = 1 \; if \; GE_z > 0$ for 2001, 2002 and 2003 in succession and $g = 0 \; if \; GE_z \leq 0$. Age, sales, labor productivity and total factor productivity of a firm are included as controls. Additionally, we control for firm ownership by including ownership dummies for state, private and foreign owned enterprises.

Overall, our estimates suggest considering government efficiency at a more detailed level than the provincial one has strong effects of the estimated relation between exports and government efficiency. Neglecting the impact of proximity to the capital city regarding the efficiency of a government leads to upward biased estimates. Nevertheless, high efficiency is positively related to exports, but the effect is less pronounced taking government efficiency at the zip-code instead of the provincial level into account.