China Economic Review 59 (2020) 101364

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

The productivity impacts of energy efficiency programs in developing countries: Evidence from iron and steel firms in China

Massimo Filippini[a,b], Thomas Geissmann[c], Valerie J. Karplus[c,*], Da Zhang[c,d]

[a] Department of Management, Technology, and Economics, ETH Zurich, Switzerland
[b] Department of Economics, Università della Svizzera Italiana, Switzerland
[c] Massachusetts Institute of Technology, Cambridge, MA, USA
[d] Tsinghua University, Beijing, China

ARTICLE INFO

Keywords:
Energy efficiency
Firms
China
Productivity
Cost function

JEL classification:
Q48
D24

ABSTRACT

We study the impact of a national energy efficiency program on total factor productivity (TFP) growth in firms in China's iron and steel industry. Using detailed firm-level survey data and multiple approaches to estimate program effects, we find participating firms experienced greater annualized TFP change. Our base specification estimates the program increased annual TFP growth by 3.1 percentage points, implying an annual private benefit of 148.7 million RMB/year per firm, with approximately equal contributions from technical change and scale efficiency change. Our results suggest that firms undervalued energy efficiency investments prior to the start of the program.

1. Introduction

Public- and private-led efforts to increase the efficiency of energy use in industrial production are common worldwide. Yet there is little evidence of how these interventions affect the productivity growth of targeted firms, especially in developing countries. We quantify the effect of a national energy efficiency program in China on the short-run total factor productivity growth (TFPG) of firms in the iron and steel industry. Using multiple empirical approaches, we find modest positive productivity effects of the program in the years following its implementation. We interpret these positive short-run effects as evidence that the program interacted with pre-existing investment inefficiencies, defined as the undervaluation of discounted savings relative to upfront costs, in managers' private valuation of energy-saving investments (Allcott, Hunt, & Greenstone, 2012). If investment inefficiencies are particularly important in developing countries, as our research suggests is the case in China's iron and steel industry, there is scope for informational and target-setting programs to improve both firms' energy decisions and productivity outcomes.

A growing body of scholarship empirically estimates the impact of energy and environmental regulation on firms. A recent review of the literature finds broad evidence of small, statistically significant negative effects in the short run on productivity, especially in sectors most exposed to regulatory costs (Dechezleprêtre & Sato, 2017). By contrast, there exists little empirical support for the notion that regulation can *enhance* productivity and, more broadly, firm competitiveness (for more discussion see Porter and Van der Linde

* Corresponding author at: Sloan School of Management, Massachusetts Institute of Technology, 100 Main Street, Building E62-482, Cambridge, MA 02142, USA.

E-mail address: vkarplus@mit.edu (V.J. Karplus).

https://doi.org/10.1016/j.chieco.2019.101364
Received 22 January 2019; Received in revised form 16 September 2019; Accepted 7 October 2019
Available online 11 November 2019
1043-951X/

(1995) and Ambec, Cohen, Elgie, and Lanoie (2013)).[1]

However, two questions remain underexplored. First, most prior work focuses on firms in developed countries; it is possible that some causes of investment inefficiencies, such as information barriers, credit constraints, subsidized energy pricing, and general allocative distortions (e.g., as described in Hsieh & Klenow, 2009), may be more salient for firms operating in developing country settings. Second, regulations differ in the ways that affect firms' private costs and benefits. Energy efficiency programs theoretically imply both private costs and benefits, because initial investments, while costly, also reduce a firm's effective cost of energy services per unit of real output. By contrast, many environmental policies require protective measures, such as end-of-pipe treatment technologies, which are costly to both install and operate. Thus, in theory it is more plausible that firms would reap productivity benefits from an efficiency-oriented intervention, relative to a regulation that uniformly imposes clean-up costs and yields distributed benefits to society that firms do not capture.

When it comes to energy efficiency in particular, there is an ongoing debate on whether or not firms optimally invest. Industry, government, and think tank analysis has claimed that energy efficiency has great potential to address environmental and security concerns, and has called for interventions to lower the barriers to realizing these gains, as described in Allcott et al. (2012). Building on Hausman (1979), prior work has largely focused on characterizing investment decisions by individuals; a limited body of research in applied microeconomics has empirically examined energy-related investments by firms (see for example DeCanio, 1993; DeCanio & Watkins, 1998; Ryan, 2018). Proposed explanations for this so-called “energy efficiency gap” include information barriers, transaction or ancillary costs, credit constraints, principal-agent conflicts, behavioral explanations, or a combination of multiple factors (see Gillingham and Palmer (2014) for a review of the literature). These factors may be more prevalent or pronounced in developing countries. Examining effects on firm productivity are one way of measuring the existence and extent of these inefficiencies.

Whether or not mandating increases in a firm energy efficiency helps or harms productivity growth is ultimately an empirical question. In this analysis, we estimate the impact of a program in China designed to raise energy efficiency, the Top 1000 Firms Energy Conservation Program (here “T1000P”), on firm productivity in the iron and steel industry. The program represented the first attempt of China's state leadership to mandate increases in industrial energy efficiency to support newly introduced, legally binding energy intensity targets for industries and provinces during the Eleventh Five-Year Plan (2006–2010) (Cao, Garbaccio, & Ho, 2009; Zhang, Aunan, Seip, & Vennemo, 2011). As is typical of many energy efficiency programs, the T1000P targeted the country's largest energy users. Within the industrial sector, the Chinese iron and steel industry has been both a major engine of economic expansion and a significant source of local air pollution and carbon dioxide emissions due to its high direct use of coal (He, Zhang, Lei, Fu, & Xu, 2013; Lin, Wu, & Zhang, 2011). We first estimate total factor productivity change in the population of Chinese iron and steel firms using both cost function and production function approaches at the sub-industry (four-digit industry code) level. Then, in a difference-in-difference (DID) design, we evaluate the program's effect on the productivity growth of treated firms. For identification, we use three main approaches (propensity score matching, variation in the inclusion threshold for provincial programs, and an instrument for program treatment) to isolate the effect of the program from potential confounders.

Our results are consistent with the existence of investment inefficiencies in China's iron and steel firms prior to the start of the program. On average, firms included in the program experienced faster TFP growth. We estimate that TFP grew on average by 6.4% in the industry as a whole. In our benchmark specification, we find that the average treated firm's TFP change is higher by 3.1 percentage points per year over the first three years of the program (2006 to 2008). The evidence that T1000P firms benefitted is robust to a range of alternative specifications and tests that explicitly consider attrition, the influence of a potentially confounding program, and the use of a production function instead of a cost function to estimate productivity. Our cost function results further distinguish the contributions of technical and scale efficiency change, which are positive, significant, and contribute about equally to the overall effect of the program on TFP change in our main specification.

Our findings imply several lessons for the future design of energy efficiency interventions in developing countries. First, the targets assigned by the T1000P may have raised the average treated firm's perceived net benefit of mobilizing external resources to support energy efficiency investments. Prior work has suggested that when investment inefficiencies exist, “mandating or subsidizing” energy efficiency could increase welfare (Allcott et al., 2012). In our study, all firms had the opportunity to apply for financial assistance to support energy efficiency investments. However, only treated firms faced binding targets and incentives to meet them. Our results suggest that targets prompted T1000P firms to direct more internal effort to evaluating their organization's energy use and to pursuing external sources of support to raise efficiency, relative to firms that did not face targets. Second, in addition to raising the perceived net benefit of energy efficiency, the program may have made payoffs to investments more certain. Managers undertaking energy efficiency investments, especially for the first time, face the risk that realized savings could be lower than anticipated at the outset, a phenomenon consistent with empirical evidence in other settings (Burlig, Knittel, Rapson, Reguant, & Wolfram, 2017; Fowlie et al., 2018; Davis, Fuchs, & Gertler, 2014). Under the T1000P, firms received benefits in the form of accolades and avoided unfavorable treatment by participating in the program and achieving targets. The relative certainty of these near-term benefits may have increased the *expected* net benefit associated with energy efficiency investments in treated relative to control firms.

[1] The Porter Hypothesis proposes that, in the medium to long run, firms facing environmental regulation would accelerate technological progress and gain a positive productivity advantage.

2. Literature and empirical setting

2.1. Impact of energy efficiency on firm productivity

The vast majority of studies that estimate the impact of energy efficiency programs examine developed country settings, typically focus on individual or household decisions, and do not consider productivity effects. These studies have found evidence that the engineering projections of recurring savings used to justify energy efficiency programs have been overestimated. One study of a weatherization program in the U.S. estimated that model-projected savings exceeded actual savings by 2.5 times, while the actual rate of return to the program was negative (Fowlie et al., 2018). A recent study of housing rental markets in the U.S. found different rates of energy efficiency investment, depending on whether tenants or owners paid for energy (Myers, 2018). Such studies in developing countries are rare, but similar to Fowlie et al. (2018) suggest savings are initially overstated. For instance, Davis et al. (2014) found that a large-scale appliance replacement program in Mexico achieved only one-quarter of the energy savings projected *ex ante*. All of these studies involve individual or household decisions; in a study involving schools, Burlig et al. (2017) found that energy efficiency projects resulted in savings of 2–5% on average, while realized savings were 50% lower than *ex ante* projected savings.

In the case of firms, a multiple studies have found negative productivity effects of environmental regulations (e.g. pollution limits) on firms (Gollop & Roberts, 1983; Gray & Shadbegian, 2003; Greenstone, List, & Syverson, 2012; Koźluk & Zipperer, 2013). Gollop and Roberts (1983) focus on sulfur dioxide emissions limits in the US electric power industry by estimating a cost function using observations of 56 electric utilities between 1973 and 1979. They find a negative effect of the regulation on TFP growth of 0.59 percentage points per year, mainly due to the higher cost of low sulfur fuel. Gray and Shadbegian (2003) study the effect of the US Clean Air and Clean Water Acts of the early 1970s on 116 pulp and paper mills in the United States from 1979 to 1990. They find that higher pollution abatement operating costs lowered TFP levels by about 2.6% annually, and that the magnitude of this effect depended on a plant's technology. Greenstone et al. (2012) study the effect of the Clean Air Act Amendment on TFP levels in a large sample of US manufacturing plants over the period from 1972 to 1993. TFP levels of polluting plants located in non-attainment counties (subject to more stringent pollution limits) are found to be significantly negatively affected in the range of 2.6 to 4.8% on average (Greenstone et al., 2012).

There is little empirical evidence of the magnitude and direction of the firm-level productivity impact of efforts to raise industrial energy efficiency anywhere in the world. Despite this, energy efficiency programs have been widely adopted in developing countries on the basis that they will not impede economic growth, because these programs index energy saving targets to economic activity. Indeed, the term “energy saving” can be misleading because economic growth at a rate that exceeds the rate of decline in energy intensity in a given period will result in an increase in absolute energy use. Many developing countries, including China and India, have adopted efficiency or intensity-based targets to implement their national climate change mitigation goals. The closest setting to ours, which uses a different (experimental) methodology, is Ryan (2018), which finds manufacturing plants in India respond to an energy efficiency treatment by hiring more skilled labor and increasing output, resulting in higher energy use in the treatment relative to the control group.

To study the productivity impacts of state-mandated increases in energy efficiency in a developing country setting, we choose to focus on China in the mid-2000s, when its central government introduced an energy efficiency program for the first time. China experienced rapid economic growth averaging 10% in real terms between 1980 and 2010, and emerged by the end of this period as the world's largest energy-using nation. Prior studies of China's energy efficiency programs have quantified changes in energy efficiency or emissions levels (see for example Xu and Lin (2016) and Zhou and Yang (2016)), but not firm productivity. Our work also contributes to a large literature that estimates trends and probes determinants of productivity growth in China (see Tian and Yu (2012) and Wu (2011) for an extensive review).

2.2. Intuition

The impact of an energy efficiency program on firm productivity depends on how firm managers trade off the upfront cost of energy efficiency investments against discounted energy savings. Below, we define investment inefficiencies with respect to energy efficiency capital, and show how positive productivity effects could result from reducing the magnitude of these inefficiencies. We begin with the basic framework from Allcott et al. (2012), which suggests that firms will invest in energy efficiency if the upfront cost c is lower than the discounted benefits, formalized in Eq. (1):

$$\frac{g_i p m_i (e_0 - e_1)}{1 - r} - \xi > c \tag{1}$$

where e_0 is the baseline energy intensity and e_1 is the energy intensity of the efficient variant of a production technology. Energy prices p and unobserved opportunity costs or benefits ξ are assumed constant across firms, while a firm's taste for an energy efficient good is captured by m_i. The parameter g_i captures the extent to which a firm under- or over-values energy savings, and serves as a sufficient statistic for a broad range of investment inefficiencies (Allcott et al., 2012). The only difference from the earlier Allcott and Greenstone (2012) framework is that here we explicitly allow g_i to be firm specific, similar to the setting in Allcott, Mullainathan, and Taubinsky (2014). For instance, agents may misperceive expected energy cost savings; savings may be genuinely uncertain for a fixed amount of investment; agents may not reap savings because buyers of their products lower prices to reflect the reduction in costs; or firms may lack access to credit at the risk-adjusted discount rate.

We are interested in how the productivity impact of mandating an increase in energy efficiency depends on a firm's value of g. Governments do not observe g when designing energy efficiency programs. If $g = 1$ or $g > 1$, increasing energy efficiency will lead to overinvestment in energy efficiency and TFP loss. However, common assumptions supported by empirical findings suggest $g \le 1$ (Allcott et al., 2014). Therefore, if an energy efficiency target requires a firm for which $g < 1$ to raise its energy efficiency by investing at a level corresponding to $g = 1$, its costs would fall because its additional investment has unlocked discounted energy savings that, contrary to the firm's original assessment, more than offset the firm's upfront cost, hence increase the TFP. If the efficiency target exceeds the optimal level, the firm will overinvest in energy efficiency, and TFP may decrease compared to the level without policy.

In practice, g varies across firms. An intervention that mandates improvements in energy efficiency will therefore affect different firms differently. In a population of heterogeneous targeted firms, the net impact of mandating increased efficiency on the productivity of the average firm is ultimately an empirical question.

2.3. Industry context

We estimate the impact of an energy efficiency program on firm productivity in China's iron and steel industry in the mid-2000s. Our research setting typifies the conditions of an energy-intensive industry in a developing country during a period of rapid economic expansion, although we recognize that the scale, pace, and drivers of China's economic expansion are in many respects unique. Between 1985 and 2013, the industry's output grew on average by 10.8%, and constituted 49.8% of global iron and steel output in 2013 (IISI, 1986, 2002; WSA, 2014). Total energy consumption in the industry increased at a slightly slower rate of 8.7% per year between 1985 and 2010 (Lin & Wang, 2014). In 2013, the iron and steel industry accounted for 29% of the total energy used in China's manufacturing sector and 23.6% of total energy used in industry (NBS, 2014). Iron and steel production is also one of the country's major sources of pollution (He et al., 2013). During the study period, the industry ranked third as a source of carbon dioxide emissions in China (after power generation and cement), accounting for roughly 10% (Zeng, Lan, & Huang, 2009). The iron and steel industry's high energy consumption is in large measure attributable to the characteristics of its production processes. Compared to the iron and steel industries of developed nations, aggregate measures (above the firm level) suggest that China's industry uses energy more intensively in production (He et al., 2013; Ross & Feng, 1991; Zhang & Wang, 2008).

2.4. Energy efficiency treatment

The central government launched the national Top 1000 Firms Energy Conservation Program (T1000P) at the start of the Eleventh Five-Year Plan (FYP) in April 2006 (Zhou, Levine, & Price, 2010). It required the country's largest 1008 energy-consuming industrial firms, defined as firms consuming a minimum of 180,000 tons of coal equivalent (tce) in 2004, in nine industries to substantially reduce their energy intensity (the ratio of energy used to output produced) (Price, Wang, & Yun, 2010). The energy consumption cutoff was the sole criterion used to determine program participation. The program was designed to support the Eleventh FYP's (2006–2010) target of reducing national energy intensity (energy use per GDP) by 20% (State Council, 2006). The government rolled out the program rapidly in response to the observation that by 2004, a long decline in the energy intensity of China's economy, largely attributed to the gradual dismantling of central planning since the 1980s, appeared to have reversed (Ke et al., 2012). One study notes that energy conservation was not an industry priority before the mid-2000s (Zhang & Wang, 2008). The program was administered by the National Development and Reform Commission (NDRC), China's major economic planning body, and was the first of its kind in China.

Energy intensity targets for firms participating in the T1000P were set in agreements between the provincial government and the firm (Price et al., 2010; Zhao, Li, Wu, & Qi, 2014). Targets were initially assigned in proportion to each province's energy-saving targets, which during the Eleventh FYP were very similar and corresponded closely to the national target of reducing energy intensity by 20% by 2010, relative to 2005 levels. Targets were then adjusted based on a firm's pre-regulation share in the energy use of all firms included in the T1000P (Zhao et al., 2014; Zhao, 2016). As the program was set up very rapidly, the target setting process was not based on an assessment of individual firms' energy conservation potential or costs (Price et al., 2010; Price et al., 2011).

In each program year, covered firms self-reported their progress toward energy intensity targets directly to the Chinese National Bureau of Statistics (NBS) following predefined reporting standards (Zhou et al., 2010). For many firms, this was the first time reporting was required. Provincial governments evaluated data quality and firm compliance on an annual basis. Assessment included short on-site inspections, but was mainly based on firms' self-reporting, due to limited inspection capacity and the complexity of calculating the energy conservation indicator (Li, Zhao, Yu, Wu, & Qi, 2016; Zhao et al., 2014, 2016). If discovered, fraudulent reporting could lead to a criminal investigation (Zhou et al., 2010).

Penalties for non-compliance were both financial and non-financial. While the nature and extent of penalties varied at the provincial level, all firms faced heightened public scrutiny. The government publicly released the list of firms in the T1000P, drawing heightened media attention (Price et al., 2011). Some firms reportedly implemented incentive payments for their staff conditional on the achievement of energy conservation targets, which also included salary limits if compliance was not achieved (Zhao et al., 2014). Furthermore, state administrators evaluated leaders of state-owned enterprises (SOEs) and local government officials based on their achievement of the T1000P energy-saving targets (Li et al., 2016; State Council, 2007; Zhao et al., 2014). For the first time, energy conservation achievement was included in the personnel appraisal system (Zhou et al., 2010). Leaders of state-owned firms or local governments could be denied promotions and honorary titles if targets were not met (Koakutsu, Usui, Watarai, & Takagi, 2012). Non-compliant firms were required to submit a written report to government authorities specifying a plan for rectifying non-compliance

(Zhao et al., 2014).

The T1000P was introduced on top of a program that provided government support for energy-efficiency investments to all firms, not only those participating in the T1000P (Ke et al., 2012; Price et al., 2010; Zhao et al., 2014). These resources included consulting support and energy audits, information benchmarking firm performance to peers, and information on how to apply for financial assistance for approved energy retrofits. Provincial energy conservation centers were established to make these resources available to firms (Qi, 2013).

By its initial benchmarks, the program exceeded expectations. At the outset, the program targeted a "reduction" of 100 mtce by 2010, relative to estimated energy use assuming no change in firms' energy intensity (NDRC, 2006). Target achievement was reported early in 2008 when the NDRC announced savings of approximately 106 mtce (Ke et al., 2012), similar to Price et al. (2011), who estimated savings of 124 mtce by 2008. By 2010, estimates of energy conservation for the entire five-year period far exceeded the initial target: Zhao (2016) reported savings of 165 mtce, while Ke et al. (2012) found savings of 150 mtce. Program compliance was reportedly high: at the conclusion of T1000P in 2010, only 1.7% of the firms did not comply (NDRC, 2011).[2] Citing these measures of progress achievement, the NDRC expanded the T1000P to the Top 10,000 Enterprises Program in the Twelfth FYP (2011-2015) (Zhao, 2016).

3. Data

3.1. Industrial survey

Our data set is assembled using firm entries in the Chinese Annual Industrial Survey (CAIS) for the years 2003 to 2008, which is compiled by the NBS. The CAIS represents the most extensive source of firm level information on the Chinese manufacturing sector available to researchers. It contains yearly observations of the balance sheet, income statement, and other non-financial information for all industrial firms registered in China with a yearly sales value higher than 5 million Chinese renminbi (RMB), which corresponds to approximately 800,000 US dollars, and all SOEs (independently of their sales value). Most firms are single plant firms (Brandt, Van Biesebroeck, & Zhang, 2012). All costs and output values are deflated to a reference year (1998) using four-digit industry-specific input and output deflators, which were used by Brandt et al. (2012) and provided by Johannes Van Biesebroeck at KU Leuven. Information on firms participating in the T1000P was obtained from the NDRC. Of these firms, 1001 out of 1008 (99.3%) were successfully matched with the CAIS. While the prices of labor and capital are derived from information contained in the CAIS, the price of material was not provided. The subindustry- (iron and steel, steel rolling, and ferroalloy smelting) and province-specific annual price of material is calculated based on information on subindustry inputs and outputs obtained from the NBS (2007), coal prices and electricity prices extracted from CEIC (2015) and iron ore prices from CCM (2015). These prices then are adjusted for inflation using an overall price deflator provided by the NBS (2013).

3.2. Characteristics of treated and non-treated firms

The CAIS records a total of 13,278 firms in the iron and steel industry (or more precisely, in the ferrous metal smelting and rolling industry) over the period of 2003 to 2008. In this sample, 5340 firms are considered for the empirical analysis, after cleaning and panel construction as described in Appendix A. The panel of firms is unbalanced with 2047 observations (or 38.3%) forming a balanced panel, while 37.3% of the sample was observed for five years, 18.4% for four years, 5.0% for three years, and 0.9% for two years. Descriptive statistics of the 5340 firms for the full sample period are given in Table 1 in columns 1 to 4. Firm heterogeneity with respect to several of these variables is large. For example, the 25th percentile of gross output value is 7.3 times smaller than the 75th percentile value, while it is 4.5 for the number of people employed. The iron- and steelmaking subindustry accounts for 18.3% of the observations, steel rolling for 64.3%, and ferroalloy smelting subindustry for 17.4%. Furthermore, 0.6% of the observations are central SOEs, 9.4% are local SOEs, and 90.0% are non-SOEs, whereby a firm is defined as being an SOE if its controlling (minority or majority) shareholder is the state. In China's iron and steel industry, state ownership fell dramatically during the period of economic reform, resulting in only a limited number of SOEs remaining by our study period.

4. Empirical strategy and identification

In our base specification, we implement a two-stage approach to estimate the effect of the T1000P on firm productivity growth. We first specify a cost function as the basis for computing total factor productivity (TFP) change and for further resolving its subcomponents, technical change and scale efficiency change.[3] Second, a difference-in-difference (DID) approach is used to estimate

[2] 881 firms remained in the program at the end of the T1000P in 2010, and 15 firms were found non-compliant. The share of non-compliant firms was 3.9%, 3.1%, and 1.7% in 2008, 2009 and 2010, respectively (NDRC, 2009; NDRC, 2010; NDRC, 2011).

[3] TFP changes cumulatively translate into differences in a firm's TFP (see Ehrlich, Gallais-Hamonno, Liu, and Lutter (1994) for more discussion). Using the same underlying data set as ours, find TFP change to be more informative than TFP levels, in the sense that between 1998 and 2007 surviving entrants in the Chinese manufacturing sector were found to be selected based on TFP change rather than TFP levels. An additional rationale for using changes rather than absolute levels is that the values of changes across treatment and control firms are more comparable, as discussed in Wagner et al. (2014). For robustness, we estimate program effects on TFP levels using a production function, and results are similar.

Table 1

Descriptive statistics of firms. All values are annualized.

	Years 2003 to 2008				Years 2003 to 2005 (pre-regulation period)		
	All firms				Treatment group	Control group	Difference
	Mean	Std. dev.	Min.	Max.	Mean	Mean	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gross output (mRMB)	353.8	2226.1	0.016	89,784.2	4795.8	115.4	4680.4
Employees	506.2	3202.1	8	120,628	9009.8	225.0	8784.8
Total assets (mRMB)	340.0	3113.6	0.324	127,167.6	5989.5	88.9	5900.6
Current assets (mRMB)	129.4	1013.2	−2.181	38,334.2	2263.3	47.1	2216.2
Intermediate inputs (mRMB)	298.1	1810.3	0.001	73,139.0	3909.5	98.7	3810.8
Age	7.85	8.78	0	108	22.19	6.48	15.71
Exporter (1 if exporting)	0.096	0.295	0	1	0.273	0.067	0.206
Total costs C (mRMB)	335.9	2158.5	0.482	90,363.0	4595.1	107.3	4487.7
Capital price P_K (kRMB / K)	0.245	1.297	0.000	93.831	0.145	0.232	−0.087
Labor price P_L (kRMB / L)	15.77	13.96	0.030	618.37	20.62	12.79	7.83
Intermed. inputs price P_M (index)	156.14	38.12	68.31	313.80	133.15	137.13	−3.98
Profitability	0.030	0.091	−2.722	2.102	0.046	0.027	0.019
# firms / # observations	5340 / 27,076				148 / 410	5192 / 12,173	
Sub-industry shares in [%]: iron- and steelmaking / steel rolling / ferroalloy smelting:	18.3 / 64.2 / 17.4				44.9 / 45.9 / 9.3	18.2 / 64.5 / 17.3	
Share in [%] of central SOE / local SOE / non-SOE:	0.6 / 9.4 / 90.0				3.7 / 40.5 / 55.9	0.5 / 4.0 / 95.5	
Share in [%] of regions East / Central / West:	59.2 / 23.4 / 17.4				45.6 / 34.1 / 20.2	59.5 / 23.5 / 17.0	
Distribution of firm size (number of employees) in [%] of observations in intervals [0;50], (50;100], (100;500], (500;1,000], (1,000;5,000] and more than 5,000:	24.4 / 24.6 / 39.9 / 5.5 / 4.1 / 1.5				0.0/0.2/2.7/9.8/49.8/37.6	26.3/25.6/40.2/5.4/2.3/0.2	

Note: This table shows descriptive statistics of the sample (columns 1 to 4) for the period 2003 to 2008 and conditional on treatment (columns 5 and 6) for the pre-regulation period of 2003 to 2005. Data is at firm level with monetary values given in real 1998 values. *Profitability* is the ratio of total profits to gross output. Column 7 shows the difference in means of the treatment and control group.

the effect of the program on TFP change.[4] The disaggregation of TFP change into its subcomponents yields further insights in terms of whether firms responded to the regulation via technical change TC (e.g., by installing new machinery) or scale efficiency SEC (e.g., by realizing economies of scale as output increased).

4.1. Formulation of the cost function and derivation of TFP change

TFP analysis can be performed using parametric and non-parametric approaches. For instance, within the parametric methods, it is possible to use a production or a cost function/cost frontier approach, whereas within the non-parametric methods, it is possible to use index numbers such as the Törnqvist index or Data Envelopment Analysis. To derive TFP change, here we use a parametric approach based on the estimation of a cost function.[5] However, as robustness check, we also compute the TFP change of the firms using two index number approaches (Törnqvist index, as described in Coelli, Estache, Perelman, and Trujillo (2003), and the index used by Allcott et al. (2012), and three methods based on the parametric estimation of a production function proposed by Wooldridge (2009), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015).

[4] As discussed in Wang and Schmidt (2002), using a two-stage approach is not always an option solution. For this reason, as a robustness test, we present in Section 5.9 the results using a one-stage approach based on the measurement of TFP change using two productivity index numbers, and they are robust.

[5] One advantage of a cost function over a production function is that fewer explanatory variables are likely to be endogeneous. In the estimation of a production function all the inputs usually are endogenous. In the estimation of a cost function the variable that potentially could be endogenous is the output. In model specification (3) we cannot exclude *a priori* that the output is endogenous. However, we can assume that the input prices are exogenous. This is supported by the fact that in China input prices are either set by the state (e.g. for electricity) or determined in commodity markets (e.g. for coal), rendering them plausibly exogenous to firm decisions. In order to solve the potential endogeneity of the output variable we could use an IV or a GMM approach. In our case, also because we don't have a powerful instrument, we think that the use of a fixed effects approach is already attenuating this potential problem. Moreover, to probe the robustness of our results, we also estimate and perform the empirical analysis using values for the TFP change obtained via a non-parametric method, as well as the methods of Wooldridge (2009), Levinsohn and Petrin (2003), and Ackerberg et al. (2015). For a review of all methods to measure the TFP change see del Gatto, Di Liberto, and Petraglia (2011). For a presentation of the use of a cost function see Lowell and Kumbhakar and Lovell (2000).

Following Coelli et al. (2003) and applying the quadratic approximation lemma of Diewert (1976), as proposed by Orea and Luis (2002), TFP change (TFPC) of firm i between two periods t and $t-1$ can be estimated as

$$TFPC_{it} = \ln\left(\frac{TFP_{it}}{TFP_{it-1}}\right)$$
$$= \underbrace{\frac{1}{2}[(1-e_{it}) + (1-e_{it-1})]\cdot(\ln Y_{it} - \ln Y_{it-1})}_{SEC_{it}}$$
$$- \underbrace{\frac{1}{2}\left(\frac{\partial \ln C_{it-1}}{\partial t} + \frac{\partial \ln C_{it}}{\partial t}\right)}_{TC_{it}} \quad (2)$$

TFP change is further disaggregated into the two subcomponents of scale efficiency change (SEC, the term on the second line) and technical change (TC, the term on the third line). Total costs are represented by C_{it} and the single output is Y. Output elasticities (which are the inverse of the returns to scale elasticity) at a data point are estimated as described in Coelli et al. (2003).

A calculation of TFP change according to Eq. (2) necessitates the empirical specification of a cost function for the Chinese iron and steel industry, which can be divided into three sub-industries s: iron- and steelmaking, steel rolling, and ferroalloy smelting. To improve comparability, we empirically estimate a separate cost function for each sub-industry. In this study, we assume a sub-industry s-specific production process as follows:

$$C_{it} = c^s(Y_{it}, P_{L,it}, P_{K,it}, P_{M,srt}, t) \quad (3)$$

Total costs C are defined as the sum of intermediate input costs (which includes all energy costs), labor costs and capital costs. Costs include depreciation and interest expenses and an assumed opportunity cost of equity of 3%.[6] The single output Y is deflated gross output.[7] The price of labor P_L is represented by the ratio of the sum of wage and welfare payments to the number of employees. The price of capital P_K is defined as capital costs divided by the real capital stock. The calculation of the real capital stock is based on the perpetual inventory method (see Appendix A). The main materials used in the production of iron and steel are coal, coke, iron and electricity. The subindustry s- and province r-specific price of material P_M is derived via a Törnqvist price index of these four main material inputs. A linear time trend t is added to the cost function in order to control for technical change. All costs and output are deflated to 1998 values using the respective input and output deflators, as described in Appendix A. Descriptive statistics of the main covariates are given in Table 1.

For the estimation of Eq. (3) we use a translog functional form (Berndt & Christensen, 1973; Christensen, Jorgenson, & Lau, 1973), since this form does not impose *a priori* restrictions on the technology parameters. The subindustry s-specific cost functions are specified as

$$c_{it} = \alpha_0^s + \alpha_i + \boldsymbol{\beta}^{s'}\mathbf{x}_{it} + \varepsilon_{it}$$
$$= \alpha_0^s + \alpha_i + \beta_Y^s y_{it} + \sum_{Z=\{K,L\}} \beta_Z^s p_{Z,it} + \beta_M^s P_{M,srt}$$
$$+ \frac{1}{2}\left(\beta_{YY}^s y_{it}^2 + \sum_{Z=\{K,L\}} \beta_{ZZ}^s p_{Z,it}^2 + \beta_{MM}^s P_{M,srt}^2 + \beta_{tt}^s t^2\right)$$
$$+ \sum_{Z=\{K,L\}} \beta_{yZ}^s y_{it} p_{Z,it} + \beta_{YM}^s y_{it} P_{M,srt} + \beta_{KL}^s p_{K,it} p_{L,it} + \sum_{Z=\{K,L\}} \beta_{ZM}^s p_{Z,it} P_{M,srt}$$
$$+ \beta_t^s t + \beta_{Yt}^s y_{it} t + \sum_{Z=\{K,L\}} \beta_{Zt}^s p_{Z,it} t + \beta_{Mt}^s P_{M,srt} t + \varepsilon_{it}, \quad (4)$$

with lower case letters y and p indicating output and prices in natural logarithms. The intercept α_0 represents total costs at the approximation point. Firm fixed effects are captured by α_i and control for firm-specific, time-invariant unobserved heterogeneity.[8] The error term is given by ε_{it}. Sub-industry-specific median values of the explanatory variables are chosen as approximation points of the translog cost functions. Eq. (4) is estimated using a fixed effects estimator, that is, running OLS on $c_{it} - \bar{c}_i = \boldsymbol{\beta}'(\mathbf{x}_{it} - \bar{\mathbf{x}}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i)$ using Huber (1967)/White (1980) cluster robust sandwich estimates at the firm level (accounting for both heteroskedasticity and serial correlation), where $\bar{c}_i = T_i^{-1}\sum_t c_{it}$. The variables $\bar{\mathbf{x}}_i$ and $\bar{\varepsilon}_i$ are constructed analogously.

[6] Opportunity costs of equity of 3% follow from assumptions of a 20% return to capital, 12% depreciation, and an interest rate of 5%. For an overview of returns to capital in China, see Bai, Hsieh, and Qian (2006).

[7] We are aware that a physical measure of TFP would be superior to a value-based measure, because quality or market power varying over time in a heterogeneous way could also influence these values. Unfortunately, as in several other studies estimating a cost or production function for industrial sectors (see e.g., Allcott et al. (2012), Gray and Shadbegian (2003), or Levinson and Petrin (2003)), we do not have quantity or quality-related information.

[8] As discussed in Lee, Stoyanov, & Zubanov (2019), when measuring TFP changes it is important to consider firm fixed effects in the econometric analysis. These fixed effects, for instance, should capture time-constant conditions affecting the outcome of a firm, e.g., geographic heterogeneity like a favorable geographic location close to iron and coal mines or ports (Greenstone, 2002), preferential political treatment, regional differences in the application and enforcement of regulation targets etc.

4.2. Identification strategy

To estimate the effect of the T1000P on TFP change using a difference-in-difference (DID) approach, two important assumptions need to be satisfied for identification (Bertrand, Duflo, & Mullainathan, 2004). First, a firm's selection into the energy conservation program should be exogenous. Second, prior to treatment, there should be a parallel trend in the outcome variable for both treatment and control firms.

Two lines of reasoning support the validity of the first assumption. First, the program originated at the central government level, without consent or input from firms regarding its creation or initial design. For instance, the energy use threshold (180,000 tce/year) that determined program inclusion was relatively arbitrary. Importantly, firms could not self-select into the program. A nationwide program focused on raising energy efficiency was new to China at the time, issued in response to unexpected evidence that China's multi-decadal decline in energy intensity had reversed in 2004, making it unlikely that firms anticipated the program's introduction; previously, many firms did not even emphasize or track energy use (Ke et al., 2012).

Second, we might worry that treated firms differed systematically from untreated firms in observed or unobserved ways that affected their response to the program. For this reason, we use several alternative approaches to construct a more comparable control group, including matching and stratification. We include firm fixed effects α_i to avoid bias in case time-invariant unobserved firm level heterogeneity is not orthogonal to the treatment effects or other covariates. We also include time fixed effects, $\boldsymbol{\theta}_t$, to capture year-specific shocks to firm performance common to all firm, for instance, output market disruptions or political shifts on a national level.

For the DID approach to yield valid results, the second assumption of parallel trends prior to treatment must be satisfied. If the parallel trend assumption holds, the average effect of the regulation on TFP change in treated firms, or the average treatment effect on the treated (ATT), can be identified as

$$TFPC_{it} = \alpha_0 + \alpha_i + \Theta_t + \beta_{ATT}\tau_i\rho_t + \boldsymbol{\gamma}'\mathbf{X}_{it} + \Theta_t\pi_i + \varepsilon_{it}. \tag{5}$$

This procedure can be followed analogously to analyze the ATT for TC and SEC by replacing TFPC with one of these other outcome variables. Treatment status is captured by the binary variable ρ_t, equal to one for all regulation periods and zero otherwise, with the first year of treatment being 2006. The binary variable τ_i indicates whether or not a firm was part of the treatment group. The ATT is estimated by the coefficient β_{ATT}. Vector $\mathbf{X}_{it}$ contains variables for ownership structure and firm size to control for time-varying heterogeneity affecting firm performance. Size effects are controlled for by the natural logarithm of the number of firm employees. Ownership-related effects are measured by a binary variable differentiating between SOEs and non-SOEs. Province-year effects $\boldsymbol{\theta}_t\pi_i$ control for province π_i- and year θ_t-specific shocks. The two-way fixed effects model in Eq. (5) is estimated again using cluster robust sandwich estimates at the firm level. This avoids a potential downward bias in the estimated standard errors of the treatment effect due to uncontrolled positive serial correlation.

We test the parallel trends assumption by estimating the following expression, with the year of implementation of the regulation indicated by T^*:

$$TFPC_{it} = \alpha_0 + \alpha_i + \beta_t t + \beta_t^{tr} t_i^{tr} + \boldsymbol{\gamma}'\mathbf{X}_{it} + \Theta_t\pi_i + \varepsilon_{it} \mid t < T^*. \tag{6}$$

Eq. (6) estimates a time trend for the treated group (indicated by "*tr*"), $t_i^{tr} = t\tau_i$, and uses observations of the pre-treatment period only. Our implementation follows Autor (2003), Lachowska and Myck (2018), and Angrist and Pischke (2008). The finding of no significant difference in the productivity trend in treatment versus control firms supports the null hypothesis of $\hat{\beta}_t^{tr} = 0$, i.e., is consistent with the parallel trends assumption.

$$TFPC_{it} = \alpha_0 + \alpha_i + \Theta_t + \theta_{i,2005}^{tr} + \beta_{ATT}\tau_i\rho_t + \boldsymbol{\gamma}'\mathbf{X}_{it} + \Theta_t\pi_i + \varepsilon_{it} \tag{7}$$

A similar test shown in Eq. (7) probes for pre-treatment effects $\theta_{i,2005}^{tr} = \tau_i\theta_{2005}$ in 2005. If $\hat{\theta}_{i,2005}^{tr} = 0$ is not rejected, our assumption of no pre-treatment effects remains unchallenged. In contrast to Eq. (6), Eq. (7) makes use of the full panel of observations. It also includes an estimation of the ATT. Firm fixed effects α_i capture the information in the covariate (treatment status) as well, and therefore it is not included in the above three specifications. Tests for a parallel trend and pre-treatment effects in TC and SEC are conducted analogously by replacing TFPC with one of these variables.

5. Results

Table 2 presents estimated values of TFP change (TFPC) and its subcomponents, technical change (TC) and scale efficiency change (SEC). Results were derived using the estimated cost function coefficients (reported in Table C.2 in the Appendix). The cost functions for the three subindustries are well behaved, i.e. monotonic (see Table C.3 in the Appendix) and quasi-concave (Table C.4 in the Appendix). Details of the monotonicity and quasi-concavity test are described in Appendix D.

TFPC is positive for all three subindustries between 2003 and 2008. In the main specification, TC contributes about 60% to the average TFP change, while the remainder is attributed to SEC. Firms in all three subindustries on average were found to exhibit positive returns to scale, see Table C.5 in the Appendix). The iron- and steelmaking subindustry shows the highest average TFP growth, followed by the steel rolling and ferroalloy smelting subindustries. TC is the dominant contributor to TFP growth in the steel rolling subindustry, while TC and SEC are roughly equally important in the ferroalloy smelting industry.

Table 2
Descriptive statistics of estimated TFPC, TC and SEC.

	Mean	Median	Std. dev.	10% perc.	90% perc.
Full period (2003–2008)					
All subindustries		[# firms: 5340 / # observations: 27,076]			
TFPC	0.064	0.056	0.108	−0.028	0.171
TC	0.041	0.042	0.039	0.001	0.085
SEC	0.023	0.015	0.098	−0.053	0.110
Iron- and steelmaking		[# firms: 1025 / # observations: 4968]			
TFPC	0.100	0.086	0.119	−0.009	0.222
TC	0.064	0.068	0.037	0.016	0.108
SEC	0.035	0.023	0.111	−0.058	0.133
Steel rolling		[# firms: 3353 / # observations: 17,391]			
TFPC	0.058	0.051	0.085	−0.016	0.141
TC	0.039	0.040	0.022	0.011	0.066
SEC	0.019	0.013	0.081	−0.048	0.094
Ferroalloy smelting		[# firms: 962 / # observations: 4717]			
TFPC	0.051	0.053	0.155	−0.102	0.203
TC	0.024	0.030	0.069	−0.069	0.106
SEC	0.028	0.019	0.134	−0.073	0.149

Note: This table shows the descriptive statistics of overall and subindustry-specific mean TFPC, TC and SEC values for the period of 2003 to 2008. The overall values (first panel "All subindustries") are based on all observations of the sample, i.e. the three subindustries are implicitly weighted by their number of observations.

Table 3
Average treatment effects on the treated (ATTs) on TFPC, TC and SEC.

DD version:	DD–1		DD–2		DD–3	
ATT on TFPC	0.029	(0.004)	0.029	(0.004)	0.031	(0.005)
ATT on TC	0.013	(0.002)	0.013	(0.002)	0.012	(0.002)
ATT on SEC	0.017	(0.003)	0.016	(0.004)	0.019	(0.004)
# firms / # obs.	5340 / 21,736		5340 / 21,736		5340 / 21,736	
R^2 (TFPC / TC / SEC)	0.368 / 0.685 / 0.300		0.373 / 0.686 / 0.307		0.399/ 0.749 / 0.324	
F-statistic (TFPC / TC / SEC)			22.63/ 2.00 / 25.45		4.70 / 21.26 / 3.07	
Size	No		Yes		Yes	
Ownership	No		Yes		Yes	
Province × Year	No		No		Yes	

Note: This table shows the ATT on TFPC, TC and SEC between 2006 and 2008 using Eq. (5). Only estimates of β_{ATT} are shown. For the sake of conciseness, estimates of $\boldsymbol{\theta}_t$, γ and province-year effects are not shown. All three model specifications (DD–1 to DD–3) control for firm fixed effects. R^2 values of the estimations with TFPC, TC or SEC as the dependent variable are unadjusted. F-statistics show the joint significance of the size, ownership and province-year variables. Robust standard errors at the firm level are reported in parenthesis.

5.1. Effect of the regulation on TFP change

In what follows, we describe the effects of the T1000P on TFP change and its subcomponents using a DID approach. Focusing first on the two tests for parallel trends, we do not reject the hypothesis of parallel trends in TFP change and its subcomponents (see Table C.1). While the visual evidence in Appendix Fig. C.1 shows higher TFPC in the control group, the evolution of TFPC before the introduction of the policy does not differ statistically between treated and control firms: The regression results find no evidence of a statistically significant difference in pre-treatment TPFC, regardless of whether specification (6) or (7) is used, as shown in Table C.1, Panel A and B, respectively.

Treatment effects are estimated based on Eq. (5), and results are shown in Table 3.[9] We estimate and compare three model specifications (DD–1 to DD–3). The most parsimonious specification is the first model (DD–1). The second model (DD–2) additionally accounts for time varying heterogeneity related to ownership and size. Finally, the third model (DD–3) allows for year-specific shocks on a provincial level as local government officials and SOE leaders are evaluated annually on the achievement of their T1000P (and other) targets. Political shocks on a provincial level could potentially affect the enforcement of the regulation in a particular year. All three models include firm fixed effects and capture shocks at the national level via year fixed effects.

[9] We only have observations for three years when the regulation was in place. Prior literature suggests firms undertook energy conservation measures immediately, because the T1000P held firms accountable to targets in each program year. Zhao et al. (2014) study the behaviour of a power plant and observe this plant to have addressed most of the internal energy management reforms, including retrofits, by 2007. See also Price et al. (2010) for a description of energy efficiency measures undertaken by firms in the first year of the T1000P.

Table 4
Average treatment effects on the treated (ATTs) on TFPC in the extended analyses.

Extended analysis	DD–1		DD–2		DD–3	
Stratification by size	0.034	(0.005)	0.033	(0.005)	0.035	(0.005)
Stratification by ownership						
SOE	0.020	(0.009)	0.020	(0.010)	0.023	(0.012)
Non-SOE	0.024	(0.006)	0.020	(0.006)	0.023	(0.008)
Stratification by matching						
Matching approach 1	0.034	(0.009)	0.033	(0.009)	0.034	(0.010)
Matching approach 2	0.028	(0.009)	0.026	(0.009)	0.026	(0.009)
Provincial Energy Conservation Program	0.010	(0.014)	0.006	(0.015)	0.013	(0.018)
Sample Attrition	0.028	(0.004)	0.028	(0.004)	0.033	(0.005)
Instrument	0.032	(0.013)	0.035	(0.013)	0.036	(0.014)
Time-varying cofounder	0.030	(0.007)	0.027	(0.008)	0.034	(0.010)
Production Function	0.041	(0.023)	0.045	(0.024)	0.044	(0.026)

Note: This table summarizes the ATTs on TFPC of the extended analyses. The list of firms that were required to shut down, retire or update part of their production capacity included 344 firms in ten provinces, including Beijing, Hebei, Shanxi, Liaoning, Jiangsu, Zhejiang, Jiangxi, Shandong, Henan, and Xinjiang (NDRC, 2007). We successfully matched 115 firms with the CAIS data. Among the ten provinces, only Shanxi and Jiangsu had more than ten firms matched (47 and 40 firms respectively, in total 87 firms). Therefore, we limit the sample for this robustness check to these two provinces and remove these 87 firms from the main analysis to avoid any potentially confounding effect from this program.

We find positive, statistically-significant effects of the program on firm TFPC. Estimated treatment effects are robust in terms of sign, magnitude and significance across all three model specifications. The third model is our preferred specification, as it most extensively controls for potential cofounding factors. Model specification DD–3 estimates a 3.1 percentage point increase in the annual TFP growth rate due to the program, equivalent to an annual private benefit to the average treated firm of 148.7 million RMB/year.[10] Both subcomponents TC and SEC are significantly affected by the program and, on average, contribute about equally to the overall treatment effect. Fig. C.1 presents yearly TFPC, TC and SEC values for the treatment and control groups.

5.2. Extended analysis

We undertake six alternative approaches to estimate the effect of the T1000P on treated firms. These include approaches to ensure comparability of our control group (stratification and matching), to examine an analogous intervention at the provincial level to address concerns about differences in firm scale, to evaluate robustness to sample attrition and a possible confounder, to address endogeneity in program targeting by instrumenting for treatment, and to test sensitivity to the approach used in productivity estimation. Taken together, estimated treatment effects using these approaches are similar in sign and magnitude to our benchmark results. Results of these extended analyses are summarized in Table 4, and described below.

5.3. Classical sample stratification by size and ownership

Firm size and ownership could plausibly affect a firm's response to the T1000P. Firms that differ in size may face divergent possibilities to reduce energy intensity by altering the scale of operations; firms that differ in ownership may face different costs of capital or receive different levels of state assistance that could affect their responses to the program. We therefore stratify the sample by size and ownership to improve the comparability of treated and control firms (Greenstone, 2002; Meyer, 1995).[11] Every stratum contains a sufficient number of observations of treated firms to support statistical inference (see Table C.6 in the Appendix). First, we re-estimate model (4) based on a sample that only includes firms in the fourth quartile of the size distribution. As the main selection criterion of the T1000P was energy consumption of at least 180,000 tce in 2004, only large firms were exposed to the regulation and thus all treated firms belong to the fourth quartile of the size distribution. Shown in Table 4, our results are similar to our benchmark results in Table 3, with treatment effects being slightly larger for the sample including only large firms (see Table C.7 in the Appendix).

Second, we stratify by firm ownership. Prior work suggests that productivity levels differ among firms of different ownership, with SOEs less productive than private (non-SOE) firms (Dougherty, Herd, & He, 2007). Responses to the program may also differ. We

[10] An annual increase in TFP change of 3.1 percentage points corresponds to a yearly increase in TFP levels of treated firms of $e^{0.031} - 1 \cong 0.031$ compared to non-treated firms. Treated firms showed an average gross output of 4795.8 mRMB in 1998 values before the introduction of the regulation. Hence, on a per firm basis, a back-of-the-envelope calculation of average annual private benefits induced by the regulation through productivity gains for the period of 2006 to 2008 yields 148.7 million RMB (in 1998 values).

[11] In addition to size and ownership, we also test whether stratification with respect to subindustry affiliation and geographic region change our results (see Table C.7-Table C.12). Results using stratified samples are in line with the results of the benchmark specification of Table 3.

define SOE firms as have a controlling minority or majority state shareholder. We find that the program has a similar effect on TFP change and its subcomponents for both SOE and non-SOE firms (see Table C.8 in the Appendix), suggesting that ownership did not play a major role in shaping firm responses.

5.4. Sample stratification by matching

Matching is conducted using propensity score matching to the nearest two neighbors with replacement in year 2005, applying a caliper of 0.25 following Rosenbaum and Rubin (1985). We implement two approaches. In matching approach 1, we match based on gross output, with exact matching based on four-digit sector code categories. Matching is only conducted on firms employing more than 100 employees, as in 2005 no treated firm employed fewer. While achieving closer similarity in the control and treatment groups on a range of observables (see Table 1 and Table C.9 in the Appendix), the resulting sample is still unbalanced on firm size (see Table C.9 in the Appendix). In a similar setting, Wagner, Muûls, Martin, and Colmer (2014) face a similar situation, in which treated firms (in the European Union's Emissions Trading System) are systematically larger than control firms after matching. Therefore, in matching approach 2, we follow Wagner et al. (2014) and match based on capital intensity, i.e. the ratio of the real capital stock to output. Matching based on capital intensity helps to ensure that both groups of firms are similar in terms of production technology. Firm size is explicitly controlled for in the subsequent difference-in-difference estimation. Matching approach 2 also includes exact matching on the four-digit sector code to further improve comparability of production technology.

After applying the above procedure, 95 and 218 matched firms form the control groups for matching strategy 1 and 2, respectively (see Table C.6 in the Appendix). A comparison of Table 1 and Table C.9 in the Appendix shows that the matched control group is indeed more similar to the treatment group compared to the population average of non-treated iron and steel firms, while matching based on capital intensity yields two statistically-similar groups (see Table C.10 in the Appendix). Results of the second stage shown in Table C.11 and Table C.12 in the Appendix are robust to our benchmark results in terms of magnitude and significance for the overall effect on TFP change. While the magnitude of the effect on technical change decreased, the effect on scale efficiency change increased.

5.5. Provincial energy conservation program

To further ensure that our results are not driven by variables correlated with the large size threshold of the national program, we evaluate the productivity impact of a variant of the T1000P introduced by several provinces to augment the national program.[12] For these provincial energy conservation programs, energy use thresholds for inclusion were much lower than the national program and, importantly, varied across provinces. Hence, we can now compare firms of similar size and activity composition that differed only in treatment status across provincial boundaries as well as within provinces on either side of the energy use cutoff that determined program coverage, although our sample size is unfortunately reduced. We focus on 19 firms that were part of a provincial energy conservation program (and excluded from the national program). These firms on average were around 7.7 times smaller in terms of gross output compared to firms treated by the national program. Results shown in Table C.13 in the Appendix indicate robustness of our benchmark results on the program's overall effect on TFP change in terms of sign, with some reduction in magnitude. Significance, however, is mostly lost, which is perhaps unsurprising given the small sample size. The statistically-significant effect of the regulation on productivity is dominated by the channel of technical change, suggesting that the scale efficiency channel is less important in smaller firms of comparable size.

5.6. Sample attrition

Firm exit could result in an upward bias of treatment effects, if firms experiencing low TFP growth leave the sample over time (Baltagi, 2008). The sample without attrition is defined by the 2047 firms observed for the full range of years 2003 to 2008 plus the 1354 firms first observed in 2004 that are older than zero years and are subsequently observed until 2008. The re-definition of the sample necessitates a re-calculation of the approximation points of the subindustry-specific translog cost functions and a subsequent re-estimation of TFPC, TC and SEC values. The estimated coefficients of the subindustry-specific cost functions are given in Table C.14 in the Appendix and the firm performance estimates given in Table C.15 in the Appendix. The null hypothesis of a parallel trend in firm performance before the introduction of the regulation is not rejected for the new sample (see Table C.16 in the Appendix). Results using the sample with no attrition yields results that are similar in terms of sign, magnitude and significance compared to those based on the corresponding benchmark specification (see Table C.17 in the Appendix). The benchmark estimates are robust to attrition bias, while the effect of technical change TC gains slightly in importance when using a more balanced panel.

5.7. Instrumenting for regulation exposure

We apply an instrumental variables (IV) approach as an additional strategy to account for potential time-varying unobserved heterogeneity not orthogonal to T1000P exposure. For example, unobserved, time-varying firm characteristics (e.g., political connections) could be correlated with both treatment status and the outcome variable(s). We therefore construct an instrument for

[12] In preceding analysis, we excluded all firms known to be part of these programs from our estimation sample.

T1000P participation that equals the distance-weighted index of the ratio of the number of treated firms to the total number of firms in the geographic cluster of the firm and its neighboring clusters. The geographic clusters within such a group are indexed by h, with an individual cluster being defined by a county q. As shown by Fig. C.2 in the Appendix, a county has on average seven neighbors. The instrument draws its validity from the observation that in China, clusters of iron and steel firms are found in particular areas of the country (see Fig. B.1 and Fig. B.2 in the Appendix). Industrial clusters capture unobserved time varying heterogeneity affecting T1000P exposure such as social, environmental, political or institutional characteristics of a formerly planned economy. As firm fixed effects are controlled for, our instrument would be expected to have only limited influence on the performance of an individual firm. The instrument $\tau_i^{IV}$is based on year 2005 observations, and for a firm i in county q is given as

$$\tau_i^{IV} = \frac{\sum_h \frac{1}{d_{qh}} \cdot \phi_h}{\sum_h \frac{1}{d_{qh}}} \tag{8}$$

where d_{qh} is the distance in kilometers between the firm's county q and neighboring counties. The distance weight of a firm's own county is 1. The ratio of treated firms to the total number of firms in a cluster is ϕ_h. Note that τ_i^{IV} does not differ between firms belonging to the same cluster q. Descriptive statistics of τ_i^{IV} are given in Table C.18 in the Appendix.

The empirical estimation is based on a panel data two-stage least squares (2SLS) within estimator, and controls for firm fixed effects and allows for a correlation of errors between the two stages. Given that τ_i is a binary variable and the outcome variable of the second stage is continuous, we follow Angrist (2001) and use a linear probability model (LPM) in the first stage. As noted in Angrist (2001), the estimation of a 2SLS model applying a LPM in the first stage bears the benefit of consistency, irrespective of whether or not the first-stage conditional expectation function is linear. As all variables included in the first stage are of limited range, the supporting restriction for LPM (i.e., of no regressor having infinite support) is satisfied. Eq. (4) is first within transformed, thereby accounting for α_i, and then a 2SLS methodology is applied instrumenting for τ_i by τ_i^{IV} in the first stage. This approach is described in detail in Baltagi (2008). Our instrument τ_i^{IV} is found to be valid.[13] Results indicate that instrumenting for T1000P selection yields overall treatment effects (see Table C.20 in the Appendix) that are very similar in terms of magnitude and significance to the benchmark results of all three model specifications (see Table 3). TC gains in magnitude, while SEC loses significance. However, the estimated effect of the T1000P on treated firms' annual TFP change is consistent with the benchmark specification.

5.8. Time-varying confounders

Our main results rely on the assumption that there are no omitted time-varying and firm-specific effects correlated with T1000P participation and the outcome variable. We have conducted an extensive review of policies potentially affecting the iron and steel sector during the study period, and found one program that could be a potential confounder. Along with the goal of reducing inefficient energy use in energy-intensive industries via the T1000P program, the national government also implemented a program to eliminate outdated production capacity during the Eleventh FYP. The program defined production technologies that would be limited or eliminated in all sectors (NDRC, 2005a). For the iron and steel sector, outdated technologies included blast furnaces for iron smelting with a capacity less than 300 cubic meters (NDRC, 2005b). A detailed implementation plan was announced in 2006 (NDRC, 2006). To test for potential confounding effects, we located a list of firms in a subset of provinces that were subject to the first phase of this program (NDRC, 2007). These firms were required to shut down, retire or update part of their production capacity. Though significance of the treatment effects drops mildly due to a much smaller sample size, the size of the effects remains very close to the benchmark results of all three model specifications listed in Table 3 (see Table C.21 in the Appendix).

5.9. TFP change based on production function and index numbers

To test the robustness of our results to the choice of productivity estimation approach, we also used alternative approaches to measure TFPC based on the estimation of a parametric production function and on the use of index numbers. In the case of the former, we implemented the production function approach of Wooldridge (2009).[14] In addition, we also applied the models of Levinsohn and Petrin (2003) and Ackerberg et al. (2015). All models are based on a value-added production function. The results of overall treatment effects on TFP change are robust in terms of magnitude and significance.

Further, we decided to use the Törnqvist index, as described in Coelli et al. (2003), and an index derived from a Cobb-Douglas production function. Since the entire population of Chinese iron and steel firms between 2003 and 2008 is observed (5340 firms or 27,076 observations, see Table 1), following Greenstone et al. (2012) we estimate transitive multilateral TFP by applying an index

[13] First stage results are shown in Table C.19 in the Appendix. The Davidson-MacKinnon test of exogeneity (Davidson & MacKinnon, 1993) rejects at a 1 percent significance level, indicating that the benchmark ATT variable may be endogenous. The Kleibergen-Paap rk LM and rk Wald F-statistics (Kleibergen & Paap, 2006) both reject at a significance level of 1%.

[14] Wooldridge (2009) shows that the moment conditions of the complicated two-step semi-parametric approach of Olley and Pakes (1996) to control for the simultaneity bias and the modification to it proposed by Levinsohn and Petrin (2003) can be implemented in a GMM framework yielding more efficient estimates among other advantages. The approach proposed by Olley and Pakes (1996) cannot be applied because

number approach.[15] We employ a Cobb-Douglas production function with homogeneity in input factors and Hicks-neutral (Hicks & John, 1963) productivity differences. The corresponding generic production function can be given as

$$TFP_{it} = y_{it} - \beta^{s}_{Kt} k_{it} - \beta^{s}_{Lt} l_{it} - \beta^{s}_{Mt} m_{it} \tag{9}$$

with small letters indicating variables in natural logarithms. Year- and subindustry s-specific input elasticities, i.e. input cost shares, are denoted by β^s. Capital costs k_{it} consist of interest and depreciation expenses. Opportunity cost of equity is assumed to be 3%. Labor costs l_{it} are the sum of wage and welfare payments and m_{it} is the cost of intermediate inputs. Subsequently, TFP change is derived by taking the first derivative of the TFP estimates. As this non-parametric approach lacks the capacity to control for noise, TFP changes outside the 0.1 and 99.9 percentiles have been excluded. The control group is matched using propensity score matching to the nearest five neighbors with replacement, resulting in 118 matched firms and a total of 538 control group observations. The assumption of a parallel trend holds for the new treatment and control samples (see Table C.22 in the Appendix). We find a positive effect of the regulation on firm level TFP change at a significance level of 10% (see Table C.23 in the Appendix) in the three model specifications. Also using a Törnqvist index we come to the same empirical results in terms of magnitude and significance of treatment effects. Finally, we include the treatment effect variable directly into the cost function of Eq. (4) and find a negative and significant (at 5% level) effect of the treatment on costs. This implies that directly using a classical difference-in-difference approach on cost confirms the results obtained through a "two-step approach" (via TFPC). Results are available upon request.

6. Conclusion

We find that the first major national energy efficiency program in China had a positive, statistically-significant effect on annualized TFP change in iron and steel firms. The gains we observe are economically significant and robust to multiple alternative empirical strategies. In our base specification, we estimate that the average annual private economic benefit due to enhanced productivity for a treated firm was 148.7 mRMB in 1998 values. A back-of-the-envelope calculation using one estimate of US $22.5 billion (IEPD, 2007) to approximate the investment in energy conservation by T1000P firms suggests that on average each firm invested approximately 171 mRMB between 2006 and 2008. Inflating our estimate of annualized benefits to ~200 mRMB in 2007 values, we find that the energy efficiency investments undertaken imply an undiscounted payback period of less than one year for the average firm. Importantly, this estimate does not account for any subsidies provided to firms, nor does it capture the cost of administering the program.

Our main estimation results suggest that technical change and scale efficiency change were about equally important in contributing to productivity improvements under the national T1000P. This implies that to some extent firms may have been able to comply with the program simply by increasing the scale of their operations. However, scaling up was not the whole story. At least half of the productivity growth is, on average, attributed to technical change. This share was even larger in the extended analysis employing a production function and in the provincial program estimation, which focuses on smaller firms and better controls for differences in firm scale. Technical change encompasses equipment adjustments or upgrades, or changes to management practices that increased emphasis on energy conservation in performance targets and personnel incentives. Consistent with a role for technical change, prior studies qualitatively describe how T1000P firms in the iron and steel industry upgraded equipment and processes to improve energy efficiency (Ke et al., 2012), and did so relatively quickly after the program began (Price et al., 2010).

An important exercise would involve quantifying the heterogeneity in investment inefficiencies across firms. In other words, we would be interested to know the distribution of g_i in Eq. (1). Firms included in the program were all large energy users, but these firms likely differed in their awareness of, and access to, support for energy efficiency upgrades. Our observation of net positive productivity effects on average mask heterogeneity in firm level effects. Some firms may have overinvested relative to the efficient level prior to the program, in which case we expect the T1000P to have negatively affected the productivity of these firms. Future research could help to characterize this heterogeneity, as well as the degree to which one-size-fits-all programs impose a high cost on some firms.

A striking feature of our estimates is that they suggest firms benefited from energy efficiency investments in a relatively short amount of time, i.e. within the first few program years. The substantial short-run benefits we estimate contrast with nearly all prior studies of the productivity effects of energy and environmental regulations, which point to short-run costs (Dechezleprêtre & Sato, 2017; Koźluk & Zipperer, 2013). These short-run benefits are further consistent with evidence of input cost savings achieved by the average treated firm: when we estimate model (1) using the log of materials cost instead of TFP growth as the independent variable (and control for scale), we find the program is associated with a reduction in materials cost of 13.5% (in log points) that is statistically significant at the 10% level. The lack of reliable sources of firm data for extending our panel beyond 2008 constrains our focus to short-run effects; however, estimating long-run effects as well as general equilibrium effects of energy efficiency programs in rapidly developing countries is an important area for future research.

The positive effects of the T1000P suggest that investment inefficiencies in developing countries may be substantial. Moreover,

(*footnote continued*)
investments of many firms in our sample are observed to be negative, whereby investments are derived from changes in the real capital stock.

[15] Given that we observe the entire population of firms, our approach of measuring TFP via four-digit industry specific cost shares used as input elasticities avoids several well-known caveats associated with estimating TFP from a subsample of a population. Examples discussed extensively in the literature are a simultaneity between unobserved productivity and either input choices or entry/exit decisions (attrition bias).

these inefficiencies do not appear to be due to capital constraints specific to particular firm categories (e.g., SOEs). Instead, it seems that raising awareness and pressure to save energy prompted firms to take up profitable energy efficiency investments. Importantly, it may have created a new form of reputational benefit associated with target achievement that reduced uncertainty in the payoff to participating firms. Future work is needed to more precisely characterize the origins of the energy efficiency gap in these settings, and derive lessons for the design of effective interventions. While our study has focused on a state-led intervention, a wide range of private initiatives exist to raise energy efficiency, e.g. between buyers and suppliers, for which the productivity consequences are not well understood. Our findings suggest that state-led programs such as the T1000P (and subsequent T10000P) in China and the analogous Perform, Achieve, and Trade energy efficiency scheme in India may, however imperfectly, be addressing investment inefficiencies, even if they do not reduce energy use in absolute terms.

Importantly, although energy efficiency programs are often justified as pro-environmental, they are not a substitute for, and at best may be complementary to, policies that directly address the environmental externalities of energy use, such as emissions taxes or cap-and-trade programs. In many countries, energy efficiency programs form a pillar of national strategies for reducing energy-related air pollution and mitigating greenhouse gas emissions. However, energy use decisions would not reflect these social costs as long as the associated externalities remain unpriced. Subsidies for energy in developing countries exacerbate the problem, to the extent that they encourage overuse of fossil energy sources. On the other hand, with prices that reflect the marginal social cost of energy use, carefully targeted energy efficiency policies that address investment inefficiencies could support the effectiveness of price-based mechanisms.

Acknowledgements

We thank participants at the Stanford Environmental and Energy Policy Analysis Center Lunch, the Young Swiss Economists Meeting 2017, the 6th Annual Summer Conference of the Association of Environmental and Resource Economists, 15th European Workshop on Efficiency and Productivity Analysis, the 23rd Annual Conference of the European Association of Environmental and Resource Economists, and the 8th Atlantic Workshop on Energy and Environmental Economics for helpful comments and discussion. This work was partially supported by a grant through the MIT Energy Initiative's Seed Fund Program and by the MIT Joint Program on the Science and Policy of Global Change through a consortium of industrial sponsors and federal grants. We further gratefully acknowledge the support of the Energy Information Administration of the U.S. Department of Energy through a cooperative agreement with MIT (DE-EI0003030). Da Zhang acknowledges the support of the National Science Foundation of China (Project No. 71690244).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chieco.2019.101364.

References

Ackerberg, D., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica, 83*(6), 2411–2451.

Allcott, Hunt, & Greenstone, M. (2012). Is there an energy efficiency gap? *The Journal of Economic Perspectives, 26*(1), 3–28.

Allcott, H., Mullainathan, S., & Taubinsky, D. (2014). Energy policy with externalities and internalities. *Journal of Public Economics, 112*, 72–88.

Ambec, S., Cohen, M. A., Elgie, S., & Lanoie, P. (2013). The Porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness? *Review of Environmental Economics and Policy, 7*(1), 2–22.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors. *Journal of Business & Economic Statistics, 19*(1), 2–28.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics, 21*(1), 1–42.

Bai, C.-E., Hsieh, C.-T., & Qian, Y. (2006). The return to capital in China. *NBER Working Paper Series No. 12755*. Cambridge, MA, USA: National Bureau of Economic Research.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data.* Chichester, UK: Hoboken, NJ: John Wiley & Sons.

Berndt, E. R., & Christensen, L. R. (1973). The translog function and the substitution of equipment, structures, and labor in US manufacturing 1929-68. *Journal of Econometrics, 1*(1), 81–113.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics, 119*(1), 249–275.

Brandt, L., Van Biesebroeck, J., & Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics, 97*(2), 339–351.

Burlig, F., Knittel, C., Rapson, D., Reguant, M., & Wolfram, C. (2017). Machine learning from schools about energy efficiency. *NBER Working Paper*.

Cao, J., Garbaccio, R., & Ho, M. S. (2009). China's 11th five-year plan and the environment: Reducing SO_2 emissions. *Review of Environmental Economics and Policy,* 1–20.

CCM. 2015. Price database for iron ore. Cheina Commodity Marketplace. Retrieved from: http://www.chinaccm.com/PriceData/default.aspx. .

CEIC (2015). *CEIC China Premium Database.* New York, NY, USA: CEIC Data.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics, 55*(1), 28–45.

Coelli, T. J., Estache, A., Perelman, S., & Trujillo, L. (2003). *A Primer on Efficiency Measurement for Utilities and Transport Regulators.* Washington DC, USA: The World Bank.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics.* New York, NY, USA: Oxford University Press.

Davis, L. W., Fuchs, A., & Gertler, P. (2014). Cash for coolers: Evaluating a large-scale appliance replacement program in Mexico. *American Economic Journal: Economic Policy, 6*(4), 207–238.

DeCanio, S. J. (1993). Barriers within firms to energy-efficient investments. *Energy Policy, 21*(9), 906–914.

DeCanio, S. J., & Watkins, W. E. (1998). Investment in energy efficiency: Do the characteristics of firms matter? *Review of Economics and Statistics, 80*(1), 95–107.

Dechezleprêtre, A., & Sato, M. (2017). The impacts of environmental regulations on competitiveness. *Review of Environmental Economics and Policy, 11*(2), 183–206.

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of Econometrics, 4*(2), 115–145.
Dougherty, S., Herd, R., & He, P. (2007). Has a private sector emerged in China's industry? Evidence from a quarter of a million Chinese firms. *China Economic Review, 18*(3), 309–334.
Ehrlich, I., Gallais-Hamonno, G., Liu, Z., & Lutter, R. (1994). Productivity growth and firm ownership: An analytical and empirical investigation. *Journal of Political Economy, 102*(5), 1006–1038.
Fowlie, M., Greenstone, M., & Wolfram, C. (2018). Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *Quarterly Journal of Economics, 133*(3), 1597–1644.
Gatto, D., Di Liberto, A., & Petraglia, C. (2011). Measuring productivity. *Journal of Economic Surveys, 25*(5), 952–1008.
Gillingham, K., & Palmer, K. (2014). Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy, 8*(1), 18–38.
Gollop, F. M., & Roberts, M. J. (1983). Environmental regulations and productivity growth: The case of fossil-fueled electric power generation. *The Journal of Political Economy,* 654–674.
Gray, W. B., & Shadbegian, R. J. (2003). Plant vintage, technology, and environmental regulation. *Journal of Environmental Economics and Management, 46*(3), 384–402.
Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures. *Journal of Political Economy, 110*(6), 1175–1219.
Greenstone, M., List, J. A., & Syverson, C. (2012). The effects of environmental regulation on the competitiveness of US manufacturing. *NBER Working Paper Series No. 18392 (asked for revisions by The Quarterly Journal of Economics)*. Cambridge, MA, USA: National Bureau of Economic Research.
Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics,* 33–54.
He, F., Zhang, Q., Lei, J., Fu, W., & Xu, X. (2013). Energy efficiency and productivity change of China's iron and steel industry: Accounting for undesirable outputs. *Energy Policy, 54*, 204–213.
Hicks, & John (1963). *The theory of wages.* London, UK: Palgrave Macmillan.
Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics, 124*(4), 1403–1448.
Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Presented at Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1. Presented at Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (pp. 221–233). Berkeley, CA, USA: University of California Press.
IEPD (2007). *Industrial Efficiency Policy Database (CN-3a:Top-1000 Energy-Consuming Enterprises Program)*.
IISI (1986). *Steel Statistical Yearbook 1986.* Brussels, Belgium: International Iron and Steel Institute.
IISI (2002). *Steel Statistical Yearbook 2002.* Brussels: International Iron and Steel Institute.
Ke, J., Price, L., Ohshita, S., Fridley, D., Khanna, N. Z., Zhou, N., & Levine, M. (2012). China's industrial energy consumption trends and impacts of the Top-1000 enterprises energy-saving program and the ten key energy-saving projects. *Energy Policy, 50*, 562–569.
Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics, 133*(1), 97–126.
Koakutsu, K., Usui, K., Watarai, A., & Takagi, Y. (2012). Measurement, reporting and verification (MRV) for low carbon development: Learning from experience in Asia. *IGES Policy Report No. 2012–03*. Hayama, Japan: Institute for Global Environmental Strategies (IGES).
Koźluk, T., & Zipperer, V. (2013). Environmental policies and productivity growth - A critical review of empirical findings. *OECD Journal: Economic Studies, 2014*(1), 155–185.
Kumbhakar, S. C., & Lovell (2000). *Stochastic frontier analysis.* New York: Cambridge University Press.
Lachowska, M., & Myck, M. (2018). The effect of public pension wealth on saving and expenditure. *American Economic Journal: Economic Policy, 10*(3), 284–308.
Lee, Y., Stoyanov, A., & Zubanov, N. (2019). Olley and Pakes-style production function estimators with firm fixed effects. *Oxford Bulletin of Economics and Statistics, 81*(1), 79–97.
Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies, 70*(2), 317–341.
Li, H., Zhao, X., Yu, Y., Wu, T., & Qi, Y. (2016). China's numerical management system for reducing national energy intensity. *Energy Policy, 94*, 64–76.
Lin, B., & Wang, X. (2014). Promoting energy conservation in China's iron & steel sector. *Energy, 73*, 465–474.
Lin, B., Wu, Y., & Zhang, L. (2011). Estimates of the potential for energy conservation in the Chinese steel industry. *Energy Policy, 39*(6), 3680–3689.
Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics, 13*(2), 151–161.
Myers, E. (2018). Asymmetric Information in Residential Rental Markets: Implications for the Energy Efficiency Gap. *E2e Working Paper*.
NBS (2007). *Input-Output Tables of China.* Beijing, China: Department of National Accounts, National Bureau of Statistics.
NBS (2013). *Chinese Statistical Yearbook 2013.* Beijing, China: National Bureau of Statistics.
NBS (2014). *Chinese Statistical Yearbook 2014.* Beijing, China: National Bureau of Statistics.
NDRC. 2005a. Guidance catalog for industry restructuring (2005 version). Beijing, China: National Development and Reform Commission. Retrieved from: http://www.sdpc.gov.cn/fzgggz/fzgh/zcfg/200512/t20051222_65963.html. .
NDRC. 2005b. Industrial policy for the iron and steel sector. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.gov.cn/flfg/2006-01/17/content_161597.htm. .
NDRC. 2006. Notice on the control of production capacity, elimination of outdated capacity and acceleration of industry restructuring. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.gov.cn/zwgk/2006-07/17/content_337825.htm. Access date.
NDRC. 2007. NDRC announces iron and steel firms subject to closing down and production capacity. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.gov.cn/gzdt/2007-09/09/content_743546.htm. .
NDRC. 2009. Notice on the evaluation results of the Top 1,000 firms energy saving target in 2008. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.sdpc.gov.cn/fzgggz/hjbh/hjjsjyxsh/200911/t20091124_315031.html. .
NDRC. 2010. Notice on the evaluation results of the Top 1,000 firms energy saving target in 2009. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.sdpc.gov.cn/zcfb/zcfbgg/201007/t20100705_358752.html. .
NDRC. 2011. Notice on the completion status of the Top 1,000 firms energy saving target during the 11th Five-Year Plan. Beijing, China: National Development and Reform Commission. Retrieved from: http://www.sdpc.gov.cn/zcfb/zcfbgg/201112/t20111227_452721.html. .
Olley, S. G., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica, 64*(6), 1263–1297.
Orea, & Luis (2002). Parametric decomposition of a generalized Malmquist productivity index. *Journal of Productivity Analysis, 18*(1), 5–22.
Porter, M. E., & Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *The Journal of Economic Perspectives,* 97–118.
Price, L., Levine, M. D., Zhou, N., Fridley, D., Aden, N., Lu, H., & Yowargana, P. (2011). Assessment of China's energy-saving and emission-reduction accomplishments and opportunities during the 11th five year plan. *Energy Policy, 39*(4), 2165–2178.
Price, L., Wang, X., & Yun, J. (2010). The challenge of reducing energy consumption of the Top-1000 largest industrial enterprises in China. *Energy Policy, 38*(11), 6485–6498.
Qi, Y. (2013). Blue book of low-carbon development. *Annual Review of Low-Carbon Development in China* (pp. 2013). Beijing, China: Social Sciences Academic Press.
Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*(1), 33–38.
Ross, M., & Feng, L. (1991). The energy efficiency of the steel industry of China. *Energy, 16*(5), 833–848.
Ryan, N. (2018). Energy productivity and energy demand: Experimental evidence from Indian manufacturing plants. *BER Working Paper Series No. 12755*. Cambridge, MA, USA: National Bureau of Economic Research.
State Council. 2006. The 11th five year plan for China's economic and social development. Beijing: Sate Council. Retrieved from http://www.gov.cn/gongbao/content/2006/content_268766.htm. Access date: October 28th 2016.
State Council. 2007. Notice of implementation scheme and methods of statistic, monitoring and evaluation of energy intensity reduction work and pollution reduction

work. Beijing: State Council. Retrieved from http://www.gov.cn/zwgk/2007-11/23/content_813617.htm. Access date: October 28th 2016.
Tian, X., & Yu, X. (2012). The enigmas of TFP in China: A meta-analysis. *China Economic Review, 23*(2), 396–414.
Wagner, U. J., Muûls, M., Martin, R., & Colmer, J. (2014). The causal effects of the European Union emissions trading scheme: Evidence from French manufacturing plants. *Presented at Fifth World Congress of Environmental and Resources Economists*. Turkey: Instanbul.
Wang, H. J., & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis, 18*, 129–144.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.
Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters, 104*, 112–114.
WSA (2014). *Steel Statistical Yearbook 2014.* Brussels, Belgium: World Steel Association.
Wu, Y. (2011). Total factor productivity growth in China: A review. *Journal of Chinese economic and business studies, 9*(2), 111–126.
Xu, B., & Lin, B. (2016). Assessing CO_2 emissions in China's iron and steel industry: A dynamic vector autoregression model. *Applied Energy, 161*, 375–386.
Zeng, S., Lan, Y., & Huang, J. (2009). Mitigation paths for Chinese iron and steel industry to tackle global climate change. *International Journal of Greenhouse Gas Control, 3*(6), 675–682.
Zhang, D., Aunan, K., Seip, H. M., & Vennemo, H. (2011). The energy intensity target in China's 11th five-year plan period—Local implementation and achievements in Shanxi Province. *Energy Policy, 39*(7), 4115–4124.
Zhang, J., & Wang, G. (2008). Energy saving technologies and productive efficiency in the Chinese iron and steel sector. *Energy, 33*(4), 525–537.
Zhao, X. (2016). Enterprise-level amount of energy saved targets in China: Weaknesses and a way forward. *Journal of Cleaner Production, 129*, 75–87.
Zhao, X., Li, H., Wu, L., & Qi, Y. (2014). Implementation of energy-saving policies in China: How local governments assisted industrial enterprises in achieving energy-saving targets. *Energy Policy, 66*, 170–184.
Zhou, N., Levine, M. D., & Price, L. (2010). Overview of current energy-efficiency policies in China. *Energy Policy, 38*(11), 6439–6452.
Zhou, K., & Yang, S. (2016). Emission reduction of China' s steel industry: Progress and challenges. *Renewable and Sustainable Energy Reviews, 61*, 319–327.