Regulating Conglomerates: Evidence from an Energy Conservation Program in China*

Qiaoyi Chen
Fudan University

Zhao Chen
Fudan University

Zhikuo Liu
Fudan University

Juan Carlos Suárez Serrato
Stanford University GSB
& NBER

Daniel Yi Xu
Duke University
& NBER

September 4, 2023

Abstract

We study a prominent energy regulation affecting large Chinese manufacturers that are part of broader conglomerates. Using detailed firm-level data and difference-in-differences research designs, we show that regulated firms cut output and shifted some production to unregulated firms in their same conglomerate instead of improving their energy efficiency. To account for conglomerate and market spillovers, we interpret these results through the lens of an industry equilibrium model featuring conglomerate production. The policy raises welfare if the per-ton benefits of carbon reduction exceed $161. Alternative polices that exploit public information on business networks can increase aggregate energy savings by 10%.

JEL Codes: Q48, L51, O44, H23.

---

[*] We are very grateful for discussions with Antung Anthony Liu, Mar Reguant, Nick Ryan, and Shaoda Wang and for comments from Hunt Allcott, Soren Anderson, Prabhat Barnwal, Raj Chetty, Julie Cullen, David Cutler, Michael Davidson, Michael Dinerstein, Matt Gentzkow, Ed Glaeser, Ken Gillingham, Josh Gottlieb, Michael Greenstone, Guojun He, Caroline Hoxby, Kelly Jones, Matthew Kahn, Louis Kaplow, Lawrence Katz, Stephanie Kestelman, Justin Kirkpatrick, Thibaut Lamadon, Ashley Langer, Shanjun Li, Neale Mahoney, Justin Marion, Leslie Martin, Robert Metcalfe, Magne Mogstad, Ben Olken, Paulina Oliva, Edson Severnini, Joe Shapiro, Felix Soliman, Michael Song, Stefanie Stantcheva, Chris Timmins, John Van Reenen, Reed Walker, Heidi Williams, Xiaodong Zhu, and seminar participants at American University, Arizona State, ASSA, Barcelona Summer Forum, Cowles Foundation Summer Conference, Fudan University, FRB of Atlanta, Harvard, HBS, Helsinki GSE, IIPF, LSE, Michigan State University, MIT, NBER Public, NBER China, NYU, Paris School of Economics, Peking University, SHUFE, STEG Conference, Stanford University, University of Chicago, University of Southern California, University of Toronto, UC Berkeley Haas, UCLA, UCSC, UCSD, Université du Québec à Montréal, University of Georgia, University of Oxford, University of Michigan, 9th Mannheim Conference on Energy and the Environment, and 3rd World Bank Tax Conference. We thank IntSig Information for providing China's Administrative Registration Data (CARD). All errors remain our own.

Balancing economic growth with the negative side effects of industrialization—such as carbon emissions and pollution—is a central problem of governments in emerging economies. Nowhere is this problem more important or consequential than in China. As Figure 1 shows, energy regulation is of national and global importance given that the industrial energy use of China overshadowed that of other leading economies in the early years of the 21st century.

This paper studies the effects of a large program aimed at curbing the energy use of Chinese industrial firms. The regulation that we study—the "Top 1,000" program—targeted the largest energy-consuming firms in the most energy-intensive industries. The regulation was designed following examples of "voluntary agreement" programs in developed countries that relied on the belief that firms could significantly reduce their energy use by improving their energy efficiency. The implementation of the program was adjusted to Chinese institutions and constraints, with the result that, in practice, lowering energy consumption became the main regulatory objective.

Understanding the effects of this regulation is central to broader questions of energy conservation for several reasons. First, the firms regulated by this program accounted for 47% of total industrial energy use in China in 2004. Additionally, as in several developing countries, industrial firms in China are often part of much larger business networks.[1] The sheer size of the regulated firms and their broader networks of related firms imply that a complete assessment of the effects of the regulation needs to account for within-conglomerate and market-level spillovers. Finally, the perceived success of the regulation led the Chinese government to significantly expand the program in later years.

This paper characterizes the effectiveness of the Top 1,000 program by combining difference-in-differences research designs with an industry equilibrium model featuring conglomerate production. Our difference-in-differences estimations show that, relative to unregulated firms in energy-intensive industries, Top 1,000 firms significantly decreased their energy use after the regulation. Regulated firms achieved these reductions by lowering output; we find no impact on their energy efficiency. Using detailed data on business networks and a second difference-in-differences design, we then show that unregulated firms in the same conglomerate as regulated firms increased both output and energy use. This result uncovers an important margin of adjustment that allowed Chinese conglomerates to shift 40% of the output decline in regulated firms to unregulated affiliates. Finally, we provide evidence of market-level spillover effects by showing that unrelated and unregulated firms in more heavily regulated industries increased their output after the regulation. These results corroborate the notion that conglomerates were not able to fully shift production across related firms.

To quantify the aggregate and welfare effects of the policy, we specify and estimate an indus-

---

[1] Ramachandran, Manikandan and Pant (2013) describe the growing importance of conglomerates in India, China, and Latin America.

try equilibrium model of conglomerate production.[2] The model accounts for conglomerate and market spillovers, matches the estimated impacts of the policy, and clarifies the interpretation of the difference-in-differences estimates in the presence of both types of spillovers. To quantify the cost of complying with the program, we use the model to compute the shadow cost of the policy. Absent the ability to shift production across related firms, the quantity restriction would have raised production costs by 11.2%. The ability of conglomerates to shift production to related firms lowered the shadow cost of the regulation to Top 1,000 firms by 30%. We also evaluate the welfare effects of the program and quantify that the Top 1,000 program improves welfare when the per-ton value of reducing carbon emissions exceeds $161.[3] Finally, we use the model to simulate the effects of expanding the program to include more firms, of relying on size-dependent and universal energy taxes, and of using public information on conglomerate networks to design conglomerate-level regulations. A conglomerate-level regulation could increase energy savings by 10% for the same welfare cost and presents a trade-off that is close to that of a universal energy tax.

We develop these results in three steps. As we discuss in Section 1, our first difference-in-differences strategy uses the fact that firms in similar industries that were regulated in later years are suitable controls for Top 1,000 firms. In Section 2, we use an event-study specification to show that regulated and unregulated firms had similar trends prior to the regulation. Relative to unregulated firms, Top 1,000 firms reduced their energy use by approximately 16% in response to the regulation. These estimates are robust to inclusion of industry-by-year and province-by-year fixed effects and controls for firm characteristics. We also document that these firms saw a decline in output of approximately 20%, and we do not find meaningful or statistically significant changes in energy efficiency.

Our second set of analyses leverages detailed business registration data to map the conglomerate networks of regulated firms. If regulated firms were able to circumvent the regulation by shifting production to related parties, we would expect to see an increase in both the output and energy use of firms linked to the regulated firms through ownership networks. We test this hypothesis in Section 3 by using a difference-in-differences strategy that compares unregulated but related firms to unregulated and unrelated firms.[4] These analyses show that after the reform,

---

[2] The Top 1,000 program had the stated goal of reducing industrial energy use to lower emissions that contribute to global warming. While energy use reductions also lower local pollution, pollution reduction was not a stated goal of the program (Price, Wang and Yun, 2010). Accordingly, our welfare analyses focus on the program's objective of reducing aggregate energy use.

[3] This value is bounded between $114 and $199 across a broad range of model extensions, including extensions incorporating heterogeneous energy efficiency levels across firms, the possibility that firms respond to the regulation by improving their energy efficiency, and alternative parameters values.

[4] To ensure that these two groups of firms are similar, we use a matching procedure based on pre-regulation characteristics to find a suitable set of control firms.

regulated conglomerates shifted production to affiliates that were not subject to the regulation.[5] We estimate an increase in the output of related firms of approximately 12%; we also find increases in energy use but no effects on energy efficiency. Importantly, we find increases in the economic activity of related firms only when their line of business coincides with the narrowly defined (4-digit) industry classification of the regulated firm. As a placebo test, we show that related firms in other industries did not see an increase in economic activity. Because related firms are smaller than regulated firms, we calculate that conglomerates were able to shift 40% of the output decline in regulated firms to related parties.

Conceptually, firms could respond to the Top 1,000 program in three ways: by increasing energy efficiency, by reducing output, or by shifting production to related parties. The fact that conglomerates adjusted their output allocation but did not improve their energy efficiency is informative about the costs of different margins of response to energy regulations. Our results are consistent with the notion that costly long-run investments would be required to improve the energy efficiency of regulated firms.[6] Additionally, the fact that regulated firms were not able to fully shift their production to related parties suggests there was no "low-hanging fruit" from a technological perspective (e.g., Allcott and Greenstone, 2012).[7] Consistent with the finding that conglomerates were not able to fully compensate their output loss by shifting production, we estimate market-level spillovers by showing that unregulated and unrelated firms in more heavily regulated industries increased their output as a result of the regulation. These results show that a complete assessment of the effects of the Top 1,000 program must take into account how declines in energy use at regulated firms can lead to within-conglomerate and market-level increases in energy use, otherwise known as "leakage."

In Section 4, we present an industry equilibrium model of conglomerate production that accounts for within-conglomerate spillovers to related firms as well as market spillovers. We estimate the model in Section 5 by matching moments of the firm size distribution and patterns

---

[5] Our measure of conglomerates in the ownership network data includes both affiliated firms—which are the Chinese counterparts of plants in a multi-establishment firm—and separate firms linked through a common owner.

[6] Since coal is the main energy source for the regulated industries, meaningful improvements in energy efficiency would require firms to adopt long-lived industrial machines that rely on electricity. Firms may have been reluctant to do this given the abundance and inexpensiveness of coal in China and the government's uncertain commitment to energy-saving policies. Zhao et al. (2016) discuss a case study of a Top 1,000 firm with old, energy-inefficient capital. Even when this firm started the process of adopting new machinery in 2007, energy efficiency gains did not materialize until 2011 due to construction and installation delays. In Section 7.1, we extend the model by allowing for investments that can improve energy efficiency and consider how these improvements can impact the aggregate and welfare effects of the policy.

[7] Gillingham, Keyes and Palmer (2018) discuss a number of reasons why firms may underinvest in energy efficiency, including imperfect information, behavioral biases, and market failures. For instance, Anderson and Newell (2004) show that even when US firms obtain information from energy audits, managers expect fast payback periods when deciding to invest in projects that improve their energy efficiency. The ability of firms to escape the burden of regulation by shifting production to related firms adds to the potential explanations for underinvestment in energy efficiency.

of the within-conglomerate allocation of production prior to the regulation. We then use our reduced-form estimates as out-of-sample validations of the model. The fact that the model does a good job of matching the differences-in-differences estimates indicates that it correctly captures the quantitative importance of within-conglomerate and market-level spillovers.[8] The model further decomposes our difference-in-differences estimates and shows that accounting for the effect of the program on unregulated firms lowers the effect on regulated firms from 20% to 14%. In contrast, the magnitude of spillovers to related firms increases from 12% to 19%. By combining the reduced-form estimates with a structural industry equilibrium model, we are able to evaluate the effects of the Top 1,000 program while taking into account spillover and equilibrium effects, which are central features of prominent energy regulations.

Section 6 uses the model to quantify the aggregate and welfare effects of the Top 1,000 program. Accounting for market and conglomerate leakage, we calculate that the program reduced aggregate energy use by 4%, an annual decrease of approximately 48 million tons of coal equivalent (tce).[9] A calibration of the government's willingness to pay (GWTP) to reduce energy use shows that the program raises welfare as long as the GWTP—which includes both the social cost of carbon (SCC) and the health damages associated with local pollution—exceeds $161 per ton of carbon.[10] Using the model, we show that expanding the program by increasing the number of regulated firms or by tightening energy savings targets leads to similar trade-offs. A government facing administrative constraints would thus prefer to tighten the stringency of the regulation rather than increase the number of regulated firms.

The model allows us to compare the aggregate and welfare effects of incomplete regulations, such as the Top 1,000 program, to those of policies that would be preferable absent political or administrative constraints, such as a universal energy tax. First, we show that the government can increase aggregate energy savings by 10% for the same welfare cost by leveraging publicly available data on the ownership networks of regulated conglomerates. By targeting conglomer-

---

[8] Additionally, we show in Section 6 that the model matches the effects of a within-conglomerate differences-in-differences estimation that compares regulated to unregulated firms in the same conglomerate. Since this design absorbs conglomerate- and market-level spillovers, this result shows that the model correctly captures the reallocation of production within business groups.

[9] Our welfare analysis focuses on the aggregate effects of the program. However, it is possible for the program to impact welfare through heterogeneous effects across more polluted or populated areas or if it shifted production to less developed regions. Appendix H provides an expanded welfare framework that incorporates these channels. This appendix also documents that the program did not significantly shift production to more polluted or populated areas and that it did not reallocate production to less developed regions in Western China. For these reasons, our model and welfare analyses abstract from spatial implications of the policy.

[10] In Appendix G, we use estimates of pollution damages in China (World Bank, 2007; Mohan et al., 2020; Ito and Zhang, 2020*b*) to calculate that between $4 and $17 of the GWTP could reflect the health benefits of reducing pollution. The remaining $144–$157 would be attributed to the SCC. The low estimates of pollution damages per ton of carbon in China contrast with estimates from richer countries; for instance, we obtain an estimate of $78 for the US based on the results of Mohan et al. (2020). Using this value would imply an SCC that rationalizes the policy of $83. For comparison, recent proposed values of the SCC range between $51 (IWG, 2021) and $125 (Carleton and Greenstone, 2021).

ates instead of firms, such a regulation would avoid distorting the within-conglomerate allocation of production. Second, the model shows that a conglomerate-level regulation closely approximates the effects of a size-dependent energy tax that applies to all affiliates in conglomerates with Top 1,000 firms. Finally, we find that this size-dependent tax is only slightly inferior to a universal energy tax. These results highlight the promise of using information on the conglomerate networks of large Chinese manufacturers to improve the design of energy regulations, particularly in contexts where monitoring capacity may be limited.

Finally, in Section 7, we show that our model results are robust to using a wide range of alternative model specifications and parameter values. First, we extend the model to consider the possibility that firms responded to the regulating by improving their energy efficiency. Consistent with our empirical results, we find that firms faced significant costs of improving their energy efficiency. An optimistic calculation suggests that, over a longer horizon, the policy may be rationalized by a GWTP of \$114. Second, we extend the model to allow for preexisting differences in energy efficiency between regulated and unregulated firms. Assuming that Top 1,000 firms are 20% less energy efficient than all other firms yields a GWTP of \$134. Finally, we show that our estimates of the GWTP that rationalizes the policy are not sensitive to alternative assumptions regarding parameter values and model specifications.

This paper contributes in three ways to our understanding of whether energy regulations and interventions aimed at improving energy efficiency are effective in developing countries (e.g., Duflo et al., 2013, 2018; Greenstone and Jack, 2015; Harrison et al., 2015; Ryan, 2018; Ito and Zhang, 2020*a*). First, our setting features a case of strict enforcement, which is given by the Chinese government's use of high-powered incentives that tie environmental performance to cadre promotion (Kahn, Li and Zhao, 2015; Jia, 2017; He, Wang and Zhang, 2020). Second, the program we study represents one of the largest efforts to curb energy use in a developing country. As Auffhammer and Gong (2015) note, the Top 1,000 program along with its expanded version in later years are the "most significant national programs" focusing on energy efficiency and energy conservation in China.[11] Official assessments of the program that compared the variation over time in the energy use of regulated firms concluded that the program was effective, which motivated the expansion of the program. Third, our results highlight the importance of accounting for both within-conglomerate and market-level leakage in an equilibrium framework. By tracing the effects of the regulation along business ownership networks, our results provide a fundamental reassessment of the effectiveness of the Top 1,000 program.[12]

---

[11] Appendix B.1 summarizes prior work studying the policy details of the Top 1,000 program, the role of local government officials in achieving energy savings targets, and the difficulty involved in evaluating energy-saving measures.

[12] Appendix B.2 summarizes prior studies of the economic effects of the program, which have relied on time-series evidence (Wang et al., 2017; Ke et al., 2012) or have focused on evaluating the effects of the regulation on firm-level patenting and productivity (Shen, Zeng and Qu, 2015; Zhang and Huang, 2019; Filippini et al.,

Our finding that the Top 1,000 program impacted economic activity in regulated and unregulated firms contributes to the literature studying the economic costs of environmental regulations. For the US, researchers have documented significant effects of environmental regulations on emissions and economic activity (e.g., Greenstone, 2002; Greenstone, List and Syverson, 2012; Walker, 2013; Shapiro and Walker, 2018; Curtis, 2018). Colmer et al. (2020) find that French firms that are subject to the European Union's emissions trading scheme do not experience significant declines in production and that their declines in energy do not spill over to unregulated firms. He, Wang and Zhang (2020) show that Chinese firms that face more stringent regulations experience significant decreases in productivity. Our paper contributes to our understanding of the economic cost of energy regulation in China, which consumes the lion's share of global industrial energy.

Researchers have also documented that regulations can have spillover effects along firm networks. For instance, Hanna (2010) finds that multinational firms respond to domestic environmental regulations by increasing their investment in foreign countries, and Gibson (2019) and Soliman (2020) find that firms may also shift economic activity to unregulated plants in counties subject to less stringent regulations. Conglomerate spillovers are particularly important in our setting since the Top 1,000 program targeted very large firms with elaborate ownership networks. Our detailed business registration data provide a unique view into how this regulation affected the production decisions of large Chinese conglomerates and how conglomerate spillovers impacted the effectiveness of the regulation. Our model leverages these spillovers to quantify the marginal cost of the regulation, using the fact that conglomerates incur a loss when they distort the within-conglomerate allocation of production (see, e.g., Anderson and Sallee, 2011).

Our paper also takes into account the roles of leakage and market competition in environmental regulation. Research has shown that emissions leakage to unregulated firms can significantly alter the effects and design of environmental policies (e.g., Fowlie, 2009; Fischer and Fox, 2012; Bushnell, Chen and Zaragoza-Watkins, 2014; Baylis, Fullerton and Karney, 2014; Fowlie and Reguant, 2022). We abstract from strategic interactions between firms in a setting with monopolistic competition since we study manufacturing industries with a large number of firms that compete in national markets.[13] This paper quantifies the aggregate and welfare effects of the Top 1,000 program by combining microdata on the operations of Chinese industrial firms, transparent research designs that identify the direct and spillover effects of a prominent energy regulation, and an industry equilibrium model that is consistent with the estimated effects of the program.

2020; Ai, Hu and Li, 2021). Relative to these studies, this paper uses firm-level data on energy consumption and production to identify the direct effects of the program as well as aggregate effects that account for conglomerate- and market-level spillovers.

[13] Studies of energy regulation with strategic interaction often focus on concentrated industries (see, e.g., Mansur, 2007; Ryan, 2012; Fowlie, Reguant and Ryan, 2016).

The combination of these approaches accounts for market competition and leakage effects and shows that conglomerate spillovers are a distinct force that plays a quantitatively important role in the context of China and that feasible conglomerate-level regulations can improve the regulation of energy.

1 Policy Background and Data

This section describes the Top 1,000 Energy Conservation Program. We corroborate the practical implementation of the program, including the reliance on energy consumption for regulation, using detailed interviews with the principal architect of the policy, government officials, and regulated companies. We provide a detailed account of these interviews in Appendix A. We also describe the different datasets that we use to measure economic activity and energy use and our strategy for mapping the ownership networks of Chinese conglomerates.

1.1 The Top 1,000 Program

To save energy and reduce related carbon emissions, the Chinese government's 11^{th} Five-Year Plan (11FYP) set the ambitious goal of reducing the country's energy intensity—defined as energy consumption per unit of GDP—by 20% between 2006 and 2011 (Price, Wang and Yun, 2010). Since the industrial sector accounts for 70% of total energy consumption, the government designed policies that focused on nine energy-intensive industries, which accounted for 80% of the country's industrial energy use. One of these key initiatives was the Top 1,000 Energy Conservation Program, which targeted the firms with the highest energy consumption in the most energy-intensive industries.

The Top 1,000 program was first announced by the National Development and Reform Commission in April 2006, and the corresponding monitoring and assessment measures were released in 2007. The name "Top 1,000" refers to the $1,008$ industrial firms in the nine energy-intensive industries with energy consumption above 180 thousand tce in 2004. The total energy consumption of these 1,008 super-firms was 670 million tce in 2004, accounting for 47% of China's industrial energy consumption and 33% of its total energy consumption. Importantly, since the policy was announced in 2006 and selected firms based on their retrospective 2004 energy consumption, it was not possible to manipulate the list of program participants. Moreover, the list of firms regulated by the program did not change during the five-year period. Table 1 reports the number of firms and their share of energy consumption in each of the regulated industries. Among the Top 1,000 firms, those in the iron and steel, chemical, and electric power industries accounted for approximately 63% of the firms and 68% of the regulated energy consumption in 2005.

The Top 1,000 program was designed based on the belief that Chinese industries could significantly increase energy efficiency at a low cost (e.g., McKinsey & Co., 2009). The program was influenced by voluntary agreement programs in developed countries and had two stated goals: to significantly increase the energy efficiency of these super-firms and to save 100 million tce in energy consumption by 2011.[14] Given the program's quick implementation, many aspects of voluntary agreement programs (such as providing technological expertise or financing energy efficiency improvements) played a relatively minor role (Price, Wang and Yun, 2010).

To implement the policy, the central government assigned an energy savings target to each provincial government. In Appendix A.2, we document that implementing this target savings measure was not feasible for either the government or the regulated firms. In practice, firms were regulated based on energy use quotas and were not primarily evaluated on energy savings or efficiency.[15] These firms were subject to annual energy audits carried out by a third party and faced potential additional audits by the Ministry of Industry and Information Technology and the National Energy Administration. Leaders of provincial governments and state-owned enterprises (SOEs) were then evaluated on whether their energy savings targets had been met. Accordingly, local government officials monitored and enforced the energy savings targets of Top 1,000 firms very closely.[16]

Due to this perceived success under the 11FYP, the Top 1,000 program was expanded into the "Top 10,000" Energy Conservation Program during the 12^{th} Five-Year Plan (12FYP) in 2012. In this case, "Top 10,000" refers to 16,078 energy-intensive firms with energy consumption above 10 thousand tce in 2010. These firms account for 60% of China's total energy consumption. As in the Top 1,000 program, firms among the Top 10,000 were required to improve their energy efficiency with a goal of saving a total of 250 million tce during the 12FYP. Our primary analysis focuses on Top 1,000 firms between 2001 and 2011. Since the industrial firms in the Top 10,000 (but not in the Top 1,000) were also energy intensive but were not regulated during the 11FYP,

---

[14] Appendix A.1 clarifies that the Top 1,000 Energy Conservation Program focused on energy conservation. Neither pollution reduction nor reallocation of economic activity across different regions was a primary goal of the program. This appendix also compares the scope of the Top 1,000 program to that of other large programs around the world.

[15] Appendix A.3 provides details on how local officials determined the individual energy use quotas of each of the Top 1,000 firms. The firms' energy quotas were based on historical energy use relative to an expected growth trajectory, which was likely set at the industry level. For these reasons, firms would not be mechanically compliant based on secular, industry-wide improvements in energy efficiency.

[16] Under the "one-vote veto" criteria, officials would not be considered for promotions or awards if the province or any of the local Top 1,000 firms did not achieve their targets. Similarly, leaders of state-owned enterprises that did not meet their targets did not receive annual bonuses. In interviews with executives of Top 1,000 firms, we confirmed that a "live energy consumption monitoring sysmtem" allowed the governments both to track energy usage and to credibly threaten to shut down production at plants that exceeded their energy use targets. In this way, the Chinese setting contrasts with other developing country settings where the design of incentives for energy auditors plays a key role (e.g., Duflo et al., 2013, 2018).

they serve as useful controls for our empirical analysis.[17]

1.2 Firm Data

Our empirical analyses combine several rich datasets that describe firm-level production and energy use. We obtain the list of firms in the Top 1,000 and Top 10,000 programs from the National Development and Reform Commission.[18] We then collect detailed information on firm energy consumption from 2001 to 2010 from China's Environmental Statistics Database (CESD) provided by China's Ministry of Environmental Protection. These data allow us to measure the effects of the regulation on the production and energy use of Top 1,000 and Top 10,000 firms. The CESD data are primarily collected by local environmental agencies and focus on polluting enterprises. These data are subject to audits by environmental protection agencies at both the local and national levels and cover the vast majority of Top 1,000 firms and a subset of Top 10,000 firms. Because the CESD reports energy consumption only from primary sources (e.g., coal, oil, gas), our analyses of energy use and energy efficiency exclude firms in industries that rely mainly on electricity.[19] We also restrict the sample to those firms with complete yearly data on coal use. After these restrictions, our estimation sample comprises of 427 Top 1,000 firms, which we observe for an average of 8 years.[20]

We complement these data with two additional datasets. First, we use data on firm characteristics from the Annual Survey of Industrial Firms (ASIF) from the National Bureau of Statistics (2001–2009 and 2011).[21] This dataset provides detailed information on a firm's industry, address, ownership, output, and financial information and covers all industrial firms with annual revenue above 5 million RMB (approximately 800,000 USD). These data are also valuable since they cover a large number of firms related to Top 1,000 firms, which allows us to estimate the spillover effects of the program. Second, we use data from the Annual Tax Survey (ATS) for 2009 and 2010 in robustness checks.

A potential concern is that firms may manipulate reported data as a way to meet compliance

---

[17] An important consideration is whether firms that were later part of the Top 10,000 expected that the Top 1,000 program would be expanded. This is unlikely to be the case since the details of the program were developed after the 12FYP by the National Development and Reform Commission, which did not announce the Top 10,000 program until 2012.

[18] Appendix C describes how we merge these lists with other firm-level datasets.

[19] In practice, we exclude industries where electricity consumption accounts for more than 30% of total energy consumption. As we show below, our results are robust to setting this threshold to between 25% and 50% and to including firms in all industries.

[20] Table A.1 shows the effects of the sample restrictions on the number of Top 1,000 firms in the data. As we show below, our results are robust to different decisions on how to construct the sample. We find similar results when we extend the sample by using administrative tax data to fill in missing values for energy use and when we restrict the sample to include firms that existed over the whole sample period.

[21] As is well known in the literature, data for the 2010 ASIF display a number of irregularities and are often excluded from statistical analyses. As we show below, our results are robust to using administrative tax data for 2009 and 2010.

requirements. Indeed, Karplus, Shen and Zhang (2020) show that self-reported data by firms is subject to manipulation. For this reason, our analyses rely on multiple measures of output and energy use from survey and administrative data that are unrelated to the government's evaluation of the program. One advantage of using multiple datasets is that we can cross-check our data to ensure our results are not driven by misreporting or other data quality issues. In Figure A.1, we show that firms report similar output and coal consumption in the CESD and tax data, which are collected independently and are not used to evaluate compliance with energy and environmental policies.[22]

Panel A of Table 2 reports summary statistics for the Top 1,000 and Top 10,000 firms in our sample. Top 1,000 firms are larger, older, more likely to be state owned, and more export oriented than Top 10,000 firms. This table also shows that Top 1,000 firms have slightly lower energy efficiency (defined as the ratio of output to energy use) than Top 10,000 firms. However, this difference is driven mostly by industry differences since Top 1,000 firms are more likely to be in energy-intensive heavy industries. As we show below, our empirical analyses are robust to controlling for these firm-level characteristics.

1.3 Mapping Conglomerate Networks

We identify firms' ownership networks using data from China's Administrative Registration Database (CARD). These data are collected by the State Administration of Industry and Commerce and list the registration information of all firms in China starting in 1980, including firm name, registration number, date of establishment, address, ownership, registered capital and related legal persons. Importantly, the data provide detailed shareholder information, which allows us to construct firm ownership networks at multiple levels.

We construct ownership networks using the four types of linkages displayed in Figure 2. First, we include wholly owned subsidiaries of regulated firms as related parties. These affiliate linkages are analogous to the concept of plants in a multi-establishment firm in the US. Second, we include firms that are at least partially owned by regulated firms. We consider firms to be related if they are owned by a regulated firm by up to two levels of investment relations. Although in practice most related firms are fully owned, we require that the regulated firm own at least 25% of the related firm at each level of investment. Third, we include shareholders of regulated firms, and we allow up to two levels of shareholder links. Finally, we include firms that are fully or partly owned by the shareholders of a regulated firm.[23] We exclude firms related only through the state-owned management committee.

---

[22] Appendix C discusses government procedures to ensure data quality, including audits and penalties for misreporting.

[23] We again allow two levels of investment, and we require ownership to be at least 25% at each level. Figure A.2 depicts all the possible links that we consider.

Panel B of Table 2 shows that we can identify 46,178 related parties of Top 1,000 firms in the CARD. Since a large number of related parties are service firms or small firms not recorded in the ASIF, we match 7,329 firms in the ASIF.[24] In our baseline regressions, we require related firms to be in the same 4-digit industry as a related Top 1,000 firm. Our main sample of related firms includes 2,466 industrial firms.[25] Since it is likely very hard to shift production to firms in other narrowly defined industries, we analyze firms within the same 2-digit industry but outside the 4-digit industry in a placebo test. A potential concern with the CARD data is that some of the related firms may not be engaged in production and may, in fact, be holding companies. By merging the CARD data with the ASIF and the CESD, we ensure that our results are driven by real economic activity in industrial firms.

Panel B of Table 2 also examines the robustness of our network definitions to alternative assumptions. Allowing for up to six levels of relations does not have a large effect on our sample of related firms in the same 4-digit industry. Decreasing the ownership requirements to 20% has a small effect on the number of related firms, and the number of related parties is similar when we increase the ownership ratio to 51%. These results suggest that, within narrowly defined industries, firm ownership networks are very compact. Importantly, our measure of firm networks uses data from 2018, after the policy was implemented. Therefore, our business networks include any firms that may have been acquired by regulated conglomerates as a result of the regulation.[26] Moreover, it is important to note that regulated firms could not escape the regulation by splitting into smaller firms. Since local policymakers face regional energy use targets, they have strong incentives to ensure that any initially regulated firm meets its energy target. If firms split, the energy use targets would accompany the firms after any such separation.

The merged CARD and ASIF data reveal some interesting patterns. First, we find that Top 1,000 firms have an average of 2.45 related parties in narrowly defined industries. Second, since Top 1,000 firms are, in most cases, the largest firms in each industry, their related parties are smaller. On average, the output of related firms is 19.3% of the output of regulated firms. These facts imply that conglomerates may have had significant scope to substitute production across related firms.[27] However, it is also unlikely that related parties could fully make up for

---

[24] As we discuss below, accounting for spillover effects to firms that are smaller than the threshold required for inclusion in the ASIF does not meaningfully impact the total spillover effect of the policy.

[25] Omitting firms in unrelated industries is unlikely to affect our results since super-firms like Top 1,000 firms would not be able to shift production to service firms or very small firms.

[26] Using the ownership change information in the CARD, we estimate that between 2007 and 2018, less than 4% of related firms experienced significant ownership changes—defined as an ownership transfer of more than 25% to or from firms not in the same conglomerate.

[27] In Chen et al. (2021), we show that most related parties of regulated firms are located in the same province as the regulated firm. For this reason, we do not expect substitution of production across related parties to significantly affect the provincial distribution of energy use or related pollution. In Appendix H, we also show that the program did not significantly alter the allocation of production across cities with different levels of emissions and population density.

production declines in Top 1,000 firms. Third, firms within conglomerates have an interesting relative size distribution. To produce Panel A of Figure 3, we compute each firm's size relative to the largest firm in the group; we then plot the average relative size by firm rank. A striking fact of this graph is that the average relative size within a conglomerate declines sharply with firm rank: the second-largest firm in a conglomerate is only 29% as large as the largest firm, on average. Interestingly, the decline in relative firm size is almost geometric, a fact that we use in our structural model. Finally, Panel B of Figure 3 shows the relation between the output of the largest firm and the number of firms in a conglomerate. The fact that conglomerates with more firms also have larger leading firms suggests that the number of firms in a conglomerate might depend on technological efficiencies shared by all firms in a conglomerate.

2 Effects of the Policy on Regulated Firms

As detailed in Section 1, the Top 1,000 program mandated that firms reduce their energy use. To study the effects of the policy, we compare the activities of regulated firms relative to those of other large firms operating in energy-intensive industries. Specifically, we use firms that became regulated after 2011 as part of the Top 10,000 program as controls. Because related firms in the same conglomerate as a regulated Top 1,000 firm may be indirectly affected by the policy, we remove these firms from the set of control firms. In addition, since control firms can be affected by market-level spillovers, the relative changes between the regulated and control firms that are identified by our difference-in-differences estimates combine the program's effects on regulated firms with its indirect effects on control firms. We interpret these effects using our model in Section 5.3, which allows us to separate and quantify these forces.

The identifying assumption of this difference-in-differences analysis is that, absent the Top 1,000 regulation, the energy use and output of Top 10,000 firms would have trended similarly to those of Top 1,000 firms. To provide evidence that these firms had similar trends prior to the implementation of this regulation, we use firm data from the CESD to estimate an event-study analysis of the form:

$$Y_{ijkt} = \sum_{\tau \neq 2006}^{2010} \beta_\tau \times Treat_i \times Year_\tau + \alpha_i + \eta_{jt} + \delta_{kt} + \varepsilon_{ijkt}, \tag{1}$$

where Y_{ijkt} is a dependent variable for firm i in industry j, province k and year t. $Treat_i$ is a treatment group indicator that equals 1 for Top 1,000 firms and 0 for Top 10,000 firms. The coefficients β_τ from this specification represent differences in the dependent variable between Top 1,000 and Top 10,000 firms in each year. Given that the policy evaluation began in 2007, we identify the effects of the policy relative to performance before 2006. We include firm-level fixed effects α_i and year fixed effects in all regressions, and we show that our results are robust to inclusion of (2-digit) industry-by-year fixed effects η_{jt} and province-by-year fixed effects δ_{kt}. We cluster standard errors at the firm level.

Figure 4 presents a visual implementation of our difference-in-differences estimation strategy. Panel A in Figure 4 displays the β_τ coefficients when the outcome variable is firm-level energy use (total coal consumption equivalent). This figure shows that, prior to the implementation of the regulation, our treatment and control firms had similar trends. Additionally, this figure makes clear that the policy did indeed succeed in lowering the energy use of regulated firms relative to that of unregulated firms.[28] Panel B of this figure compares these year-by-year effects to the overall trend in energy consumption.[29] As this figure shows, the program successfully arrested the explosive growth in the energy use of regulated firms.

We quantify the effects of the policy by estimating difference-in-differences specifications of the form:

$$Y_{ijkt} = \beta Treat_i \times Post_t + X'_{it}\gamma + \alpha_i + \eta_{jt} + \delta_{kt} + \varepsilon_{ijkt}, \tag{2}$$

where $Post_t$ is an indicator that equals one after 2006. In addition to the different fixed effects, some specifications include controls for firm characteristics X_{it}, which include indicators for state-owned firms and exporting firms, measures of profitability (e.g., return on assets), and firm age. Panel A of Table 3 shows that, on average, the total energy consumption of regulated firms decreased by 12%–16%. These estimates are stable across specifications that include different levels of fixed effects and firm controls. To interpret the magnitude of this effect, recall that the regulated firms consumed 670 million tce in 2004. Taking the coefficients in Table 3 at face value therefore implies annual reductions in energy use of close to 100 million tce or approximately 20% of the total industrial energy use of the European Union.

To discern whether this reduction in energy use was driven by changes in economic activity or in energy efficiency, we now estimate the effects of the program on firm output (i.e., revenue). Panels C–D of Figure 4 show that, after the reform, firm output in regulated firms also decreased significantly. Indeed, Panel B of Table 3 reports output declines of between 10% and 23%, depending on the specification. Accounting for the declines in output implies that the policy had limited impacts on energy efficiency. Panels E–F of Figure 4 show that we cannot reject the null hypothesis that the policy had no impact on energy efficiency. Based on the specification with both industry- and province-by-year fixed effects in column (3) of Panel C of Table 3, the 95% confidence interval rules out that the policy increased energy efficiency by more than 4%,

---

[28] One potential concern is that our results may be contaminated by mean reversion. Because firms were regulated based on their 2004 energy use, one possibility is that regulated firms had idiosyncratically large levels of energy use in 2004 that reverted to lower levels in later years. As this and other similar graphs show, the outcomes for 2004 are not significantly different from those for 2001–2003, nor do we see large differences from the outcomes for 2005–2006.

[29] For visual clarity, Panels B, D, and F in Figure 4 follow Ohrn (2018) by plotting trends for the control group that are re-scaled to have the same average level in the pre-period as the treated group.

which is significantly below the government's goal of improving energy efficiency by 20%. Our results show that, contrary to the hypothesis that there was "low-hanging fruit" to be harvested in terms of energy efficiency, regulated firms were not able to significantly improve their energy efficiency over a period of five years.[30]

We now explore the robustness of the effects of the Top 1,000 program on regulated firms. First, as we discuss in Section 1, our analysis sample excludes some firms based on missing data and firms in industries that rely primarily on electricity. Table A.4 and Figure A.3 show that our results are robust to excluding more or fewer industries based on their electricity use and even to including all regulated industries regardless of their primary energy source. This table also shows that we obtain similar results when we use tax survey data on energy use to increase the number of firms in our sample.[31] Second, our results are robust to different constructions of the sample of firms in our regression and to the exclusion of new firms and those that exit during our sample period.[32] Third, we explore whether regulated firms respond to the Top 1,000 program by changing their energy mix or by relying on different production inputs. As we show in Appendix E, we do not find evidence of other margins of substitution. Fourth, one potential concern is that our results may be influenced by other, concurrent policies. Appendix D clarifies that this is not the case by showing that our estimates are independent of the effects of other pollution monitoring policies. As we show in Table A.9, these policies did not significantly impact the operations of Top 1,000 firms, and our results are robust to excluding firms that are part of these other programs. Finally, we explore the potential for heterogeneous effects across industries. Given the small number of regulated firms in each industry, we estimate heterogeneous effects across broad industry groups. Table A.10 shows similar effects of the program across different industry groups, and Table A.11 shows that SOEs do not have statistically distinct responses to the program.[33]

---

[30] The lack of an effect on energy efficiency is also consistent with the extant view (e.g., Price, Wang and Yun, 2010) that other aspects of the reform (such as financing for energy efficiency improvements) played a relatively minor role in the short term. In Section 7.1, we consider how allowing for endogenous investments in energy efficiency impacts the aggregate outcomes and welfare effects of the policy as firms adopt energy saving technologies over a longer period of time. In addition, Appendix E.2 explores whether regulated firms responded to the reform by adjusting the mix of their production inputs and shows that the program did not have a significant impact on the input mix.

[31] While the setting of the Top 1,000 program may seem amenable to a regression discontinuity design, in practice, there are few treated and control firms at the energy use threshold, which makes such an approach unfeasible. Tables A.5–A.6 and Figures A.4–A.5 show that our main results are generally robust to restricting the sample to more comparable treatment and control firms on the basis of their rank in the Top 10,000 program or on an estimated likelihood of treatment.

[32] Table A.7 shows that the effects of the program on regulated firms are robust to narrowing the sample to include only firms that existed before 2006 and after 2010. Table A.8 shows that our results are also robust to using an almost-balanced panel where we require that firms have no more than one missing year in the data. Note that due to the survey nature of the CESD data, our sample is substantially smaller in this case. Nonetheless, these results show that our estimates are not driven by firms entering the sample or ceasing operations.

[33] We also explore the effects of the program on other outcomes. In Table A.12 and Figure A.6, we show

The effects of the policy on regulated firms paint a picture of mixed success. On the one hand, the regulation succeeded in achieving a meaningful reduction in the energy use of energy-intensive firms. However, this reduction did not come about through a significant increase in energy efficiency, which—while not directly targeted—was one of the underlying intents of the policy. The next section studies whether conglomerates avoided the burden of the regulation by shifting economic activity to related parties.

3 Spillover Effects of the Policy through Ownership Networks

Regulated firms have strong incentives to shift production to related parties. By shifting production, conglomerates can partially offset declines in economic activity in regulated firms. Such shifting also allows conglomerates to comply with the letter of the regulation—if not with its intent—without having to invest in potentially costly improvements to energy efficiency.

To measure the empirical importance of conglomerate spillovers, we use CARD data on the ownership networks of regulated firms to identify firms that may have indirectly expanded as a consequence of the Top 1,000 regulation. We then use matching methods to identify control firms that were (1) not part of the Top 1,000 program, (2) not related to a regulated firm, and (3) in the same industry and of similar size (measured in terms of output) in the years prior to the regulation. Using these firms as controls, we then conduct event-study and difference-in-differences analyses using specifications similar to those in Equations (1) and (2).[34] In this setting, the $Treat_i$ variable is now an indicator of whether a firm is related to a Top 1,000 firm. As we discuss in Section 1, we focus our study of spillovers on related firms in the same 4-digit industry as the regulated firm. This restriction follows from the logic that only firms selling products similar to those of the regulated firms may be able to make up for the production decline in Top 1,000 firms.

Figure 5 plots the results of these event-study analyses using ASIF data. Panel A shows that, prior to the regulation, related firms had output trends similar to those of unrelated firms. After the regulation, firms related to Top 1,000 firms saw significant increases in output that persisted for several years. The last column of Panel A of Table 4 shows that related firms expanded by

---

that regulated firms experienced a decline in their probability of investing after the regulation was enacted. Similarly, Figure A.7 shows that regulated conglomerates did not respond to the regulation by adding new affiliates. Additionally, we test the Porter and van der Linde (1995) hypothesis by examining whether firms became more innovative after the regulation. Figure A.8 shows no increase in the filing of patents related to energy efficiency in regulated firms.

[34] Specifically, we use one-to-one matching within 4-digit industries based on the Euclidean distance in output levels before the policy. To ensure that the matched firms are comparable to the related firms, we drop the 5% of observations with the least comparable matches. As we show below, our results are robust to using alternative matching methods.

13%, on average, after the regulation. This table also shows that we obtain very similar results across specifications with different levels of fixed effects and with firm-level controls.

To gauge the magnitude of these spillover effects, it is important that we account for the number of related parties of each regulated firm and for their relative size. On average, Top 1,000 firms have 2.45 related parties. However, since the average related firm is only 19.3% as large as its regulated counterpart, we calculate that conglomerates could shift only close to 41% of the output decline in regulated firms.[35] This result is informative for a couple of reasons. First, it shows that conglomerates were not able to fully circumvent the regulation. Second, combined with the null effect of the program on the energy efficiency of regulated firms, the result shows that firms were unable or unwilling to increase their energy efficiency in production processes even if this meant losing profits to competitors. The result that related firms display an increase in economic activity is robust across a number of checks. First, we show that we obtain similar results when we use the entropy balancing method of Hainmueller (2012) to find controls for related firms (see Figure A.10 and Tables A.15–A.16).[36] Second, we show that only those related firms operating in regulated firms' own narrowly defined industries—and that could thus possibly produce substitute output—increased their economic activity. Indeed, Panel B of Figure 5 and Panel B of Table 4 show no impact on the output of related firms operating outside the 4-digit industry of (but still in the same 2-digit industry as) the regulated firm. This placebo test rules out the possibility that firms related to large conglomerates saw increases in economic activity after 2007, say, in response to the financial crisis or other shocks or trends. Third, these results are robust to alternative definitions of ownership networks. Table A.17 shows similar spillover effects when we drop related firms with ownership changes between 2007 and 2018, when we restrict the sample by requiring 51% ownership at each link, and when we expand the sample to include 6 levels of relations and 20% ownership stakes. Fourth, these results are robust to dropping firms in power generation (see Table A.18 and Figure A.11). Finally, we assuage concerns that our results may be affected by data quality issues by showing similar effects when

---

[35] Using the estimate on related firms from column (4) of Panel A of Table 4 of 12.7%, we calculate that the overall increase in related firms amounted to $6\%(\approx 2.45 \times 19.3\% \times 12.7\%)$ of the output of regulated firms. This increase is 41% of the comparable 14.5% decrease from column (4) of Panel B of Table 3. We consider the sensitivity of this estimate to the measurement of business networks in three ways. First, using estimates from the specifications in column (3), we obtain an estimate of 27%. ($i.e., 27\% \approx 2.45 \times 19.3\% \times 11.8\%/20.4\%$). Second, if we suppose that regulated firms have an average of 3 related firms, spillovers would account for 51% of the output decline in regulated firms. Finally, we consider the role of related firms that are not in the ASIF. To obtain an upper bound of the importance of these firms, we assume that their output matches the size threshold of 5 million RMB for inclusion in the ASIF. Even in this case, the 3,081 related firms in the same 4-digit industry can account for at most 0.21% of the conglomerate's output. For this reason, our spillover calculation is largely unaffected by excluding firms not in the ASIF.

[36] Our estimates of spillover effects are also not driven by the entry and exit of related firms. To find a suitable control, our matching analysis requires firms to have existed prior to 2006. Moreover, because we map business networks in 2018, our estimates include the effects on firms that joined regulated business groups after the program.

we rely on tax data to measure the output of related firms (see Figure A.12 and Table A.19).[37]

We now explore the potential for heterogeneity in the spillovers across related firms. Table A.20 shows that related firms in higher terciles of the size distribution display larger increases in output. This result suggests that larger related firms were more able to expand or, alternatively, that these firms had larger installed production capacity. As in our analysis of regulated firms, we explore potential effect heterogeneity across industries. Table A.22 shows no significant differences in how related firms in different industries responded to the program.

We also explore the possibility that the regulation shifted the location of economic activity. First, as we show in Appendix H, the spillover effects of the regulation did not disproportionately shift production to areas with higher population density or with higher preexisting levels of industrial emissions. Similarly, Appendix F shows that, within regulated firms, the program did not lead to increased concentration of production in more energy-efficient firms.

Having established that conglomerates shifted output across related parties, we now explore whether these firms also saw changes in energy use and energy efficiency. Panels C and D of Figure 5 report these results using data from the CESD. Panel C shows that related firms saw an increase in energy use after the regulation. Panel C of Table 4 shows that energy use in related firms increased by 30%–32% after the regulation. Note that the number of observations in this panel is smaller than that in Panel A of Table 4. This is because related firms are overall smaller and only the larger related firms are included in the CESD. These larger effects are consistent with our results in Table A.20 showing larger spillover effects on larger related firms. While the available data include firms across all affected industries, caution is warranted in ascribing these increases in energy use to all related firms. Panel D of Figure 5 and Panel D of Table 4 show that these firms did not experience statistically significant changes in energy efficiency.

Overall, we find robust evidence that conglomerates shifted production across related parties. On average, this shifting behavior allowed conglomerates to recover approximately 40% of the output reduction in regulated firms. As we show in Section 6, the ability to shift production to related firms diminished the aggregate energy savings from and lowered the shadow cost of the regulation.

Market-Level Spillovers

Since related parties could not compensate the entire output loss of Top 1,000 firms, other firms in regulated industries may have been indirectly affected by the energy conservation program due to reduced competition. Intuitively, we would expect larger increases in the output of unrelated and unregulated firms in industries where the Top 1,000 program covered a larger share of industrial

---

[37] We also find positive spillover effects on other measures of economic activity. Table A.21 shows estimates of positive spillover effects on sales, profits, capital and labor (see Figure A.13 for the corresponding event studies).

energy use. To test this hypothesis, we create the variable $spillover_j$, which is the ratio of total energy savings targets of Top 1,000 firms in a 2-digit manufacturing industry j to the total energy consumption of the industry in 2004. We then estimate the following difference-in-differences specification:

$$Y_{ijt} = \beta spillover_j \times Post_t + X'_{it}\gamma + \alpha_i + \tau_t + \varepsilon_{ijt}. \tag{3}$$

To interpret the coefficient β as the average spillover effect, we normalize the $spillover_j$ variable by the average exposure across regulated industries. Since the variation in the independent variable is at the industry–year level, we do not include industry-by-year fixed effects in this regression. We instead use firm fixed effects and year fixed effects only, and we additionally control for overall output and energy use at the industry–year level.[38] Finally, to ensure that market-level spillovers are not contaminated by ownership-network spillovers, we exclude firms related to Top 1,000 firms from this specification.

Panel A of Figure 6 shows that unregulated firms in industries with stricter regulation increased their output significantly after the policy was implemented. Panel A of Table 5 shows that, across all industries, the average market-level spillover led to a 7%–8% increase in the output of unregulated firms. The regressions in the first two columns of this table include both regulated and unregulated industries. We find similar increases when we include only firms in regulated industries. In this case, the identifying variation is driven solely by differences in regulation intensity across industries.[39]

These results yield a couple of insights. First, the findings further confirm our previous estimates that related parties were not able to offset the full output loss of Top 1,000 firms. Second, a full accounting of the spillover effects of the regulation needs to include both within-conglomerate spillovers and market-level spillovers. Third, a potential limitation of the difference-in-differences analyses is that their interpretation depends on the strength of the conglomerate and market spillovers. The next section builds on these insights by proposing a model of conglomerate production. The model clarifies the interpretation of our reduced-form estimates in the presence of market and conglomerate spillovers, computes the aggregate effects of the Top 1,000 program, and allows us to consider the effects of alternative policies.

---

[38] Note that the variation in $spillover_j$ is absorbed in our previous specifications that include industry-by-year fixed effects. By controlling for industry-level aggregates, the coefficient β in Equation 3 captures the impact of the regulation on the market share of unregulated firms.

[39] Similarly to spillovers to related firms, market spillovers are concentrated in the same 4-digit industries as Top 1,000 firms; see Table A.23. We also confirm that the market spillover results are not driven by firm entry. Specifically, in Table A.24, we report similar estimates of market-level spillovers when we restrict the sample to firms in operation prior to 2006.

4 A Model of Conglomerates with Regulation

This section presents an industry equilibrium model of conglomerate production that is consistent with the cross-sectional data patterns and reduced-form responses to the energy regulation. Appendix J provides detailed derivations of the model results.

4.1 Demand and Technology

Our industry equilibrium model draws the structure of product differentiation and monopolistic competition from Melitz (2003). We consider an individual sector with an exogenous aggregate expenditure R. The representative consumer has CES preferences over a continuum of varieties $\omega \in \Omega$:

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho} d\omega \right]^{1/\rho},$$

where $q(\omega)$ represents the consumption level of variety ω and $\sigma = 1/(1-\rho) > 1$ denotes the elasticity of substitution between varieties.[40]

Utility maximization by the representative consumer yields the following residual demand curve for each variety ω:

$$q(\omega) = RP^{\sigma-1} p(\omega)^{-\sigma},$$

where $P = [\int_{\omega \in \Omega} p(\omega)^{1-\sigma} d\omega]^{\frac{1}{1-\sigma}}$ is the aggregate price index.[41]

We define a conglomerate in our model by the presence of a variety ω that can be manufactured by multiple affiliates.[42] Each conglomerate starts with a central producer—the model counterpart of a Top 1,000 firm. Conglomerates have heterogeneous production efficiencies ϕ, which are drawn from the distribution $G(\phi)$ with density $g(\phi)$.

Production at each affiliate i requires capital k_i, energy e_i, and variable inputs l_i. Energy and variable inputs are combined using Leontief technology $\tilde{l}_i = \min\{l_i, e_i \nu_i\}$, where ν_i is the affiliate's energy efficiency. The assumption that energy and variable inputs are perfect complements follows recent work in this area (e.g., van Biesebroeck, 2003; Fabrizio, Rose and Wolfram, 2007;

---

[40] Since the regulated firms produce raw and intermediate materials, one can view the representative consumer as a stand-in for the downstream industry.

[41] This market structure implicitly assumes that this industry is not characterized by dominant firms that may act strategically. This is a reasonable assumption in our setting since we study manufacturing industries that, even when narrowly defined, feature a large number of firms and that serve a national market.

[42] This assumption implies that the outputs of related firms are perfect substitutes. The assumption that related firms' output is more substitutable with regulated firms' than with unrelated firms' output is consistent with the spillover results in Table 4. To see this, note that equal degrees of substitution would imply that related and unrelated firms would have responded in equal measure to the regulation. In this case, the empirical analysis would reveal no differential response between related and unrelated firms. In Section 7.4, we explore the robustness of our results to allowing for an imperfect degree of substitution within conglomerates.

Gao and Van Biesebroeck, 2014; Ryan, 2018).[43] Production at affiliate i is then $q_i = \phi_i \tilde{l}_i^{\alpha_l} k_i^{\alpha_k}$, which is subject to decreasing returns to scale, i.e., $\alpha = \alpha_k + \alpha_l < 1$. The decreasing-returns-to-scale assumption is consistent with the literature on span of control. Intuitively, conglomerates may operate more firms as a way to escape decreasing returns to scale and as a way to share production knowledge ϕ across firms. However, as we show in Panel A of Figure 3, conglomerates are not able to replicate the same scale across related firms. To match this fact, we assume that the productivity of the i^{th} affiliated firm is $\delta^{i-1}\phi$. This assumption can be interpreted as either a limit on the span of managerial control or a measure of imperfect knowledge-sharing across firms. Finally, each manufacturing establishment incurs a fixed outlay of capital denoted by f. This assumption is motivated by the fact that conglomerates have a finite number of affiliates. Because of the fixed cost, Top 1,000 firms with larger efficiency values ϕ will also have a larger number of related firms, matching the data pattern in Panel B of Figure 3.

We consider the conglomerate's problem in two stages. Prior to the regulation, conglomerates observe their productivity ϕ and optimally choose the number of affiliated firms n and the amount of capital $\{k_i\}_{i=1}^n$ and variable inputs $\{l_i\}_{i=1}^n$ for each affiliate.[44] After the regulation, since capital is quasi-fixed, the conglomerate adjusts its variable inputs to maximize profits. We initially assume that energy efficiency is constant and fixed (i.e., $\nu_i = 1$ for all firms) but consider costly investments to improve energy efficiency and heterogeneous efficiencies in Sections 7.1 and 7.2.

4.2 Profit Maximization

The conglomerate takes the prices of energy p_e, capital r, and the variable input bundle w as given. Given the Leontief technology, the conglomerate sets $l_i = e_i$ so that the cost of intermediate inputs is $w + p_e$. Holding the number of affiliates n constant, the conglomerate maximizes

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^n, \{k_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\sum_{i=1}^{n} \phi \delta^{i-1} k_i^{\alpha_k} l_i^{\alpha_l} \right]^{\rho} - (w + p_e) \sum_{i=1}^{n} l_i - r \sum_{i=1}^{n} k_i \right\}. \tag{4}$$

For a firm i, the first-order conditions for l_i and k_i imply that $l_i = \frac{\alpha_l}{\alpha_k} \frac{r}{(w+p_e)} k_i$. Substituting this expression and comparing the first-order conditions for k_1 and k_i, we obtain the following result.

Proposition 1 (Within-Conglomerate Distribution)**.** *Absent regulation, the inputs and the output of producers in a conglomerate follow a decreasing geometric sequence given by*

$$\frac{q_i}{q_1} = \frac{k_i}{k_1} = \frac{l_i}{l_1} = \frac{e_i}{e_1} = \delta^{\frac{i-1}{1-\alpha}}. \tag{5}$$

---

[43] Fabrizio, Rose and Wolfram (2007); Gao and Van Biesebroeck (2014) adopt this assumption from van Biesebroeck (2003) in the context of energy generation. Gao and Van Biesebroeck (2014) study the case of China. Ryan (2018) estimates a production function with energy using data from India and finds that energy and unskilled labor are close to perfect complements.

[44] Conglomerates can choose $n = 0$, which we interpret as an exit decision.

The within-conglomerate distribution described in Proposition 1 is broadly consistent with the empirical pattern in Panel A of Figure 3, where the average output of the second-largest affiliate in a conglomerate is less than 30% of that of the largest one and where the output of other affiliated producers in the conglomerate decreases exponentially with their rank i. Equation 5 links this distribution to two model parameters. First, the size gap among affiliates is larger if within-group knowledge depreciation is more severe (lower δ). Second, if firms are closer to having constant-returns-to-scale production (α is closer to one), the conglomerate concentrates more activity in its top producer, which increases the dispersion of the within-group size distribution.

To consider the choice of total capital $K_n = \sum_i^n k_i$, define the conglomerate's total productivity $\phi\Delta_n = \phi[\sum_{i=1}^n (\delta^{i-1})^{\frac{1}{1-\alpha}}]^{1-\alpha}$ and the constant $C_\pi = (1-\alpha\rho)\left[\left(\frac{\rho\alpha_l}{w+p_e}\right)^{\alpha_l\rho}\left(\frac{\rho\alpha_k}{r}\right)^{\alpha_k\rho}\right]^{\frac{1}{1-\alpha\rho}}$. We reformulate Equation 4 using the results of Proposition 1 so the optimal choice of capital K_n solves

$$\pi(\phi, n) = \max_{K_n}\left\{\frac{R^{1-\rho}P^\rho C_\pi^{1-\alpha\rho}}{(1-\alpha\rho)^{1-\alpha\rho}}\left(\frac{\rho\alpha_k}{r}\right)^{-\alpha\rho}(\phi\Delta_n)^\rho K_n^{\alpha\rho} - r\left(\frac{\alpha}{\alpha_k}\right)K_n\right\}.$$

The optimal capital K_n and the firm profits for a conglomerate of size n are then

$$K_n = \frac{R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi}{(1-\alpha\rho)}\frac{\rho\alpha_k}{r}(\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \quad \text{and} \quad \pi(\phi, n) = R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi(\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}}.$$

Consider now the optimal number of affiliates. The conglomerate adds an affiliate if

$$\pi(\phi, n+1) - \pi(\phi, n) - fr = R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi \times \left[(\phi\Delta_{n+1})^{\frac{\rho}{1-\alpha\rho}} - (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\right] - fr > 0. \tag{6}$$

Adding a new affiliate can improve the conglomerate's revenue and profit by lowering its overall marginal cost curve. On the other hand, the conglomerate incurs a fixed cost of fr when adding a new affiliate. While the marginal benefit of adding a new affiliate is increasing in ϕ, it is also decreasing in the number of existing affiliates n. Since the fixed cost is the same for all affiliates, Equation 6 guarantees the existence of a cutoff value ϕ_n, where conglomerates with efficiency $\phi > \phi_n$ operate at least n affiliated producers.

Proposition 2 (Optimal Conglomerate Size)**.** *Without regulation, the optimal number of firms in a conglomerate n is nondecreasing in its fundamental efficiency ϕ. For $n > 1$, a conglomerate chooses to have n affiliated producers when $\phi_n \le \phi < \phi_{n+1}$, where*

$$\phi_{n+1} = \frac{(fr)^{\frac{1-\rho\alpha}{\rho}}}{R^{\frac{1-\rho}{\rho}}PC_\pi^{\frac{1-\rho\alpha}{\rho}}\left(\Delta_{n+1}^{\frac{\rho}{1-\rho\alpha}} - \Delta_n^{\frac{\rho}{1-\rho\alpha}}\right)^{\frac{1-\rho\alpha}{\rho}}}. \tag{7}$$

Let $\pi(\phi) = \max_n \pi(\phi, n) - nfr$ be the profit for a conglomerate of efficiency ϕ at the optimal number of affiliates. The prediction from Proposition 2 is consistent with the observation in Panel B of Figure 3 that conglomerates with higher efficiency have, on average, a larger number of affiliated firms.

4.3 Equilibrium and Welfare

The unique equilibrium of the model is characterized by product-market clearing, the zero cut-off profit condition, and the free-entry condition.

With M denoting the mass of active firms, the aggregate price index is given by

$$P = \left[\int_{\phi_1}^{\infty} p(\phi)^{1-\sigma} \frac{g(\phi)M}{1-G(\phi_1)} d\phi\right]^{\frac{1}{1-\sigma}}. \tag{8}$$

Conglomerates operate whenever

$$\pi(\phi) \geq 0 \;\Rightarrow\; \phi \geq \phi_1 = \frac{(fr)^{\frac{1-\rho\alpha}{\rho}}}{R^{\frac{1-\rho}{\rho}} P C_\pi^{\frac{1-\rho\alpha}{\rho}}}. \tag{9}$$

Equation 9 shows that only firms with $\phi > \phi_1$ choose to participate in the market.[45]

To enter the market, an entrepreneur pays an entry cost rf_e. Upon entry, the efficiency of the conglomerate ϕ is realized. Since the conglomerate operates only if $\phi > \phi_1$, the free-entry condition is given by

$$\int_{\phi_1}^{\infty} \pi(\phi) g(\phi) d\phi - rf_e = 0. \tag{10}$$

An equilibrium is given by the exit threshold ϕ_1 and the mass of active conglomerates M such that (1) conglomerates make optimal allocation and size decisions, (2) the product market clears, and (3) the zero-profit and free-entry conditions (Equations 9–10) are satisfied.

Welfare depends on consumption utility and on the utility costs of energy use. The CES preferences of the representative consumer imply that indirect utility is given by $\frac{R}{P}$, where R is total expenditure. Utility decreases in total carbon emissions $\beta_0 E$, where E denotes aggregate energy use and β_0 captures the carbon dioxide emitted per unit of energy, as well as in pollution $\beta_1 E$, where β_1 captures the composite effect of energy use on pollution. We assume that welfare takes the form

$$W = \left(\frac{R}{P}\right)^{1-\kappa} \left(\frac{1}{\beta_0 E}\right)^{\kappa_0} \left(\frac{1}{\beta_1 E}\right)^{\kappa_1}, \tag{11}$$

where the parameter κ_0 captures the social welfare loss from carbon emissions, κ_1 captures the welfare loss associated with pollution in China, and $\kappa = \kappa_0 + \kappa_1$.[46]

---

[45] ϕ_1 is the minimum efficiency for a single-firm conglomerate, so that $\pi(\phi_1) = 0$.

[46] See Shapiro (2016, 2021) for similar formulations of social welfare. Note that total energy use, E, in Equation 11 impacts welfare through its overall effect on pollution and through global externalities. Since we find that the regulation does not significantly shift the geographic distribution of energy use, our welfare measure does not account for the location of emissions. See Appendix H for a discussion of our results in the context of an extended model with localized effects of emissions reductions.

4.4 Effects of the Top 1,000 Program

We denote outcomes in the unregulated equilibrium with an asterisk to differentiate them from those in the regulated equilibrium. Since the Top 1,000 program targeted very large firms, we assume that only conglomerates with ϕ above an efficiency level $\tilde{\phi}$ are subject to the regulation. The regulation sets a proportional input quota for the largest firm in each conglomerate, which is the model counterpart of a Top 1,000 firm. Specifically, the energy use of regulated firms cannot exceed $\bar{e}_1(\phi) = \xi e_1^*(\phi)$, where $\xi < 1$ and e_1^* is the unregulated optimal energy use. At the time of the regulation, the conglomerate's capital allocations $\{k_i^*\}_{i=1}^n$ are quasi-fixed, but it can respond by adjusting its use of inputs $\{l_i, e_i\}_{i=1}^n$. Our model characterizes firm-level, conglomerate-level, and industry-wide effects of the program.

We first study how the regulation impacts firm-level production decisions. To do so, we substitute the result from Proposition 1 that $k_i = \delta^{\frac{i-1}{1-\alpha}} k_1$ into Equation 4, define $\phi^* = \phi (k_1^*)^{\alpha_k}$, and let λ be the Lagrange multiplier associated with the regulatory constraint.[47] The first-order conditions for l_i $(1 \leq i \leq n)$ are then

$$\frac{\partial \pi}{\partial l_i} = \underbrace{R^{1-\rho} P^{\rho}}_{\text{Market Demand}} \underbrace{\rho \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi^* \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \alpha_l (l_i)^{\alpha_l - 1}}_{\text{Marginal Product}} = w + p_e + \underbrace{\lambda(\phi) \mathbb{I}[i=1]}_{\substack{\text{Shadow Cost} \\ \text{of Regulation}}} . \quad (12)$$

An important insight of this expression is that conglomerates internalize the marginal product of inputs across firms through the residual revenue term, which is common to all firms in the conglomerate. The impact of energy regulations on the residual revenue term is key to understanding the difference between within-conglomerate and market-level spillovers.

This equation shows that the regulation distorts the allocation of inputs within a conglomerate by adding a shadow cost $\lambda(\phi)$ to the input of the regulated firm. Because conglomerates with more affiliates can shift more production to related parties, conditional on being regulated, more efficient conglomerates (those with a higher ϕ) are subject to a smaller shadow cost $\lambda(\phi)$. Since only conglomerates with $\phi > \tilde{\phi}$ are part of the Top 1,000 program, the regulation also distorts input use across conglomerates.

The following proposition shows that the regulation leads conglomerates to allocate more inputs to the unregulated firms than in the case without the regulation.

Proposition 3 (Within-Conglomerate Distribution under Regulation)**.** *Under the Top 1,000*

---

[47] Note that $k_1^* = K_n^* (\Delta_n)^{\frac{-1}{1-\alpha}}$.

regulation, the inputs and the output of producers follow the sequences given by

$$\frac{e_j}{e_2} = \frac{l_j}{l_2} = \frac{q_j}{q_2} \;=\; \delta^{\frac{j-2}{1-\alpha}} \text{ for } j > 2,$$

$$\frac{e_i}{e_1} = \frac{l_i}{l_1} = \delta^{\frac{i-1}{1-\alpha}} \times \left[1 + \frac{\lambda(\phi)}{w+p_e}\right]^{\frac{1}{1-\alpha_l}} \text{ and } \frac{q_i}{q_1} \;=\; \delta^{\frac{i-1}{1-\alpha}} \times \left[1 + \frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}} \text{ for } i > 1.$$

Even though conglomerates substitute production across firms, the regulation leads to an overall reduction in the conglomerate's output. The following proposition describes the conglomerate-level effects of the regulation on output and energy use.

Proposition 4 (Conglomerate-Level Distortions from the Regulation)**.** *Under the Top 1,000 regulation, the energy use* $e(\phi, n)$ *and the output* $q(\phi, n)$ *of regulated conglomerates are given by*

$$\frac{e(\phi,n)}{e^*(\phi,n)} = \underbrace{\frac{\xi\left[1 + \left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)\left[1 + \frac{\lambda(\phi)}{w+p_e}\right]^{\frac{1}{1-\alpha_l}}\right]}{\Delta_n^{\frac{1}{1-\alpha}}}}_{=\xi_e(\phi)} \text{ and } \frac{q(\phi,n)}{q^*(\phi,n)} = \underbrace{\frac{\xi^{\alpha_l}\left[1 + \left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)\left[1 + \frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}}\right]}{\Delta_n^{\frac{1}{1-\alpha}}}}_{=\xi_q(\phi)},$$

where $e^*(\phi, n)$ *and* $q^*(\phi, n)$ *are the unregulated counterparts of energy use and output and* $\xi_e(\phi)$ *and* $\xi_q(\phi)$ *describe the effective input and output wedges.*

The term $\xi_e(\phi)$ captures the net effect on energy use by combining the reduction in energy use in the regulated firm (ξ) with the increase in related firms, which is governed by $\lambda(\phi)$. The denominator follows from the insight of Proposition 1 that, in the unregulated case, the conglomerate-level input and output are $\Delta_n^{\frac{1}{1-\alpha}}$ times the input and output of the largest firm. The term $\xi_q(\phi)$ has a similar intuition, and it translates the effects of input changes on output through the exponent α_l.

We now characterize the equilibrium effects of the regulation.

Proposition 5 (Equilibrium under Regulation)**.** *The equilibrium price level under the Top 1,000 regulation solves the following system of nonlinear equations:*

$$\left(\frac{P}{P^*}\right)^{-\rho} \;=\; (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha_l \rho^2}{1-\alpha_l \rho}} + s_{\tilde{\phi}} \mathbb{E}_e\left[\xi_q(\phi)^{\rho} \,\middle|\, \phi > \tilde{\phi}\right] \tag{13}$$

$$1 + \frac{\lambda(\phi)}{w+p_e} \;=\; (\xi)^{\alpha_l - 1}\left(\frac{P}{P^*}\right)^{\rho} \xi_q(\phi)^{\rho-1}, \tag{14}$$

where $s_{\tilde{\phi}}$ *is the share of energy in regulated conglomerates prior to the regulation and* $\mathbb{E}_e$ *denotes the expectation with respect to the energy-use distribution from the unregulated equilibrium. Additionally, the aggregate change in energy use is given by*

$$\frac{E}{E^*} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l \rho}} + s_{\tilde{\phi}} \mathbb{E}_e\left[\xi_e(\phi) \,\middle|\, \phi > \tilde{\phi}\right]. \tag{15}$$

Equation 13 shows that the equilibrium price depends on two forces. First, prices increase as regulated firms reduce their output by $\xi_q(\phi)$. Second, unregulated firms respond to this price increase by increasing their output. The relative importance of these forces depends on the share of energy in regulated conglomerates $s_{\tilde{\phi}}$.

Equation 14 describes the shadow cost of the regulation in terms of the equilibrium price effect $\frac{P}{P^*}$ and the conglomerate-level output wedge $\xi_q(\phi)$. This equation follows from the first-order conditions of both the regulated and unregulated cases and from the results of Proposition 3. Given $\frac{P}{P^*}$, Equation 14 and Proposition 4 define an implicit function for $\lambda(\phi)$. Interestingly, the shadow cost $\lambda(\phi)$ and the conglomerate-level wedge $\xi_q(\phi)$ are step functions of ϕ. While these functions depend on the number of affiliates in a conglomerate n, they are constant across conglomerates of the same size but with different values of ϕ.[48] Intuitively, this result is a consequence of the fact that the energy cap in the regulation is proportional to the firm's prior energy use, which itself depends on ϕ.

The equilibrium under the regulation is then determined by a single shadow cost for every value of n along with the equilibrium price $\frac{P}{P^*}$, which greatly facilitates the computation of the new equilibrium. Equation 15 then shows that the equilibrium effect on energy depends on the net change in conglomerate energy use $\xi_e(\phi)$ and the market leakage to unregulated firms.

These results characterize the welfare effects of the program since Equation 11 implies that

$$\frac{d\ln W}{1-\kappa} = -\ln\left(\frac{P}{P^*}\right) - \frac{\kappa}{1-\kappa}\ln\left(\frac{E}{E^*}\right). \tag{16}$$

Recall that Equation 11 captures the welfare costs of carbon emissions through the parameter κ_0 and the welfare losses associated with pollution in China through the parameter κ_1. Since $\kappa = \kappa_0 + \kappa_1$, Equation 16 shows that this framework captures both of these effects.

Propositions 4 and 5 show that the equilibrium effects of the regulation on the industry-level price $\frac{P}{P^*}$ and on aggregate energy use $\frac{E}{E^*}$ are closely related to the conglomerate-level distortions ($\xi_q(\phi)$ and $\xi_e(\phi)$), which themselves depend on the shadow cost to regulated firms $\lambda(\phi)$. As we show in Section 6, these model quantities are closely related to our empirical estimates from Sections 2–3. This framework also allows us to study the effects of alternative policies. For instance, a universal energy tax would have $s_{\tilde{\phi}} = 1$ and a constant ξ_q for all firms. In Section 6, we compare the Top 1,000 program to a universal energy tax, a size-dependent energy tax (i.e., $s_{\tilde{\phi}} < 1$), and alternative forms or regulations, including ones that mirror the Top 10,000 program.

---

[48] As we show in Appendix J.2, this result follows by substituting $\xi_q(\phi)$ into the expression for $\lambda(\phi)$ in Proposition 5 and noting that this expression varies across firms depending on the term Δ_n only.

5 Model Estimation

This section estimates the key parameters of the model to quantitatively match the data patterns for the period prior to the regulation. We validate our estimated model by showing that it matches the untargeted difference-in-differences estimates of the effects of the Top 1,000 program.

5.1 Parameterization and Estimation

We briefly describe the set of structural parameters of the model and how they are identified by the data. We start by setting the values of two parameters based on previous estimates. We follow the literature by calibrating the elasticity of substitution $\sigma = 4$ (Melitz and Redding, 2015, i.e., $\rho = 0.75$). We use the estimate of returns to scale of $\alpha = 0.9$ from Burnside, Eichenbaum and Rebelo (1995), who use energy data to proxy for utilized capital, and set $\alpha_l = 0.8$ to match the cost share of variable inputs in the data.[49] Finally, we parameterize the conglomerate efficiency distribution $G(\phi)$ with a log-normal distribution with mean zero and standard deviation σ_m.

The model is characterized by the three parameters that we estimate: $(\delta, \phi_1, \sigma_m)$, which include the within-conglomerate size depreciation δ, the conglomerate-level survival threshold ϕ_1, and the dispersion of the efficiency distribution σ_m. Given values of ϕ_1 and the market expenditure R, Equations 8 and 9 pin down f. The entry cost f_e is then determined by the conglomerate free-entry condition.[50]

We estimate the parameters $\theta = (\delta, \phi_1, \sigma_m)$ using the method of moments. For a candidate value of θ, we solve the model and compute the following moments: (1) the share of firms in three bins of firm revenue (5–20 million RMB, 20–100 million RMB, and greater than 100 million RMB); (2) the share of firm output in the same three bins; (3) the average output of the second, third, and fourth largest affiliates relative to that of the top firm in the conglomerate; and (4) the fraction of firms with revenue below 1 million RMB. Our data moments describe the equilibrium prior to the regulation using the ASIF and industrial census data for 2004. Intuitively, the parameter σ_m is pinned down by the moments (1) and (2) describing the firm size and firm output distribution. The parameter δ is determined by the within-conglomerate output distribution moments (3). The last moment (4) helps pin down ϕ_1. Our estimate of θ is

---

[49] Conventional estimates of returns to scale range from 0.85 to 0.95, depending on aggregation and time period. In Section 7.3, we show that the aggregate and welfare effects of the program are robust to reestimating the model on the basis of different values of ρ and α.

[50] To pin down f_e, first note that Equations 7 and 9 imply that $\phi_{n+1} = \phi_1 / \left(\Delta_{n+1}^{\frac{\rho}{1-\rho\alpha}} - \Delta_n^{\frac{\rho}{1-\rho\alpha}}\right)^{\frac{1-\rho\alpha}{\rho}}$ and that $\pi(\phi) = \left[\left(\frac{\Delta_n \phi}{\phi_1}\right)^{\frac{\rho}{1-\rho\alpha}} - n\right] rf$. The conglomerate free-entry condition is then $f_e = \int_{\phi_1} \left[\left(\frac{\Delta_n \phi}{\phi_1}\right)^{\frac{\rho}{1-\rho\alpha}} - n\right] f g(\phi) d\phi$, which is determined by our fixed and estimated parameters.

given by

$$\hat{\theta} = \arg\min_{\theta \in \Theta} [m_d - m(\theta)]' W [m_d - m(\theta)],$$

where m_d are the data moments, $m(\theta)$ are the moments generated by the model, and W is the identity matrix.[51]

Table 6 reports the results of the estimation. We estimate that $\delta = 0.90$, which means that the productivity of the second largest firm in the conglomerate is close to 90% of that in the largest firm. Recall that Equation 5 shows that the output of affiliates depreciates in rank by the factor $\delta^{\frac{1}{1-\alpha}}$. This relation implies that the output of the second largest firm is close to 35% of the largest firm's (cf., 29% in the data) and that of the third largest is close to 13% (cf. 20% in the data), which matches the pattern in Panel A of Figure 3. We also estimate that $\sigma_m = 1.24$ and $\phi_1 = 0.61$. To interpret these estimates, note that they imply a conglomerate entry cost of $f_e = 8.9$ million RMB (or approximately 1.1 million USD), which is reasonably commensurate with average profit in the economy. The per-firm operating fixed cost is determined by the average sales per conglomerate in the data, which implies that $f = 44,000$ RMB. Panel A of Figure 7 shows that our model does a good job of fitting both the observed firm-size distribution and the concentration of output prior to the regulation.

5.2 Model Response to the Top 1,000 Program

We need two additional parameters to implement the Top 1,000 program in our model. As discussed in Section 4, our version of the regulation targets conglomerates with efficiency level ϕ above $\tilde{\phi}$. We choose the threshold $\tilde{\phi}$ to match the share of total energy consumed by regulated firms within energy-intensive industries. Given our estimated parameters, the model implies a value of $\tilde{\phi} = 9.29$, which reproduces the fact that regulated firms account for 56% of total energy consumption in energy-intensive industries. Finally, we take the policy intensity ξ from the 11FYP, which targeted an energy reduction of 20%. For this reason, we set $\xi = 0.8$. Table 6 collects the model parameters.

We now use our estimated model to compute the effects of the Top 1,000 program. As in Section 4.4, we assume conglomerates take the number of affiliates and capital allocation as given. The new industry equilibrium ensures that (1) regulated conglomerates allocate variable inputs optimally (as in Equation 12), (2) unregulated firms increase output to respond to the increase in market prices, and (3) the product market clears (as in Equation 8).

Panel B of Figure 7 compares our difference-in-differences estimates to their model-simulated analogues. The model does a remarkable job of matching the estimated effects on firm output.

---

[51] We use the identity matrix since the sample size for the moments describing the size and output distribution is much larger than the sample size for the moments describing the relative size of firms within conglomerates. Nonetheless, we calculate standard errors using a bootstrap covariance matrix of the moments that incorporates this information.

This is true for regulated firms, related firms, and market-level spillovers. The model prediction of the change in input use of regulated firms is within the 95% confidence interval of our empirical estimate, but the model has a hard time fitting the effect on the energy use of related firms. This may reflect the fact that, as we discuss in Section 3, this estimate is based on a smaller sample of larger firms and may not be representative of the overall response. However, the model does a good job of matching the effects of the program on the energy efficiency of both regulated and related firms. Overall, these results show that our model can reproduce the effects of the regulation on the output of regulated, related, and unrelated firms, which is remarkable since these are all out-of-sample predictions of the model.

5.3 Using the Model to Interpret Our Difference-in-Differences Estimates

An important force in the model is that unregulated firms are impacted by the regulation through the market spillover. This force contributes to the effects of the program on the equilibrium price and aggregate energy use. We now use our model to understand how this market spillover impacts our difference-in-differences estimates.

To see how the regulation in our model connects to our difference-in-differences analysis, note that we can write conglomerate j's revenue from affiliate i as follows:

$$\ln \text{Revenue}_{ij} = \underbrace{\ln(\text{Production Share}_{ij})}_{\text{Allocation Effect}} + \underbrace{\rho \ln\left(\sum_{i \in j} q_{ij}\right)}_{\text{Residual Revenue}} + \underbrace{\ln(R^{1-\rho} P^{\rho})}_{\text{Market Demand}}, \tag{17}$$

where $\text{Production Share}_{ij} = q_{ij}/\sum_{i\in j} q_{ij}$.[52] Equation 17 clarifies the three ways in which the Top 1,000 program impacts the revenue of regulated firms. First, when firm i is regulated, the conglomerate is forced to reallocate inputs to other firms, which lowers the production share in regulated firms. Panel A of Table 7 reports that, in our model, the share of production in regulated firms within a conglomerate decreases by 12.9%. Second, since the marginal cost goes up at the conglomerate level, the market share of the conglomerate's variety decreases, which lowers the group's residual revenue. Table 7 shows that regulated conglomerates see their residual revenue decrease by 3.7%. Finally, the Top 1,000 program impacts the industry-level price P. This price increase has a countervailing effect on the revenue of the regulated firm and lessens the overall decline by 2.6%. Combining these three forces, our model implies that regulated firms decreased their output by 14%.

Equation 17 also characterizes the impact of the regulation on the control firms in our difference-in-differences analyses. Since these firms are not regulated or related to Top 1,000

---

[52] Equation 17 follows by multiplying conglomerate j's inverse residual demand by affiliate i's production.

firms, the regulation does not impact the within-conglomerate allocation of production. Control firms see an increase in their residual and firm-level revenue as the market reallocates demand. Table 7 shows that the residual revenue of control firms increases by 3.9%. As in the case of regulated firms, unregulated firms also benefit from the equilibrium impact on market demand. This analysis highlights the importance of interpreting the market-level spillovers in Panel A of Table 5 as the combination of quantity and price effects.

This discussion clarifies that our difference-in-differences estimates differ from the total effect on Top 1,000 firms along two margins. First, the difference-in-differences estimator captures both the within- and across-conglomerate reallocation of production. This leads to an overestimation of the effect of the program on regulated firms of 3.9%. Second, since the market effect cancels out, the difference-in-differences estimator does not capture the countervailing effect on the industry-level price and thus further overestimates the effect of the program by 2.6%.[53]

Similarly, our model allows us to decompose the estimates of the spillover effects of the regulation through ownership networks. Panel B of Table 7 shows that related firms share the residual revenue and market effect terms but have a positive allocation effect as their share of production within the conglomerate increases. The total effect on related firms is an output increase of 19.3%. The effect on control firms is the same as that in Panel A. By ignoring the positive market effect and subtracting the residual revenue effect on control firms, the difference-in-differences estimator understates the spillover effect on related firms by 6.5%.

In addition to clarifying the interpretation of our reduced-form estimates, our model motivates an alternative approach that does not depend on the residual revenue or market effects. Specifically, consider a within-conglomerate difference-in-differences estimator where the treated firms are the regulated Top 1,000 firms and the control firms are unregulated firms in the same conglomerate. Because Equation 17 shows that the residual revenue and market effect are common to a given conglomerate, this estimator captures only the allocation effects of the program. Panel B of Figure 6 implements this within-conglomerate difference-in-differences approach. This figure plots the results from an event-study specification similar to Equation 1 but where we additionally include conglomerate-by-year fixed effects. Consistent with our previous results, we find a significant decline in the output of Top 1,000 firms relative to that of other firms in their same conglomerates. Panel B of Table 5 reports estimates of these relative declines of between -31.5% and -36.7%. As with our previous reduced-form effects, Panel B of Figure 7 shows that the model matches this within-conglomerate effect very well. Moreover, Panel C of Table 7 confirms that this effect is a combination of the allocation effects on regulated and related firms.

These insights highlight the importance of interpreting quasi-random estimates through the

---

[53] Note that the first channel arises from the impact of the regulation on the control firms. The second channel is an aggregate effect that is not identified by a difference-in-differences research design.

lens of a model that accounts for within- and across-conglomerate reallocation of production and equilibrium impacts on industry-level prices.

6 Policy Analysis

This section uses our estimated model to capture the effects of the Top 1,000 program by quantifying the shadow cost to regulated firms, the aggregate effects on prices and energy use, and the implied welfare trade-off of the program. We then consider the effects of alternative policies including program expansions (e.g., the Top 10,000 program) and the possibility of the government's using information on business networks to improve energy regulation.

6.1 Effects of the Top 1,000 Program

Given that the shadow cost is the fundamental building block of our model outcomes, we first quantify this cost of the policy.

6.1.1 Shadow Cost of the Policy

To compute the shadow cost of the policy, we solve for the regulated equilibrium as in Proposition 5. Panel A of Figure 8 plots the implied shadow cost as a function of efficiency ϕ. The blue line plots the shadow cost of our computed Top 1,000 program. This shadow cost is zero for firms with $\phi < \tilde{\phi}$ and jumps to an average of 8.7% for regulated firms. Since the shadow cost has the same scale as the cost of variable inputs, we can interpret this value as an equivalent tax on variable inputs. While 8.7% might seem like a small number, recall that inputs constitute a large tax base, especially relative to profits.[54]

A somewhat surprising feature of Panel A of Figure 8 is that the shadow cost appears to be constant with respect to conglomerate productivity ϕ. Panel B of Figure 8 zooms in to show the different shadow costs for regulated firms (i.e., $\phi > \tilde{\phi}$). As we discuss in Section 4, the shadow cost is constant for conglomerates with the same number of related firms. This result follows from the fact that the regulation is based on previous energy use, which is proportional to firm productivity. We see that, when ϕ crosses the thresholds that define conglomerate size (Proposition 2), the shadow cost drops, as conglomerates with more affiliates are more able to escape the burden of the regulation. However, the differences in shadow costs are very small in comparison to the overall difference between regulated and unregulated conglomerates. This small impact is driven by the decay in within-conglomerate size, which implies that the marginal

---

[54] Indeed, in models with constant marginal cost and with a similar value of σ, inputs are $(\sigma - 1) = 3-$times as large as profits. An equivalent profit tax would then be 26.1%.

(n^{th}) affiliate may not be able to compensate a large fraction of the combined activity of all the other affiliates (1 through $n-1$) in the conglomerate.

We now validate the magnitude of the shadow cost using an additional implication of the model. Recall the insight from Proposition 3 that the shadow cost is related to the within-conglomerate output distribution after the regulation. Based on this insight, we can write the output difference between the Top 1,000 firm and related firms as follows:

$$\begin{aligned}\ln \text{Revenue}_{\text{Top1000},jt} - \ln \text{Revenue}_{\text{Related},jt} &= \ln\left(q_{\text{Top1000},j}\right) - \ln\left(\sum_{i\neq \text{Top1000},j} q_{ij}\right) \\ &= -\frac{\alpha_l}{1-\alpha_l}\ln\left[1+\frac{\lambda(\phi)}{w+p_e}\right] - \ln\left(\Delta_n^{\frac{1}{1-\alpha}}-1\right).\end{aligned}$$

By definition, $\lambda(\phi)=0$ prior to the regulation, and Δ_n is constant over time. Therefore, taking a time difference of this expression shows that the within-conglomerate difference-in-differences estimation identifies the term $-\frac{\alpha_l}{1-\alpha_l}\ln\left[1+\frac{\lambda(\phi)}{w+p_e}\right]$. Using the estimate in the first column of Panel B of Table 5 of -34.3% and our value of $\alpha_l = 0.80$, we estimate a shadow cost of 8.95%.[55] Given that the shadow cost plays an important role in our quantification exercises on the aggregate effects of the program, it is reassuring that our estimated magnitude is consistent with the reduced-form pattern of within-conglomerate reallocation of production.

Our model allows us to consider how different mechanisms impact the shadow cost of the program. The top line of Panel A of Figure 8 shows that shutting down the market and conglomerate spillovers would increase the shadow cost to 15.6%. That is, the equilibrium price increase and the ability to shift production to related firms lowered this mechanical effect by almost 50%. The model also allows us to isolate how the ability of conglomerates to shift production to related firms lowered the shadow cost of the regulation. The second line from the top in Panel A of Figure 8 plots the shadow cost under the assumption that market prices adjust but that regulated firms are not able to shift production to related parties. In this case, the shadow cost of the regulation would be 11.2% of input costs, which is approximately 30% larger than the level in the baseline case.[56] These calculations showcase the importance of accounting for both market and conglomerate spillovers in the measurement of the shadow cost of the regulation.

6.1.2 Aggregate and Welfare Effects of the Program

The analysis so far has focused on the distortionary aspects of the regulation. We now evaluate the aggregate and welfare effects of the policy by considering the social welfare function of a

---

[55] This calculation follows from $1+\frac{\lambda(\phi)}{w+p_e} = \exp\{-\beta\frac{1-\alpha_l}{\alpha_l}\} = \exp\{0.343\frac{1-0.8}{.8}\} = 1.0895$. In principle, we identify the average value of this quantity across firms. However, as we show in Figure 8, there is little variation in $\lambda(\phi)$.

[56] This counterfactual assumes that, similar to unregulated firms, the production of related firms responds to equilibrium price increases. Restricting production in related firms to pre-regulation levels further increases the shadow cost to 14%.

government concerned with both decreases in energy use–related emissions and the impact of distortions to production on consumption.[57]

We compute the aggregate effects of the program by solving the equilibrium conditions in Proposition 5. Panel A of Figure 9 plots the effects of the Top 1,000 program in the space of price increase and energy reduction. The red diamond in this figure shows that the Top 1,000 program led to a price increase of approximately 3.5% and an aggregate energy use reduction of close to 4%. Equation 15 helps us understand how we obtain a 4% aggregate reduction in energy use. First, we find that—including within-conglomerate reallocation—regulated conglomerates reduced their energy use by $1-\mathbb{E}_e\left[\xi_e(\phi)\,\middle|\,\phi > \tilde{\phi}\right] = 5.8\%$. Second, the unregulated conglomerates increased their energy use by close to 6.5%. Finally, we obtain the aggregate 4% decline by using the fact that the share of energy in regulated conglomerates is $s_{\tilde{\phi}} = 86\%$.[58] Thus, even though we find in Section 2 that Top 1,000 firms reduced their energy use by close to 100 million tce, the annual aggregate reduction—including conglomerate and market leakage—was closer to 48 million tce.[59]

To determine whether the program increased welfare, we compare the effects of the program on energy reduction and output losses. Using the Cobb–Douglas structure of the welfare function, we can relate $\kappa = \kappa_1 + \kappa_0$ to the shares of aggregate income that the government would like to spend on reducing carbon emissions (κ_0) and pollution (κ_1):

$$\kappa_0 = \frac{\text{Social Cost of Carbon} \times \text{Carbon Emissions}}{\text{Aggregate Income} \times 0.8} \quad \text{and} \quad \kappa_1 = \frac{\text{Total Pollution Damages}}{\text{Aggregate Income} \times 0.8}.$$

In both cases, the adjustment factor 0.8 comes from the fact that the Chinese government's 11FYP recognized its underspending in reducing energy and adopted a goal to reduce energy use by 20%. We first consider the combined effect of these two motives to estimate the government's total willingness to pay (GWTP) to reduce emissions:[60]

$$\kappa \;=\; \frac{\text{GWTP} \times \text{Carbon Emissions}}{\text{Aggregate Income} \times 0.8}.$$

---

[57] To match the short-run nature of our empirical analysis, we focus our discussion on the short-run effects of the policy, ignoring entry of new conglomerates. Later changes to regulations and the overall environment also complicate the simulation of long-run impacts.

[58] The energy increase for unregulated firms is given by $\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} = (1.035)^{\frac{0.75}{1-0.8*0.75}} \approx 1.065$. Recall that regulated firms account for 56% of energy use; accounting for related firms in the same conglomerate raises this fraction to 86%. The aggregate effect is then $-4\% = \ln(1.06 * 0.14 + 0.942 * 0.86)$.

[59] In Appendix J.2.3, we connect the model solution to our reduced-form estimates by showing that we can solve an approximate version of the equilibrium using the value of λ that is implied by the within-conglomerate difference-in-differences results.

[60] This expression follows by considering the pollution damage per ton of carbon, so that $\kappa = \kappa_0 + \kappa_1 = \frac{(\text{Social Cost of Carbon}+\text{Pollution Damage Per Ton of Carbon})\times\text{Carbon Emissions}}{\text{Aggregate Income}\times 0.8} = \frac{(\text{GWTP})\times\text{Carbon Emissions}}{\text{Aggregate Income}\times 0.8}$, where the GWTP includes both the SCC and the pollution damages associated with a ton of carbon. In practice, pollution damages per ton of carbon may vary according to the energy source. In Appendix G, we allow all of the pollution damages to be due to usage of coal—the predominant energy source of Top 1,000 firms—and show that our results are not sensitive to this alternative assumption.

According to Equation 16, welfare increases when the aggregate price-to-energy use elasticity $\left(\text{i.e., } \frac{-\ln\left(\frac{P}{P^*}\right)}{\ln\left(\frac{E}{E^*}\right)}\right)$ is smaller than $\frac{\kappa}{1-\kappa}$. Given our aggregate estimates of the effects of the Top 1,000 program, we find that the program raises welfare as long as κ is greater than the threshold value $\bar{\kappa}$, given by the condition $\frac{\bar{\kappa}}{1-\bar{\kappa}} = 0.875$, which implies that $\bar{\kappa} = 0.47$. That is, the Top 1,000 program raises welfare as long as the government is willing to accept an 0.875% output loss for every 1% reduction in energy use. We obtain an estimate of the GWTP based on this threshold value $\bar{\kappa}$. Using 2006 data on overall emissions in China (6.38 billion tons of carbon) and national income (2.752 trillion USD), we calculate a GWTP value of $161.

To visualize the government's trade-off between energy reduction and output losses, Figure 9 plots black indifference curves that correspond to different hypothetical GWTP values, which correspond to different values of κ. For a given GWTP value, these lines plot combinations of price and energy use changes that yield the same effect on welfare. The red diamond in this figure shows that the Top 1,000 program lies on the indifference curve that corresponds to a GWTP of $161.

We now estimate the social cost of carbon (SCC) implied by the Top 1,000 program by subtracting the pollution damages associated with a ton of carbon emissions from the GWTP. We rely on four estimates of pollution damages from the World Bank (2007), Mohan et al. (2020) and Ito and Zhang (2020*b*), which encompass varied methodologies spanning calculations of gross external damages, willingness-to-pay analyses, and adjusted human capital approaches. Across these different methodologies, we estimate that the pollution damages associated with a ton of carbon can be valued at between $4 and $17.[61] Relative to our estimated GWTP of $161, these estimates imply that the policy would increase welfare if the SCC exceeded the $144–$157 range.[62]

Our estimates of the health benefits of reducing pollution may contrast with intuition derived from the context of high-income countries. In Appendix G, we show that the GED estimate of Mohan et al. (2020) implies a cost of pollution damages per ton of carbon of $78. If the Chinese

---

[61] The gross external damages (GED) approach of Mohan et al. (2020) implies total pollution damages in China of $108 billion. Relative to carbon emissions in 2006, the pollution damage per ton of carbon is valued at $17. The World Bank (2007) produces two independent estimates of air pollution damages. The first uses a willingness-to-pay methodology and yields a value of pollution damage per ton of carbon of $13. The second estimate uses an adjusted human capital approach, which implies a value of pollution damage per ton of carbon of $4. Finally, Ito and Zhang (2020*b*) use a willingness-to-pay approach for reducing exposure to air particulates that relies on quasi-experimental variation in pollution exposure due to the Huai River policy. Combined with estimates of the mortality effects of pollution (e.g., Ebenstein et al., 2017), the estimates in Ito and Zhang (2020*b*) imply a value of pollution damage per ton of carbon of $6. To be more conservative, our analyses rely on the GED and willingness-to-pay measures. See Appendix G for details.

[62] In practice, the health benefits of reducing energy use of Top 1,000 firms could be larger than average if these firms rely on dirtier energy sources. In Appendix G, we obtain an upper bound on the health benefits from reducing emissions by assuming that all of the air pollution in China is generated by coal, the main source of energy of Top 1,000 firms. The estimates of the value of health benefits increase to between $5 and $22, implying a range of values of the SCC that rationalizes the policy of between $139 and $156.

government used this valuation, the SCC that would justify the policy would then be \$83, which can be compared to SCC values used in high-income countries. For instance, in the US, the Biden administration has recently proposed an SCC value of \$51 (IWG, 2021), and researchers have recently argued for a higher value of \$125 (Carleton and Greenstone, 2021).

6.2 Alternative Policies

We now use the model to consider alternative policies. We first explore different ways in which the Top 1,000 program could be expanded or contracted. This exercise is motivated by the fact that the Chinese government expanded the program to include more than 16,000 firms in the Top 10,000 program in 2012. We then explore the effects of alternative regulations and energy taxes to examine the degree to which the government can improve the regulation of energy.

We explore two ways to change the scope of the Top 1,000 program. First, we consider the effect of varying the regulation threshold $\tilde{\phi}$, which changes the number of firms affected by the program. The blue dots in Panel A of Figure 9 show the effects of changing the size threshold, $\tilde{\phi}$. The first blue dot (to the left of the red diamond) considers the effect of decreasing the number of regulated firms to cover only 50% of the energy use in the regulated industry (relative to the current 56%). The second blue dot lowers $\tilde{\phi}$ so that the regulation instead covers 60% of the industry's energy use.[63] As is to be expected, we find larger energy decreases when the program covers a larger fraction of overall energy use. However, Figure 9 shows that expanding or contracting the number of firms in the program does not alter the fundamental trade-off that the government faces between price increases and reductions in energy use.

An alternative way to change the scope of the Top 1,000 program is to increase or decrease the energy use quota ξ. The maroon squares in Panel A of Figure 9 plot the effects of policies where $1 - \xi$ varies in 5% increments between 5% and 30%. Larger values of $1 - \xi$ lead to both larger price increases and larger energy reductions. Taking both changes into account, we find that the implied GWTP increases with the required energy reduction and equals \$167 when $1 - \xi = 30\%$. This result is valuable since the government may be concerned about the administrative costs of regulating a larger number of firms. Since increasing ξ and lowering $\tilde{\phi}$ have similar welfare effects, it may be desirable to place stricter energy use limits on fewer firms if the government lacks the capacity or the funds to conduct additional energy audits.

We now consider the effects of an alternative policy that targets the energy use of all firms in a given conglomerate.[64] The orange crosses in Panel B of Figure 9 plot the effects of this

---

[63] This alternative regulation mirrors that of the Top 10,000 program by increasing the program's coverage. We study the empirical effects of the Top 10,000 program in Appendix I. As with our simulation, we find that a more complete regulation reduces the scope for within-conglomerate and market leakage.

[64] We derive equilibrium conditions under these alternative regulations in Appendix K. To make this case comparable, we model the effects of a regulation that limits the conglomerate-level energy use to the levels under

type of regulation at different values of ξ. These policies have the benefit of not distorting the within-conglomerate distribution of production. Panel A of Figure 8 shows that regulating conglomerates has a lower shadow cost of 4.9% (instead of 8.7%). Such a policy is preferable to the Top 1,000 program from a welfare perspective since it can achieve larger energy use reductions for a given price increase. As we show in Panel B of Figure 9, this type of regulation can yield a 4.36% reduction in aggregate energy use for the same price increase as the Top 1,000 program. This is a 10% increase from the energy reduction under the Top 1,000 program, corresponding to additional energy savings of 5 million tce. This policy improves welfare as long as the GWTP $\geq$ \$154. While this program would involve monitoring additional firms, the number of firms related to Top 1,000 firms is less than 20% of the number of firms in the Top 10,000 program. These results show that the government can improve the regulation of energy by using publicly available data on business networks to target conglomerates and that doing so would be more effective than regulating additional unrelated firms as with the Top 10,000 program.

Finally, we consider the effects of energy taxes. We first model the effects of a size-dependent energy tax that affects all firms in conglomerates with Top 1,000 firms (i.e., with $\phi > \tilde{\phi}$). Panel A of Figure 8 shows that the policymaker could obtain the same energy reduction as that under the Top 1,000 program by taxing inputs at 4.9%.[65] The green circles in Panel B of Figure 9 show that the effects of this energy tax are very close to those of the conglomerate-level regulation.[66] We further consider the effects of a universal energy tax that impacts all firms in the economy. We model this tax by setting $s_{\tilde{\phi}} = 100\%$ instead of 86%. Panel B of Figure 9 shows that, while a universal energy tax yields a slight improvement over the size-dependent tax (GWTP= \$153), both the size-dependent tax and the conglomerate-level regulations imply very similar welfare trade-offs.

The preferred policy solution for most economists on the regulation of carbon emissions related to energy use is a universal carbon tax. In practice, this policy may not be feasible given legal, administrative, or political constraints. The results in this section are informative about the efficacy and design of a prominent real-world policy that regulates quantities and has

---

the Top 1,000 program. That is, we set ξ to values corresponding to $\xi_e(\phi)$ in the Top 1,000 program. While conglomerate-level regulations may face political pressure from business groups, such policies have been used in other countries, where, for example, subsidy eligibility criteria may depend on conglomerate-level characteristics (see, e.g., the details on the UK's R&D tax credit for small and medium-sized enterprises in the study by Dechezleprêtre et al., 2023).

[65] As with the Top 1,000 regulation, the shadow cost of the conglomerate-level regulation decreases slightly as the number of related firms increases. In contrast, the shadow cost of the size-dependent energy tax is constant for all firms affected by it. Since energy costs are close to 15% of variable input costs for Top 1,000 firms, the equivalent energy tax would be closer to $32.7\% (\approx \frac{4.9\%}{15\%})$.

[66] It is worth noting that our quantification lacks two features that often motivate the use of taxes over regulation. First, in our calculations, the revenue from the tax is not rebated to consumers; this calculation ignores potential "double dividend" effects. Second, firms in our setting have homogeneous abatement costs; in a setting with heterogeneous abatement costs, a tax would additionally reallocate production to "cleaner" firms.

incomplete coverage. We find that the government could achieve similar aggregate effects by either expanding the program through stricter regulations for current firms or increasing the number of firms in the program. While the former option has narrower coverage and generates larger inequities between regulated and unregulated firms, the latter may require an increase in administration costs. We also find that the government can improve the regulation of energy by targeting the ownership networks of regulated firms. This policy increases aggregate energy savings by 10% without increasing welfare costs. Moreover, this policy can be implemented with publicly available data, has a lower administrative cost than the Top 10,000 program, and implies a welfare trade-off close to that under a universal energy tax.

7 Extensions and Robustness

This section explores four extensions of our model. First, we consider the possibility of firms responding to the regulation by investing in energy efficiency. Second, we consider how preexisting differences in energy efficiency across firms may alter the overall energy savings from the program. Third, we study the robustness of our welfare calculations to alternative model parameterizations. Finally, we extend our model to allow imperfect substitution between products produced by firms in a given conglomerate.

7.1 Endogenous Energy Efficiency

Our baseline analysis assumes that, in the short run, firms do not make any investments to improve their energy efficiency. This assumption is consistent with our empirical results in Sections 2–3. Following the intent of the Top 1,000 program, we now extend our model to allow firms to respond by adjusting their energy efficiency. Appendix M.1 provides additional details of this model extension.

We assume that the conglomerate can improve energy efficiency at firm i, ν_i, by spending $l_i c(\nu_i)$, where $c'(\nu_i) > 0$ and $c''(\nu_i) \geq 0$. We can then restate the regulated conglomerate's problem as

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^n, \{\nu_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho} - \sum_{i=1}^{n} l_i \left(w + \frac{p_e}{\nu_i} + c(\nu_i) \right) \right\},$$

where we omit the cost of fixed capital. Absent the regulation, the conglomerate sets $c'(\nu^*)\nu^{*2} = p_e$ for all firms. This result implies that Propositions 1–2 continue to describe the equilibrium prior to the regulation. To simplify the exposition, we assume that $c(\nu) = \frac{\nu^\gamma}{1+\gamma}$, where $\gamma \geq 1$. This implies that the effective price of energy inclusive of investments in energy efficiency is $\frac{p_e}{\nu^*} + c(\nu^*) = c'(\nu^*)\nu^* + c(\nu^*) = (\nu^*)^\gamma$.

Consider now the effects of the regulation. First, note that the Top 1,000 regulation does not impact the choice of ν_i for unregulated firms. We then use these results and the fact that $\nu_i = \frac{l_i}{e_i}$ to restate the conglomerate problem as

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^{n}} \left\{ R^{1-\rho}P^{\rho}\left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l}\right]^{\rho} - (w + (\nu^*)^{\gamma}) \sum_{i=1}^{n} l_i - l_1 \left[\frac{1}{1+\gamma}\left(\frac{l_1}{\xi e_1^*}\right)^{\gamma} - (\nu^*)^{\gamma}\right]\right\},$$

where we substitute the regulatory constraint into the cost of energy efficiency and where we abstract from the cost of the regulated energy, $p_e \xi e_1^*$, since it is a constant.

The conglomerate's first-order conditions for l_i $(1 \leq i \leq n)$, i.e., $\frac{\partial \pi}{\partial l_i}$, are then

$$\underbrace{R^{1-\rho}P^{\rho}}_{\text{Market Demand}} \underbrace{\rho\left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l}\right]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi^* \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \alpha_l (l_i)^{\alpha_l - 1}}_{\text{Marginal Product}} = w + (\nu^*)^{\gamma} + \underbrace{\left[\left(\frac{l_1}{\xi e_1^*}\right)^{\gamma} - (\nu^*)^{\gamma}\right]}_{\substack{\text{Shadow Cost} \\ \text{of Regulation}}} \mathbb{I}[i = 1].$$

Interestingly, this extension of the model yields very similar results to those from Equation 12. For the case of unregulated firms (i.e, l_i for $i > 1$), we simply substitute p_e with the effective price of energy: $(\nu^*)^{\gamma}$. The first-order condition for the regulated firm implies that the shadow cost of the regulation is given by

$$\frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} = s_e \left[\left(\frac{\nu_1}{\nu^*}\right)^{\gamma} - 1\right],$$

where s_e is the share of variable input costs accounted for by energy and where energy efficiency at the Top 1,000 firm improves by $\frac{\nu_1}{\nu^*}$ relative to that in the unregulated case.[67] The results of Proposition 3 continue to hold when we use this definition of the shadow cost of the regulation, which captures the incremental cost of improving energy efficiency in the regulated firm.

This framework allows us to infer how costly it would have been for regulated firms to have improved their energy efficiency. We implement this calculation using our estimated model parameters. Our estimated model fundamentals remain valid since Propositions 1–2 continue to describe the unregulated equilibrium and since we estimated the model with data from prior to the regulation. To implement this model extension, we set $s_e = 15\%$ based on our data and calibrate γ based on the estimated effects of the program on the energy efficiency of regulated firms. Since we estimate statistically insignificant effects of the program on the revenue-to-energy ratio of regulated firms, we calibrate γ using the upper bound of the 95% confidence interval.[68]

---

[67] The functional form assumption for $c(\nu)$ only simplifies the derivation. Absent this assumption, one can replace p_e in Equation 12 with the effective cost of energy $\nu^* c'(\nu^*) + c(\nu^*)$. Similarly, the shadow cost would be $\lambda(\phi) = [\nu_1 c'(\nu_1) + c(\nu_1)] - (\nu^* c'(\nu^*) + c(\nu^*))$.

[68] We also find that the Top 1,000 program did not lead to improvements in the variable costs–to–energy ratio. Figure A.14 and Table A.25 show a null effect on this measure of energy efficiency for regulated firms, and Figure A.15 and Table A.26 reveal the same result for related firms.

In Appendix M.1, we show that when $\gamma = 7.4$, the improvement in the revenue-to-energy ratio of Top 1,000 firms in our model equals 11.6%, which matches the 95% confidence interval of column (1) in Panel C of Table 3.[69] The ability to improve the energy efficiency of the regulated firm lowers the shadow cost of the regulation, as it loosens the energy use restriction on the Top 1,000 firm. In this case, the shadow cost is 6.5% instead of the 8.7% in our baseline scenario.

This model extension also allows to quantify the aggregate effects of the policy under different assumptions of the costs of improving energy efficiency. Panel A of Figure 10 plots the aggregate effects of the regulation at different values of γ. This model nests our baseline model under the assumption that $\gamma \to \infty$. In this case, firms do not improve their energy efficiency, which is consistent with our empirical results. When the cost of improving energy efficiency decreases (lower values of γ), the Top 1,000 program achieves greater energy reductions and results in smaller price increases. Both of these forces imply that the GWTP that rationalizes the program also increases with γ. For our calibrated value of $\gamma = 7.4$, the Top 1,000 program reduces aggregate energy use by 5.6% for a price increase of 2.8%. This calculation also yields a bound on the GWTP since, as we show in Panel A of Figure 10, the program raises welfare at this value of γ as long as the GWTP$\geq$ \$114.[70]

We can also use our model to calculate the value of γ that would have allowed regulated firms to improve their revenue-to-energy ratio by 20%. We find that firms would have met this energy efficiency improvement target if $\gamma = 0.55$. In this case, the shadow cost of the regulation would have been 1.6%. This scenario represents the "low-hanging fruit" perspective, according to which small investments can lead to large gains in energy efficiency. While our results show that firms did not expect to recoup the costs of improving energy efficiency over a five-year period, it is possible that firms may find ways to improve their energy efficiency in the long run.[71]

7.2 Heterogeneous Energy Efficiency

The previous section showed that allowing endogenous investments in energy efficiency does not significantly alter our results. We now explore the possibility that—even prior to the Top 1,000 program—regulated, related, and unrelated firms operated under heterogeneous energy efficiencies.

One possibility is that the government targeted Top 1,000 firms because they are particularly

---

[69] Alternatively, our model implies a value of $\gamma = 12.8$ if we target the 99% confidence interval of column (4) of Panel C of Table 3.

[70] The result that the conglomerate-level regulation would increase energy savings by 10% at the same welfare cost is robust to allowing firms to improve their energy efficiency.

[71] The fact that conglomerates were able to shift production to related affiliates lowered the incentive for regulated firms to invest in energy efficiency. Using our baseline calibration of $\gamma = 7.4$, we calculate that regulated firms would have increased their revenue-to-energy ratio by 13.8% if they had not been able to shift production to related firms.

energy inefficient. Similarly, the production increase in unrelated firms may have smaller effects on overall energy use if these firms are more energy efficient. In both of these cases, the regulation may be more effective to the extent that it shifts production to more energy-efficient firms. Alternatively, the Top 1,000 program may lead to smaller reductions in energy use if Top 1,000 firms are more energy efficient than other firms.

Appendix M.2 generalizes our analysis to allow the energy efficiency of related firms to differ from that of Top 1,000 firms by a factor of ν_R. Differences in energy efficiency would alter the pattern of production within a conglomerate even prior to the regulation. This is because energy efficiency influences the unit cost of related firms. Differences in energy efficiency would then influence the allocation of production within the conglomerate. We can additionally assume that other unrelated firms have an energy efficiency that differs from that of Top 1,000 firms by a factor of ν_O. These differences impact both the response of related and other firms to the regulation and overall energy use. To explore the sensitivity of our results to differences in energy efficiency, we solve the model under different values of ν_R and ν_O. We allow ν_O and ν_R to reflect energy efficiency that is up to 20% lower (which would exacerbate energy leakage) or up to 20% greater than that of the Top 1,000 firms (leading to negative leakage; e.g., Baylis, Fullerton and Karney, 2014).[72]

Panel B of Figure 10 shows the effect of heterogeneity in energy efficiency on our welfare calculations. We first assume that firms related to Top 1,000 firms are 20% less energy efficient, i.e., set $\nu_R = 0.80$. This case implies a larger welfare loss due to a larger price increase and smaller energy savings. Under this assumption, the GWTP that rationalizes the program increases to \$193. Further assuming that other firms are also 20% less energy efficient, i.e., setting $\nu_O = 0.80$, also increases prices and decreases energy savings, implying a GWTP that would rationalize the program of \$199. These calculations show that the within-conglomerate energy leakage is an important contributor to the overall energy effects of the program. If we alternatively assume that Top 1,000 firms are particularly energy inefficient, i.e., set $\nu_R = \nu_O = 1.2$, the program can be rationalized with GWTP values as low as \$134. To the best of our knowledge, Top 1,000 firms are not particularly inefficient relative to their related firms. Nonetheless, this calculation provides an interesting bound for the welfare effects of the program.

---

[72] Unfortunately, there are no comprehensive survey data covering a large sample of manufacturing firms that allow us to precisely measure differences in the energy efficiency of regulated, related, and competing firms in the market. Using the CESD, we observe in Panel A of Table 2 that Top 1,000 firms are 10% less energy efficient than Top 10,000 firms. Based on this statistic, the 20% deviations that we explore below allow for considerable differences in energy efficiency. Additionally, we report in Figure A.16 the SCC that rationalizes the policy when energy efficiencies differ by up to 50%.

7.3 Robustness to Alternative Parameters and Timing Assumptions

The structural model in Section 4 used calibrated values for the decreasing-returns-to-scale parameter α and for ρ, which determines the elasticity of substitution $\sigma = \frac{1}{1-\rho}$. This section discusses how our results are affected by changing these parameter values. To do so, we first vary the values of these parameters. We then reestimate the structural parameters following the same procedure as in Section 4. Finally, we solve for the regulated equilibrium that is implied by every pair of values of α and ρ.[73]

Panel C of Figure 10 shows how varying these parameters affects the estimated aggregate effects of the Top 1,000 program and the implied GWTP. The red diamond plots the effects of the program under our baseline parameterization that sets $\alpha = 0.90$ and $\rho = 0.75$. Consider first the effects of fixing $\alpha = 0.90$ and setting ρ to either 0.70 (so that $\sigma = 3.33$) or 0.90 (so that $\sigma = 10$). These two cases are denoted in Panel C of Figure 10 by the green circles, which show that the aggregate effects of the policy are barely altered by changing σ. This result is driven by the fact that when σ is larger, the distributions of firm size and output imply a smaller variance of firm productivity ϕ. Thus, even though we would expect a larger market spillover for a larger value of σ, this effect is offset by the decrease in the dispersion in firm productivity.

Consider now the effect of fixing $\rho = 0.75$ and setting α to either 0.85 or 0.95. These two cases are denoted in Panel C of Figure 10 by the blue triangles. In this case, we find that lower values of α lead to both larger energy use reductions and larger price increases. The intuition for this result is that when production faces more decreasing returns to scale, conglomerates are less able to substitute production across related firms. Similarly, unregulated firms are less able to respond to the price increase by increasing their own production. Both of these forces lead to larger energy use reductions and price increases.

Across these cases, the reduction in aggregate energy use ranges between 2.9% ($\alpha = 0.95$) and 4.9% ($\alpha = 0.85$). In terms of the total reduction in energy use, these values imply annual aggregate energy savings of between 37 and 57 million tce. Since the parameter α governs the extent of energy leakage to related and unregulated firms, it is reasonable that uncertainty in this parameter would generate uncertainty in the aggregate energy use reduction. Interestingly, the change in the aggregate price covaries with the reduction in energy use. As a result, the ratio of the price change to the energy change varies very little. We denote this in the graph by plotting gray lines that correspond to the implied values of the GWTP for each case. These lines show that across all these different parameterizations, the implied GWTP that rationalizes the policy varies only between \$156 and \$167.

---

[73] The value of α_l is determined by the value of α and the cost share of variable inputs. Additionally, we solve for a new regulation threshold $\tilde{\phi}$ to match the share of energy in regulated firms. Table A.27 reports the estimated model parameters across different specifications.

Our baseline model assumes that conglomerates were not able to respond to the Top 1,000 program by adjusting their capital. While this is a convenient assumption in the short-run, as we discuss in Appendix E.2, it is also possible that some conglomerates may have partially responded to the regulation by adjusting their capital. Perhaps more importantly, it is likely that conglomerates will be able to adjust their capital over time. We extend our model to allow for this possibility in Appendix J.3. Since the model is estimated based on pre-reform data, this extension does not require us to re-estimate the model. The X-mark in Panel C of Figure 10 reports the aggregate and welfare effects of the program under this alternative timing assumption. When conglomerates can adjust their capital, prices increase less, but the energy savings are also smaller. In this case, the GWTP that rationalizes the program increases slightly to \$170. Intuitively, allowing for capital adjustment enhances the role of within-conglomerate spillovers, as conglomerates are less constrained in their response to the regulation.

Overall, the values of α and ρ or assumptions about the timing of capital adjustment do not significantly affect our quantitative assessment of the fundamental trade-off faced by the government.

7.4 Imperfect Substitution within Conglomerates

The previous section shows that the welfare effects of the Top 1,000 program are robust to the use of a range of values for the parameters ρ and α. This section explores the robustness of our results to allowing the outputs of firms within a conglomerate to be imperfect substitutes. We now assume that conglomerates produce a composite good $q(\omega) = \left(\sum_i q(\omega, i)^{\rho_c}\right)^{1/\rho_c}$, where $0 < \rho < \rho_c < 1$. This assumption implies that consumers have a larger elasticity of substitution between products of firms in a given conglomerate than across goods produced by different conglomerates, i.e., $\frac{1}{1-\rho} < \frac{1}{1-\rho_c}$. As we show in Appendix M.3, many of the results of our baseline model extend to this case after slight modifications.[74] For instance, we redefine the total productivity of a conglomerate as $\phi\Delta_n^C$, where $\Delta_n^C = \left[\sum_{i=1}^{n} \delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}}\right]^{\frac{1-\alpha\rho_c}{\rho_c}}$.

To operationalize this extension, we first reestimate the model assuming that $\rho_c = 0.90$. To gain intuition into how this extension impacts our model parameters, note that the share of the conglomerate's revenue from the i^{th} affiliate is now $\left(\frac{\delta^{i-1}}{\Delta_n^C}\right)^{\frac{\rho_c}{1-\alpha\rho_c}}$. Since $\rho_c < 1$, the within-conglomerate distribution of output in Panel A of Figure 3 implies a lower value of δ. Intuitively, since firms related to Top 1,000 firms are now less productive, the firm faces a greater productivity loss when shifting output to related firms. Accordingly, we estimate that $\delta = 0.80$ when we reestimate the model, which implies a larger shadow cost of the regulation of approximately 11%.

---

[74] Indeed, we obtain our baseline when $\rho_c = 1$. Table A.27 reports the estimated model under this extension.

The black square in Panel C of Figure 10 plots the aggregate effects of the program under the assumption that $\rho_c = 0.90$. Two features of the model lead to larger decreases in energy use. First, the imperfect substitution of output within the conglomerate limits the extent to which regulated conglomerates can shift production to related firms. Second, the lower value of δ limits the extent of these spillovers. Both of these forces limit within-conglomerate leakage of energy use. However, these forces also lead to a larger price increase, which exacerbates market leakage by shifting production to unrelated firms. On the whole, we find larger price effects and energy savings. As the black square in Panel C of Figure 10 shows, these effects imply a GWTP that is quite close to our baseline estimate.

Finally, we compare our model to a hypothetical scenario without conglomerates. In this model, related firms respond to market forces similarly to any other, unrelated firms, but conglomerates do not jointly optimize production decisions across their related firms. The red cross in Panel C of Figure 10 shows that a version of the model without conglomerates yields larger aggregate changes. This result reflects the fact that the within-conglomerate spillovers are stronger than the market-level spillovers. The GWTP implied by this model is very similar to our baseline estimate.[75]

8 Conclusion

This paper studies the effects of a prominent energy conservation program in China. We combine detailed data on energy use and business networks to study the effects of the regulation on both regulated firms and unregulated firms within the same conglomerate. While the program led regulated firms to decrease their energy use, this decrease was driven by a decline in production output and not by an increase in energy efficiency. We show that the program led to large increases in the output and energy use of unregulated firms in the same conglomerate. By shifting production to related firms, the regulated conglomerates escaped close to 40% of the regulation-driven output reduction. The facts that regulated conglomerates were unable to fully shift lost output to related firms and that we find no impacts on the energy efficiency of regulated firms imply that regulated firms found it costly to increase their energy efficiency.

We calculate the shadow cost of the regulation using a model of conglomerate production that matches our setting and the reduced-form effects of the regulation. The model shows that, even with the ability on the part of regulated firms to shift some production to related firms, the regulation increased the cost of conglomerate production by 8.7%.

A welfare analysis of the aggregate effects of the policy on consumption and energy use characterizes the government's willingness to pay (GWTP) to reduce carbon emissions that would

---

[75] Appendix M.4 derives this alternative version of the model, shows that the results of Propositions 4 and 5 continue to hold when we set $\Delta_n = 1$, and provides additional discussion.

be required for the Top 1,000 program to raise welfare. Importantly, the GWTP includes both global externalities from reducing carbon emissions and the local health benefits from reducing pollution. Our results suggest that the program increases welfare as long as the government is willing to pay at least \$161 to reduce a ton of carbon emissions. We characterize the uncertainty in this estimate by exploring a number of alternative model specifications and parameter values. Across these wide-ranging assumptions, we find that the GWTP value that rationalizes the policy lies between \$114 and \$199. On the basis of calibrations of the health benefits from reducing pollution, we estimate that approximately \$4–\$17 of this GWTP could be justified by local health benefits, with the remainder being justified by the social cost of carbon. We also show that the government can improve the regulation of energy by targeting the energy use of conglomerates. Such policies have lower shadow costs and are more effective from a welfare perspective.

Our analysis of the Top 1,000 program improves our understanding of the trade-offs involved in reducing energy use and related emissions in an important context. Indeed, the firms regulated by this program are some of the largest emitting firms in the world and understanding their behavior is crucial for the global control of carbon emissions. While it is important to recognize that our analyses are grounded in specific Chinese regulatory institutions, some of the economic mechanism we highlight—such as the importance of leakage within conglomerates—may also be important in other settings.

Overall, this paper shows that the economic effects and the efficacy of policies that target large firms are modulated by substitution along ownership networks. Since ownership networks are public information, the results of our paper reveal a potential avenue for improving existing energy regulations.

References

Ai, Hongshan, Yang Hu, and Ke Li. 2021. “Impacts of environmental regulation on firm productivity: evidence from China’s Top 1000 Energy-Consuming Enterprises Program.” *Applied Economics*, 53(7): 830–844.

Allcott, Hunt, and Michael Greenstone. 2012. “Is There an Energy Efficiency Gap?” *Journal of Economic Perspectives*, 26(1): 3–28.

Anderson, Soren T., and James M. Sallee. 2011. “Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards.” *American Economic Review*, 101(4): 1375–1409.

Anderson, Soren T., and Richard G Newell. 2004. “Information programs for technology adoption: the case of energy-efficiency audits.” *Resource and Energy Economics*, 26(1): 27–50.

Auffhammer, Maximilian, and Yazhen Gong. 2015. “China’s Carbon Emissions from Fossil Fuels and Market-Based Opportunities for Control.” *Annual Review of Resource Economics*, 7(1): 11–34.

Baylis, Kathy, Don Fullerton, and Daniel H. Karney. 2014. “Negative Leakage.” *Journal of the Association of Environmental and Resource Economists*, 1(1/2): 51–73.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 1995. “Capacity Utilization and Returns to Scale.” *NBER Macroeconomics Annual*, 10: 67–110.

Bushnell, James, Yihsu Chen, and Matthew Zaragoza-Watkins. 2014. “Downstream regulation of CO2 emissions in California’s electricity sector.” *Energy Policy*, 64: 313–323.

Cao, Kaihu. 2010. “Power Limit Affect Electrolytic Aluminum Industry, Jiaozuo Wanfang Aluminum Co,. Ltd. Reduced Production by 1/3.” *China Business News*.

Carleton, Tamma, and Michael Greenstone. 2021. “Updating the United States Government’s Social Cost of Carbon.” EPIC Working Paper 2021-04.

CCCPC. 2005. “Proposal of 11th Five Year Plan on National Economic and Social Development from the Central Committee of the Communist Party of China (中共中央关于制定国民经济和社会发展第十一个五年规划的建议).” Central Committee of the Communist Party of China Central Government.

Chen, Qiaoyi, Zhao Chen, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. 2021. “Industrial Energy Regulation: The Role of Business Conglomerates in China.” *AEA Papers and Proceedings*, 111: 396–400.

Colmer, Jonathan, Ralf Martin, Mirabelle Muûls, and Ulrich J. Wagner. 2020. “Does pricing carbon mitigate climate change? Firm-level evidence from the European Union emissions trading scheme.” Center for Economic Performance Discussion Paper.

CSBTS. 1991. “Method of Calculating Energy Saved in Industrial Enterprise (企业节能量计算方法).” China State Bureau of Technical Supervision Central Government.

Curtis, E. Mark. 2018. “Who Loses under Cap-and-Trade Programs? The Labor Market Effects of the NOx Budget Trading Program.” *The Review of Economics and Statistics*, 100(1): 151–166.

Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen. 2023. “Do Tax Incentives Increase Firm Innovation? An RD Design for R&D, Patents, and Spillovers.” *American Economic Journal: Economic Policy (forthcoming).*

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India.” *The Quarterly Journal of Economics*, 128(4): 1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. “The Value of Regulatory Discretion: Estimates from Environmental Inspections in India.” *Econometrica*, 86(6): 2123–2160.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. “New evidence on the impact of sustained exposure to air pollution on life expectancy from China’s Huai River Policy.” *Proceedings of the National Academy of Sciences*, 114(39): 10384–10389.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. “Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency.” *American Economic Review*, 97(4): 1250–1277.

Filippini, Massimo, Thomas Geissmann, Valerie J Karplus, and Da Zhang. 2020. “The productivity impacts of energy efficiency programs in developing countries: Evidence from iron and steel firms in China.” *China Economic Review*, 59: 101364.

Fischer, Carolyn, and Alan K. Fox. 2012. “Comparing policies to combat emissions leakage: Border carbon adjustments versus rebates.” *Journal of Environmental Economics and Management*, 64(2): 199–216.

Fowlie, Meredith L. 2009. “Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage.” *American Economic Journal: Economic Policy*, 1(2): 72–112.

Fowlie, Meredith L, and Mar Reguant. 2022. “Mitigating emissions leakage in incomplete carbon markets.” *Journal of the Association of Environmental and Resource Economists*, 9(2): 307–343.

Fowlie, Meredith L., Mar Reguant, and Stephen P. Ryan. 2016. “Market-Based Emissions Regulation and Industry Dynamics.” *Journal of Political Economy*, 124(1): 249–302.

Gao, Hang, and Johannes Van Biesebroeck. 2014. “Effects of Deregulation and Vertical Unbundling on the Performance of China’s Electricity Generation Sector.” *The Journal of Industrial Economics*, 62(1): 41–76.

Gibson, Matthew. 2019. “Regulation-Induced Pollution Substitution.” *The Review of Economics and Statistics*, 101(5): 827–840.

Gillingham, Kenneth, Amelia Keyes, and Karen Palmer. 2018. "Advances in Evaluating Energy Efficiency Policies and Programs." *Annual Review of Resource Economics*, 10(1): 511–532.

Greenstone, Michael. 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, Michael, and B. Kelsey Jack. 2015. "Envirodevonomics: A Research Agenda for an Emerging Field." *Journal of Economic Literature*, 53(1): 5–42.

Greenstone, Michael, John A List, and Chad Syverson. 2012. "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing." National Bureau of Economic Research Working Paper 18392.

Hainmueller, Jens. 2012. "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis*, 20(1): 25–46.

Hanna, Rema. 2010. "US Environmental Regulation and FDI: Evidence from a Panel of US-Based Multinational Firms." *American Economic Journal: Applied Economics*, 2(3): 158–89.

Harrison, Ann, Benjamin Hyman, Leslie Martin, and Shanthi Nataraj. 2015. "When do Firms Go Green? Comparing Command and Control Regulations with Price Incentives in India." National Bureau of Economic Research Working Paper 21763.

He, Guojun, Shaoda Wang, and Bing Zhang. 2020. "Watering Down Environmental Regulation in China." *The Quarterly Journal of Economics*, 135(4): 2135–2185.

IEA. 2023. "IEA Policies Database." International Energy Agency. Accessible at `https://www.iea.org/policies?sector%5B0%5D=Industry&topic%5B0%5D=Energy%20Efficiency` (Accessed: April 12, 2023).

Ito, Koichiro, and Shuang Zhang. 2020*a*. "Reforming Inefficient Energy Pricing: Evidence from China." National Bureau of Economic Research Working Paper 26853.

Ito, Koichiro, and Shuang Zhang. 2020*b*. "Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China." *Journal of Political Economy*, 128(5): 1627–1672.

IWG. 2021. "Technical Support Document: Social Cost of Carbon, Methane, and Nitrous Oxide Interim Estimates under Executive Order 13990." Interagency Working Group on Social Cost of Greenhouse Gases United States Government.

Jia, Ruixue. 2017. "Pollution for Promotion." *21st Century China Center Research Paper No. 2017-05*, (March 21, 2017).

Kahn, Matthew E., Pei Li, and Daxuan Zhao. 2015. "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives." *American Economic Journal: Economic Policy*, 7(4): 223–42.

Karplus, Valerie J, Xingyao Shen, and Da Zhang. 2020. "Herding cats: Firm noncompliance in China' s industrial energy efficiency program." *The Energy Journal*, 41(4).

Ke, Jing, Lynn Price, Stephanie Ohshita, David Fridley, Nina Zheng Khanna, Nan Zhou, and Mark Levine. 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.

Kostka, Genia, and William Hobbs. 2012. “Local energy efficiency policy implementation in China: bridging the gap between national priorities and local interests.” *The China Quarterly*, 211: 765–785.

Li, Bin. 2010. “Tangshan: the Energy Eaving Dilemma of a Steel City.” *China Business*.

Li, Ruoxin, and Xiaohui Li. 2010. “Energy Saving Sprint Set Production Limit, Energy Intensive Firms All Suffered.” *China Securities Journal*.

Ma, Liang, and Jiaqi Liang. 2018. “The effects of firm ownership and affiliation on government’ s target setting on energy conservation in China.” *Journal of Cleaner Production*, 199: 459–465.

Mansur, Erin T. 2007. “Prices vs. Quantities: Environmental Regulation and Imperfect Competition.” National Bureau of Economic Research Working Paper 13510.

McKinsey & Co. 2009. “Unlocking Energy Efficiency in the U.S. Economy.” McKinsey & Co.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71(6): 1695–1725.

Melitz, Marc J., and Stephen J. Redding. 2015. “New Trade Models, New Welfare Implications.” *American Economic Review*, 105(3): 1105–46.

Mohan, Aniruddh, Nicholas Z Muller, Akshay Thyagarajan, Randall V Martin, Melanie S Hammer, and Aaron van Donkelaar. 2020. “The Growth of Nations Revisited: Global Environmental Accounting from 1998 to 2018.” National Bureau of Economic Research Working Paper 27398.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus. 2011. “Environmental Accounting for Pollution in the United States Economy.” *American Economic Review*, 101(5): 1649–75.

NDRC. 2006. “Announcement of the Implementation of the Top1000 Energy Conservation Program (关于印发千家企业节能行动实施方案的通知).” National Development and Reform Commission Central Government.

NPC. 1989. “Environmental Protection Law of the People’s Republic of China.” National People’s Congress. Accessible at `https://www.cecc.gov/resources/legal-provisions/prc-environmental-protection-law` (Accessed: May 3, 2023).

Ohrn, Eric. 2018. “The Effect of Corporate Taxation on Investment and Financial Policy: Evidence from the DPAD.” *American Economic Journal: Economic Policy*, 10(2): 272–301.

Porter, Michael E., and Claas van der Linde. 1995. “Green and Competitive.” *Harvard Business Review*, 120–134.

Price, Lynn, Lynn Price, Xuejun Wang, and Jiang Yun. 2008. "China's Top-1000 Energy-Consuming Enterprises Program:Reducing Energy Consumption of the 1000 Largest Industrial Enterprises in China." Lawrence Berkeley National Laboratory.

Price, Lynn, Mark D Levine, Nan Zhou, David Fridley, Nathaniel Aden, Hongyou Lu, Michael McNeil, Nina Zheng, Yining Qin, and Ping Yowargana. 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, Lynn, Xuejun Wang, and Jiang Yun. 2010. "The Challenge of Reducing Energy Consumption of the Top-1000 Largest Industrial Enterprises in China." *Energy Policy*, 38(11): 6485–6498.

Ramachandran, J, K S Manikandan, and Anirvan Pant. 2013. "Why Conglomerates Thrive (Outside the U.S.)." *Harvard Business Review*, Dec 1, 2013.

Ritchie, Hannah, Max Roser, and Pablo Rosado. 2020. "China: CO2 country profile."

Ryan, Nicholas. 2018. "Energy Productivity and Energy Demand: Experimental Evidence from Indian Manufacturing Plants." National Bureau of Economic Research Working Paper 24619.

Ryan, Stephen P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica*, 80(3): 1019–1061.

Shapiro, Joseph S. 2016. "Trade Costs, CO2, and the Environment." *American Economic Journal: Economic Policy*, 8(4): 220–54.

Shapiro, Joseph S. 2021. "The Environmental Bias of Trade Policy." *The Quarterly Journal of Economics*, 136(2): 831–886.

Shapiro, Joseph S., and Reed Walker. 2018. "Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade." *American Economic Review*, 108(12): 3814–54.

Shen, Meng, Yanping Zeng, and Ruxiao Qu. 2015. "Environmental Regulation and Exportation: Micro Evidence from China's Top–1000 Policy." *Journal of International Trade (In Chinese)*, 8: 43–50.

Soliman, Felix Samy. 2020. "Intrafirm Leakage." CESifo CESifo Working Paper Series 8619.

the People's Government of Chongqing City. 2007. "Announcement of the People's Government of Chongqing City on the Energy Saving Task Decomposition in 2007 (重庆市人民政府办公厅关于 2007 年节能降耗目标任务分解的通知)." the People's Government of Chongqing City Local Government.

the People's Government of Hainan Province. 2006. "Announcement of the People's Government of Hainan Province on the Energy Consumption per unit of GDP Reduction Target for Cities, Counties and Energy Consumption Reduction Target for Energy-Intensive Firms during the 11th Five Year Plan (海南省人民政府办公厅关于下达十一五期间市县单位生产总值能源消耗降低指标计划和重点耗能企业能源消耗降低指标计划的通知)." the People's Government of Hainan Province Local Government.

van Biesebroeck, Johannes. 2003. “Productivity Dynamics with Technology Choice: An Application to Automobile Assembly.” *The Review of Economic Studies*, 70(1): 167–198.

Walker, W. Reed. 2013. “The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce.” *The Quarterly Journal of Economics*, 128(4): 1787–1835.

Wang, Jie. 2010. “Luo Bingsheng, Vice Executive President of CISA Oppose the ”One Size Fits All” Policy.” *21st Century Business Herald.*

Wang, Qinyong. 2008. “China Aluminum Group Co., Ltd. Guizhou Branch Faces the Largest Power Limit Crisis.” *China Nonferrous Metals News.*

Wang, Xuejun, Zhihua Fu, Wei Zhang, and Youbin Ye. 2008. “A Study of Approaches and Methods to Meet China’s 20% Energy Efficient Target for 2010 (中国可持续能源实施“十一五”20% 节能目标的途径与措施研究).” Chapter 5. Top1000 Energy Conservation Program and the 20% Energy Saving Target Achievement in the 11th Five Year Plan (《千家企业节能行动》与“十一五”20% 节能目标的实现). China Sustainable Energy Project, Central Government.

Wang, Yufei, Huimin Li, Qijiao Song, and Ye Qi. 2017. “The consequence of energy policies in China: A case study of the iron and steel sector.” *Resources, Conservation and Recycling*, 117: 66–73.

World Bank. 2007. “Cost of pollution in China : economic estimates of physical damages (English).” Chapter 4. World Bank Group.

Xiao, Yuxian, Haitao Yin, and Jon J Moon. 2023. “Coping with Externally Imposed Energy Constraints: Competitiveness and Operational Impact of China’ s Top-1000 Energy-Consuming Enterprises Program.” *The Energy Journal*, 44(2).

Xu, Min. 2010. “From dilemma to Path – A New Perspective on the Energy Conservation Supervision and Audit.” National Audit Office of the People’s Republic of China United States Government.

Yang, Mian, Dalia Patino-Echeverri, Fuxia Yang, and Eric Williams. 2015. “Industrial energy efficiency in China: achievements, challenges and opportunities.” *Energy Strategy Reviews*, 6: 20–29.

Yu, Hong. 2011. “Fertilizer firm show a wry face to the Power Limit.” *China Chemical Industry News.*

Zhang, Yanyun, and Tao Huang. 2019. “Environmental regulation and enterprises’ innovation: Based on the analysis of China’ s top-1000 energy-saving program.” *Journal of Chongqing University of Technology(Social Science)(In Chinese)*, 6: 43–49.

Zhao, Xiaofan, Huimin Li, Liang Wu, and Ye Qi. 2014. “Implementation of energy-saving policies in China: How local governments assisted industrial enterprises in achieving energy-saving targets.” *Energy Policy*, 66: 170–184.

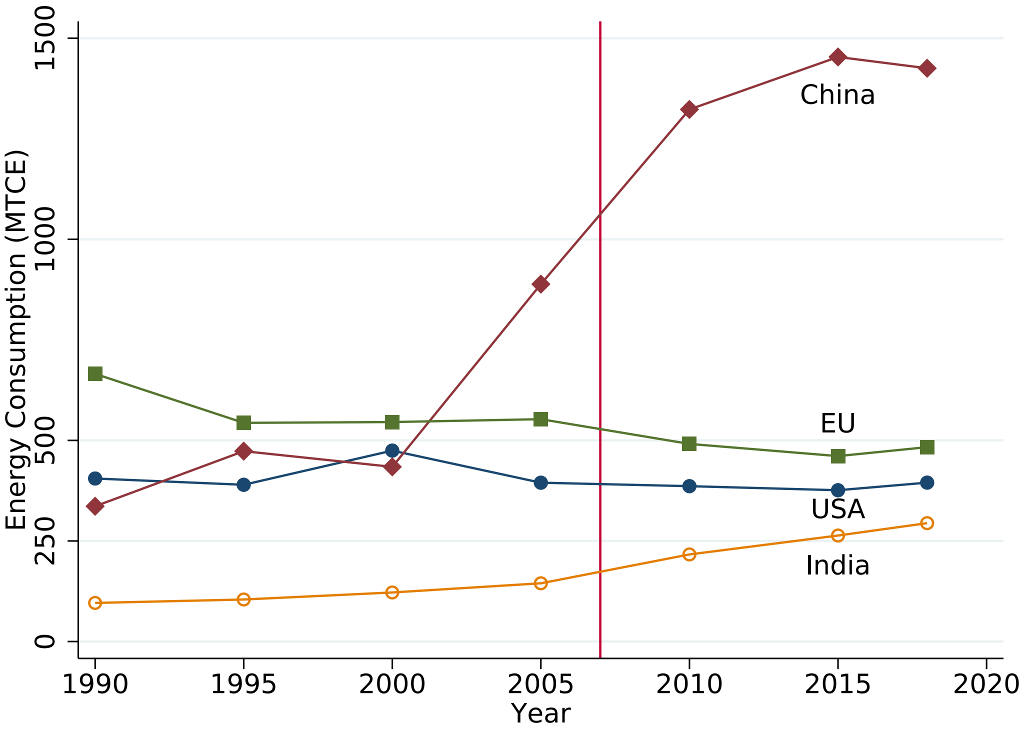
Zhao, Xiaofan, Huimin Li, Liang Wu, and Ye Qi. 2016. “Enterprise-level amount of energy saved targets in China: weaknesses and a way forward.” *Journal of Cleaner Production*, 129: 75–87.

Zhou, Nan, Mark D Levine, and Lynn Price. 2010. "Overview of current energy-efficiency policies in China." *Energy policy*, 38(11): 6439–6452.

Zhu, Junming, Limin Niu, Matthias Ruth, and Lei Shi. 2018. "Technological change and energy efficiency in large Chinese firms." *Ecological Economics*, 150: 241–250.

Figures

Figure 1: Cross-Country Differences in Industrial Energy Use

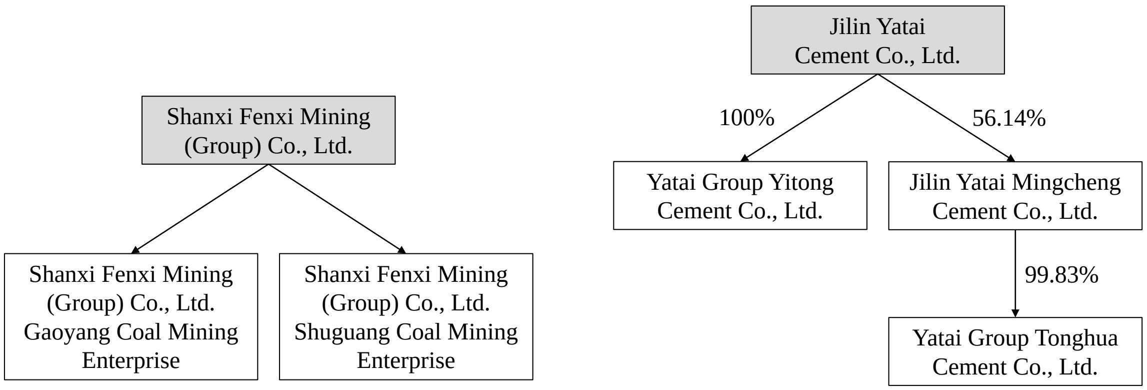


Notes: Authors' calculations using data from the IEA. This figure plots aggregate industrial energy consumption in China, the US, the EU and India from 1990 to 2018 using units of million tons of coal equivalent (mtce). The industrial energy consumption of China increased dramatically after 2000, by more than threefold, while the industrial energy consumption of the US and EU remained relatively stable with a slight downward trend. The red line marks the start year of the Top 1,000 Energy Conservation Program.

Figure 2: Examples of Firm Relations

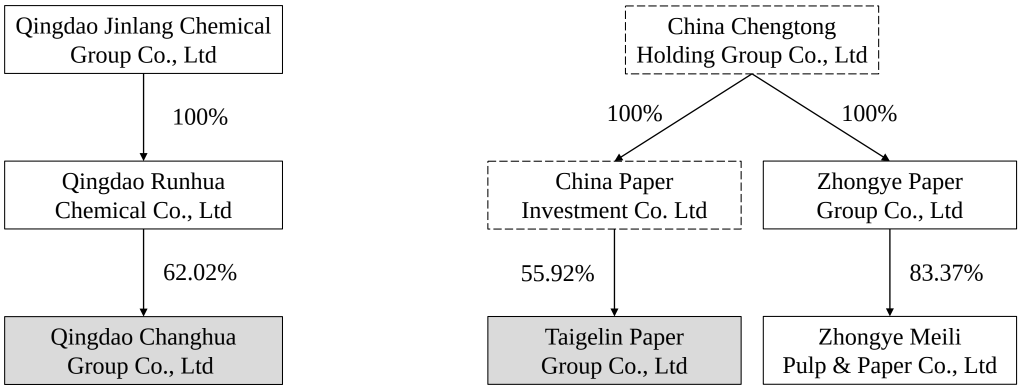
A. Subsidiary

B. Investment



C. Shareholder

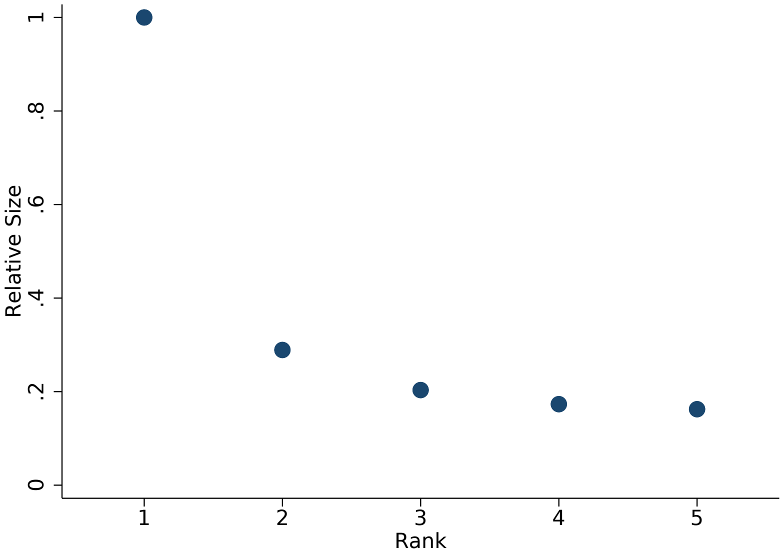
D. Shareholder Investment



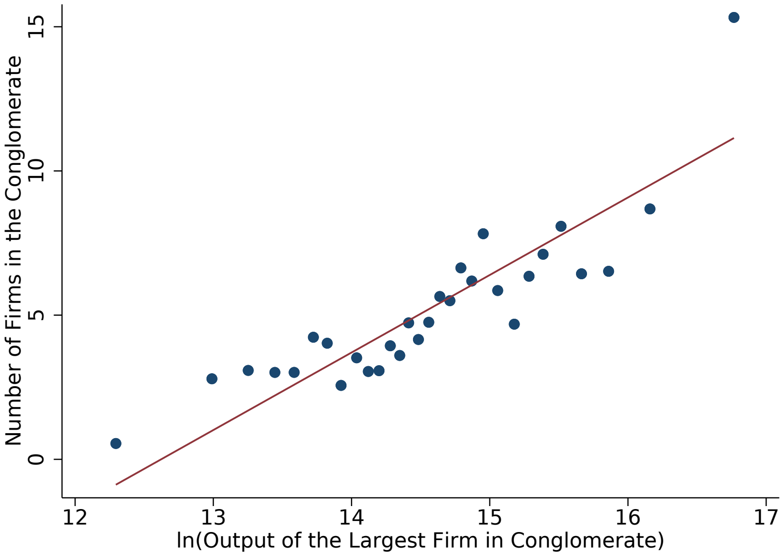
Notes: Authors' calculations using data from CARD. This figure shows examples of related firms, including wholly owned subsidiary firms in Panel A, investment firms in Panel B, shareholder firms in Panel C and shareholder investment firms in Panel D. Firms shaded in gray are part of the Top 1,000 program; firms without shading are part of the same conglomerate. In Panel D, we denote firms not in the same 4-digit industry as the Top 1,000 firm with dashed lines. Ownership share is reported next to each link. See Section 1.3 for the definition of related firms.

Figure 3: Conglomerate Size and Production Allocation

A. Relative Firm Size

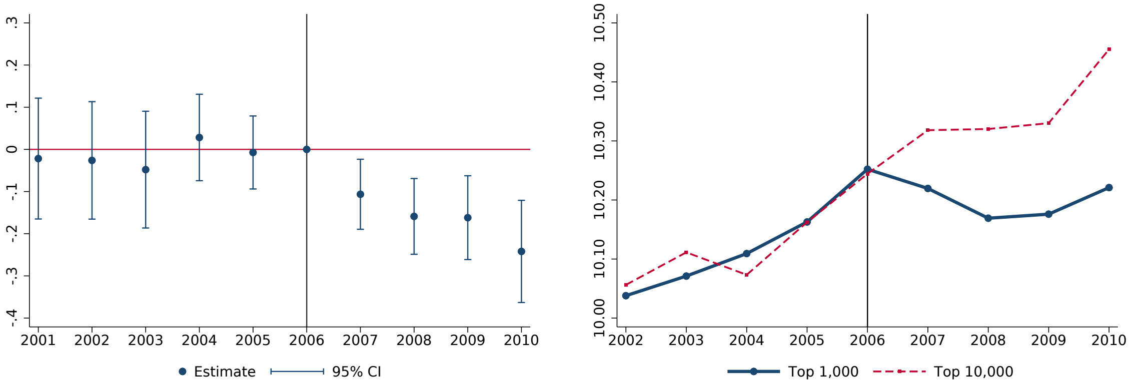


B. Output and Conglomerate Size

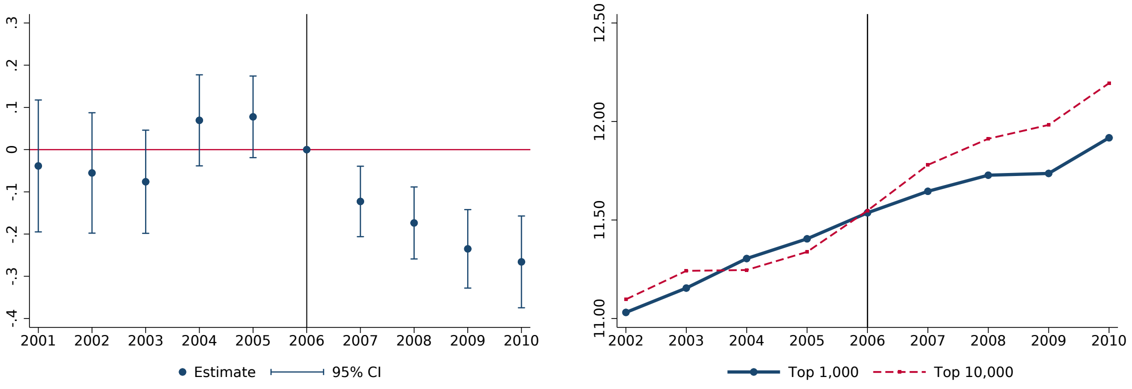


Notes: Authors' calculations using data from ASIF and CARD. This figure shows stylized facts about conglomerate size and relative firm size within conglomerates. Panel A plots the average relative size within a Top 1,000 conglomerate (each firm's size relative to the largest firm in the conglomerate). Firms are ranked by size from the largest to the smallest, and size is measured by industrial output. This figure shows that firm size declines very quickly in a conglomerate, with the second largest firm being only 29% of the size of the largest. Panel B plots the results of a regression of firm number on log output of the largest firm in a Top 1,000 conglomerate. It shows that conglomerates with larger leading firms usually have more firms. See Section 1.3 for additional discussion.

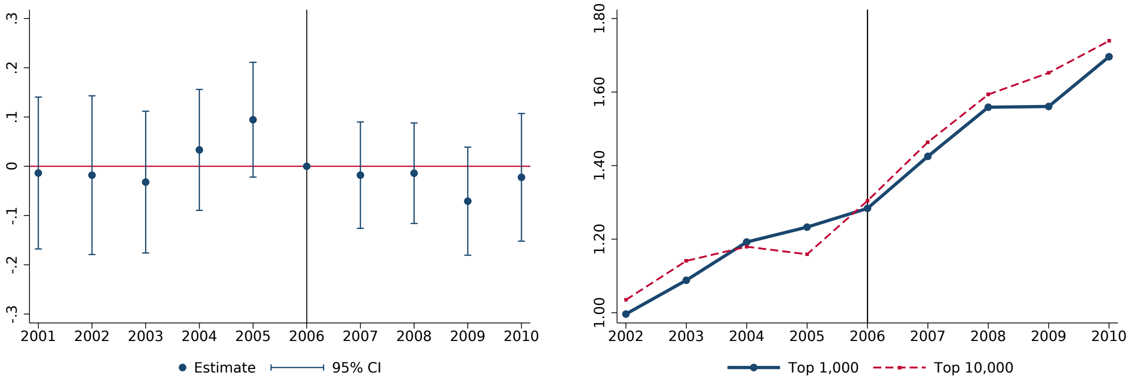
Figure 4: Effects of the Program on Regulated Firms


A. Energy Use: Coefficients

B. Energy Use: Event Study


C. Output: Coefficients

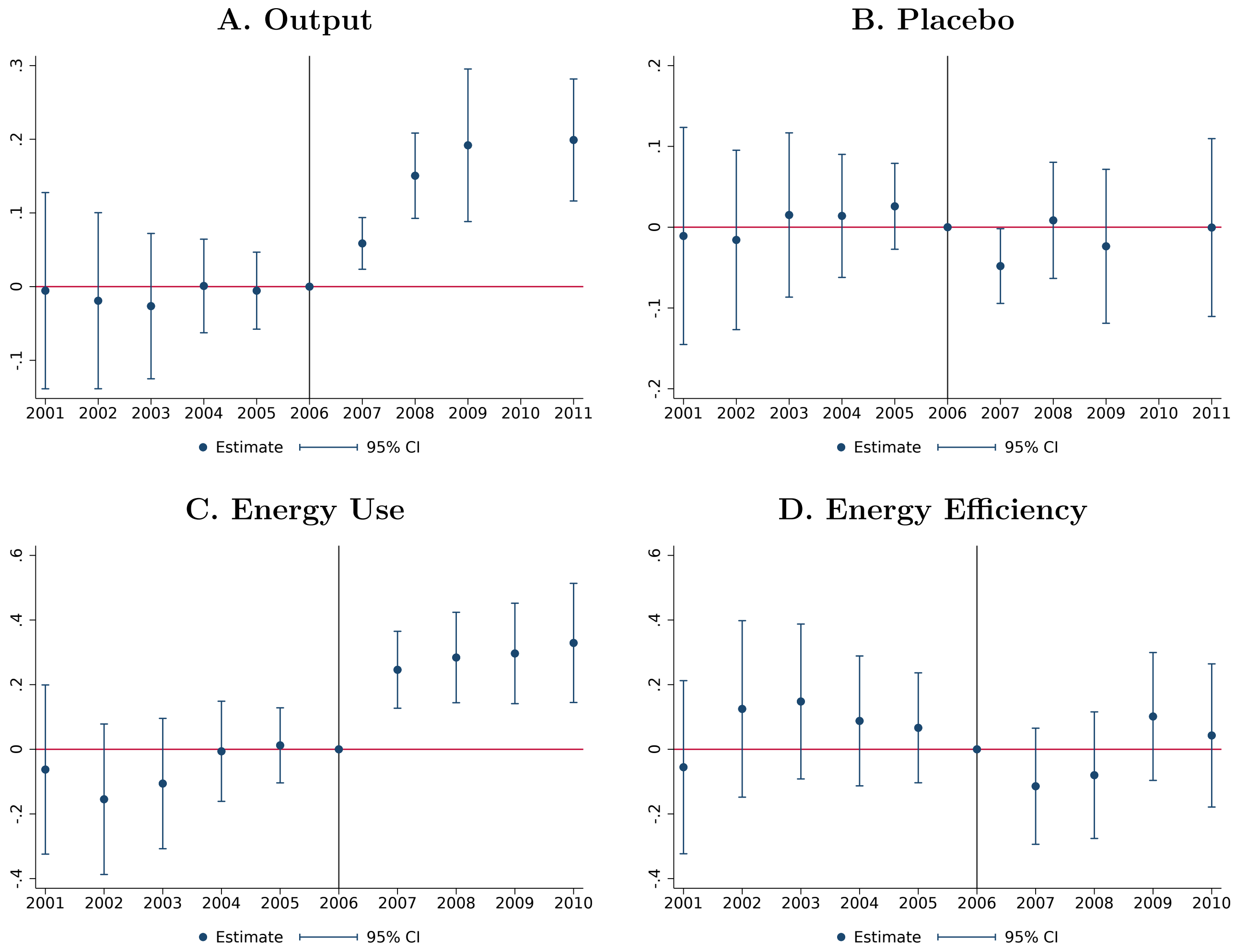
D. Output: Event Study


E. Energy Efficiency: Coefficients

F. Energy Efficiency: Event Study

Notes: Authors' calculations using data from the CESD. This figure shows estimates of Equation 1 where the dependent variable is log firm energy consumption in Panels A and B, log firm output in Panels C and D, and log firm energy efficiency in Panels E and F. Energy efficiency is defined as output per unit of energy consumption. This figure shows that regulated firms (Top 1,000 firms) decreased their energy consumption and output substantially relative to similar control firms (Top 10,000 firms not related to Top 1,000 firms) after the regulation, while no improvement in energy efficiency in these regulated firms can be found. Point estimates are displayed in Table 3. See Section 2 for additional discussion. Standard errors are clustered at the firm level.

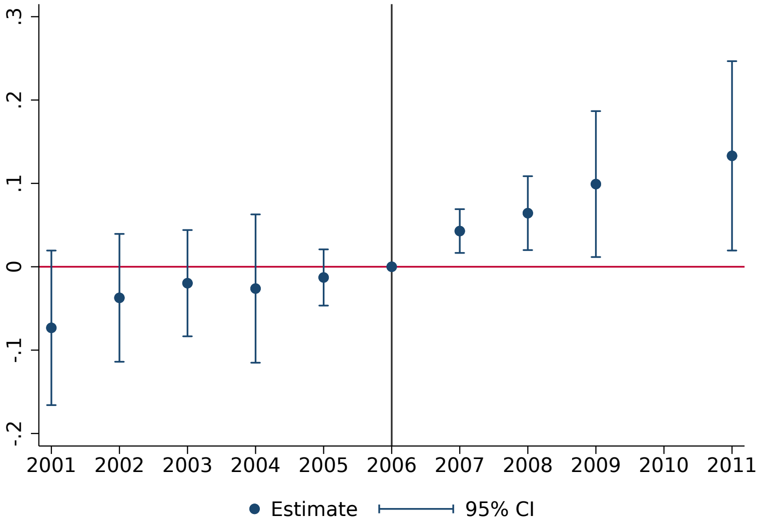
Figure 5: Spillover Effects on Related Firms



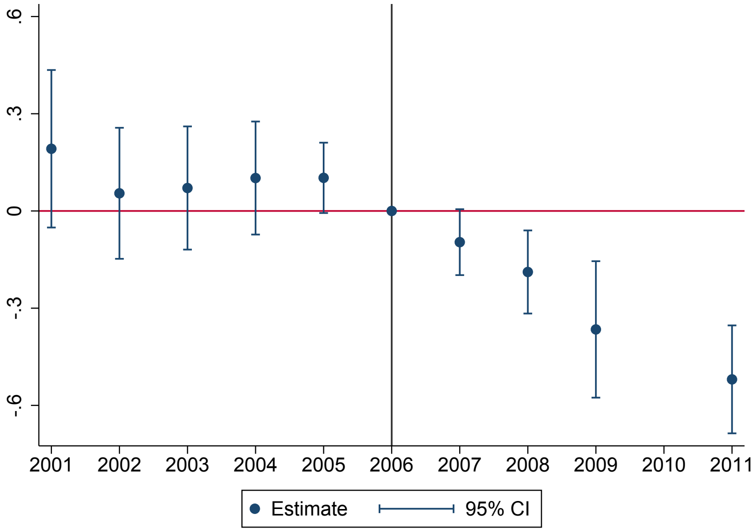
Notes: Authors' calculations using data from ASIF. This figure shows the effects of the Top 1,000 Energy Conservation Program on the related parties of regulated firms. Panel A shows that related firms in the same 4-digit industry as regulated firms increased their output significantly after the policy implementation relative to similar control firms and that this effect persisted during the policy period. See Section 3 for a description of the procedure used to identify the comparison firms. The point estimate for Panel A is displayed in Panel A of Table 4. Panel B plots the output results for placebo firms (related firms in the same 2-digit industry but outside the 4-digit industry of regulated firms). This graph shows that placebo firms were not affected by the regulation. The point estimate for Panel B is displayed in Panel B of Table 4. Panels C and D show that related firms in the same 4-digit industry increased their energy consumption after the regulation but did not improve their energy efficiency relative to similar control firms. The point estimates for Panels C and D are displayed in Table 4. See Section 3 for additional discussion. The results of robustness checks using an alternative matching method are shown in Figure A.10. Standard errors are clustered at the firm level.

Figure 6: Industry-Level Spillovers and Within-Conglomerate Effects

A. Industry-Level Spillovers



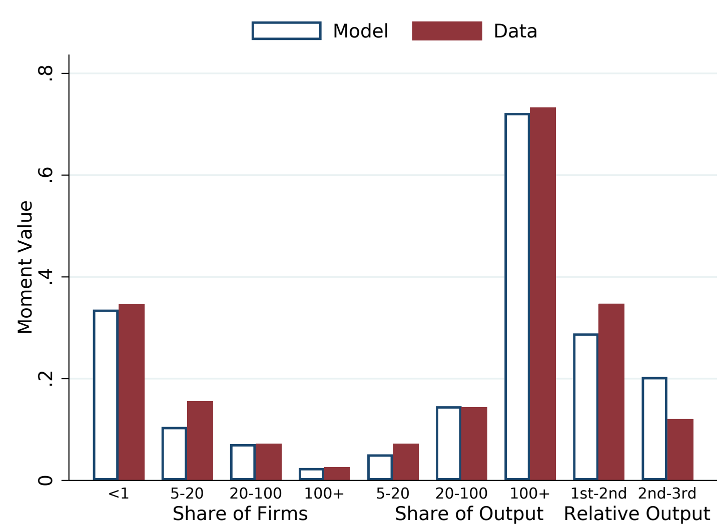
B. Within-Conglomerate Difference-in-Differences



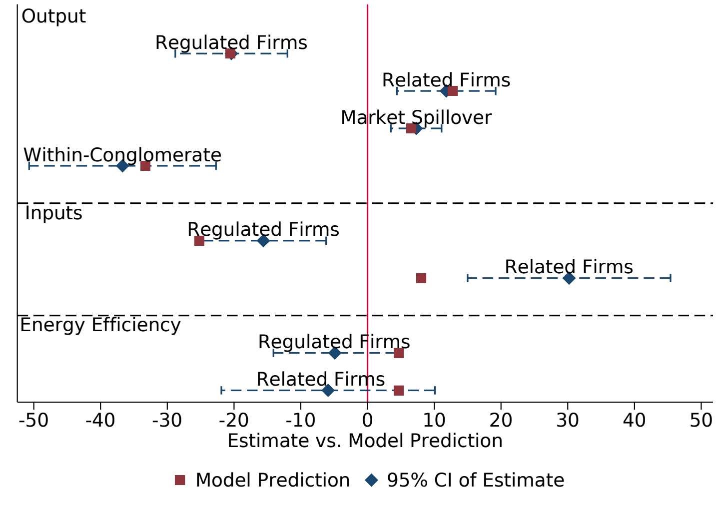
Notes: Authors' calculations using data from the ASIF. Panel A of this figure shows estimates of Equation 3 where the dependent variable is log firm output. Consistent with the market spillover hypothesis, we see that unregulated firms in industries with stricter regulation increased their output significantly after the policy was implemented. Coefficient estimates and robustness checks are shown in Panel A of Table 5 and Table A.24. See Section 3 for additional discussion. Standard errors are clustered at the firm level. Panel B of this figure plots the output change of regulated firms relative to their related firms (in the same 4-digit industry) within the same conglomerate. We see a strong and persistent output reallocation following the regulation from regulated firms to their related firms. Point estimates are displayed in Panel B of Table 5. See Section 5.3 for additional discussion. Conglomerate-by-year fixed effects are included, and standard errors are clustered at the conglomerate level.

Figure 7: Structural Model Fit and Out-of-Sample Validation

A. Moments: Data vs. Model



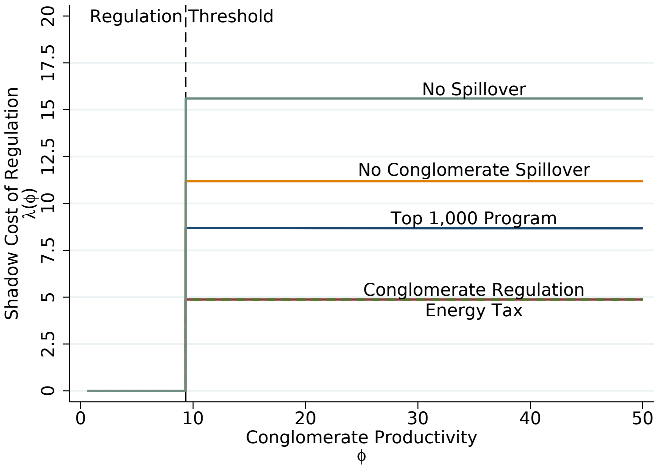
B. Out-of-Sample Validation: Difference-in-Differences Effects of the Program



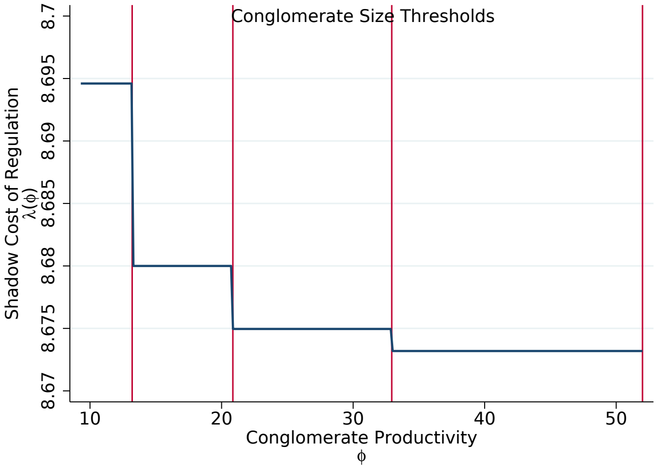
Notes: This figure shows the model fit for both the firm size distribution prior to the policy and the firm response after the policy. Panel A plots the size distribution of firms predicted by our model in blue bars and the size distribution calculated from the ASIF and economic census of 2004 in red bars. It shows that our model fits the data well in terms of both the observed firm size distribution and the concentration of output prior to the regulation. Panel B plots the firm response predicted by the model in red squares and the firm response obtained from our previous difference-in-differences estimates in blue diamonds. The blue lines span the 95% confidence interval for our difference-in-differences estimates. This graph shows that our model does a good job of fitting the output, input, and efficiency response of firms, with almost all model-predicted values lying within the 95% confidence intervals. See Section 5.2 for additional discussion.

Figure 8: Model-Based Estimates of the Shadow Costs of Regulation

A. Shadow Costs of Alternative Regulations



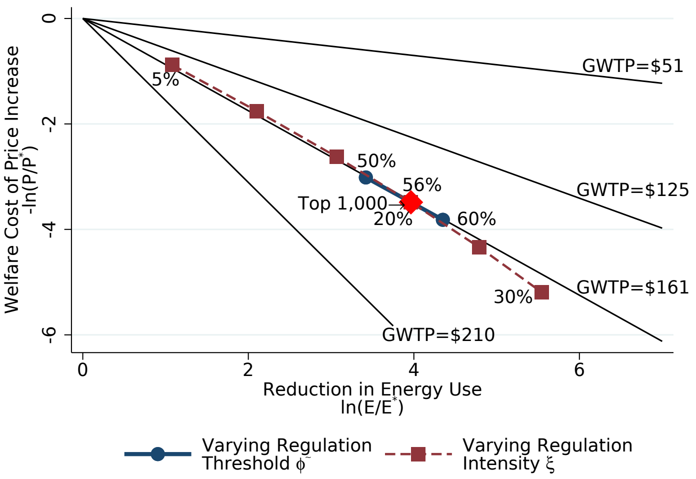
B. Size Distortions in Top 1,000 Program



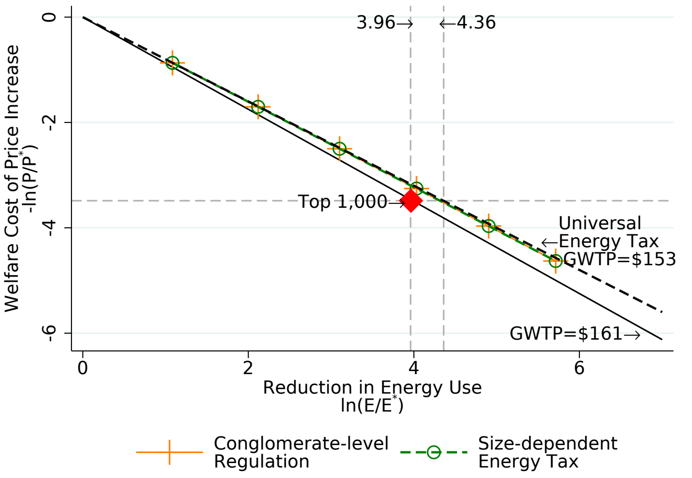
Notes: This figure shows the implied shadow cost of different regulations estimated by our model. Panel A plots the shadow cost of our baseline regulation (the Top 1,000 Energy Conservation Program) with the blue line, the shadow cost of the counterfactual in which both conglomerate spillovers and market spillovers are shut down with the gray line, the shadow cost of the counterfactual in which only conglomerate spillovers are shut down with the yellow line, the shadow cost of regulating conglomerates with the same energy saving amount with the green line, and the shadow cost of imposing an energy tax with the same energy saving amount with the dashed red line. For each scenario, we solve the model and calculate the corresponding shadow cost. See Appendices K and L for the equilibrium conditions under these alternative regulations. We can see that both market spillovers and conglomerate spillovers lower the shadow cost substantially while regulating conglomerates and imposing an energy tax can further lower the shadow cost by over 40% with the same amount of energy saving. Panel B zooms in to show the shadow cost under our baseline regulation. We see that the shadow costs are no longer constant among different conglomerates, unlike in Panel A. Conglomerates with more firms have a slightly lower shadow cost, while conglomerates with the same number of firms have the same shadow cost. See Sections 6.1 and 6.2 for additional discussion.

Figure 9: Welfare Effects of Alternative Regulations

A. Changes to Top 1,000 Program

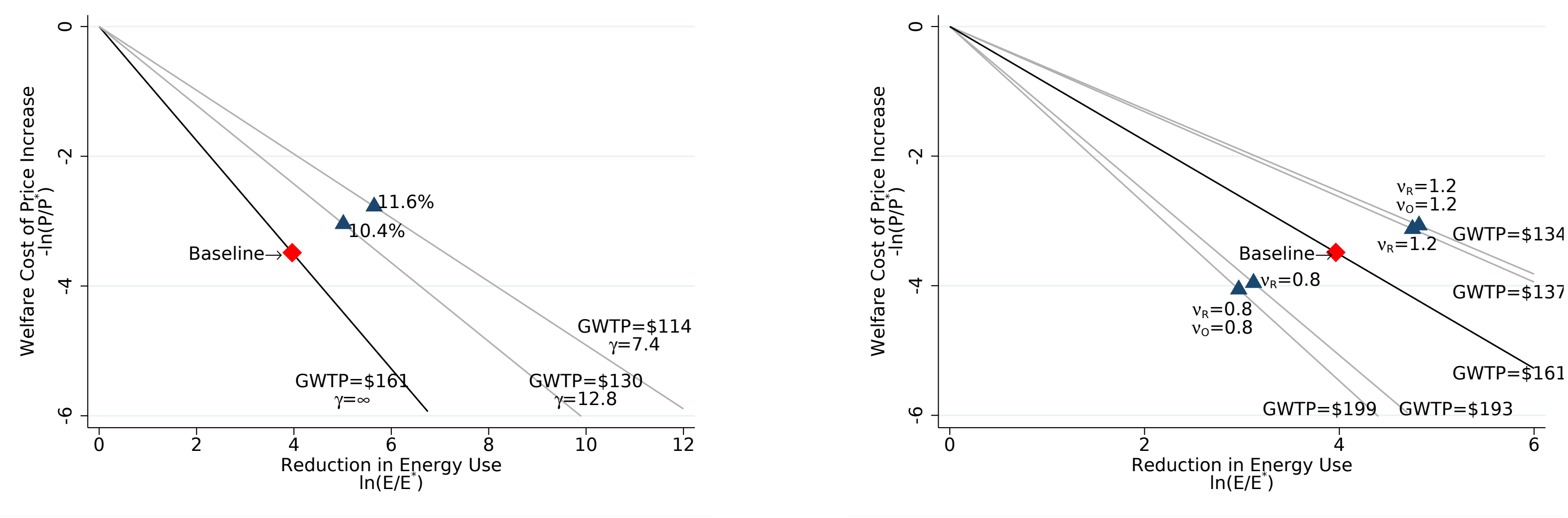


B. Alternative Regulations



Notes: This figure shows the welfare effects of different regulations measured by the trade-off between reductions in energy use and the welfare cost of price increases. Panel A shows the welfare effects of the Top 1,000 program. The black lines are indifference curves for different GWTPs. The red diamond shows that the Top 1,000 program led to an aggregate energy consumption reduction of close to 4% and a price level increase of approximately 3.5%, which can be rationalized with a GWTP=$161. The navy line indicates that expanding or contracting policy coverage to cover between 50% and 60% of an industry's energy use does not change the fundamental trade-off between reductions in energy consumption and price increases. The crimson line indicates that increasing the input reduction quota from 5% to 30% makes this trade-off slightly worse. See Section 6.1 for additional discussion. Panel B shows the welfare effects under alternative types of regulations. Regulating conglomerates and imposing a size-dependent energy tax show a similar trade-off at a GWTP=$154, which corresponds to a better performance than that of the original policy. A universal energy tax performs even better, with a GWTP=$153. See Section 6.2 for additional discussion.

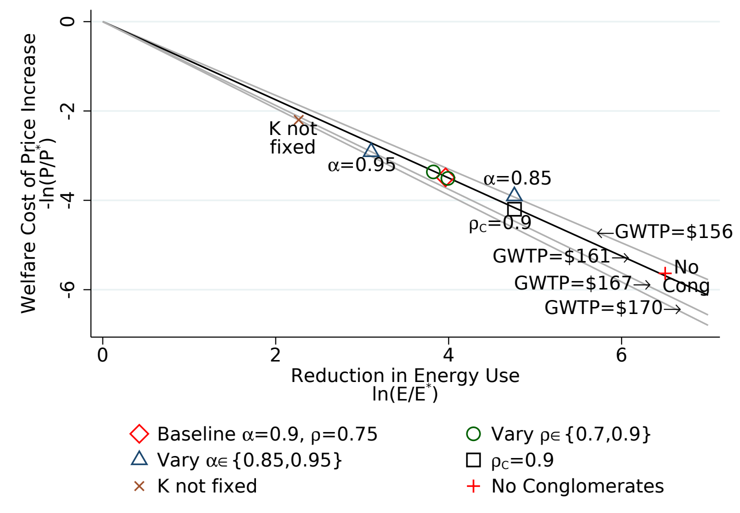
Figure 10: Welfare Effects: Robustness



A. Endogenous Energy Efficiency

B. Heterogeneous Energy Efficiency

C. Alternative Parameters



Notes: This figure shows the robustness of the estimated welfare effects of the Top 1,000 program. The black lines plot indifference curves for the baseline value of GWTP=\$161, and the light gray lines plot values of the GWTP according to different extensions. Panel A shows the effects under different values of the parameter γ, which determines the cost of improving energy efficiency. See Section 7.1 for details. Panel B shows the effects when we assume preexisting differences in energy efficiency. $\nu_R > 1$ denotes that related firms are more energy efficient than Top 1,000 firms, while $\nu_O < 1$ denotes that unregulated and unrelated firms are less efficient than Top 1,000 firms and vice versa. See Section 7.2 for details. Panel C shows the effects when we assume different values of the calibrated parameters. Blue triangles hold $\alpha = 0.9$ and vary $\rho \in \{0.7, 0.9\}$, while green circles hold $\rho = 0.75$ and vary $\alpha \in \{0.85, 0.98\}$. The brown X-mark shows the results when we assume that firms adjust their capital. See Section 7.3 for details. The black square in Panel C plots the effects when we assume that the outputs of firms in a conglomerate are imperfect substitutes (i.e., $\rho_c = 0.9$). The red cross shows results under an alternative model without conglomerates (which sets $s_{\tilde{\phi}} = 0.56$). See Section 7.4 for details. Across these wide-ranging assumptions, the GWTP that rationalizes the Top 1,000 program lies between \$114 and \$199.

Tables

Table 1: Energy Consumption of Top 1,000 Firms in Different Industries

Industry	Energy Consumption (10, 000 ton coal equiv.)	Proportion (%)	Firm Number
Iron and Steel	22528.63	30.72	249
Electric Power	16249.64	22.16	144
Chemical	10909.29	14.88	238
Petroleum and Petrochemical	10581.76	14.43	98
Mining	5278.77	7.20	60
Nonferrous	2993.08	4.08	70
Construction Materials	2913.19	3.97	93
Pulp and Paper	961.36	1.31	24
Textile	917.57	1.25	22

Notes: This table reports the number of firms and energy consumption of Top 1,000 firms in each industry in 2005 according to the National Development and Reform Commission (NDRC) and National Bureau of Statistics (NBS) of China (Bulletin on Top 1,000 Firms' Energy Consumption, 2007). The first column shows the industry name, the second column shows the aggregate energy consumption of Top 1,000 firms in each industry in 2005, the third column shows the proportion of energy consumption, and the last column shows the number of firms. A total of 998 of the 1008 Top 1,000 firms are included in this report.

Table 2: Summary Statistics

A. Firm-Level Data

Source	Variables	Top 1,000			Top 10,000 (Excluding Top 1,000)		
		Obs	Mean	SD	Obs	Mean	SD
CESD	ln(Energy)	3,419	12.37	1.48	20,207	9.84	1.58
	ln(Output)	3,381	13.69	1.66	20,076	11.26	1.58
	ln(Efficiency)	3,381	1.31	1.45	20,076	1.42	1.71
ASIF	SOE	3,405	0.26	0.44	19,386	0.08	0.28
	ROA	3,236	0.04	0.09	17,790	0.07	0.15
	Age	3,253	23.95	19.61	18,202	12.64	12.18
	Export	3,415	0.30	0.46	20,151	0.11	0.31

B. Conglomerate Networks: Related Parties

Datasets	Two Levels			Six Levels
	25%	20%	51%	20%
CARD	46,178	50,846	30,096	77,783
CARD&ASIF	7,329	7,907	5,061	9,832
CARD&ASIF (same 2-digit industry)	3,992	4,137	2,941	4,800
CARD&ASIF (same 4-digit industry)	**2,466**	2,514	1,963	2,827

Notes: This table reports summary statistics for Top 1,000 firms, Top 10,000 firms, and the conglomerate networks of Top 1,000 firms. Panel A shows the mean characteristics and firm counts with nonmissing data from the ASIF and CESD for Top 1,000 and Top 10,000 firms. We exclude related firms from the same 4-digit industry as Top 10,000 firms for both datasets. Additionally, for the CESD data, we exclude all industries whose electricity consumption accounts for more than 30% of total industry energy consumption. Output value is in thousands of RMB. Energy is measured in tons of coal equivalent. Energy efficiency is defined as thousand RMB of output per ton of energy input. See Section 1.2 for a detailed data description and the cleaning procedure. Panel B shows the total number of related firms that Top 1,000 firms have under different definitions of related parties. With 2 levels and a 25% ownership requirement, Top 1,000 firms have 3,992 related firms in the same 2-digit industry in the ASIF and 2,466 related firms in the same 4-digit industry in the ASIF. See Section 1.3 for additional discussion.

Table 3: Effects of the Program on Regulated Firms

A. Energy Use

Variables	ln(Energy Use)			
Treat × Post	-0.125***	-0.156***	-0.156***	-0.128***
	(0.042)	(0.045)	(0.047)	(0.048)
Observations	23,607	23,602	23,151	20,571
R^2	0.887	0.890	0.892	0.898

B. Output

Variables	ln(Output)			
Treat × Post	-0.096**	-0.226***	-0.204***	-0.145***
	(0.040)	(0.041)	(0.042)	(0.042)
Observations	23,435	23,430	22,991	20,446
R^2	0.881	0.887	0.889	0.893

C. Energy Efficiency

Variables	ln(Energy Effiency)			
Treat × Post	0.032	-0.069	-0.049	-0.019
	(0.042)	(0.044)	(0.046)	(0.047)
Observations	23,435	23,430	22,991	20,446
R^2	0.837	0.840	0.842	0.848
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the CESD and ASIF. This table shows estimates of Equation 2 where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variable is log firm energy consumption in Panel A, log firm output in Panel B, and log firm energy efficiency in Panel C. The estimates in this table correspond to a pooled version of the regression displayed in Figure 4. The coefficient in column (4) means that regulated firms decreased energy consumption by 12.8% and output by 14.5%, while no significant energy efficiency improvement after the policy implementation can be found. See Section 2 for additional discussion and Table 2 for more information about the data and variables. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Spillover Effects on Related Firms

A. Output

Variables	ln(Output)			
Related × Post	0.152***	0.147***	0.118***	0.127***
	(0.037)	(0.037)	(0.037)	(0.035)
Observations	18,423	18,420	18,418	17,905
R^2	0.865	0.873	0.881	0.889

B. Placebo Test on Output

Variables	ln(Output)			
Related × Post	-0.026	-0.025	-0.015	-0.003
	(0.040)	(0.039)	(0.039)	(0.038)
Observations	8,923	8,921	8,905	8,730
R^2	0.898	0.903	0.911	0.919

C. Energy Use

Variables	ln(Energy Use)			
Related × Post	0.322***	0.320***	0.302***	0.318***
	(0.075)	(0.073)	(0.076)	(0.094)
Observations	3,759	3,759	3,705	2,823
R^2	0.916	0.919	0.927	0.926

D. Energy Efficiency

Variables	ln(Energy Efficiency)			
Related × Post	-0.077	-0.077	-0.059	-0.087
	(0.078)	(0.077)	(0.080)	(0.099)
Observations	3,724	3,722	3,668	2,801
R^2	0.866	0.870	0.880	0.867
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the ASIF and CESD. This table shows estimates of Equation 2 where Related × Post is an indicator for related firms interacted with an indicator for years after 2006. The estimates in this table correspond to a pooled version of the regression displayed in Figure 5. This table shows that related firms in the same 4-digit industries increased output by 11.8%–15.2% and energy consumption by 30.2%–32.2% after the policy implementation while their energy efficiency and that of related firms outside the same 4-digit industry as the regulated firms (but still in the same 2-digit industry) were not significantly affected by the policy. See Section 3 for additional discussion. The results of robustness checks with additional matching methods are shown in Tables A.15 and A.16. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Industry-Level Spillovers and Within-Conglomerate Effects

A. Industry-Level Spillovers

Variables	ln(Output)			
	All Sample		Energy-Intensive Industries	
Spillover × Post	0.081***	0.073***	0.083***	0.084**
	(0.022)	(0.019)	(0.023)	(0.027)
Observations	2,557,940	2,557,940	843,313	843,313
R^2	0.840	0.856	0.831	0.848
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry-Level Controls	Y	Y	Y	Y
Firm-Level Controls		Y		Y

B. Within-Conglomerate Difference-in-Differences

Variables	ln(Output)			
Treat × Post	-0.343***	-0.350***	-0.367***	-0.315***
	(0.067)	(0.068)	(0.070)	(0.067)
Observations	15,174	15,149	15,146	14,745
R^2	0.530	0.535	0.582	0.626
Treat	Y	Y	Y	Y
Conglomerate × Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the ASIF. Panel A of this table shows estimates of Equation 3 where Spillover × Post is an indicator for industry-level exposure to the Top 1,000 program interacted with an indicator for years after 2006 and the dependent variable is log firm output. Exposure to the Top 1,000 program is defined as the proportion of total energy savings targets of Top 1,000 firms relative to the total energy consumption in 2004 for each industry. The estimates in this table correspond to a pooled version of the regression displayed in Panel A of Figure 6. The results show that the average market-level spillover led to a 7.3%–8.4% increase in the output of unregulated firms. See Section 3 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. Panel B of this table shows the output change of regulated firms relative to that of their related firms (in the same 4-digit industry) within the same conglomerate. Treat × Post is an indicator for regulated firms (Top 1,000 firms) interacted with an indicator for years after 2006, and the dependent variable is log firm output. The estimates in this table correspond to a pooled version of the regression displayed in Panel B of Figure 6 and show that regulated firms experienced a 31.5%–36.7% output decrease relative to the output of their related firms in the same conglomerate. See Section 5.3 for additional discussion. Conglomerate-by-year fixed effects are included, and standard errors clustered at the conglomerate level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Structural Model Parameters

Parameter		Value	Target
1. Fixed Values			
Elasticity of substitution	$\sigma = \frac{1}{1-\rho}$	4.00	Melitz and Redding (2015)
Returns to scale	α	0.90	Burnside, Eichenbaum and Rebelo (1995)
Returns to scale (Labor Share)	α_l	0.80	Cost share of variable inputs
2. Method of Moments			
Efficiency depreciation	δ	0.900	Within-conglomerate distribution
		(0.003)	
Dispersion of ln-ability ϕ	σ_m	1.239	Firm size distribution
		(0.055)	
Survival threshold	ϕ_1	0.609	Share of small firms
		(0.166)	
3. Policy Parameters			
Policy threshold	$\tilde{\phi}$	9.29	Energy share of Top 1,000 firms
Input quota	$1-\xi$	0.20	11th Five-Year Plan

Notes: This table summarizes the parameters that we set or estimate to solve the model. Standard errors are calculated by means of a bootstrapped variance-covariance matrix of data moments. See Section 5.1 for the detailed estimation procedure.

Table 7: Model Decomposition of Difference-in-Differences Estimates

	Allocation Effect	Residual Revenue Effect	Market Effect	Total Effect
A. Effect on Regulated Firms				
Top 1,000 Firms	-0.129	-0.037	0.026	-0.140
Control Firms	0	0.039	0.026	0.065
Difference-in-Differences	-0.129	-0.076	0	-0.205
B. Effect on Related Firms				
Related Firms	0.204	-0.037	0.026	0.193
Control Firms	0	0.039	0.026	0.065
Difference-in-Differences	0.204	-0.076	0	0.128
C. Within-Conglomerate Effect				
Difference-in-Differences	-0.333	0	0	-0.333

Notes: This table reports the decomposition results for difference-in-differences estimates according to the model. Panels A, B, and C in this table correspond to Panel B of Table 3, Panel A of Table 4, and Panel B of Table 5 separately by decomposing the difference-in-differences estimates first into the effects on treated and control firms and then further into allocation effects, residual revenue effects and market effects. See Section 5.3 for additional discussion.

Online Appendix: Not For Publication

This appendix contains multiple additional analyses. Appendix A provides a detailed account of the intended design of the Top 1,000 program and its practical implementation. Appendix B reviews prior research into the effects of the Top 1,000 program. Appendix C describes in more detail the data construction. Appendix D shows that our results are robust to accounting for competing policies in this time period. Appendix E explores other potential margins of substitution for regulated firms. Appendix F shows that our welfare analysis is not confounded by reallocation across regulated firms of different size. Appendix G describes how we can incorporate the health benefits from pollution reduction in our policy analysis. Appendix H shows that the program did not lead to a significant geographic reallocation of economic activity. Appendix I estimates the effects of the Top 10,000 program. Appendix J provides a detailed derivation of our baseline model. Appendix K describes how we model alternative regulations. Appendix L examines the effects of the program when we shut down market- and conglomerate-level spillovers. Finally, Appendix M details the extensions of the model.

A The Top 1,000 Program in Theory and Practice

This appendix provides a detailed account of the Top 1,000 program relying on information from policy documents, audit reports, relevant newspaper accounts, and semistructured interviews with a number of stakeholders. The interviews include the main policy designer, two local government officials charged with implementing the program, and executives from six Top 1,000 firms. We first use this information to describe the intended goals of the program, the initial policy design, and how the program was implemented in practice. We then include detailed accounts of these interviews in the last section of this appendix.

A.1 Policy Goals

A key focus of the 11^{th} Five-Year Plan was to significantly improve energy efficiency. Specifically, the goal of the plan was to reduce the ratio of aggregate energy to GDP by approximately 20% (CCCPC, 2005). The Top 1,000 program was part of this overarching goal. Given the importance of industrial energy use in the Chinese economy, the Top 1,000 program was expected to be a major contributor to achieving this goal and had a specific energy savings target of 100 million tonnes of coal equivalent by 2011 (NDRC, 2006).[76]

---

[76] In addition to the Top 1,000 program, the 11^{th} Five-Year Plan included government support to encourage energy savings for all firms. Zhou, Levine and Price (2010), Price et al. (2011), and Zhao et al. (2014) discuss other forms of government support.

The government's motivation for the Top 1,000 program is explained by the program's designers in a report prepared for the Chinese government (Wang et al., 2008). The report mentions three issues: (1) the rapid development of the heavy manufacturing industry, (2) the potential for energy shortages to become a key bottleneck for economic development, and (3) the fact that the energy efficiency of the Chinese economy lagged that of developed economies.[77] According to these official documents, pollution reduction and the reallocation of economic activity across different regions were not primary goals of the program.[78] While Wang et al. (2008) briefly mention environmental benefits from the program, the focus is on the potential benefits from reducing energy use.

Comparison with Other Policies

This subsection compares the Top 1,000 program to other policies that target industrial production by incentivizing improvements in energy efficiency. We use the International Energy Agency's database of policies to survey 547 policies that focus on the energy efficiency of the industrial sector (IEA, 2023). Of these policies, 439 target firms, and 58 of them have clear numerical objectives. The main objectives of these policies are to improve energy efficiency, decrease energy consumption, or reduce greenhouse gas emissions. Most of the policies offer firms incentives such as subsidies, funding, low-interest loans, tax deductions, training, free energy audits, and technological assistance. However, only 35 policies have penalties for noncompliance, including penalty fees or charges, more stringent permitting, or decreased energy supply.

Policies vary in their described goals, with some policies focusing on absolute reductions in carbon emissions while others mention percentage reductions. The policies also vary in length. To make the policies comparable, we calculate the implied yearly reduction in carbon dioxide emissions required to achieve each policy's objective. In the case of the Top 1,000 program, the goal was to reduce energy use by 100 mtce, which is equivalent to a reduction of 240.38 million tonnes in carbon dioxide emissions. Since the program lasted 5 years, the implied average yearly reduction would be 48.08 million tonnes of carbon emissions. Below, we compare the Top 1,000

---

[77] Specifically, Section 3.1 of Wang et al. (2008) notes that:

> Due to extensive economic growth, energy shortages have become a major bottleneck for further economic development and environmental issues. The reliance of economic growth on heavy manufacturing industries makes China's energy intensity much higher than that of developed countries, even higher than that of some developing countries. Furthermore, China is still facing heavy industrialization, accelerated urbanization, consumption upgrading and the relocation of international manufacturing activity. Energy demand in China will therefore continue to grow for a long period. If no effective measures are taken to reduce energy consumption, China's economic growth will become unsustainable.

[78] We also document empirically that the program did not lead to disproportionate production increases in more polluted or more populated areas (see Table A.28) and that it did not lead to significant reallocation of economic activity across different geographic regions (see Table A.29).

program to some of the largest policies in this database by standardizing their objectives using similar calculations:

1. "Turning the Corner: An Action Plan to Reduce Greenhouse Gases and Air Pollution" from Canada (2007–2020) implies a total reduction of 150 million tonnes over 14 years, for an average yearly reduction of 10.7 million tonnes.
2. The "2030 Climate and Energy Framework" from the European Union (2014–2030) implies a carbon emission reduction of 859.44 million tonnes over 17 years.[79] The average yearly reduction for this program is 50.56 million tonnes.
3. The "Law on Energy Transition for Green Growth (LTECV)" from France (2015–present) implies a carbon emission reduction of 135.71 million tonnes over 16 years.[80] The average yearly reduction for this program is 8.48 million tonnes.
4. The "Agreement with Industry on CO_2 Emissions Cuts" from Germany (2000–2012) implied a carbon emission reduction of 224.27 million tonnes over 13 years.[81] The average yearly reduction for this program is 17.25 million tonnes.
5. The "Climate VISION (Voluntary Innovative Sector Initiatives: Opportunities Now)" from the United States (2003–2012) implied a carbon emission reduction of 500 million tonnes over 10 years for an average yearly reduction of 50 million tonnes.

Comparing the Top 1,000 program to these programs above, we find that this single program in China is comparable in its objectives to the overall goals of the EU and the United States and is larger than programs in other leading economies, including Canada, France, and Germany. Importantly, the actual impact of these programs will depend on both the strictness of their implementation and the magnitude of leakage in the economy.

A.2 Intended Program Implementation

As described by Price, Wang and Yun (2010), the Top 1,000 program was designed following examples of "voluntary agreement" programs in other countries. As originally designed, the program first divided the national energy saving target across provinces based on the initial share of energy use of Top 1,000 firms. Each local government was then tasked with signing an

---

[79] The program's objectives are to reduce the 2014 level of emissions of 3537.80 million tonnes to 60% of the 1990 level of 4463.94 million tonnes, which implies a reduction of 859.44 million tonnes.

[80] The program's objectives are to reduce the 2015 level of emissions of 433.22 million tonnes to 60% of the 1990 level of 495.85 million tonnes, which implies a reduction of 135.71 million tonnes.

[81] The program's objectives are to reduce the 2000 level of emissions of 957.53 million tonnes to 65% of the 1990 level of 1128.09 million tonnes, which implies a reduction of 224.27 million tonnes.

agreement with each Top 1,000 firm committing the firms to achieving specific energy savings targets.

The policy relied on a conceptual measure of energy savings that the government had previously described in an official document in 1991 (CSBTS, 1991). The historical government guidance allowed five different formulas for energy saving, which could be based on production quantity, output value, product mix, changes in the types of technologies used in production, or changes in the types of energy utilized.

The use of these metrics was unreliable for a number of reasons. First, local governments did not have the ability to monitor or audit inputs into these formulas, including data on product mix, benchmark energy use for a given product and production technique, or actual energy consumption per unit of production. Second, firms could choose which of the five calculations to use and could also change which metric they were evaluated on in any given year. In addition, each of these calculations could be computed based on a fixed base year or on a cumulative basis, which further allowed firms flexibility in choosing the evaluation metric that would show the greatest energy savings.[82]

Given that provincial governments could not monitor the inputs into these energy saving formulas, the initial design made it hard for local governments to monitor firms or to enforce specific goals. For these reasons, as we document in our narrative and interview evidence below, the original energy saving formulas in CSBTS (1991) were not used by local governments to assess firm performance.[83]

A.3 Practical Implementation

Despite the drawbacks in the initial program design, provincial governments were still liable to the central government for achieving their overall energy savings targets. To achieve these targets, local governments instead regulated Top 1,000 firms' total energy consumption, which they could both monitor and credibly enforce.

From regional policy documents such as those of the People's Government of Hainan Province (2006) and the People's Government of Chongqing City (2007) and the interviews with Top 1,000 firms and government officials described in Appendix A.4, we can confirm that each Top 1,000 firm was given an energy consumption quota. To design these quotas, local governments relied on three factors: (1) the initial energy use of a given firm, (2) the overall energy saving target, and (3) the expected growth of each regulated firm's industry.

Relying on energy use quotas was desirable for local governments for three reasons. First, the

---

[82] Zhao et al. (2016) discuss the limitations of the metrics used to measure energy savings.

[83] For this reason, the output of these formulas measuring firms' "energy savings" in government reports may not be used to directly assess firm compliance with the program.

quotas allowed local governments to map their assigned energy saving objective to specific tasks for each firm. Second, it was possible for local governments to monitor each firm's energy use. Local governments could monitor firms by relying on the "live energy consumption monitoring system" in the case of electricity use and could also use shipments of coal or other primary fuels to track other energy sources. Finally, as recounted by both firms and local officials, local governments had the capacity to shut down production at regulated firms if they exceeded their energy use quotas.

Our interviews with firms and local officials also reveal how these quotas were determined. An interesting feature of the energy use targets is that they recognized that energy use would depend on expected industry growth. That is, the benchmark for the quotas took into account the potential growth in energy use in the absence of the policy. For our empirical analysis, this feature confirms that, if we abstract from spillover effects, the effectiveness of the program can be appropriately evaluated relative to a set of unregulated firms in the same industry, as in our difference-in-differences analysis.

A.4 Narrative and Interview Evidence

To understand the nature of the Top 1,000 Energy Conservation Program and to obtain more details on the policy implementation, we interviewed a number of stakeholders including the policy designer, local government officials in charge of implementation in two different cities, and executives of six Top 1,000 firms. For confidentiality reasons, we refer to interviewees by their location and organization.

We complement these sources with information from the official report of the National Audit Office (the audit office is responsible for energy audits of Top 1,000 firms) and with newspaper interviews of local government officials and Top 1,000 firm executives published during the policy period. Figure A.17 displays our interview list and how each stakeholder is connected to the program.

A.4.1 Interview with Policy Designer

The designer of the Top 1,000 Energy Conservation Program—number (1) in Figure A.17—wrote many reports and books about the energy saving target allocation and assessment of the Top 1,000 Program during the 11^{th} Five-Year Plan (see, e.g., Wang et al., 2008; Price et al., 2008; Price, Wang and Yun, 2010).

Below, we include extracts from our conversation with the policy designer. We first asked how the targets for the Top 1,000 program were determined. He explained the following:

> The central government first set a total goal of saving 100 million tce for the Top

> 1,000 program and then allocated this target to each province according to the energy consumption share of all Top 1,000 firms in that province. After that, provincial governments coordinated with each Top 1,000 firm within their jurisdiction to set the firm-specific target. At this stage, the central government did not interfere with the provincial government's decisions any more, but provincial governments usually used an allocation rule similar to the central government's. They allocated their local targets based on each Top 1,000 firm's original energy consumption, while industry energy saving potential and development would also be taken into account.

We then asked how the policy was implemented. He replied:

> After the target is set, Top 1,000 firms need to carry out energy conservation audits and make energy conservation plans. The plan should include energy saving measures the firm would take and the energy saving goals the firm would achieve. Moreover, Top 1,000 firms should submit energy utilization reports to the provincial NDRC regularly, including their energy saving performance. Then, the provincial NDRC is responsible for tracking, guidance and supervision of the Top 1,000 firms' energy saving behavior.

Finally, we asked about the role of the firm-level calculations on energy saving. He noted:

> The calculation of the amount of energy saved (AES) is not my job, so I don't know much about the details. I remember there is an official document talking about that, the GB/T 13234-1991 [see CSBTS (1991)]; you can refer to that. But as far as I know, the AES is not directly related to the firm's total energy consumption, which the provincial government really monitored and assessed in practice.

This interview clarified two important facets of the program's implementation. First, the energy saving target was decentralized from the central to local governments. Second, while policy documents describe the concept of the amount of energy saved (AES) (CSBTS, 1991), it does not appear that this concept was used by local governments to implement the program. To understand which metrics were used to implement the program in practice, we interviewed local officials in charge of the implementation.

A.4.2 Interviews with Government Officials

The first official whom we interviewed works in the Department of Energy Conservation and Environmental Protection (ECEP) of the National Development Reform Commission in Municipality A (NDRC, listed as number (2) in Figure A.17). Municipality A has an administrative status similar to that of a provincial government. The ECEP is responsible for implementing and assessing the energy saving policy in Municipality A.

We first asked this official how the local government set energy consumption targets. This official told us:

> We set both the energy consumption quotas and energy efficiency targets for Top 1,000 firms during the 11^{th} Five-Year Plan.
>
> The total 100 million tce energy saving target for the Top 1,000 firms was first allocated to us [the provincial government], and then we coordinated with each Top 1,000 firm to set its targets. The energy saving target mentioned in the official document is the amount of energy saved (AES). However, it was impossible for us to monitor the efficiency improvement of every product of a firm at that time. Because monitoring the AES of a firm is not feasible in practice, we instead use a more transparent method: we monitor the firm's total energy consumption.
>
> We gave them an incremental energy consumption quota based on their energy consumption before the policy, their potential growth and our total target [the provincial target]. You can consider their quota as their expected energy consumption during the policy period minus their current energy consumption, and further minus the energy saving target that they are assigned. And that's the total amount by which [their energy use] could increase during the 11^{th} Five-Year Plan.

We also asked this official how the energy targets were benchmarked and monitored. He explained:

> As to the energy efficiency target, we take the average industry improvement into account, and we also consider the firms in the sector with the most advanced energy efficiency as the benchmark. So, in theory, the Top 1,000 firms could only achieve the energy efficiency target if they improved faster than other firms in the same industry, which means that no Top 1,000 firms could easily achieve the energy efficiency target.
>
> But during the policy implementation, we really could not verify those large firms' energy efficiency accurately and in a timely way since they produce too many kinds of goods and their production process is so complicated. We cannot be experts for every industry. So, in the daily supervision, we monitor their energy consumption. We have a "live energy consumption monitoring system," which is especially timely and accurate on firm's electricity consumption. With this system, we can monitor each firm's energy consumption and intervene if needed. If the firm ultimately uses

> less energy than the incremental quota that we set, then we assume that the firm has achieved its energy saving target.

The second official whom we interviewed is the director of the Economic and Technological Development Zone in City B (listed as number (3) in Figure A.17). The Economic and Technological Development Zone in City B has an administrative status below that of a province and is closer to a county. There are five Top 1,000 firms in City B. The director whom we interviewed took office in 2019. While he was not in office during the 11^{th} Five-Year Plan, he is familiar with the current energy consumption quota, which he mentioned has been in use for a long time.

We first asked this official how the county and provincial government cooperated in the implementation of the Top 1,000 program. He mentioned the following:

> We [the Economic and Technological Development Zone in City B] receive an energy consumption quota every year. The quota is allocated from top to bottom: from the central government to provincial governments, from provincial governments to municipal governments, and from municipal governments to county governments. It is our [the county government's] job to monitor firms' energy consumption within our jurisdiction. The central government cannot directly send someone to supervise a firm—it's all local work. Although our development zone is not a real county, it is treated as a county in administration. The municipal government decides our quota. They won't coordinate with us, but they will consider our contribution, such as GDP and fiscal revenue. The more we contribute, the larger quota that we will receive.

We then asked this official how the energy consumption target is enforced. He noted:

> To complete our total energy consumption target, we set quotas for each energy-intensive firm within our jurisdiction, including both electricity and coal quotas. As to the energy efficiency you mentioned, we cannot monitor it directly. You need to be more realistic. Firms that become more energy efficient will never tell us that they can use less energy. All firms want as much energy as they can get. So the only thing we can do is to control their total energy consumption, and we really have power to do that. We can cut their electricity if necessary.
>
> Here is a real case for your reference. After we first received our energy consumption quota, we coordinated with the energy-intensive firms immediately. But no firms listened to us. They continued to overuse energy. So one day we cut their electricity without telling them. After that, all firms followed our instructions immediately, because it is very dangerous to stop those high-temperature high-pressure large equipment when it is operating, and they got a huge amount of scrap instantly.

> Some of the products were already in the last step of production but were all gone in a flash. From then on, all firms complied with our energy consumption limit.

Finally, we asked the official how each individual firm received their energy consumption quota. He explained:

> As to the quota allocation, we allocate our quota fairly based on each firm's energy consumption. We also consider the energy growth rate in different industries; for example, some industries are promising, while others are declining, but we usually ignore the firm difference. If we allocate the quota unfairly within an industry, firms will blow up. They will say that we took money from the other firms in exchange for a larger quota. So, to avoid trouble, we don't consider firm differences during the quota allocation process.

From our interviews with local officials, we conclude the following. First, for feasibility reasons, local governments relied on measures of energy consumption, not energy saved. Second, local officials had the capacity to strictly enforce the energy consumption targets. Finally, the energy use targets were set taking into account expected industry growth.

A.4.3 Report of the National Audit Office of the People's Republic of China

We relied on an audit report from the National Audit Office in Shanghai (Xu, 2010, (4) in Figure A.17) to corroborate the interviews with local officials.

First, the auditor's report confirms that it was not feasible to rely on energy saving formulas from CSBTS (1991) to implement the Top 1,000 program. Xu (2010) notes the following:

> The energy saving amount varies with calculation methods. It is acceptable that different industries use different calculation methods since they produce different goods, but firms in the same industries or even different years of a same firm use different calculation methods. This attracts the attention of auditors. For example, a cement firm calculated their energy saving amount with different methods in 2007, 2008 and 2009 separately. In one year, they adjust for inflation of output prices, while in others they didn't. Similarly, in some years they calculated the aggregate output of the clinker and cement products, while in other years they calculated them separately. Changing calculation methods makes the energy saving amount not comparable.

Second, the auditor's report notes that local government officials instead focused on the total energy consumption of Top 1,000 firms. The report notes:

> When local governments signed voluntary agreement with Top 1,000 firms, their only focus is the total energy consumption, and the National Economic Council (NEC) and the National Development Reform Commission (NDRC) didn't correct them either. Local governments divide the total energy saving amount by 5 to get the target for each year. After firms invest in energy saving equipment, their energy saving amount may change. However, during assessment, the regulators do not take these changes into account, they just use the predetermined total energy consumption quota as the only assessment criterion, making the original assessment rule (the AES target) useless.

A.4.4 Interviews with Top 1,000 Firms

We asked the same set of questions to each of the Top 1,000 firms. These questions were as follows:

1. What kinds of energy savings targets did your firm receive in the Top 1,000 Energy Conservation Program?
2. Do you know how your energy saving target was set? Did you have bargaining power?
3. How did government assess your energy saving performance?
4. Was your production affected by the program?
5. Did you make energy-saving investments during the 11^{th} Five-Year Plan? What kinds of investments did you make?

Below, we summarize how executives from each of these firms answered these questions.

Interview (5) with a Top 1,000 Cement Company in Zhejiang Province

The interview was conducted with the administrative staff member who is responsible for energy saving in the company. He noted:

> Our company received both a total energy consumption quota and energy consumption per unit product requirement. The total energy consumption quota is determined by the local government; we do not know where this target comes from, and we don't have any bargaining power. The benchmark of the energy consumption per unit product is usually the industry advanced level.
>
> In the assessment, we are also evaluated by both total energy consumption and energy consumption per unit product. But the latter is not easy to show since we have multiple products.

> Our company is a high-tech firm, so we could make some energy-saving investments. But even with these technique improvements, we still could not produce and sell at full capacity during the 11^{th} Five-Year Plan. For a cement firm, the reasonable payback period of investment is 3–5 years, and operation can usually begin after one year of the investment, but it also depends on different projects.

Interview (6) with a Top 1,000 Paper Company in Hunan Province

This interview was conducted with the company manager. He noted:

> We signed an energy saving agreement with the local government during the 11^{th} Five-Year Plan. The agreement included both a total energy consumption limit and an energy efficiency improvement requirement. We don't know how these two targets are set. The local government won't coordinate with us about our target.
>
> While the efficiency target is easy for us, the total energy consumption control is the hard part. The local government cannot determine the energy efficiency of various products in our company, so they focus more on energy consumption. They have a "live energy consumption monitoring system," which can monitor our real-time energy consumption, especially electricity.
>
> We face severe punishment if we exceed the total energy consumption quota. Our firm will be forced to stop production. That's fatal for business.
>
> A merger won't change anything as long as the energy consumption quota is set. You just merge, you cannot move. As long as the firm is still in the same province, its energy consumption target is the same, and the administrative office in charge won't change, either.

Interview (7) with a Top 1,000 Textile Company in Shandong Province

This interview was conducted with the administrative staff member who is responsible for energy saving in the company. This administrative staff member joined the company in 2014, so he does not know much about the energy conservation program during 11^{th} Five-Year Plan. His responses were based on the firm's energy saving in recent years. He noted:

> Our company receives a total energy consumption quota each year. The quota is determined by the local government according to our energy consumption and potential growth, but we don't know whether there is a formula to calculate this quota; the local government just tells us its final decision. I have little impression about the energy efficiency target. We have a wide variety of products, so energy efficiency is hard to measure even for ourselves.

> In terms of assessment, we have a self-report system, and the local government sends a spot check team 1–2 times a year, mainly to monitor our energy consumption.
>
> Our production is affected a lot by the energy consumption control, but in some sense, it may also be an incentive for the firm [wry smile]. We are asked to submit an energy saving plan each year, so we have to make some investment.

Interview (8) with a Top 1,000 Cement Company in Shandong Province

This interview was conducted with the administrative staff member who is responsible for energy saving in the company. He noted:

> I was not in charge of energy saving during the 11^{th} Five-Year Plan, but the energy saving policy during the 11^{th} Five-Year Plan is quite similar to the present one. It's always the same trick.
>
> We receive a total energy consumption limit each year, including both a coal consumption limit and an electricity consumption limit. The local government just suddenly informs us to comply with a certain energy consumption limit, and there is no place to appeal. We also receive a requirement on the energy consumption per ton of cement.
>
> To reduce costs, we do have incentives to reduce energy consumption ourselves. We are also asked to attend some energy saving conferences and update our equipment.
>
> In terms of production, we are highly affected by the energy consumption limit. In recent years, we stopped production for about 161 days a year, although excess capacity was also part of the reason.

Interview (9) with a Top 1,000 Chemical Company in Shandong Province

This interview was conducted with the chemical engineer in the company. The chemical engineer whom we interviewed is familiar with the energy conservation policy during the 12th Five-Year Plan. (The target under the "Top 10,000 Energy Conservation Program" during the 12th Five-Year Plan is very similar to that under the "Top 1,000 Energy Conservation Program" during the 11^{th} Five-Year Plan, asking energy-intensive firms to improve their energy efficiency and save a total of 250 mtce.) He noted:

> Our company got both a total energy consumption target and an energy efficiency target. However, in practice, the government assessment mainly focuses on energy consumption because efficiency compliance is actually difficult to test. The efficiency indicator we got is measured by energy consumption per output. However, the output value may fluctuate with the price cycle. And when we divide the production process

> into different units, we can double-count the output value, which makes it look like efficiency has improved a lot. So the efficiency indicator can be largely manipulated and cannot represent our true energy saving performance. In comparison, the energy consumption is easier for government to monitor and also more binding for us. Not only is our firm development affected by the energy consumption control—sometimes our production even has to be limited.
>
> We did make some improvements in energy saving to meet the target, but the feasibility varies with the payback period. For a project with a payback period of less than one year, we will do it; for a project with payback period of between 1 and 3 years, we need to consider; but for a project with a payback period over 3 years, we rarely do it. The limit on energy-saving investment not only comes from financing problems but also technical problems including lacking enough physical space or significant production disruptions.

Interview (10) with a Top 1,000 Iron & Steel Company in Zhejiang Province

The last firm interview was conducted with the manager of the company. He showed us the firm's record of its energy saving plan for the 12^{th} Five-Year Plan. The plan mentioned three types of energy saving:

> First, Managed Energy Saving, which included:
>
> - Implement a live report system on firm energy consumption and control the total energy consumption. Change the firm energy consumption from "calculate after use" to "calculate to use."
> - Use "early warning system" on firm's energy consumption. Track the energy consumption of each energy-intensive production units. Warn and regulate the projects or production units that increase their energy consumption fast. Never consume energy without limit any more.
> - Furthermore, allocate the total energy consumption quota to each production units, implement target responsibility system, and improve supervision.
>
> Second, Structural Energy Saving. Continue improving the energy consumption structure. Study the optimal input structure of iron ore, coking coal, sintering materials and steel craps, as well as the optimal product structure of pig iron, bessemer steel and electric steel, etc. Continue increasing industrial output per unit of energy consumption. Make more efforts in high value-added industries.
>
> Third, Technical Energy Saving. Upgrade the industrial furnace, optimize combustion control, reduce the surface heat loss, apply digital and regenerative heating

> technology, increase the temperature of hot charging and hot delivery, and improve the thermal efficiency of industrial furnaces. Make full use of waste heat generated in the production process of iron and steel, improve energy efficiency and self-generated electricity.

A.4.5 Interviews Reported in News Media

As a second way to corroborate the information gathered in our interviews, we researched contemporary news articles that cited interviews with Top 1,000 firms and local government officials. These interviews confirm (1) that local governments relied on measures of energy consumption to implement the policy and (2) that local government officials had the capacity to stop production at Top 1,000 firms to ensure that energy use targets were met.

(11) Director and Deputy Director of Resource Conservation and Environmental Protection Department of the Tangshan Development and Reform Commission

Li (2010), writing for *China Business*, interviewed Li Dazheng, director of the Resource Conservation and Environmental Protection Department of the Tangshan Development and Reform Commission. Li Dazheng is quoted as saying:

> To complete the energy saving target of the 11^{th} Five Year Plan is our first priority, and limiting production is the only way. But to stop a steel mill's blast furnace is not easy. There are molten steel in the blast furnace, so if you want to stop, you need to inform the firm four or five days in advance. And then they can have time to clear the raw materials and molten steel in the blast furnace. Finally, on the day of shutdown, municipal leaders, county leaders and the firm manager will come together and close the fan, feed port and discharge port slowly. If you just cut the electricity, the molten steel in the blast furnace will turn into an iron lump and the blast furnace cannot be used any more.

Li (2010) also interviewed Lu Yanan, deputy director of the Resource Conservation and Environmental Protection Department of the Tangshan Development and Reform Commission. Lu Yanan is quoted as saying:

> Limiting the production and energy consumption of iron and steel and coking industry is the most efficient way to complete the energy saving target. Tangshan government has dispatched 92 officials from various relevant departments to form 19 inspection groups to carry out on-site supervision on counties and key firms from October. All groups are led by deputy county-level leaders. They live in these firms.

(12) Official from Wu'an Local Government

Li and Li (2010), writing for *China Securities Journal*, cite a Wu'an government official who said that:

> Energy saving is managed locally, and the firms' energy saving assessment is done by county government. There are some county governments who cut the electricity of the energy-intensive firms in their jurisdiction to meet the county energy consumption quota during the 11th Five Year Plan. However, the municipal government sometimes doesn't even know about it.

(13) China Aluminum Group Co., Ltd Guizhou Branch

Writing for *China Nonferrous Metal News*, Wang (2008) describes the experience of China Aluminum Group Co. in dealing with power cuts related to the enforcement of the Top 1,000 program:

> Since mid-December in 2007, China Aluminum Group Co., Ltd Guizhou Branch suffered an unprecedented power limit. The electrolytic aluminum plant was forced to shut down, and the electrolytic production was seriously affected. By the end of December, 300 electrolytic tanks have been stopped, and production capacity plummeted by 1/3. The firm suffered from heavy losses.
>
> Abnormal stop will seriously affect the life span of the tank, and when the blackout is over, the restart cost is huge. Every time a tank stops, it is necessary to move 20 to 30 tons of electrolyte and other materials out of the tank manually, which means high labor cost. Moreover, now the firm has stopped all production lines of different products and focus on general aluminum production to digest the backlog primary aluminum in the stopped tank. At the same time, clean the tank and the furnace for restart.

(14) Jiangsu Sha Steel Group Co., Ltd.

In reporting for the *21st Century Business Herald*, Wang (2010) interviewed Shen Wenming, the executive director of the Top 1,000 firm Jiangsu Sha Steel Group Co., Ltd., who described how, due to the requirements of the 11th Five-Year Plan, the production of Jiangsu Sha Steel Group Co., Ltd., would be limited in September and October. "Power cut always affects, but we haven't decided whether to stop the rolling part or the making part. It depends on the market condition. So, for now, we do not know the exact impact on production."

(15) Jiangsu Yong Steel Group Co., Ltd.

Wang (2010) also mentions that the "electricity quota of Jiangsu Yong Steel Group Co., Ltd (Top 1,000) is 10 million watt in September, and the expected production loss of line screw is 70,000-80,000 tons."

(16) Nanjing Iron & Steel Group Co., Ltd.

Wang (2010) also notes that Nanjing Iron & Steel Group Co., Ltd, was forced to stop production for 15 days a month due to the energy saving requirement: "The affected products are mainly high quality steel, such as bearing steel and spring steel, with a production loss of 50,000 tons."

(17) Jiaozuo Wanfang Aluminum Co., Ltd.

Writing in *China Business News*, Cao (2010) cites a news release from Jiaozuo Wanfang Aluminum Co., Ltd.:

> Jiaozuo Wanfang Aluminum Co., Ltd. announced today that in response to the power supply quota, the firm will reduce electrolytic aluminum capacity of about 140,000 tons from today, accounting for one-third of the total electrolytic aluminum capacity of the firm. This means a reduction of 30 thousand ton of aluminum, accounting for 7.1% of the annual production plan. Yesterday, the stock price of Jiaozuo Wanfang Aluminum Co., Ltd decreased by 7.64%.

(18) Guizhou Kailin Group Co., Ltd.

Writing for *China Chemical Industry News*, Yu (2011) interviewed Tong Xianglong, deputy director of public relations of Guizhou Kailin Group Co., Ltd, who said that:

> It's not appropriate to ask firms to reduce electricity consumption to complete the energy saving target. It cure the symptoms, not the disease. It is also against the central government's original intention on scientific development. Because when the limit is over, the energy intensity of these industrial firms doesn't really decline, it's just the same as before. The right thing to do is to improve energy efficiency through innovation.

(19) Jinfeng Coal Chemical Co., Ltd.

Yu (2011) also interviewed the manager of Jinfeng Coal Chemical Co., Ltd, who said that:

> The government wants to save energy, but it actually causes the depletion of equipment and the waste of resources. If you really want to save energy, you should assess the energy efficiency. Like for fertilizer firm, if our energy efficiency of urea production reaches the efficiency target, you should allow us to produce. But now, some firms have already made a huge amount of investments to improve their technique and equipment, and their energy efficiency is already leading, but they still face electricity limit. It's not fair.

B Prior Research on the Top 1,000 Program

This section summarizes prior work studying the Top 1,000 Energy Conservation Program. We discuss related papers by focusing on four aspects of the program: (1) administrative details of the program, (2) how local governments help Top 1,000 firms achieve the energy savings targets, (3) verification of the official reported energy saving performance, and (4) the economic effects of the program.

B.1 Policy Details of the Top 1,000 Program

The Top 1,000 program marked the first time that the Chinese government directly imposed an energy saving regulation on firms. Researchers have discussed the policy details of the program based on official documents, including the rule for selecting firms into the program, how targets were set, and the functioning of government assessment and firm compliance (e.g., Price, Wang and Yun, 2010; Yang et al., 2015; Zhu et al., 2018; Ma and Liang, 2018). In Appendix A, we provide a detailed discussion of the official program details and the practical on-the-ground aspects of the program's implementation.

Role of Local Governments in Achieving the Energy Savings Targets

Starting with the 11^{th} Five-Year Plan, local government officials had strong incentives to meet energy savings targets in their own jurisdiction. Local officials could incentivize firms to comply with the program in several ways. Based on 53 interviews with firms in Shanxi Province, Kostka and Hobbs (2012) document three ways in which local government can ensure that firms do not exceed their energy quota: (1) cutting off electricity and water, (2) forcing firms to close for several months (a practice known as "sleeping management"), and (3) implementing regional investment restrictions that limit the approval of all new energy-intensive projects. Crucially, these compliance strategies all target the total energy consumption of energy-intensive firms instead of their energy efficiency.

Evaluation of Energy Savings Measures

As we discuss in Appendix A, the government did not specify the exact procedure for computing measures of energy savings. Two sets of researchers have studied how firms computed measures of energy savings and how these related to the stated targets.

Zhao et al. (2016) survey 6 Top 1,000 firms and 4 other energy-intensive firms to understand how they calculated their reported energy savings. Across these 10 firms, 4 different calculation methods were used. This result shows that it is hard to replicate the energy savings reported to the government, which, as Zhao et al. (2016) note, are likely an overestimation. Zhao et al.

(2016) also note that local governments are not equipped with sufficient capacity or resources to verify this firm energy saving amount.

Karplus, Shen and Zhang (2020) analyze data on reported energy savings and document stark "bunching" patterns, with many firms reporting energy savings immediately above the target level. Similarly to Zhao et al. (2016), they suggest that this is evidence that Top 1,000 firms deliberately exaggerate their energy saving performance.[84]

Both of these studies corroborate that firms had significant flexibility in how they reported energy savings, consistent with the fact that the local governments that were responsible for implementing the program were unable to verify the calculations in energy saving formulas. As we discuss in Appendix A, these facts led local governments to instead rely on energy consumption in implementing the program.

B.2 Economic Effects of the Top 1,000 Program

A number of papers study the effects of the Top 1,000 program on energy efficiency and other measures of firm performance. We first survey the results of two studies that use time-series data at the aggregate and industry levels to assess the effects of the program on energy efficiency. We then discuss papers that use microdata to study the effects of the program on firm-level measures of exporting, patenting, and productivity.

In the absence of firm-level data on energy use and energy efficiency, two sets of researchers have relied on time-series data to assess the effects of the program. Wang et al. (2017) use data on industry-level energy consumption and output for the iron and steel sector to assess the trend in energy efficiency during this period. Using these data, they compute two measures of energy efficiency based on either value added or production output. They argue that energy per unit of value added increased faster than energy per unit of production, and they attribute this difference to structural improvements in production techniques.

In a second attempt at evaluating the effects of the Top 1,000 program, Ke et al. (2012) use time-series data on energy consumption and value added of the Chinese manufacturing sector to study the effects of the Top 1,000 program. They assume that 47% of the energy savings is achieved by Top 1,000 firms. Given the performance of the overall sector, they argue that if 47% of the energy savings were indeed achieved by Top 1,000 firms, this would be consistent with the government's stated goal for the program. As with Wang et al. (2017), since Ke et al. (2012) rely only on time-series data, it is not possible to separate the effects of the program from underlying

---

[84] In Table A.3, we document compliance with the Top 1,000 program using official statistics. Given the results from prior work, these numbers are best interpreted as reported rather than real compliance. In Table A.14, we explore whether reported noncompliance leads to heterogeneity in the effects of the policy on regulated firms. The point estimates indicate that noncompliant firms did not experience reductions in energy use, although these estimates are imprecisely estimated given the small number of noncompliant firms in the sample.

trends.

Relative to Wang et al. (2017) and Ke et al. (2012), our paper contributes to this area by relying on firm-level microdata. In contrast to the data in these papers, the microdata allow us to evaluate the performance of Top 1,000 firms relative to that of a control group not directly affected by the program. This allows us to separate the overall trends in energy use and energy efficiency from the effects of the program.

We now discuss four papers that evaluate the effects of the program on measures of firm performance. All four papers focus on the Porter hypothesis, which posits that firms may respond to regulation through innovation that leads to improvements in firm performance. To evaluate this hypothesis, Shen, Zeng and Qu (2015) study the impact of the Top 1,000 program on firms' export volume with ASIF data from between 2005 and 2007 using a difference-in-differences design. They estimate a positive treatment effect of 10.3% on the exporting volume of 240 Top 1,000 firms relative to that of unregulated firms in the same industry. Zhang and Huang (2019) also evaluate the Porter hypothesis using patent application data. They use data for 122 Top 1,000 firms from between 2004 and 2007 and argue that the number of patent applications of Top 1,000 firms increased by 60% after the policy was implemented in 2006. Finally, Filippini et al. (2020) use ASIF data from 2003–2008 on iron and steel firms and find a positive effect of the Top 1,000 program on estimates of firm productivity. In contrast, Ai, Hu and Li (2021) use the same data and time window but focus on chemical firms and find that the program decreased firm annual TFP growth. While Ai, Hu and Li (2021) attempt to evaluate the effects of the program on energy consumption, the lack of firm-level energy data leads them to rely on province-level data. They show that provinces where Top 1,000 firms accounted for a larger share of output in the chemical industry also experienced relative decreases in energy consumption. Since the circulation of our paper, other researchers have adopted our research strategy and replicated our results on the effects of the regulated firms (Xiao, Yin and Moon, 2023), but not on those of related firms.

Overall, the literature investigating the effects of the Top 1,000 program mostly relies on aggregate data or focuses on assessing the Porter hypothesis. Relative to these papers, we focus on the role of conglomerate- and market-level spillovers. Using various datasets covering longer time periods, this paper also evaluates the firm-level effects of the Top 1,000 program on energy use and output. While we do not focus on the Porter hypothesis, we show in Figure A.8 that we do not find significant increases in energy-related patents for Top 1,000 firms.

C Data Quality and Data Merge

We rely on firm-level data from the CESD to evaluate the effects of the Top 1,000 program. As we discuss in Section 1.2, these data are not used by program administrators to implement the Top 1,000 program. For this reason, firms do not have a direct incentive to manipulate the data.

To allay other potential concerns about data quality, we interviewed a government official who works in the Ecology and Environment Bureau. This official explained the government's efforts—based on the Environmental Protection Law of the People's Republic of China (NPC, 1989)—to ensure that the data that it collects is reliable, including through audits and penalties. The official explained that the government takes several steps to ensure that data are reported accurately, including conducting annual comprehensive firm inspections and spot checks, using an automatic pollution monitoring system for real-time monitoring, and auditing pollution emission permits. The government can also enact the following punitive measures for polluting firms that falsify data:

1. Impose economic penalties, including fines, as a deterrent.
2. Suspend or revoke firms' pollutant discharge permit to limit their discharge behavior.
3. Implement mandatory pollution reduction measures through forced closures or through upgrading of pollution control facilities.
4. In cases of serious violations, refer firms to the public security department for criminal prosecution.

In terms of implementation, the official explained that the environmental protection department has a number of tools at its disposal if it discovers that a firm has falsified data, including:

1. Instructing the firm to immediately stop falsifying data.
2. Requiring the firm to remonitor and report accurate data.
3. Punishing the firm for falsifying data through fines or revocation of relevant licenses.
4. In cases where the firm has not taken corrective measures, take administrative enforcement measures to monitor the firm and collect data directly to ensure the accuracy of environmental supervision.

The evidence from this interview is supportive of using data from the CESD for the analysis of the effects of the Top 1,000 program both because firms' program compliance is not evaluated on the basis of these data and because the government has procedures in place to ensure the quality of the data.

Details of the Data Merge

Table A.2 shows the results of our data construction. We merge the lists of regulated firms using both the firm name and unique legal identifier. Since the Top 1,000 and Top 10,000 firms are all large firms, the match rate with the ASIF is very high. We match over 99% of Top 1,000 firms and over 97% of Top 10,000 firms. We also have a fairly good match rate with the CESD, where we match nearly 80% of Top 1,000 and over 70% of Top 10,000 firms. Overall, our combined datasets capture the majority of the economic activity in Top 1,000 and Top 10,000 firms.

D Robustness of Effects on Regulated Firms to Controlling for Competing Policies

This appendix explores the robustness of our results to accounting for competing polices. Specifically, in the same time period when the Top 1,000 program was implemented, the Chinese government also adopted the National Specially Monitored Firms (NSMF) program. This is an environmental policy targeting over 6,500 firms in China listed as high polluters in 2007 with a selection rule for firms' chemical oxygen demand (COD), NH_4^+, SO_2, smoke and industrial dust emissions in 2005. Half of the Top 1,000 firms and 14% of the Top 10,000 firms were included on the NSMF list.

We now show that our estimated effects of the Top 1,000 program are not driven by this competing policy. Consider first the effect of the NSMF program within the group of firms in the Top 1,000 program. Panel A of Table A.9 reports estimates of a difference-in-differences model of the effects of the NSMF program and shows that it had little effect on the energy use, output, and energy efficiency of Top 1,000 firms. We also consider whether the NSMF program impacts our estimates of the effects of the Top 1,000 program on regulated firms. Panel B reports estimates of the effects of the Top 1,000 program when we exclude all treated firms included under both polices. Panel B shows that we obtain results similar to those of our baseline regressions when we leave out the all Top 1,000 firms included under the NSMF program. These findings suggest that taking the NSMF program into account does not affect our main results.

E Potential Margins of Substitution for Regulated Firms

This appendix explores whether regulated firms can respond to the Top 1,000 program by changing their energy mix or by using other inputs in production.

E.1 Substitution of Energy Mix

Our baseline results show that the Top 1,000 program led to significant declines in output and energy use but did not lead to significant changes in energy efficiency. Because these results rely mostly on coal use and because our data do not measure electricity use, a potential concern is that regulated firms may increase their electricity use, which would result in a smaller decrease in total energy consumption.[85]

We conduct three exercises to explore whether regulated firms changed their energy mix in response to the Top 1,000 program. First, we use the fact that Top 1,000 firms were responsible for close to 60% of the energy consumption of the regulated industries. While the CESD data do not measure electricity use at the firm level, we can use data from the China Statistical Yearbook to measure the aggregate energy mix of these industries. Given the importance of the Top 1,000 firms, we would expect that aggregate quantities would reflect any significant shifts in the energy consumption of these firms. However, as Figure A.18 shows, the energy mix of the regulated industries was quite stable during this period. This result is a first indication that Top 1,000 firms may not have significantly changed their energy mix.

Second, while the CESD data do not measure electricity use in all years, our ATS data allow us to measure electricity use starting from 2007. As a second exercise, we use these data to construct a measure of electricity intensity for each firm. We then test whether the regulation had heterogeneous effects on firms with higher electricity consumption. Intuitively, firms that found it easier to use electricity could have experience smaller declines in output and larger declines in our coal-based measure of energy use. However, as Table A.30 shows, our point estimates are not consistent with this hypothesis. There are no statistical differences in the effects of the regulation for firms that had higher electricity consumption in 2007.

As a third exercise, we use data from the CESD to estimate whether Top 1,000 firms changed their mix among the energy sources that we do observe: coal, petroleum, and gas. Table A.31 shows that Top 1,000 firms did not experience meaningful or statistically significant changes in their energy mix following the regulation.

Combined, these three results are consistent with the notion that changes to their energy mix did not represent a primary margin of substitution for Top 1,000 firms.

E.2 Substitution of Other Production Inputs

Our baseline model assumes that firms can respond to the policy by adjusting their energy use and their variable inputs but cannot adjust their production technology in terms of input mix.

---

[85] Note that, if this were the case, our estimates using measured energy efficiency (primarily based on coal as an energy source) would be an upper bound on the effects of the program on energy efficiency.

This appendix explores these assumptions empirically by estimating the effects of the program on the capital–labor, energy–labor, and energy–capital ratios. Table A.13 and Figure A.9 report the results of our analyses of the effects of the Top 1,000 program on these ratios for regulated firms. These analyses show that the program did not lead to significant changes in the capital–labor, energy–labor, or energy–capital ratios.

These null effects carry interesting economic implications. The fact that the energy–labor ratio does not change is consistent with our model assumptions that, absent improvements in energy efficiency, firms will reduce their variable inputs in a way commensurate to the reduction in energy use.

On the other hand, the result that the energy–capital ratio does not change suggests that firms were able to partially adjust their capital levels in response to the policy. We learn two things from this result. First, firms did not respond to the policy by investing in more capital. Together with the lack of an effect on energy efficiency (see Table 3), this result indicates that, in the short run, it was hard for firms to respond to the program by replacing old capital with new, energy-saving capital. This result follows because we would expect to see increases in the output–energy ratio had firms replaced their capital with more energy-efficient vintages. Second, the fact that capital levels move in tandem with energy use indicates that firms were able to partially respond to the policy by reducing their capital stock. To account for this possibility, Appendix J.3 develops an extension of our model that allows firms to adjust their capital in response to the program. As we discuss in Section 7.3, this alternative assumption leads to a comparable albeit slightly larger estimate of the government willingness to pay (GWTP) (see Panel C of Figure 10).

F Potential Reallocation across Regulated Firms

A key objective of our model is to combine reduced-form estimates of the firm-level effects of the Top 1,000 program on regulated, related, and unregulated firms to compute the aggregate effects of the program on total energy use and output. One potential concern is that the policy may have led to reallocation of economic activity among treated firms that would not be captured by our firm-level results. Specifically, if the policy had heterogeneous effects on larger firms or those with higher baseline energy use, the policy's aggregate effect on energy use may differ from our estimates of the firm-level effects. In this appendix, we first show how heterogeneity in the effects of the policy can influence its aggregate effects and then explore whether the policy had such heterogeneous effects.

Following the notation in our model, we first write the total energy use of Top 1,000 firms as follows:

$$E^{\text{Top 1,000}} = \int\limits_{\tilde{\phi}} e_1(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi,$$

where we take the integral for regulated conglomerates with $\phi > \tilde{\phi}$ and with respect to the truncated distribution of active firms, $\frac{g(\phi)M}{1-G(\phi_1)}$. Similarly, the total output of Top 1,000 firms is given by:

$$Q^{\text{Top 1,000}} = \int\limits_{\tilde{\phi}} q_1(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi.$$

Using Δ to denote percentage changes, we have that:

$$\Delta Q^{\text{Top 1,000}} = \int\limits_{\tilde{\phi}} \Delta q_1(\phi, n) s^q(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi,$$

where $s^q(\phi, n) = \frac{q_1(\phi,n)}{Q^{\text{Top 1,000}}}$ is the initial output share of the Top 1,000 firm ϕ relative to all Top 1,000 firms. This equation says that—as long as the firm-level effects of the policy on firm output, $\Delta q_1(\phi, n)$, do not vary with the firm's initial size—the firm-level effects of the policy will map directly to the overall effects of the policy on the overall output of Top 1,000 firms.

To study the total change in energy use of Top 1,000 firms, we consider the role of output and energy efficiency:

$$E^{\text{Top 1,000}} = \int\limits_{\tilde{\phi}} \frac{e_1(\phi, n)}{q_1(\phi, n)} q_1(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi = \int\limits_{\tilde{\phi}} \nu_1(\phi, n) q_1(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi,$$

where $\nu_1(\phi, n)$ is energy efficiency. We now decompose the percentage change in the total energy use of Top 1,000 firms:

$$\Delta E^{\text{Top 1,000}} = \int\limits_{\tilde{\phi}} \left[\Delta \nu_1(\phi, n) + \Delta q_1(\phi, n)\right] s^e(\phi, n) \frac{g(\phi)M}{1 - G(\phi_1)} d\phi,$$

where $s^e(\phi, n)$ is the initial energy consumption share among Top 1,000 firms. Under the assumption that the effects of the policy on energy efficiency, $\Delta\nu_1(\phi, n)$, and output, $\Delta q_1(\phi, n)$, are uncorrelated with firm initial energy consumption, this equation shows that the average firm-level effects of the policy can be used to calculate the overall change in energy use among Top 1,000 firms.

The derivations above show that the firm-level effects of the program can be used to study its aggregate impacts whenever the effects are not heterogeneous by initial firm size and energy use. We now explore whether this is the case. Table A.32 estimates the effects of the Top 1,000 program on regulated firms as in Table 3 but allows the effects of output to vary by firm size and both output and energy efficiency to vary by initial energy consumption. Panel A shows that large firms—those with above-median initial output—do not experience output declines significantly different from those in the baseline results in Table 3. We also obtain average effects that are very similar to those in Table 3 when we incorporate this heterogeneity. Specifically, using the fact that firms with size greater than the median are responsible for 88% of the output, the average effect of the policy according to column (3) of Table A.32 is $0.88 \times (-.22) + 0.12 \times (-.167) = 0.214$, while column (3) of Table 3 reports a value of 0.204. Our derivation above shows that the aggregate effects on energy use also depend on whether there is heterogeneity in the program's effects on output and energy efficiency across firms with different initial levels of energy consumption. Panels B and C estimate the effects of the program on output and energy efficiency and show that the program does not have heterogeneous effects along this margin. Overall, the results in this appendix show that the program did not lead to significant reallocation of economic activity and energy use across regulated firms, suggesting that we can interpret the mean firm-level effects as the effects on the regulated segment of firms when estimating the aggregate effects of the policy within our framework.[86]

G Incorporating Pollution Damages

Our analysis of the welfare effects of the program capture the government's trade-off between the benefits of reducing energy use and the costs of distorting the market through price increases. Conceptually, reducing energy use has global externalities that reduce the amount of CO_2 in the atmosphere. If this were the only benefit, our estimates would correspond to the social cost of carbon (SCC). However, a reduction in energy use also leads to reductions in local pollution, which reduces local pollution damages.

This section details how we incorporate pollution damages into the welfare effects of the policy. As we discuss in Section 6.1.2, in our model, the government's willingness to pay (GWTP) for reducing energy emissions includes both of these channels. Specifically, the government's trade-off between production distortions and reductions in energy use is governed by parameters that measure the social cost of carbon (SCC)—κ_0—and the local health costs of associated pollution—κ_1. As we discuss in the text, these parameters can be related to values of the SCC and to estimates of pollution damages as follows:

$$\kappa_0 = \frac{\text{Social Cost of Carbon} \times \text{Carbon Emissions}}{\text{Aggregate Income} \times 0.8} \quad \text{and} \quad \kappa_1 = \frac{\text{Total Pollution Damages}}{\text{Aggregate Income} \times 0.8}.$$

---

[86] One potential reason for these results is that, as discussed in Appendix A.1, the energy savings targets of Top 1,000 firms were allocated mainly based on each firm's initial energy consumption, so their quota could be considered proportional. This may explain why there is no effect heterogeneity among Top 1,000 firms with different initial sizes or energy consumption.

The overall GWTP for energy use reductions combines both of these forces and is defined as:

$$\kappa \;=\; \frac{\text{GWTP} \times \text{Carbon Emissions}}{\text{Aggregate Income} \times 0.8},$$

where the GWTP includes both the SCC and the pollution damage per ton of carbon.

To implement these expressions, we rely on four estimates of pollution damages. We first rely on the estimates of gross external damage (GED) from Mohan et al. (2020). These estimates include the social cost of carbon and the local damage from pollution derived from the value of a statistical life. Mohan et al. (2020) estimate that the GED for China was close to 12.27% of GDP in 2006. Given that China's GDP in 2006 was \$2.752 trillion, the GED equaled \$337.67 billion. Since Mohan et al. (2020) rely on an estimate of the SCC of \$36, the damages from China's 2006 emissions of 6.38 billion tons of CO_2 are then \$229.68 billion. Thus, the estimate from Mohan et al. (2020) implies total pollution damages of \$107.99 billion (= \$337.67 – \$229.68).[87]

Our second estimate of pollution damages comes from the World Bank (2007) and relies on a willingness-to-pay measure of the costs of pollution. Using this methodology, the World Bank (2007) reports an estimate of total health costs from air pollution of 519.9 billion yuan in 2003. Converting to US dollars with an exchange rate of 8.28 yuan per dollar yields a value of total health damages of \$62.79 billion.

As a third estimate of pollution damages, we rely on an estimate from the World Bank (2007) based on an adjusted human capital approach. The World Bank (2007) reports an estimate of total health costs from air pollution of 157.3 billion yuan in 2003. Converting to dollars with an exchange rate of 8.28 yuan per dollar yields a value of total health damages of \$19 billion.

Finally, we rely on a measure of willingness to pay from quasi-experimental variation in pollution exposure due to the Huai River policy. Ebenstein et al. (2017) show that total pollution emissions north of the Huai river are larger due to a larger reliance on coal for heating, which they show reduces life expectancy. Ito and Zhang (2020*b*) use this discontinuity and the purchase of air purifiers to estimate households' willingness to pay to reduce their exposure to air particulates (PM_{10}). They estimate that, across China, the total willingness to pay to lower PM_{10} by $1\mu g/m^3$ is \$0.45 billion. On the basis of the national average of PM_{10} cited in Ito and Zhang (2020*b*) of $124\mu g/m^3$, total damages in 2013 according to this measure were \$55.8 billion.

Table A.33 collects these estimates of total pollution damages in China in column (2). Column (4) lists estimates of total CO_2 emissions in China from Ritchie, Roser and Rosado (2020). In column (6), we report the ratio of total damages from column (2) to the total CO_2 emissions

---

[87] As Muller, Mendelsohn and Nordhaus (2011) note, the external damages from pollution may be higher or lower depending on where it is emitted. As we show in Appendix H, we do not find evidence that the program led to reallocation of production toward more polluted or populated places or toward regions in Western China. For this reason, we do not account for changes in the location of production in this calculation.

from column (4). Across these four methodologies, we estimate that reducing CO_2 by one ton has associated health benefits for the local population valued at between \$4 and \$17.

An assumption implicit in column (6) of Table A.33 is that the energy use forgone under the Top 1,000 program would have imposed the same pollution damages as those associated with all other energy sources. However, since the firms in the Top 1,000 program relied primarily on coal, it is possible that reductions in their energy use could have larger marginal impacts on pollution damages. To account for this possibility, we assume that all pollution damages are generated by coal, and we report the estimates of CO_2 emissions from coal in column (5) (Ritchie, Roser and Rosado, 2020). Dividing the total pollution damages in column (2) by the CO_2 emissions from coal in column (5) gives an upper bound of the pollution damages associated with reducing the energy use of Top 1,000 firms of between \$5 and \$22 per ton of carbon. It is noteworthy that these upper-bound estimates are not much larger than the estimates in column (6). This result is a consequence of the fact that approximately 75% of CO_2 emissions are due to coal, so that attributing all of the pollution damages to coal usage increases the estimates from column (6) to column (7) by a factor of approximately $\frac{1}{0.75} = 1.33$. This pattern may be different in other countries where coal represents a smaller fraction of overall carbon emissions.

The estimates of the reduction in pollution damages help us interpret our estimate of the government's total benefits from reducing energy use. In our baseline model, we find that the policy is justified if the total benefits from energy reduction are greater than \$161. Subtracting the estimates of the local health benefits from pollution reduction in column (6) of Table A.33 of between \$4 and \$17 implies that the social cost of carbon that rationalizes the policy is between \$144 and \$157. Using the upper-bound estimates in column (7) implies that the social cost of carbon that rationalizes the policy is between \$139 and \$156.

The finding that the health benefits from reducing pollution may be quantitatively small in the case of China contrast with intuitions from richer countries. To reconcile this, note that Mohan et al. (2020) estimate larger health benefits from reducing pollution in richer countries. In the case of the US, Mohan et al. (2020) estimate GED of close to 5% in 2006. Given US GDP at the time of \$13.8 trillion and emissions of 6.05 billion tones of CO_2, we compute that reducing a ton of CO_2 has a GED of \$114. Subtracting their assumed SCC of \$36 implies a value of health benefits of \$78 per ton of CO_2. If the Chinese government valued the health benefits from reducing pollution in China as much as these estimates from the US, the social cost of carbon that rationalized the policy would then be \$83.

H Potential Reallocation across Geographic Regions

To understand our motivation for studying geographic reallocation, recall that our model of welfare in Equation 11 incorporates the effects on both economic output and emissions. One possibility is that the government cares about reducing emissions in particular locations, which may have higher levels of pollution or population density.

To allow for this possibility, we consider the extended indirect utility function:

$$W = \left(\frac{R}{P}\right)^{1-\kappa} \left(\frac{1}{\beta_0 E}\right)^{\kappa_0} \prod_{n=1}^{N} \left(\frac{1}{\beta_1 E_n}\right)^{\lambda_n},$$

where $\kappa = \kappa_0 + \sum_n \lambda_n$. β_0 captures the conversion of energy into carbon emissions. In this formulation, κ_0 captures the global externalities of carbon emissions on global warming. β_1 captures the effects of energy use on pollutants.[88] The terms λ_n capture the welfare costs of emissions in particular locations—including health effects from local pollution—and may vary depending on the government's preferences.

The overall change in welfare is now:

$$\begin{aligned} \frac{d\ln W}{1-\kappa} &= -\ln\left(\frac{P}{P^*}\right) - \frac{\kappa_0}{1-\kappa}\ln\left(\frac{E}{E^*}\right) - \sum_{n}^{N} \frac{\lambda_n}{1-\kappa}\ln\left(\frac{E_n}{E_n^*}\right) \\ &= -\ln\left(\frac{P}{P^*}\right) - \frac{\kappa}{1-\kappa}\ln\left(\frac{E}{E^*}\right) - \sum_{n}^{N} \frac{\lambda_n}{1-\kappa}\left[\ln\left(\frac{E_n}{E_n^*}\right) - \ln\left(\frac{E}{E^*}\right)\right], \end{aligned} \tag{H.1}$$

where the second line adds and subtracts $\sum_{n}^{N} \frac{\lambda_n}{1-\kappa}\ln\left(\frac{E}{E^*}\right)$.

The last term in this expression shows that the policy has larger effects on welfare if it leads to larger emissions reductions in locations with higher values of λ_n. Even if there is no aggregate decline in emissions, i.e., $\ln\left(\frac{E}{E^*}\right) = 0$, the policy may raise welfare as long as the reallocation effects in the third term exceed the effect on production in the first term.

The rest of this appendix explores empirically whether the policy reallocated emissions across locations. We find that the spillover effects of the policy did not lead to reallocation of emissions to locations with higher levels of pollution or population density and were not targeted to reallocate activity toward Western regions. These results suggest that the first-order welfare effects of the policy can be captured by the aggregate changes in energy use.

---

[88] For simplicity, we consider a composite pollutant. An extended model could incorporate J different pollutants by including the term $\prod_{j=1}^{J}\prod_{n=1}^{N}\left(\frac{1}{\beta_1^j E_n}\right)^{\lambda_n^j}$ with differing values of β_1^j and λ_n^j. Such a model would lead to similar insights as those drawn from Equation H.1.

H.1 Heterogeneous Effects of Spillovers by Local Density and Pollution

One important concern is that the spillovers that we identify may have shifted the location of production and of related emissions to more populated areas and areas with higher levels of pre-existing industrial emissions. To address this concern, we first use data on city-level sulphur dioxide emissions and population density to generate the following measure of exposure: $\frac{City\ SO_2\ Emission \times Population}{City\ Area}$. We then calculate the difference in this exposure measure between each pair of regulated and related firms. Next, we split our sample by terciles of this measure. Table A.28 shows that related firms in places with lower or similar exposure increased their output by 10% while those in more exposed areas saw larger increases of close to 22%. However, because a higher share of the output of related firms was concentrated in less exposed areas (58% in relatively less exposed vs. 28% in relatively more exposed areas), we find similar increases in related firm production across more and less exposed areas. Overall, the spillover effects of the regulation did not disproportionately shift production to areas with higher population density or higher pre-existing levels of industrial emissions.

H.2 Heterogeneous Effects of Spillovers between Eastern and Western Regions

During this time period, the Chinese government was interested in reallocating economic activity to the less developed areas in the Western regions of China. This appendix studies whether the within-conglomerate spillovers of the Top 1,000 program contributed to this goal. To do so, we first categorize each regulated and related firm pair into one of three categories: (1) regulated firms located in Eastern or mid-China regions and related firms in Western regions, (2) both types of firms in the same region, and (3) regulated firms in Western regions and related firms in Eastern or mid-China regions. In this setting, regions in Western China include Xinjiang Uygur Autonomous Region, Tibet Autonomous Region, Qinghai Province, Gansu Province, Sichuan Province, Yunnan Province, Inner Mongolia Autonomous Region, Ningxia Hui Autonomous Region, Shaanxi Province, Chongqing Municipality, Guizhou Province and Guangxi Zhuang Autonomous Region. This grouping reveals that more than 80% of the production of related firms was located in the same region as their regulated counterparts. However, the policy could still have led to reallocation if it had larger spillover effects for related firms that were located in the Western regions. Table A.29 explores this possibility by estimating heterogeneous spillover effects across these three groups. This table shows that the spillover effects are, if anything, smaller for related firms in Western regions. Relating these heterogeneous effects to their relevant output shares, this analysis reveals that the largest effect of the policy was within region and that the cross-region effects mostly offset each other. Given these results and the fact that official documents do not explicitly mention reallocation of production to Western regions as a policy objective, our policy analysis does not incorporate this margin.

I Effects of the Top 10,000 Program

The main objective of our paper is to estimate the effects of the initial regulation of the Top 1,000 program. In particular, access to firm-level data on the production and energy use of regulated, related, and competing firms allows us to estimate important margins of response to this program.

Given the perceived success of the Top 1,000 program, the Chinese government expanded this initiative through the Top 10,000 program. The program was put in place as part of the 12^{th} Five-Year Plan starting in 2012. The program expanded the regulation by targeting firms that consumed energy in excess of 10,000 tonnes of coal equivalent in 2010. This requirement meant that about 17,000 firms could be included in the program.

This appendix explores the effects of this program expansion. We study the effects of the Top 10,000 program using a similar difference-in-differences approach. However, relative to our analyses of the Top 1,000 program, we face two constraints in terms of data availability. First, we can rely on ASIF data only for the years 2009 and 2011 (pre-period) and 2012 and 2013 (post-period). Second, the ASIF does not include information on energy use.

Panel A of Table A.34 shows that the ASIF data cover 14,300 firms in the manufacturing sector that are regulated by the Top 10,000 program. This panel also shows that we can find 9,787 firms that are part of the business networks of regulated firms. To study the effects of the Top 10,000 program, we exclude from our analyses firms that were previously regulated by the Top 1,000 program (including firms related through their ownership networks). Panel B of Table A.34 shows that our data cover 12,524 Top 10,000 firms not previously regulated and 8,522 firms related to them.

Relative to Top 1,000 firms, Top 10,000 firms have a smaller number of related firms, and these firms are also smaller. The newly regulated firms have 0.7 related firms, on average—or 0.5 related firms, on average, if we require that related firms operate in the same 4-digit industry. Figure A.19 plots the size distribution of firms related to Top 1,000 and Top 10,000 firms. This figure shows that the 90^{th} percentile value of the log-output distribution of Top 10,000–related firms is close to the median value of that for Top 1,000–related firms. The fact that Top 10,000 firms have fewer and smaller related firms suggests that there may be more limited scope for within-conglomerate spillover effects.

We first estimate the effects of the Top 10,000 program. We use a specification similar to that of Equation 2. In this case, the set of reference firms includes all nonrelated firms in the ASIF. Panel A of Table A.35 shows that firms that are newly regulated as part of the Top 10,000 program experience a reduction in output of approximately 13–16% relative to that of nonregulated firms. These results are consistent with those of the Top 1,000 program in both direction and magnitude.

We now explore whether firms related to Top 10,000 programs experienced increases in output as a result of the regulation. As in our analyses of withing-conglomerate spillovers of the Top 1,000 program, we estimate these spillover effects by matching firms related to Top 10,000 firms to unregulated firms with similar characteristics in the pre-period. Panel B of Table A.35 shows that we only find spillover effects for related firms with output above the 90th percentile of the size distribution. While we do not estimate statistically significant spillover effects for all firms, these results are consistent with those of the Top 1,000 program since Table A.20 shows that we estimate larger spillover effects for larger related firms.

Finally, we explore whether the Top 10,000 program had spillover effects at the market level. To do so, we estimate a regression similar to that in Equation 3. Recall that, in our analysis of the Top 1,000 program, we measure market-level exposure to the regulation by computing the aggregate energy savings target of regulated firms relative to the energy used in a given industry. In the case of the Top 10,000 program, the main variable, $Incremental\ Spillover_j$ measures the increase in this share relative to the value under the Top 1,000 program. Panel C of Table A.35 shows the effects of these analyses. The last two columns of this table show small and statistically insignificant spillover effects in energy-intensive industries. This result is intuitive since these industries were already highly regulated by the Top 1,000 program and the Top 10,000 program did not significantly increase the stringency of regulation. In contrast, the Top 10,000 introduced regulations to other industries that had not been covered by the Top 1,000 program. The first two columns of Panel C of Table A.35 show that unregulated and unrelated firms in newly regulated industries experienced significant increases in output. While these effects are larger than those for firms in energy-intensive industries, they are also smaller than the market-level spillover effects of the Top 1,000 program (see Panel A of Table 5).

One interpretation of these results is that the magnitude of within-conglomerate and market-level spillover effects depends on the policy coverage. Since the Top 10,000 program is a more complete version of the regulation—i.e., it covers a larger share of the market—there is less scope for both within-conglomerate and market-level leakage.

J Model Appendix

This appendix provides detailed model derivations.

J.1 Model Equilibrium

Recall that the conglomerate takes the prices of energy p_e, capital r, and the variable input bundle w as given. Given the Leontief technology, the conglomerate sets $l_i = e_i$ so that the cost of intermediate inputs is $w + p_e$. Holding the number of affiliates n constant, the conglomerate maximizes

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^n, \{k_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\sum_{i=1}^{n} \phi \delta^{i-1} k_i^{\alpha_k} l_i^{\alpha_l} \right]^{\rho} - (w + p_e) \sum_{i=1}^{n} l_i - r \sum_{i=1}^{n} k_i \right\}.$$

For a firm i, the first-order conditions for l_i and k_i imply that $l_i = \frac{\alpha_l}{\alpha_k} \frac{r}{(w+p_e)} k_i$.

Substituting this expression, we can write the profit maximization problem as

$$\pi(\phi, n) = \max_{\{k_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\sum_{i=1}^{n} \phi \delta^{i-1} k_i^{\alpha} \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w + p_e)} \right)^{\alpha_l} \right]^{\rho} - \left(\frac{\alpha}{\alpha_k} r \right) \sum_{i=1}^{n} k_i \right\}.$$

Comparing the first-order conditions for k_1 and k_i, we find that $\frac{k_i}{k_1} = \delta^{\frac{i-1}{1-\alpha}}$. The final result from Proposition 1 follows since

$$q_i = \phi \delta^{i-1} k_i^{\alpha} \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w + p_e)} \right)^{\alpha_l} = \phi \delta^{\frac{i-1}{1-\alpha}} k_1^{\alpha} \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w + p_e)} \right)^{\alpha_l} = \delta^{\frac{i-1}{1-\alpha}} q_1.$$

Using these results, we can write the profit maximization problem in terms of $K_n = \sum_i^n k_i$. To do so, we define the conglomerate's total productivity $\phi\Delta_n = \phi[\sum_{i=1}^n (\delta^{i-1})^{\frac{1}{1-\alpha}}]^{1-\alpha}$ and the constant $C_\pi = (1 - \alpha\rho) \left[\left(\frac{\rho\alpha_l}{w+p_e} \right)^{\alpha_l \rho} \left(\frac{\rho\alpha_k}{r} \right)^{\alpha_k \rho} \right]^{\frac{1}{1-\alpha\rho}}$ to obtain

$$\pi(\phi, n) = \max_{K_n} \left\{ \frac{R^{1-\rho} P^{\rho} C_\pi^{1-\alpha\rho}}{(1 - \alpha\rho)^{1-\alpha\rho}} \left(\frac{\rho\alpha_k}{r} \right)^{-\alpha\rho} (\phi\Delta_n)^{\rho} K_n^{\alpha\rho} - r \left(\frac{\alpha}{\alpha_k} \right) K_n \right\}.$$

The optimal capital K_n and the firm profits for a conglomerate of size n are then

$$K_n = \frac{R^{\frac{1-\rho}{1-\alpha\rho}} P^{\frac{\rho}{1-\alpha\rho}} C_\pi}{(1 - \alpha\rho)} \frac{\rho\alpha_k}{r} (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \quad \text{and} \quad \pi(\phi, n) = R^{\frac{1-\rho}{1-\alpha\rho}} P^{\frac{\rho}{1-\alpha\rho}} C_\pi (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}}.$$

The optimal profit given ϕ is then

$$\pi(\phi) = \max_n \pi(\phi, n) - rfn = \max_n R^{\frac{1-\rho}{1-\alpha\rho}} P^{\frac{\rho}{1-\alpha\rho}} C_\pi \times (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} - rfn.$$

Solving for indifference points ϕ_n yields the result of Proposition 2 and the zero-profit condition (Equation 7).

We now compute the price level. For given quantities $q(\phi, n)$, the price level is given by

$$P^{-\rho} = R^{-\rho} \int_{\phi_1} q(\phi, n)^{\rho} \frac{g(\phi) M}{1 - G(\phi_1)} d\phi.$$

To differentiate from the regulated case below, we use starred variables to denote the prices and quantities in the unregulated case. To derive $q^*(\phi, n)$, note from Proposition 1 that $K_n = k_1 \Delta_n^{\frac{1}{1-\alpha}}$, and also recall that $l_1 = \frac{\alpha_l}{\alpha_k}\frac{r}{(w+p_e)}k_1$. We then have

$$\begin{aligned} q^*(\phi, n) &= \Delta_n^{\frac{1}{1-\alpha}} q_1^*(\phi, n) \\ &= \Delta_n^{\frac{1}{1-\alpha}} \phi \left(\frac{\alpha_l}{\alpha_k}\frac{r}{(w+p_e)}\right)^{\alpha_l} k_1^{*\alpha} \\ &= (\phi\Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w+p_e}\right)^{\alpha_l} \left(\frac{\alpha_k}{r}\right)^{\alpha_k}\right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q}. \end{aligned} \tag{J.1}$$

The price level in the absence of the regulation is then

$$\begin{aligned} P^{*-\rho} &= R^{-\rho} \int_{\phi_1} \left((\phi\Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} C_Q\right)^{\rho} \frac{g(\phi)M}{1-G(\phi_1)} d\phi \\ P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} &= \int_{\phi_1} (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \frac{g(\phi)M}{1-G(\phi_1)} d\phi. \\ P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1-G(\phi_1)}{M} &= \int_{\phi_1} (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi \\ P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1-G(\phi_1)}{M} &= \sum_n \Delta_n^{\frac{\rho}{1-\alpha\rho}} \underbrace{\int_{\phi_n}^{\phi_{n+1}} (\phi)^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi}_{=\pi_n} = \sum_n \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n. \end{aligned}$$

Given the assumption that ϕ follows a log-normal distribution, we can express

$$\pi_n = \int_{\phi_n}^{\phi_{n+1}} (\phi)^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi = \exp\left\{\frac{\tilde{\sigma}^2}{2}\right\} [\Phi(b_{n+1} - \tilde{\sigma}) - \Phi(b_n - \tilde{\sigma})],$$

where $\tilde{\sigma} = \frac{\rho}{1-\alpha\rho}\sigma_\phi$ and $b_n = \frac{\rho}{1-\alpha\rho}\frac{\ln(\phi_n)}{\tilde{\sigma}}$. Note also that Equations 6 and 7 imply that

$$\phi_{n+1} = \frac{\phi_1}{\left(\Delta_{n+1}^{\frac{\rho}{1-\rho\alpha}} - \Delta_n^{\frac{\rho}{1-\rho\alpha}}\right)^{\frac{1-\rho\alpha}{\rho}}}.$$

Thus, π_n depends on ϕ_1; it does not directly depend on equilibrium prices P^*.

J.2 Response to Top 1,000 Program

The Top 1,000 program limits energy use at the largest firm e_1 to a fraction $\xi < 1$ of the energy use in the unregulated case e_1^*; recall that we use starred variables to denote the optimal choices in the unregulated case. We assume that the number of firms n and the capital allocations $\{k_i^*\}_{i=1}^n$ are quasi-fixed but that the conglomerate can adjust $\{l_i\}_{i=1}^n$.

Regulated Conglomerates

Using the fact that $k_i^* = k_1^* \delta^{\frac{i-1}{1-\alpha}}$, we can write the profit maximization problem as

$$\max_{\{l_i\}_1^n} \left\{ R^{1-\rho} P^{\rho} \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho} - (w+p_e) \sum_{i=1}^{n} l_i - r \sum_{i=1}^{n} k_i^* \right\} \quad \text{subject to} \quad l_1 \leq \xi l_1^*,$$

where $\phi^* = \phi (k_1^*)^{\alpha_k}$. The first-order conditions for l_i $(1 \leq i \leq n)$ are then

$$\frac{\partial \pi}{\partial l_i} = \underbrace{R^{1-\rho} P^{\rho}}_{\text{Market Demand}} \underbrace{\rho \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi^* \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \alpha_l (l_i)^{\alpha_l - 1}}_{\text{Marginal Product}} = w + p_e + \underbrace{\lambda(\phi) \mathbb{I}[i=1]}_{\substack{\text{Shadow Cost} \\ \text{of Regulation}}}.$$

These conditions yield the results in Proposition 3.

We now show that, given n, $\lambda(\phi)$ does not depend on ϕ. To do so, we note that an implication of Proposition 3 is that we can write total conglomerate production under the regulation as

$$q(\phi, n) = \sum_{i}^{n} q_i(\phi, n) = q_1(\phi, n) \left[1 + \sum_{i>1}^{n} \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{w + p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right].$$

In contrast, recall from Proposition 1 that, in the unregulated case, total conglomerate output is

$$q^*(\phi, n) = \sum_{i}^{n} q_i^*(\phi, n) = q_1^*(\phi, n) \sum_{i}^{n} \delta^{\frac{i-1}{1-\alpha}}.$$

To connect these expressions, note that, since k_1^* is fixed and $l_1 = \xi l_1^*$, we have that

$$q_1(\phi, n) = \phi (k_1^*)^{\alpha_k} (l_1)^{\alpha_l} = \phi (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l} \xi^{\alpha_l} = q_1^*(\phi, n) \xi^{\alpha_l}.$$

Together, the last three expression imply that

$$q(\phi, n) = q^*(\phi, n) \xi^{\alpha_l} \frac{\left[1 + \sum\limits_{i>1}^{n} \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right]}{\sum\limits_{i}^{n} \delta^{\frac{i-1}{1-\alpha}}} = q^*(\phi, n) \underbrace{\left[\xi^{\alpha_l} \frac{1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right]}_{= \xi_{q,n}}, \tag{J.2}$$

where $\xi_{q,n}$ captures the impact of the regulation on conglomerate output.

Using this expression and the fact that $l_1 = \xi l_1^*$, we can rewrite the first-order condition for l_1 in terms of the capital and labor choices in the unregulated case:

$$\underbrace{R^{1-\rho} P^{*\rho} \rho \left[\phi \sum_{i=1}^{n} \delta^{\frac{i-1}{1-\alpha}} (k_i^*)^{\alpha_k} (l_i^*)^{\alpha_l} \right]^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l - 1}}_{\text{FOC Unregulated Case}}$$

$$\times (\xi)^{\alpha_l - 1} \left(\frac{P}{P^*} \right)^{\rho} \left[\xi^{\alpha_l} \frac{1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right]^{\rho-1} = w + p_e + \lambda(\phi),$$

where $\frac{P}{P^*}$ is the equilibrium change in the industry-level price. Using the fact that the first-order condition in the unregulated case equals $w + p_e$, we have

$$(\xi)^{\alpha_l-1}\left(\frac{P}{P^*}\right)^{\rho}\left[\xi^{\alpha_l}\frac{1+(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}}\right]^{(\rho-1)} = 1+\frac{\lambda(\phi)}{w+p_e}. \tag{J.3}$$

This expression shows that, conditional on n, the shadow cost does not depend on ϕ, and so we now write λ_n. The expression above is equivalent to

$$\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{1}{1-\rho}}\left[1+(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}}\right] = \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\rho}}\xi^{\frac{-(1-\alpha_l\rho)}{1-\rho}}\Delta_n^{\frac{1}{1-\alpha}}. \tag{J.4}$$

This expression does not have a general closed-form solution. Consider, however, the special case where $\frac{1}{1-\rho} = \frac{\alpha_l}{1-\alpha_l}$.[89] In this special case, we can write this expression as

$$(\Delta_n^{\frac{1}{1-\alpha}}-1)x^2 + x - \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\rho}}\xi^{\frac{-(1-\alpha_l\rho)}{1-\rho}}\Delta_n^{\frac{1}{1-\alpha}} = 0,$$

where $x = \left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{1}{1-\rho}} = \left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}}$, which allows us to solve for λ_n using the quadratic formula. Focusing on the positive root implies

$$1+\frac{\lambda_n}{w+p_e} = \left\{\frac{-1+\sqrt{1+4(\Delta_n^{\frac{1}{1-\alpha}}-1)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\rho}}\xi^{\frac{-(1-\alpha_l\rho)}{1-\rho}}\Delta_n^{\frac{1}{1-\alpha}}}}{2(\Delta_n^{\frac{1}{1-\alpha}}-1)}\right\}^{1-\rho}.$$

Note that λ_n depends on the equilibrium price P in both the expression above and in Equation J.4. In this case, we also have

$$\xi_{q,n} = \xi^{\alpha_l}\frac{1+\sqrt{1+4(\Delta_n^{\frac{1}{1-\alpha}}-1)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\rho}}\xi^{\frac{-(1-\alpha_l\rho)}{1-\rho}}\Delta_n^{\frac{1}{1-\alpha}}}}{2\Delta_n^{\frac{1}{1-\alpha}}}.$$

Unregulated Conglomerates

Unregulated conglomerates are affected by the policy through the price adjustment in the product market. The first-order conditions for l_i ($1 \leq i \leq n$) are then

$$\frac{\partial\pi}{\partial l_i} = R^{1-\rho}P^{\rho}\rho\left[\phi^*\sum_{i=1}^{n}\delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}}l_i^{\alpha_l}\right]^{\rho-1}\phi^*\delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}}\alpha_l(l_i)^{\alpha_l-1} = w+p_e.$$

---

[89] While this is a knife-edge case, it holds in the empirically relevant case of $\rho = .75$ and $\alpha_l = 0.8$, so that $\frac{1}{1-\rho} = \frac{\alpha_l}{1-\alpha_l} = 4$.

Since these firms do not face a shadow cost, the conditions of Proposition 1 continue to hold. Using these results to solve for l_1, we obtain

$$l_1 = \left[\frac{R^{1-\rho}P^{\rho}\rho[\phi(k_1^*)^{\alpha_k}]^{\rho}\Delta_n^{\frac{\rho-1}{1-\alpha}}\alpha_l}{w+p_e}\right]^{\frac{1}{1-\alpha_l\rho}}.$$

This further implies that

$$l_1 = l_1^*\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} \tag{J.5}$$

and both

$$q_1(\phi,n) = q_1^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}} \quad \text{and} \quad q(\phi,n) = q^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}. \tag{J.6}$$

J.2.1 Product Market Equilibrium

We now derive the equilibrium price under the Top 1,000 program. Recall that only conglomerates with $\phi \geq \tilde{\phi}$ have a regulated firm. Under the definition of $\xi_{q,n}$ in Equation J.2 for regulated firms and Equation J.6 for unregulated firms, the price level under the regulation is then

$$\begin{aligned} P^{-\rho} &= R^{-\rho}\int_{\phi_1}^{\tilde{\phi}}\left(\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}(\phi\Delta_n)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi \\ &+ R^{-\rho}\int_{\tilde{\phi}}\left(\xi_{q,n}(\phi\Delta_n)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi. \end{aligned}$$

As in the case of the unregulated equilibrium, we use π_n to denote the output of conglomerates with n affiliates. However, the regulation threshold $\tilde{\phi}$ in general does not line up with the size thresholds ϕ_n that define π_n. Let $\tilde{n}$ denote the smallest firm size for regulated conglomerates. We split $\pi_{\tilde{n}}$ into $\tilde{\pi}_1$ and $\tilde{\pi}_2$ as follows:

$$\tilde{\pi}_1 = \int_{\phi_{\tilde{n}}}^{\tilde{\phi}}(\phi)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi \quad \text{and} \quad \tilde{\pi}_2 = \int_{\tilde{\phi}}^{\phi_{\tilde{n}+1}}(\phi)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi.$$

We can then manipulate the expression for the equilibrium price as follows:

$$\begin{aligned} P^{-\rho}P^{*\frac{-\rho^2\alpha}{1-\alpha\rho}}R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}C_Q^{-\rho}\frac{1-G(\phi_1)}{M} &= \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}\left(\sum_{n=1}^{\tilde{n}-1}\Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n + \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1\right) \\ &+ \xi_{q,\tilde{n}}^{\rho}\Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2 + \sum_{n=\tilde{n}+1}\xi_{q,n}^{\rho}\Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n. \end{aligned}$$

Note that this equation holds in the case without the regulation if we set $\xi_{q,n} = 1$ and $P = P^*$. Let $s_{\tilde{\phi}} = \frac{\Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2 + \sum\limits_{n=\tilde{n}+1} \Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n}{\sum\limits_n \Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n}$ be the output share of the Top 1,000 conglomerates prior to the regulation. Additionally, we introduce the notation:

$$\mathbb{E}_e\left[x_n \middle| \phi > \tilde{\phi}\right] = x_{\tilde{n}}\left(\frac{\Delta_n^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2}{\Delta_n^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2 + \sum\limits_{n=\tilde{n}+1}\Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n}\right) + \sum_{n=\tilde{n}+1} x_n\left(\frac{\Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n}{\Delta_n^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2 + \sum\limits_{n=\tilde{n}+1}\Delta_n^{\frac{\rho}{1-\alpha\rho}}\pi_n}\right)$$

to denote the average of a size-dependent variable x_n conditional on being part of the Top 1,000 program with respect to the distribution of energy use in the unregulated equilibrium.

Assuming that the regulation does not impact r and w, taking the ratio of price levels before and after the regulation, we obtain

$$\left(\frac{P}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[\xi_{q,n}^{\rho} \middle| \phi > \tilde{\phi}\right]. \tag{J.7}$$

Equations J.4 and J.7 jointly determine the shadow costs of the regulation $\{\lambda_n\}_{n\geq\tilde{n}}$ and the increase in the price level $\frac{P}{P^*}$.

J.2.2 Characterizing Energy Use

Recall that the energy use in firm 1 before the regulation is

$$e_1^*(\phi, n) = l_1^*(\phi, n) = \left(\frac{\alpha_l}{\alpha_k}\frac{r}{(w + p_e)}\right)k_1^*(\phi, n) = \left(\frac{\alpha_l}{\alpha_k}\frac{r}{(w + p_e)}\right)\left(\frac{K_n^*}{\Delta_n^{\frac{1}{1-\alpha}}}\right).$$

The energy use for the conglomerate before the regulation is then

$$\begin{aligned} e^*(\phi, n) &= \left(\frac{\alpha_l}{\alpha_k}\frac{r}{(w + p_e)}\right)K_n^* \\ &= (\phi\Delta_n)^{\frac{\rho}{1-\alpha\rho}} R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \underbrace{\rho^{\frac{1}{1-\alpha\rho}}\left[\left(\frac{\alpha_l}{w + p_e}\right)^{1-\alpha_k\rho}\left(\frac{\alpha_k}{r}\right)^{\alpha_k\rho}\right]^{\frac{1}{1-\alpha\rho}}}_{=C_E}. \end{aligned} \tag{J.8}$$

The total energy use prior to the regulation is then

$$\begin{aligned} E^* &= \int\limits_{\phi_1} e^*(\phi, n)\frac{g(\phi)M}{1 - G(\phi_1)}d\phi \\ &= R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}}\frac{C_E M}{1 - G(\phi_1)}\sum_n (\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n. \end{aligned}$$

Regulated Conglomerates

We now characterize the change in energy use for regulated conglomerates. The fact that $e_i = l_i$ and the results of Proposition 1 imply that the energy use for an unregulated conglomerate is

$$e^*(\phi, n) = \sum_i e_i^*(\phi, n) = e_1^*(\phi, n) \sum_i^n \delta^{\frac{i-1}{1-\alpha}}.$$

Proposition 3 implies that the energy use for a regulated conglomerate is

$$e(\phi, n) = \sum_i e_i(\phi, n) = e_1(\phi, n) \left[1 + \sum_{i>1}^n \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda_n}{w + p_e} \right]^{\frac{1}{1-\alpha_l}} \right].$$

Using the fact that $e_1 = \xi e_1^*$, we then have

$$e(\phi, n) = e^*(\phi, n) \underbrace{\frac{\xi \left[1 + \sum\limits_{i>1}^n \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda_n}{w+p_e} \right]^{\frac{1}{1-\alpha_l}} \right]}{\sum\limits_i^n \delta^{\frac{i-1}{1-\alpha}}}}_{=\xi_{e,n}},$$

where $\xi_{e,n}$ captures the effect of the regulation on the energy use of a regulated conglomerate.

Unregulated Conglomerates

Proposition 1 and Equation J.5 imply that, for unregulated conglomerates,

$$e_1(\phi, n) = e_1^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l \rho}}$$

and additionally that

$$e(\phi, n) = e^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l \rho}}.$$

Aggregate Change in Energy

Putting the above together, total energy use after the regulation is now

$$\begin{aligned} E &= \int\limits_{\phi_1} e(\phi, n) \frac{g(\phi) M}{1 - G(\phi_1)} d\phi = \int\limits_{\phi_1}^{\tilde{\phi}} \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l \rho}} e^*(\phi, n) \frac{g(\phi) M}{1 - G(\phi_1)} d\phi + \int\limits_{\tilde{\phi}} \xi_{e,n} e^*(\phi, n) \frac{g(\phi) M}{1 - G(\phi_1)} d\phi \\ &= R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \frac{C_E M}{1 - G(\phi_1)} \left[\left(\sum_{n=1}^{\tilde{n}-1} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n + \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1 \right) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l \rho}} + \xi_{e,\tilde{n}} \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \xi_{e,n} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n \right]. \end{aligned}$$

This implies that

$$
\begin{aligned}
\frac{E}{E^*} &= \frac{\left(\sum_{n=1}^{\tilde{n}-1} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n + \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1\right) \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + \xi_{e,\tilde{n}} \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \xi_{e,n} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n}{\sum_n \left(\Delta_n\right)^{\frac{\rho}{1-\alpha\rho}} \pi_n} \\
&= \left(1 - s_{\tilde{\phi}}\right) \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{e,n} \middle| \phi > \tilde{\phi}\right]. \qquad \text{(J.9)}
\end{aligned}
$$

Equation J.9—along with the equilibrium price increase and shadow costs determined by Equations J.4 and J.7—allows us to compute the effect of the regulation on welfare.

J.2.3 Solving the New Equilibrium with the Reduced-Form Estimates

In Section 6, we present the full solution to the model using the derivations above. This appendix shows that we can also solve for an approximation of the equilibrium based on our reduced-form estimates. To do so, we make the assumption that λ and Δ are constant for regulated firms. Taking the value of λ implied by our reduced-form estimate of 8.95% and $\Delta^{\frac{1}{1-\alpha}} = 1.6$ (which approximates the value for $n = 6$), we use Proposition 4 to compute the production distortion $\xi_q = 0.9648$. Equation 14 delivers the equilibrium price change: $\ln\left(\frac{P}{P^*}\right) = 4.2\%$. We then compute that regulated conglomerates lower their energy use by 5.6%. Finally, we use these numbers to implement Equation 15, where we find that $\ln\left(\frac{E}{E^*}\right) = -3.65\%$. The advantage of this calculation is that it relies on only a handful of calibrated parameter values and the result of the within-conglomerate difference-in-differences estimation. In particular, this calculation does not rely on distributional assumptions for $G(\phi)$. It is thus reassuring that we obtain aggregate quantities close to those from the full model solution.

J.3 Response to Top 1,000 Program: With Capital Adjustment

This appendix extends our baseline model to allow the conglomerate to adjust both its capital and labor allocations.

Regulated Conglomerates

The profit maximization problem as

$$
\max_{\{l_i\}_1^n, \{k_i\}_1^n} \left\{ R^{1-\rho} P^{\rho} \left[\phi \sum_{i=1}^{n} \delta^{i-1} k_i^{\alpha_k} l_i^{\alpha_l}\right]^{\rho} - (w + p_e) \sum_{i=1}^{n} l_i - r \sum_{i=1}^{n} k_i \right\} \quad \text{subject to} \quad l_1 \leq \xi l_1^*.
$$

The first-order conditions for l_i $(1 \leq i \leq n)$ are then

$$\frac{\partial \pi}{\partial l_i} = \underbrace{R^{1-\rho}P^{\rho}}_{\text{Market Demand}} \underbrace{\rho\left[\phi\sum_{i=1}^{n}\delta^{i-1}k_i^{\alpha_k}l_i^{\alpha_l}\right]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi\delta^{i-1}\alpha_l(k_i)^{\alpha_k}(l_i)^{\alpha_l-1}}_{\text{Marginal Product}} \quad = \quad w+p_e+\underbrace{\lambda(\phi)\mathbb{I}[i=1]}_{\substack{\text{Shadow Cost}\\ \text{of Regulation}}}.$$

Capital's first-order conditions imply that $l_1 = \frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e+\lambda(\phi)}k_1$ and $l_i = \frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}k_i, \forall i > 1$. Recall that the Lagrangian function is

$$\max_{\{l_i\}_1^n,\{k_i\}_1^n} R^{1-\rho}P^{\rho}\left[\phi\left(\frac{\alpha_l}{\alpha_k}\right)^{\alpha_l}\left(k_1^{\alpha}\left(\frac{r}{w+p_e+\lambda(\phi)}\right)^{\alpha_l}+\sum_{i=2}^{n}\delta^{i-1}k_i^{\alpha}\left(\frac{r}{w+p_e}\right)^{\alpha_l}\right)\right]^{\rho}$$
$$-(w+p_e+\lambda(\phi))l_1 - rk_1 - (w+p_e)\sum_{i=2}^{n}l_i - r\sum_{i=2}^{n}k_i + \lambda(\phi)\xi l_1^*.$$

Substituting this expression back into the objective function, we can write the profit maximization problem as

$$\max_{\{k_i\}_{i=1}^n}\left\{R^{1-\rho}P^{\rho}\left[\phi\left(\frac{\alpha_l}{\alpha_k}\right)^{\alpha_l}\left(k_1^{\alpha}\left(\frac{r}{w+p_e+\lambda(\phi)}\right)^{\alpha_l}+\sum_{i=2}^{n}\delta^{i-1}k_i^{\alpha}\left(\frac{r}{(w+p_e)}\right)^{\alpha_l}\right)\right]^{\rho}-\left(\frac{\alpha}{\alpha_k}r\right)\sum_{i=1}^{n}k_i\right\}.$$

Comparing the first-order conditions for k_1 and k_i, we now have $\frac{k_i}{k_1} = \delta^{\frac{i-1}{1-\alpha}}\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{\frac{\alpha_l}{1-\alpha}}$. In other words, both capital and labor at the Top 1,000 firms are now distorted from their initial equilibrium.

We can again write the total conglomerate production under regulation as

$$q(\phi,n) = \sum_{i}^{n}q_i(\phi,n) = q_1(\phi,n)\left[1+\sum_{i>1}^{n}\delta^{\frac{i-1}{1-\alpha}}\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha}}\right].$$

We now again connect the total conglomerate production under regulation to its unregulated optimum:

$$q_1(\phi,n) = \phi(k_1)^{\alpha_k}(l_1)^{\alpha_l} = \phi\left[\xi\left(1+\frac{\lambda(\phi)}{w+p_e}\right)k_1^*\right]^{\alpha_k}(\xi l_1^*)^{\alpha_l} = q_1^*(\phi,n)\xi^{\alpha}\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{\alpha_k}.$$

Together, the last two expressions imply that

$$q(\phi,n) = q^*(\phi,n)\xi^{\alpha}\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{\alpha_k}\frac{\left[1+\sum\limits_{i>1}^{n}\delta^{\frac{i-1}{1-\alpha}}\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha}}\right]}{\sum\limits_{i}^{n}\delta^{\frac{i-1}{1-\alpha}}}$$
$$= q^*(\phi,n)\underbrace{\left[\xi^{\alpha}\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{\alpha_k}\frac{\left(1+(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha}}\right)}{\Delta_n^{\frac{1}{1-\alpha}}}\right]}_{=\xi_{q,n}(\phi)},$$

where $\xi_{q,n}$ captures the impact of the regulation on conglomerate output.

We now rewrite the first-order condition for l_1 in terms of capital and labor in the unregulated case:

$$\underbrace{R^{1-\rho}P^{*\rho}\rho(q^*(\phi,n))^{\rho-1}\times\phi\alpha_l(k_1^*)^{\alpha_k}(l_1^*)^{\alpha_l-1}}_{\text{FOC Unregulated Case}}\times$$

$$(\xi)^{\alpha-1}\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{\alpha_k}\left(\frac{P}{P^*}\right)^{\rho}[\xi_{q,n}(\phi)]^{\rho-1}=w+p_e+\lambda(\phi).$$

Using the fact that the first-order condition in the unregulated case equals $w+p_e$, we have

$$(\xi)^{\alpha-1}\left(\frac{P}{P^*}\right)^{\rho}[\xi_{q,n}(\phi)]^{\rho-1}=\left(1+\frac{\lambda(\phi)}{w+p_e}\right)^{1-\alpha_k}.$$

This equation again shows that, conditional on n, the shadow cost does not depend on ϕ. Thus, we can write the equation as

$$(\xi)^{\rho\alpha-1}\left(\frac{P}{P^*}\right)^{\rho}\left[\frac{\left(1+(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha}}\right)}{\Delta_n^{\frac{1}{1-\alpha}}}\right]^{\rho-1}=\left(1+\frac{\lambda_n}{w+p_e}\right)^{1-\rho\alpha_k}.$$

Hence, we can express the shadow cost of regulation as a nonlinear function of the equilibrium price changes and the regulation parameter ξ

$$\left(1+\frac{\lambda_n}{w+p_e}\right)^{\frac{1-\rho\alpha_k}{1-\rho}}\left[1+(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{\alpha_l}{1-\alpha}}\right]=\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\rho}}(\xi)^{\frac{-(1-\alpha\rho)}{1-\rho}}\Delta_n^{\frac{1}{1-\alpha}}.$$

Unregulated Conglomerates

The unregulated conglomerates are affected by the policy only through price:

$$k_1=\frac{R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi}{(1-\alpha\rho)}\frac{\rho\alpha_k}{r}(\phi)^{\frac{\rho}{1-\alpha\rho}}(\Delta_n)^{\frac{\rho-1}{(1-\alpha\rho)(1-\alpha)}}$$

$$l_1=\frac{R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi}{(1-\alpha\rho)}\frac{\rho\alpha_l}{w+p_e}(\phi)^{\frac{\rho}{1-\alpha\rho}}(\Delta_n)^{\frac{\rho-1}{(1-\alpha\rho)(1-\alpha)}}.$$

This implies that $l_1=l_1^*\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha\rho}}$ and $k_1=k_1^*\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha\rho}}$ and $q(\phi,n)=q^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\rho\alpha}{1-\alpha\rho}}$.

Product Market Equilibrium

Under the new definition of $\xi_{q,n}$, the product market equilibrium becomes

$$\left(\frac{P}{P^*}\right)^{-\rho}=(1-s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha\rho^2}{1-\alpha\rho}}+s_{\tilde{\phi}}\mathbb{E}_e\left[\xi_{q,n}^{\rho}\middle|\phi>\tilde{\phi}\right].$$

J.3.1 Characterizing Energy Use

Regulated Conglomerates

Energy use follows the variable inputs such that

$$\frac{e_i(\phi,n)}{e_1(\phi,n)} = \frac{l_i(\phi,n)}{l_1(\phi,n)} = \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda_n}{w+p_e}\right]^{\frac{1-\alpha_k}{1-\alpha}}.$$

The energy use for a regulated conglomerate is then

$$e(\phi,n) = e_1(\phi,n)\left[1 + \sum_{i>1}^{n} \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda_n}{w+p_e}\right]^{\frac{1-\alpha_k}{1-\alpha}}\right] \equiv e^*(\phi,n)\underbrace{\frac{\xi\left[1 + \sum_{i>1}^{n} \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda_n}{w+p_e}\right]^{\frac{1-\alpha_k}{1-\alpha}}\right]}{\sum_i^n \delta^{\frac{i-1}{1-\alpha}}}}_{\xi_{e,n}}.$$

Unregulated Conglomerates

We again have that the unregulated conglomerate responds only to market price changes:

$$e(\phi,n) = e^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha\rho}}.$$

Aggregate Change in Energy

Under the new definition of $\xi_{e,n}$, the change in aggregate energy use is

$$\frac{E}{E^*} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[\xi_{e,n}\,\middle|\,\phi > \tilde{\phi}\right].$$

K Alternative Regulations

Conglomerate-Level Regulation

Suppose that, instead of regulating the energy use of the top firm, the government restricted the energy use of all firms in a conglomerate to at most ξ of the energy use at the Top 1,000 firm plus the energy use at its related firms. The regulatory constraint would be

$$e(\phi,n) \leq \xi_{e,n} e^*(\phi,n).$$

Arguments similar to the derivation for Equation J.9 imply that

$$\frac{E^C}{E^*} = (1 - s_{\tilde{\phi}})\left(\frac{P^C}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[\xi_{e,n} \middle| \, \phi > \tilde{\phi}\right], \tag{K.1}$$

where we use the superscript C to denote the case of the conglomerate-level regulation.

Since energy use in all firms contributes equally to the regulatory constraint, this regulation does not distort the allocation of inputs across related firms; i.e., Proposition 1 continues to hold. This implies that $l_i = \xi_{e,n} l_i^*$ for all firms i in the conglomerate. It further implies that

$$q(\phi, n) = (\xi_{e,n})^{\alpha_l} q^*(\phi, n).$$

Arguments similar to the derivation for Equation J.7 imply that

$$\left(\frac{P^C}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) \left(\frac{P^C}{P^*}\right)^{\frac{\alpha_l \rho^2}{1-\alpha_l \rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[(\xi_{e,n})^{\alpha_l \rho} | \, \phi > \tilde{\phi} \right]. \tag{K.2}$$

We now derive the shadow cost of this regulation. Substituting l_i into the first-order condition for firm 1 implies that

$$\underbrace{R^{1-\rho} P^{*\rho} \rho \left[\phi \sum_{i=1}^{n} \delta^{\frac{i-1}{1-\alpha}} (k_i^*)^{\alpha_k} (l_i^*)^{\alpha_l} \right]^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l - 1}}_{\text{FOC Unregulated Case}}$$
$$\times (\xi_{e,n})^{\alpha_l - 1} \left(\frac{P^C}{P^*}\right)^{\rho} [(\xi_{e,n})^{\alpha_l}]^{\rho-1} = w + p_e + \lambda^C(\phi).$$

Using the fact that the first-order condition in the unregulated case equals $w + p_e$, we obtain

$$\left[1 + \frac{\lambda_n^C}{w + p_e}\right] = \left(\frac{P^C}{P^*}\right)^{\rho} (\xi_{e,n})^{-(1-\alpha_l \rho)}.$$

Size-Dependent Energy Tax

Suppose that the government instituted a per-unit energy tax for all the affiliates of conglomerates with $\phi > \tilde{\phi}$. As in the case above, this policy would not impact the within-conglomerate allocation of inputs of regulated firms, and Proposition 1 would continue to hold. That is, related firms would all reduce their energy use by the same proportion. Let λ_ξ^τ be the tax associated with a proportional energy use reduction of $1 - \xi_\tau$. The first-order condition for firm 1 is then

$$\underbrace{R^{1-\rho} P^{*\rho} \rho \left[\phi \sum_{i=1}^{n} \delta^{\frac{i-1}{1-\alpha}} (k_i^*)^{\alpha_k} (l_i^*)^{\alpha_l} \right]^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l - 1}}_{\text{FOC Unregulated Case}}$$
$$\times (\xi_\tau)^{\alpha_l - 1} \left(\frac{P^\tau}{P^*}\right)^{\rho} [\xi_\tau^{\alpha_l}]^{\rho-1} = w + p_e + \lambda_\xi^\tau,$$

where we use the superscript τ to denote this case. Using the fact that the first-order condition in the unregulated case equals $w + p_e$, we obtain

$$\left[1 + \frac{\lambda_\xi^\tau}{w + p_e}\right] = \left(\frac{P^\tau}{P^*}\right)^{\rho} (\xi_\tau)^{-(1-\alpha_l \rho)}.$$

Since all related firms reduce their energy use by the same proportion, it follows that $e(\phi, n) = \xi_\tau e^*(\phi, n)$ for regulated firms. Arguments similar to the derivation for Equation J.9 imply that

$$\frac{E^\tau}{E^*} = (1 - s_{\tilde{\phi}}) \left(\frac{P^\tau}{P^*}\right)^{\frac{\rho}{1-\alpha_l \rho}} + s_{\tilde{\phi}} \xi_\tau. \tag{K.3}$$

Noting that $q(\phi, n) = \xi_\tau^{\alpha_l} q^*(\phi, n)$ then implies that

$$\left(\frac{P^\tau}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) \left(\frac{P^\tau}{P^*}\right)^{\frac{\alpha_l \rho^2}{1-\alpha_l \rho}} + s_{\tilde{\phi}} (\xi_\tau)^{\alpha_l \rho}. \tag{K.4}$$

To make this case comparable to that of the Top 1,000 regulation, we implement a tax that leads to the same average energy reduction:

$$\left[1 + \frac{\lambda_\xi^\tau}{w + p_e}\right] = \left(\frac{P^\tau}{P^*}\right)^\rho \left(\mathbb{E}_e \left[\xi_{e,n} \middle| \phi > \tilde{\phi}\right]\right)^{-(1-\alpha_l \rho)};$$

that is, $\xi_\tau = \mathbb{E}_e \left[\xi_{e,n} \middle| \phi > \tilde{\phi}\right]$. Note that the aggregate effects differ to the extent that we obtain different price responses (and therefore different responses from unregulated firms).

L Inspecting the Effect Mechanisms of the Top 1,000 Regulation

Shutting down Market Leakage

In this case, firms believe that prices do not adjust. The perceived shadow cost is given by the solution to

$$\left[1 + \frac{\lambda_n}{w + p_e}\right]^{\frac{1}{1-\rho}} \left[1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + p_e}\right]^{\frac{\alpha_l}{1-\alpha_l}}\right] = \xi^{\frac{-(1-\alpha_l \rho)}{1-\rho}} \Delta_n^{\frac{1}{1-\alpha}}.$$

We recompute $\xi_{e,n}$ and $\xi_{q,n}$ based on these shadow costs. Aggregate energy use is then given by

$$\frac{E}{E^*} = (1 - s_{\tilde{\phi}}) + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{e,n} \middle| \phi > \tilde{\phi}\right].$$

The new price is given by

$$\left(\frac{P}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{q,n}^\rho \middle| \phi > \tilde{\phi}\right].$$

The actual shadow cost follows Equation 14 by incorporating the equilibrium price adjustment.

Shutting down Conglomerate Leakage

In this case, we set $e_1(\phi,n) \leq \xi e_1^*(\phi,n)$. We further assume that firms related to regulated firms do not take into account the reduction in $e_1(\phi,n)$ but do respond to the market price increase, so that $e_i(\phi,n) = \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} e_i^*(\phi,n)$ for $i \geq 2$. We then have

$$\begin{aligned} e(\phi,n) &= \xi e_1^*(\phi,n) + \sum_{i=2} e_i^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} = e_1^*(\phi,n)\left(\xi + \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)\right) \\ &= e^*(\phi,n)\frac{\xi + \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}}. \end{aligned}$$

Aggregate energy use is then

$$\frac{E}{E^*} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[\left.\frac{\xi + \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}}\right| \phi > \tilde{\phi}\right].$$

Similarly, the effect on total production is

$$\begin{aligned} q(\phi,n) &= \xi^{\alpha_l} q_1^*(\phi,n) + \sum_{i=2} q_i^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}} = q_1^*(\phi,n)\left(\xi^{\alpha_l} + \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)\right) \\ &= q^*(\phi,n)\frac{\xi^{\alpha_l} + \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}}. \end{aligned}$$

Aggregate prices are then

$$\left(\frac{P}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[\left.\left(\frac{\xi^{\alpha_l} + \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}}\right)^{\rho}\right| \phi > \tilde{\phi}\right].$$

We now derive the shadow cost of the regulation for the Top 1,000 firm. Substituting l_i into the first-order condition for firm 1 implies that

$$\underbrace{R^{1-\rho}P^{*\rho}\rho\left[\phi\sum_{i=1}^{n}\delta^{\frac{i-1}{1-\alpha}}(k_i^*)^{\alpha_k}(l_i^*)^{\alpha_l}\right]^{\rho-1} \times \phi\alpha_l(k_1^*)^{\alpha_k}(l_1^*)^{\alpha_l-1}}_{\text{FOC Unregulated Case}}$$

$$\times (\xi)^{\alpha_l-1}\left(\frac{P}{P^*}\right)^{\rho}\left[\frac{\xi^{\alpha_l} + \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}\left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}}\right]^{\rho-1} = w + p_e + \lambda(\phi).$$

Using the fact that the first-order condition in the unregulated case equals $w + p_e$, we obtain

$$\left[1 + \frac{\lambda_n}{w + p_e}\right] = \left(\frac{P}{P^*}\right)^{\rho} \xi^{-(1-\alpha_l)} \left(\frac{\xi^{\alpha_l} + \left(\frac{P}{P^*}\right)^{\frac{\alpha_l \rho}{1-\alpha_l \rho}} \left(\Delta_n^{\frac{1}{1-\alpha}} - 1\right)}{\Delta_n^{\frac{1}{1-\alpha}}} \right)^{-(1-\rho)} .$$

Shutting down Both Market and Conglomerate Leakage

In this case, energy use and production at regulated conglomerates are the same as in the case in which only the conglomerate leakage is shut down. Unregulated firms assume that there will be no price increase, such that the aggregate energy use is then

$$\frac{E}{E^*} \;=\; (1 - s_{\tilde{\phi}}) + s_{\tilde{\phi}} \mathbb{E}_e \left[\left. \frac{\xi - 1 + \Delta_n^{\frac{1}{1-\alpha}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right| \phi > \tilde{\phi} \right] .$$

Aggregate prices are then

$$\left(\frac{P}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) + s_{\tilde{\phi}} \mathbb{E}_e \left[\left. \left(\frac{\xi^{\alpha_l} - 1 + \Delta_n^{\frac{1}{1-\alpha}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right)^{\rho} \right| \phi > \tilde{\phi} \right] .$$

The actual shadow cost follows Equation 14 using the definition of $\xi_q(\phi) = \frac{\xi - 1 + \Delta_n^{\frac{1}{1-\alpha}}}{\Delta_n^{\frac{1}{1-\alpha}}}$ and the equilibrium price above.

M Model Extensions

M.1 Endogenous Energy Efficiency

Assume that the conglomerate can improve energy efficiency at firm i, ν_i, by spending $l_i c(\nu_i)$, where $c'(\nu_i) > 0$ and $c''(\nu_i) \geq 0$. The conglomerate's problem is then

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^n, \{\nu_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho} - \sum_{i=1}^{n} l_i \left(w + \frac{p_e}{\nu_i} + c(\nu_i) \right) \right\},$$

where we omit the cost of fixed capital. In the absence of the regulation, the conglomerate sets $c'(\nu^*) \nu^{*2} = p_e$ for all firms, so that Proposition 1 continues to hold.

We assume that $c(\nu) = \frac{\nu^{\gamma}}{1+\gamma}$, where $\gamma \geq 1$, so the effective price of energy is $(\nu^*)^{\gamma}$. Additionally, note that the Top 1,000 regulation does not impact the choice of ν_i for unregulated firms. Using these results and the fact that $\nu_i = \frac{l_i}{e_i}$, we can restate the conglomerate problem as

$$\pi(\phi, n) = \max_{\{l_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho} - (w + (\nu^*)^{\gamma}) \sum_{i=1}^{n} l_i - l_1 \left[\frac{1}{1+\gamma} \left(\frac{l_1}{\xi e_1^*} \right)^{\gamma} - (\nu^*)^{\gamma} \right] \right\},$$

where we substitute the regulatory constraint into the cost of energy efficiency and abstract from the cost of the regulated energy.

Deriving the Shadow Cost of the Regulation

The conglomerate's first-order conditions for l_i $(1 \leq i \leq n)$, i.e., $\frac{\partial \pi}{\partial l_i}$, are

$$R^{1-\rho} P^{\rho} \rho \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} \right]^{\rho-1} \phi^* \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \alpha_l (l_i)^{\alpha_l - 1} = w + (\nu^*)^{\gamma} + \left[\left(\frac{l_1}{\xi e_1^*} \right)^{\gamma} - (\nu^*)^{\gamma} \right] \mathbb{I}[i=1].$$

We can write the binding energy use constraint as $\xi e_1^* = e_1 = \frac{l_1}{\nu_1}$ so that $\nu_1 = \frac{l_1}{\xi e_1^*}$. It is also useful to write the share of variable input costs accounted for by energy as $s_e = \frac{(\nu^*)^{\gamma}}{w + (\nu^*)^{\gamma}}$. The shadow cost of the policy as a fraction of variable inputs is then

$$\frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} = \frac{1}{w + (\nu^*)^{\gamma}} \left[\left(\frac{l_1}{\xi e_1^*} \right)^{\gamma} - (\nu^*)^{\gamma} \right] = s_e \left(\left(\frac{\nu_1}{\nu^*} \right)^{\gamma} - 1 \right).$$

Using these expressions, we can then write the ratios of these first-order conditions between $j \geq 2$ and the Top 1,000 firm as

$$\frac{l_j}{l_1} = \delta^{\frac{j-1}{1-\alpha}} \left[1 + s_e \left(\left(\frac{\nu_1}{\nu^*} \right)^{\gamma} - 1 \right) \right]^{\frac{1}{1-\alpha_l}} = \delta^{\frac{j-1}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} \right]^{\frac{1}{1-\alpha_l}},$$

which confirms that the results of Proposition 3 extend to the case with endogenous energy efficiency. We can then write the first-order condition for the Top 1,000 firm as

$$\underbrace{R^{1-\rho} (P^*)^{\rho} \rho \left[\phi^* \sum_{i=1}^{n} \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} (l_i^*)^{\alpha_l} \right]^{\rho-1} \phi^* \alpha_l (l_1^*)^{\alpha_l - 1}}_{\text{FOC Unregulated}} \left(\frac{P}{P^*} \right)^{\rho} \tag{M.1}$$

$$\left[\underbrace{\left(\frac{\xi \nu_1}{\nu^*} \right)^{\alpha_l} \frac{\left[1 + \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} \right]^{\frac{\alpha_l}{1-\alpha_l}} \sum_{j>1} \delta^{\frac{j-1}{1-\alpha}} \right]}{\sum_i \delta^{\frac{j-1}{1-\alpha}}}}_{\xi_{q,n}^{\nu}} \right]^{\rho-1} \times \left(\frac{\xi \nu_1}{\nu^*} \right)^{\alpha_l - 1} = (w + (\nu^*)^{\gamma}) \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} \right].$$

This equation also defines the relevant decrease in conglomerate-level production $\xi_{q,n}^{\nu}$, which we use to solve the new product market equilibrium (cf. Propositions 4–5). Noting that the first-order condition for the unregulated firm equals $w + (\nu^*)^{\gamma}$, we can derive the following expression

$$\left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} \right]^{\frac{1}{1-\rho}} \left[1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^{\gamma}} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right] = \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\rho}} \left(\frac{\xi \nu_1}{\nu^*} \right)^{\frac{-(1-\alpha_l \rho)}{1-\rho}} \Delta_n^{\frac{1}{1-\alpha}}. \tag{M.2}$$

This equation has one difference from Equation J.4. Since conglomerates can avoid the impact of the regulation by increasing the energy efficiency of the Top 1,000 firm, the effective regulation is now $\frac{\xi \nu_1}{\nu^*}$. That is, conglomerates have less of a need to reduce their energy use if they increase ν_1. This also implies that $\lambda(\phi)$ is decreasing in ν_1. Similarly to in Equation J.4, note that, in this equation, the ratio $\frac{\nu_1}{\nu^*}$ is an implicit function of quantities that are common for firms with the same number of affiliates n. This implies that firms with different values of ϕ have the same shadow cost and improvement to energy efficiency as long as they belong to a conglomerate of the same size. We then write λ_n and $\nu_{1,n}$ to signify the dependence of these variables on n.

Aggregate Energy Use

Prior to the regulation, we have that $e^*(\phi,n) = \frac{l^*(\phi,n)}{\nu^*}$ for both regulated and unregulated firms. For unregulated firms, we still have $e(\phi,n) = \frac{l(\phi,n)}{\nu^*}$ since these firms do not change their investment in energy efficiency. However, for regulated conglomerates, we have

$$\begin{aligned} e(\phi,n) &= \frac{l_1}{\nu_{1,n}} + \frac{1}{\nu^*}\sum_{i>1} l_i = \frac{l_1}{\nu_{1,n}}\left(1+\frac{\nu_{1,n}}{\nu^*}(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+(\nu^*)^{\gamma}}\right]^{\frac{1}{1-\alpha_l}}\right)\\ &= e_1(\phi,n)\left(1+\frac{\nu_{1,n}}{\nu^*}(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+(\nu^*)^{\gamma}}\right]^{\frac{1}{1-\alpha_l}}\right)\\ &= e_1^*(\phi,n)\xi\left(1+\frac{\nu_{1,n}}{\nu^*}(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+(\nu^*)^{\gamma}}\right]^{\frac{1}{1-\alpha_l}}\right)\\ &= e^*(\phi,n)\underbrace{\left[\xi\frac{1+\frac{\nu_{1,n}}{\nu^*}(\Delta_n^{\frac{1}{1-\alpha}}-1)\left[1+\frac{\lambda_n}{w+(\nu^*)^{\gamma}}\right]^{\frac{1}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}}\right]}_{\equiv\xi^{\nu}_{e,n}}. \end{aligned}$$

The term $\xi^{\nu}_{e,n}$ incorporates the insight that the shifting of production to related firms now leads to a larger increase in total energy use since these firms do not improve their energy efficiency, i.e., $\frac{\nu_{1,n}}{\nu^*} > 1$. Using this term in Equation J.9 yields the aggregate change in energy use when firms can respond to the regulation by improving their energy efficiency.

Relation to Empirical Measures of Energy Efficiency

We now discuss how we connect the model to our difference-in-differences estimate of the effect of the Top 1,000 program on the energy efficiency of regulated firms. In the data, we measure energy efficiency as $\frac{R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-1}q_1(\phi,n)}{e_1(\phi,n)}$. Note that, for Top 1,000 firms,

$$\frac{R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-1}q_1(\phi,n)}{e_1(\phi,n)} = \frac{R^{1-\rho}P^{*\rho}(q(\phi,n)^*)^{\rho-1}q_1^*(\phi,n)}{e_1^*(\phi,n)} \times \frac{\left(\frac{P}{P^*}\right)^{\rho}(\xi^{\nu}_{q,n})^{\rho-1}\left(\xi\frac{\nu_{1,n}}{\nu^*}\right)^{\alpha_l}}{\xi}.$$

Note also that the first term after the equation is the energy efficiency prior to the regulation. We can then manipulate the second term using Equation M.1 as follows:

$$\frac{\left(\frac{P}{P^*}\right)^{\rho}(\xi^{\nu}_{q,n})^{\rho-1}\left(\xi\frac{\nu_{1,n}}{\nu^*}\right)^{\alpha_l}}{\xi} = \left(\frac{\nu_{1,n}}{\nu^*}\right)\left[1+s_e\left(\left(\frac{\nu_{1,n}}{\nu^*}\right)^{\gamma}-1\right)\right].$$

The log time difference in energy efficiency for a given regulated firm is then

$$\ln\left(\frac{\nu_{1,n}}{\nu^*}\right) + \ln\left[1 + s_e\left(\left(\frac{\nu_{1,n}}{\nu^*}\right)^{\gamma} - 1\right)\right].$$

Note that the second term in this equation is equal to $\ln\left[1 + \frac{\lambda_n}{w+(\nu^*)^{\gamma}}\right]$.

Since unregulated firms do not have an incentive to invest in energy efficiency, their energy efficiency depends only on the output price. Recall from above that we have

$$\frac{e_1(\phi,n)}{e_1^*(\phi,n)} = \frac{e(\phi,n)}{e^*(\phi,n)} = \left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} \quad \text{and} \quad \frac{q_1(\phi,n)}{q_1^*(\phi,n)} = \frac{q(\phi,n)}{q^*(\phi,n)} = \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}.$$

We then have

$$\begin{aligned}
\frac{R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-1}q_1(\phi,n)}{e_1(\phi,n)} &= \frac{R^{1-\rho}(P^*)^{\rho}(q(\phi,n)^*)^{\rho-1}q_1^*(\phi,n)\left(\frac{P}{P^*}\right)^{\rho}\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}}{e_1^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}} \\
&= \frac{R^{1-\rho}(P^*)^{\rho}(q(\phi,n)^*)^{\rho-1}q_1^*(\phi,n)}{e_1^*(\phi,n)}.
\end{aligned}$$

That is, the Top 1,000 program does not impact the energy efficiency of unregulated firms.

Letting β^{EE} denote the difference-in-differences estimate of the effect of the Top 1,000 program on energy efficiency, we then have

$$\beta^{EE} = \mathbb{E}\left[\ln\left(\frac{\nu_{1,n}}{\nu^*}\right)\middle|\, \phi > \tilde{\phi}\right] + \mathbb{E}\left[\ln\left[1 + s_e\left(\left(\frac{\nu_{1,n}}{\nu^*}\right)^{\gamma} - 1\right)\right]\middle|\, \phi > \tilde{\phi}\right]. \tag{M.3}$$

Calibration of γ

For a range of values of γ, we compute the following:

1. Solve for the values of $\{\nu_{1,n}\}_{n\geq\tilde{n}}$ and $\frac{P}{P^*}$ that jointly satisfy Equation M.2 and

$$\left(\frac{P}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\mathbb{E}_e\left[(\xi_{q,n}^{\nu})^{\rho}\,\middle|\,\phi > \tilde{\phi}\right].$$

2. Implement the right-hand side of Equation M.3.

We then choose the value of γ that is consistent with our empirical estimates of β^{EE}. Since we estimate zero or negative values for β^{EE}, we can bound γ by choosing the value that implies the upper bound of the confidence interval of β^{EE}.

M.2 Heterogeneous Energy Efficiency

We now explore the possibility that regulated firms might differ in their energy efficiency from other firms in the economy. We assume that regulated firms have energy efficiency ν_1, that related firms in the same conglomerate have ν_R, and that other firms in the economy have ν_O.

Unregulated Conglomerates

Firms in these conglomerates face an effective price for variable inputs of $w + \frac{p_e}{\nu_O}$. The results of Propositions 1–3 continue to hold for these firms. We therefore have that in this case

$$\begin{aligned} q^*(\phi, n) &= (\phi\Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_O}} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}} \\ &= (\phi\Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_1}} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q} (d_O)^{\frac{-\alpha_l}{1-\alpha\rho}}, \end{aligned}$$

where $d_O = 1 + s_e \frac{\nu_1 - \nu_O}{\nu_O}$ and $s_e = \frac{\frac{p_e}{\nu_1}}{w + \frac{p_e}{\nu_1}}$ is the share of energy in variable inputs for Top 1,000 firms. The optimal choice of l_1 is now

$$l_1 = \left[\frac{R^{1-\rho} P^{\rho} \rho [\phi (k_1^*)^{\alpha_k}]^{\rho} \Delta_n^{\frac{\rho-1}{1-\alpha}} \alpha_l}{w + \frac{p_e}{\nu_O}} \right]^{\frac{1}{1-\alpha_l\rho}} \quad \text{so that} \quad l_1 = l_1^* \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}}.$$

That is, this difference in energy efficiency does not impact how unregulated firms respond to changes in the output price in terms of their use of intermediate inputs. Moreover, since $l_1 = \nu_O e_1$ and $l_1^* = \nu_O e_1^*$, we have that $\frac{l_1}{l_1^*} = \frac{e_1}{e_1^*}$. These results then imply that

$$q_1(\phi, n) = q_1^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}} \quad \text{and} \quad q(\phi, n) = q^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}. \tag{M.4}$$

Regulated Conglomerates

Since $\nu_1 \neq \nu_R$, the effective cost of inputs differs across regulated and related firms in the same conglomerate. This difference in input costs influences the within-conglomerate distribution of production. Within a given firm, we have that

$$l_i = \frac{\alpha_l}{\alpha_k} \frac{r}{w + \frac{p_e}{\nu_R}} k_i = \frac{\alpha_l}{\alpha_k} \frac{r}{w + \frac{p_e}{\nu_1}} k_i \frac{1}{1 + s_e \frac{\nu_1 - \nu_R}{\nu_R}} = \frac{\alpha_l}{\alpha_k} \frac{r}{w + \frac{p_e}{\nu_1}} k_i \frac{1}{d_i},$$

where $d_i = 1 + s_e \frac{\nu_1 - \nu_i}{\nu_i}$. The choice of capital across firms is now

$$\pi(\phi, n) = \max_{\{k_i\}_{i=1}^n} \left\{ R^{1-\rho} P^{\rho} \left[\sum_{i=1}^{n} \phi \delta^{i-1} k_i^{\alpha} \left(\frac{\alpha_l}{\alpha_k} \frac{r}{\left(w + \frac{p_e}{\nu_1} \right) d_i} \right)^{\alpha_l} \right]^{\rho} - \left(\frac{\alpha}{\alpha_k} r \right) \sum_{i=1}^{n} k_i \right\}.$$

Comparing the first-order conditions for k_1 and k_i, we find that $\frac{k_i}{k_1} = \delta^{\frac{i-1}{1-\alpha}} d_i^{\frac{-\alpha_l}{1-\alpha}}$. We then have

$$\frac{l_i}{l_1} = \delta^{\frac{i-1}{1-\alpha}} d_i^{\frac{-\alpha_l}{1-\alpha} - 1} = \delta^{\frac{i-1}{1-\alpha}} d_i^{\frac{-(1-\alpha_k)}{1-\alpha}}.$$

Production is then

$$q_i = \phi\delta^{i-1}k_i^{\alpha}\left(\frac{\alpha_l}{\alpha_k}\frac{r}{\left(w+\frac{p_e}{\nu_1}\right)d_i}\right)^{\alpha_l} = \phi\delta^{\frac{i-1}{1-\alpha}}k_1^{\alpha}\left(\frac{\alpha_l}{\alpha_k}\frac{r}{\left(w+\frac{p_e}{\nu_1}\right)}\right)^{\alpha_l} d_i^{\frac{-\alpha_l}{1-\alpha}} = \delta^{\frac{i-1}{1-\alpha}}d_i^{\frac{-\alpha_l}{1-\alpha}}q_1.$$

Let $d_R = d_i$ for $i > 1$, and recall that $d_1 = 1$. Total capital is then

$$K_n = k_1\left(1 + d_R^{\frac{-\alpha_l}{1-\alpha}}\sum_{i>1}\delta^{\frac{i-1}{1-\alpha}}\right) = k_1\left(1 + d_R^{\frac{-\alpha_l}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}} - 1)\right).$$

Define $\hat{\Delta}_n = \left(1 + d_R^{\frac{-\alpha_l}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}} - 1)\right)^{1-\alpha}$. The analysis for the optimal choice of K_n now holds with $\hat{\Delta}_n$ in place of Δ_n. In the case of regulated firms, we have

$$q^*(\phi, n) \;=\; \left(\phi\hat{\Delta}_n\right)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{\alpha}{1-\alpha\rho}}\left[\left(\frac{\alpha_l}{w+\frac{p_e}{\nu_1}}\right)^{\alpha_l}\left(\frac{\alpha_k}{r}\right)^{\alpha_k}\right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q}. \qquad \text{(M.5)}$$

Moreover, the thresholds defining the optimal number of conglomerates in Proposition 2 continue to hold using $\hat{\Delta}_n$. Note, however, that our assumption that energy costs change discontinuously at $\tilde{\phi}$ implies that regulated and unregulated conglomerates have different thresholds ϕ_n. Let $\{\phi_n^O\}$ be the set of size thresholds for unregulated conglomerates and $\{\phi_n^R\}$ be the set of related conglomerates. Let n' be the largest firm size for unregulated firms (so that $\phi_{n'+1}^O > \tilde{\phi}$) and n'' be the smallest size of regulated conglomerates (so that $\phi_{n''}^R < \tilde{\phi}$). The combined set of size thresholds is then $\{\{\phi_n^O\}_{n=1}^{n'}, \tilde{\phi}, \{\phi_n^R\}_{n=n''+1}\}$. We include $\tilde{\phi}$ in this list since it is possible that the change in energy efficiency for regulated firms leads to a change in firm size, though (depending on the differences in energy efficiency) this may not always be the case.

Consider now the response of the firms to the regulation. Using the fact that $k_i^* = k_1^*\delta^{\frac{i-1}{1-\alpha}}d_i^{\frac{-\alpha_l}{1-\alpha}}$, we can write the profit maximization problem as

$$\max_{\{l_i\}_1^n}\left\{R^{1-\rho}P^{\rho}\left[\phi^*\sum_{i=1}^{n}\delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}}l_i^{\alpha_l}d_i^{\frac{-\alpha_l\alpha_k}{1-\alpha}}\right]^{\rho} - \left(w+\frac{p_e}{\nu_1}\right)\sum_{i=1}^{n}d_il_i - r\sum_{i=1}^{n}k_i^*\right\} \text{ subject to } l_1 \leq \xi l_1^*,$$

where $\phi^* = \phi(k_1^*)^{\alpha_k}$. The first-order conditions for l_i $(1 \leq i \leq n)$ are then

$$\underbrace{R^{1-\rho}P^{\rho}}_{\text{Market Demand}}\underbrace{\rho\left[\phi^*\sum_{i=1}^{n}\delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}}l_i^{\alpha_l}d_i^{\frac{-\alpha_l\alpha_k}{1-\alpha}}\right]^{\rho-1}}_{\text{Residual Revenue}}\underbrace{\phi^*\delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}}\alpha_l(l_i)^{\alpha_l-1}d_i^{\frac{-\alpha_l\alpha_k}{1-\alpha}}}_{\text{Marginal Product}} \;=\; \left(w+\frac{p_e}{\nu_1}\right)d_i + \underbrace{\lambda(\phi)\mathbb{I}[i=1]}_{\substack{\text{Shadow Cost}\\ \text{of Regulation}}}.$$

Taking the ratio of the conditions for l_1 and $l_i (i > 1)$, we have

$$\begin{aligned}
\left(\frac{l_i}{l_1}\right)^{1-\alpha_l} &= \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right] d_i^{\frac{-\alpha_l \alpha_k}{1-\alpha} - 1} \\
\left(\frac{l_i}{l_1}\right)^{1-\alpha_l} &= \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right] d_i^{\frac{-(1-\alpha_l)(1-\alpha_k)}{1-\alpha}} \\
\frac{l_i}{l_1} &= \delta^{\frac{(i-1)}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right]^{\frac{1}{1-\alpha_l}} d_i^{\frac{-(1-\alpha_k)}{1-\alpha}}.
\end{aligned}$$

The residual revenue term now becomes

$$\phi (k_1^*)^{\alpha_k} l_1^{\alpha_l} \left[1 + \sum_{i>1} \delta^{\frac{i-1}{1-\alpha}} d_i^{\frac{-\alpha_l}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right]^{\frac{\alpha_l}{1-\alpha_l}}\right] = \phi (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l} \xi^{\alpha_l} \left[1 + (\hat{\Delta}_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right]^{\frac{\alpha_l}{1-\alpha_l}}\right].$$

The first-order condition for the regulated firm is then

$$\underbrace{R^{1-\rho} P^{*\rho} \rho \left[\phi \sum_{i=1}^{n} \delta^{\frac{i-1}{1-\alpha}} (k_i^*)^{\alpha_k} (l_i^*)^{\alpha_l}\right]^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l - 1}}_{\text{FOC Unregulated Case}}$$

$$\times (\xi)^{\alpha_l - 1} \left(\frac{P}{P^*}\right)^{\rho} \left[\underbrace{\xi^{\alpha_l} \frac{1 + (\hat{\Delta}_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)}\right]^{\frac{\alpha_l}{1-\alpha_l}}}{\hat{\Delta}_n^{\frac{1}{1-\alpha}}}}_{\equiv \xi_{q,n}^d}\right]^{\rho-1} = \left(w + \frac{p_e}{\nu_1}\right) + \lambda(\phi),$$

where $\frac{q(\phi,n)}{q^*(\phi,n)} = \xi_{q,n}^d$. Using the fact that the first-order condition in the unregulated case equals $\left(w + \frac{p_e}{\nu_1}\right)$, we obtain

$$1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1}\right)} = (\xi)^{\alpha_l - 1} \left(\frac{P}{P^*}\right)^{\rho} (\xi_{q,n}^d)^{(\rho-1)}. \tag{M.6}$$

Product Market Equilibrium

The price level absent the regulation is then

$$\begin{aligned}
P^{*-\rho} &= R^{-\rho}\int_{\phi_1}^{\tilde{\phi}}\left(\left(\phi\Delta_n\right)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\left(d_O\right)^{\frac{-\alpha_l}{1-\alpha\rho}}\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi \\
&+ R^{-\rho}\int_{\tilde{\phi}_1}\left(\left(\phi\hat{\Delta}_n\right)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi \\
P^{*\frac{-\rho}{1-\alpha\rho}}R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}C_Q^{-\rho}\frac{1-G(\phi_1)}{M} &= \int_{\phi_1}^{\tilde{\phi}}\left(\phi\Delta_n d_O^{-\alpha_l}\right)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi+\int_{\tilde{\phi}}\left(\phi\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi \\
P^{*\frac{-\rho}{1-\alpha\rho}}R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}C_Q^{-\rho}\frac{1-G(\phi_1)}{M} &= \sum_{n=1}^{\tilde{n}-1}(\Delta_n d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}}\pi_n+(\Delta_{\tilde{n}}d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1 \\
&+ (\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2+\sum_{n=\tilde{n}+1}(\hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n.
\end{aligned}$$

After the regulation, the equilibrium is then

$$\begin{aligned}
P^{-\rho} &= R^{-\rho}\int_{\phi_1}^{\tilde{\phi}}\left(\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}\left(\phi\Delta_n\right)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\left(d_O\right)^{\frac{-\alpha_l}{1-\alpha\rho}}\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi \\
&+ R^{-\rho}\int_{\tilde{\phi}}\left(\xi_{q,n}^d\left(\phi\hat{\Delta}_n\right)^{\frac{1}{1-\alpha\rho}}R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\phi_1)}d\phi
\end{aligned}$$

$$\begin{aligned}
P^{-\rho}P^{*\frac{-\rho^2\alpha}{1-\alpha\rho}}R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}C_Q^{-\rho}\frac{1-G(\phi_1)}{M} &= \int_{\phi_1}^{\tilde{\phi}}\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}\left(\phi\Delta_n d_O^{-\alpha_l}\right)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi+\int_{\tilde{\phi}}(\xi_{q,n}^d)^{\rho}\left(\phi\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi \\
P^{-\rho}P^{*\frac{-\rho^2\alpha}{1-\alpha\rho}}R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}}C_Q^{-\rho}\frac{1-G(\phi_1)}{M} &= \sum_{n=1}^{\tilde{n}-1}\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}(\Delta_n d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}}\pi_n+\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}(\Delta_{\tilde{n}}d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1 \\
&+ (\xi_{q,\tilde{n}}^d)^{\rho}(\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2+\sum_{n=\tilde{n}+1}(\xi_{q,n}^d)^{\rho}(\hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n.
\end{aligned}$$

Comparing the regulated and unregulated equilibrium conditions, we then have

$$\left(\frac{P}{P^*}\right)^{-\rho}=(1-s_{\tilde{\phi}}^d)\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}}+s_{\tilde{\phi}}^d\mathbb{E}_e\left[(\xi_{q,n}^d)^{\rho}\middle|\phi>\tilde{\phi}\right],$$

where $s_{\tilde{\phi}}^d=\frac{(\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2+\sum_{n=\tilde{n}+1}(\hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n}{(d_O)^{\frac{-\alpha_l\rho}{1-\alpha\rho}}\left[\sum_{n=1}^{\tilde{n}-1}(\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n+(\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1\right]+(\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_2+\sum_{n=\tilde{n}+1}(\hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n}$ and where $\mathbb{E}_e$ takes the expectation of the distribution of energy use among regulated firms. This case differs from our baseline case in that we use the $\hat{\Delta}_n$ expressions to calculate $\mathbb{E}_e$.

Aggregate Energy Use

Total intermediate inputs for a regulated conglomerate are

$$\frac{l(\phi,n)}{l^*(\phi,n)} = \xi \frac{\left[1+(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\left[1+\frac{\lambda(\phi)}{\left(w+\frac{p_e}{\nu_1}\right)}\right]^{\frac{1}{1-\alpha_l}}\right]}{1+(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}}.$$

To compute total energy use, however, we need to take into account differences in energy efficiency across firms in the conglomerate. In the absence of the regulation, we have

$$e^*(\phi,n) = \frac{l_1^*}{\nu_1} + \frac{1}{\nu_R}\sum_{i>1} l_i^* = \frac{1}{\nu_1}\left(l_1^* + \frac{\nu_1}{\nu_R}\sum_{i>1} l_i^*\right) = \frac{1}{\nu_1} l_1^*\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right).$$

After the regulation, we have

$$e(\phi,n) = \frac{1}{\nu_1} l_1\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\left[1+\frac{\lambda(\phi)}{\left(w+\frac{p_e}{\nu_1}\right)}\right]^{\frac{1}{1-\alpha_l}}\right).$$

We then have

$$\frac{e(\phi,n)}{e^*(\phi,n)} = \xi \frac{1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\left[1+\frac{\lambda(\phi)}{\left(w+\frac{p_e}{\nu_1}\right)}\right]^{\frac{1}{1-\alpha_l}}}{1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}} \equiv \xi_{e,n}^d.$$

To obtain an expression for $e^*(\phi,n)$ for regulated firms, recall that

$$\begin{array}{rcl} l^*(\phi,n) & = & l_1^*\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right) \\ l^*(\phi,n) & = & \left(\phi\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}} R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \underbrace{\rho^{\frac{1}{1-\alpha\rho}}\left[\left(\frac{\alpha_l}{w+\frac{p_e}{\nu_1}}\right)^{1-\alpha_k\rho}\left(\frac{\alpha_k}{r}\right)^{\alpha_k\rho}\right]^{\frac{1}{1-\alpha\rho}}}_{=C_E} \end{array}$$

so that

$$e^*(\phi,n) = \left(\phi\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}} R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} C_E \frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right)},$$

where the last terms adjust for the composition of energy use across establishments with different energy efficiency.

Consider now the unregulated firms. Since all firms in unregulated conglomerates have the same energy efficiency, we have

$$\frac{e(\phi,n)}{e^*(\phi,n)}=\frac{l(\phi,n)}{l^*(\phi,n)}=\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}$$

and

$$e^*(\phi,n)=\frac{l^*(\phi,n)}{\nu_O}=\frac{1}{\nu_O}\left(\phi\Delta_n\right)^{\frac{\rho}{1-\alpha\rho}}R^{\frac{1-\rho}{1-\alpha\rho}}P^{*\frac{\rho}{1-\alpha\rho}}\underbrace{\rho^{\frac{1}{1-\alpha\rho}}\left[\left(\frac{\alpha_l}{w+\frac{p_e}{\nu_1}}\right)^{1-\alpha_k\rho}\left(\frac{\alpha_k}{r}\right)^{\alpha_k\rho}\right]^{\frac{1}{1-\alpha\rho}}}_{=C_E}\left(d_O\right)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}.$$

The total energy use prior to the regulation is then

$$\begin{aligned}E^* &= \int\limits_{\phi_1} e^*(\phi,n)\frac{g(\phi)M}{1-G(\phi_1)}d\phi\\ &= R^{\frac{1-\rho}{1-\alpha\rho}}P^{*\frac{\rho}{1-\alpha\rho}}\frac{C_E M}{1-G(\phi_1)}\left[\frac{(d_O)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}}{\nu_O}\left(\sum_{n=1}^{\tilde{n}-1}(\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n+(\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1\right)\right.\\ &+ \left.\left(\hat{\Delta}_{\tilde{n}}\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)\right)}\tilde{\pi}_2+\sum_{n=\tilde{n}+1}\left(\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right)}\pi_n\right]\end{aligned}$$

The total energy use after the regulation is then

$$\begin{aligned}E &= \int\limits_{\phi_1} e(\phi,n)\frac{g(\phi)M}{1-G(\phi_1)}d\phi\\ &= R^{\frac{1-\rho}{1-\alpha\rho}}P^{*\frac{\rho}{1-\alpha\rho}}\frac{C_E M}{1-G(\phi_1)}\left[\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}\frac{(d_O)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}}{\nu_O}\left(\sum_{n=1}^{\tilde{n}-1}(\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\pi_n+(\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\tilde{\pi}_1\right)\right.\\ &+ \xi^d_{e,\tilde{n}}\left(\hat{\Delta}_{\tilde{n}}\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)\right)}\tilde{\pi}_2\\ &+ \left.\sum_{n=\tilde{n}+1}\xi^d_{e,n}\left(\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right)}\pi_n\right].\end{aligned}$$

The change in aggregate energy use is then

$$\frac{E}{E^*}=(1-s^e_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}+s^e_{\tilde{\phi}}\mathbb{E}_e\left[\xi^d_{e,n}\middle|\phi>\tilde{\phi}\right],$$

where

$$\frac{1-s^e_{\tilde{\phi}}}{s^e_{\tilde{\phi}}} = \frac{\frac{(d_O)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}}{\nu_O}\left(\sum_{n=1}^{\tilde{n}-1}(\Delta_n)^{\frac{\rho}{1-\alpha\rho}}\,\pi_n + (\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}}\,\tilde{\pi}_1\right)}{\left(\hat{\Delta}_{\tilde{n}}\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)\right)}\tilde{\pi}_2 + \sum_{n=\tilde{n}+1}\left(\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right)}\pi_n}$$

and where we evaluate $\mathbb{E}_e$ using the conditional probabilities

$$\mathbb{P}r\left[n=n'\middle|\phi>\tilde{\phi}\right] = \frac{\left(\hat{\Delta}_{n'}\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_{n'}^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_{n'}^{\frac{1}{1-\alpha}}-1)\right)}\pi_{n'}}{\left(\hat{\Delta}_{\tilde{n}}\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}}-1)\right)}\tilde{\pi}_2 + \sum_{n=\tilde{n}+1}\left(\hat{\Delta}_n\right)^{\frac{\rho}{1-\alpha\rho}}\frac{\frac{1}{\nu_1}\left(1+\frac{\nu_1}{\nu_R}(\Delta_n^{\frac{1}{1-\alpha}}-1)d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}\right)}{\left(1+d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}(\Delta_n^{\frac{1}{1-\alpha}}-1)\right)}\pi_n}.$$

M.3 Imperfect Substitution within Conglomerates

We now consider the possibility that the outputs of affiliates within the same conglomerate are not perfect substitutes. In particular, we assume that the conglomerate-level composite good $q(\omega)$ can be represented as affiliate output, such that

$$q(\omega) = \left(\sum_i q(\omega,i)^{\rho_c}\right)^{1/\rho_c}, \quad \text{where } \ 0<\rho<\rho_c<1.$$

The residual demand for the i^{th} affiliate of conglomerate ω, i.e., (ω,i), is

$$p(\omega,i) = R^{1-\rho}P^{\rho}q(\omega)^{\rho-\rho_c}q(\omega,i)^{\rho_c-1}.$$

The profit maximization problem for the conglomerate is now

$$\begin{aligned} & \max_{\{l_i\}_{i=1}^n,\{k_i\}_{i=1}^n} R^{1-\rho}P^{\rho}q(\phi)^{\rho-\rho_c}\sum_i q(\phi,i)^{\rho_c} - (w+p_e)\sum_i l_i - r\sum_i k_i \\ = \quad & \max_{\{l_i\}_{i=1}^n,\{k_i\}_{i=1}^n} R^{1-\rho}P^{\rho}\left[\sum_i q(\phi,i)^{\rho_c}\right]^{\rho/\rho_c} - (w+p_e)\sum_i l_i - r\sum_i k_i, \end{aligned}$$

where $q(\phi,i)=\phi\delta^{i-1}k_i^{\alpha_k}l_i^{\alpha_l}$. The first-order conditions for $l_1,\ldots l_n$ imply that $\frac{l(\phi,i)}{l(\phi,1)} = \left(\frac{q(\phi,i)}{q(\phi,1)}\right)^{\rho_c}$. Since the capital–labor ratio remains $\frac{l_i}{k_i} = \frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}$, substituting into $q(\phi,i)$ allows us to express the first-order conditions as

$$\frac{k_i}{k_1} = \frac{l_i}{l_1} = \frac{e_i}{e_1} = \delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}} \quad \text{and} \quad \frac{q_i}{q_1} = \delta^{\frac{(i-1)}{1-\alpha\rho_c}}.$$

These expressions reduce to Proposition 1 when $\rho_c = 1$. To understand how values of $\rho_c < 1$ impact the use of inputs within the conglomerate, note that, in the exponent for the first expression, a lower value of ρ_c decreases the numerator and increases the denominator. Both forces work to reduce the magnitude of the exponent such that the use of inputs depreciates more slowly in firm rank given the same δ.

We now define $\Delta_n^C = \left[\sum_{i=1}^n (\delta^{(i-1)})^{\frac{\rho_c}{1-\alpha\rho_c}}\right]^{\frac{1-\alpha\rho_c}{\rho_c}}$, so that $\frac{k_1}{K_n} = \frac{l_1}{L_n} = \frac{1}{(\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}}}$. Total composite output is then

$$\begin{aligned} q(\phi)^{\rho_c} &= \left(\frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}\right)^{\alpha_l\rho_c} \phi^{\rho_c} \sum_i (\delta^{(i-1)\rho_c})(\delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}} k_1)^{\alpha\rho_c} \\ &= \left(\frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}\right)^{\alpha_l\rho_c} \phi^{\rho_c} k_1^{\alpha\rho_c} (\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}} = \left(\frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}\right)^{\alpha_l\rho_c} (\phi\Delta_n^C)^{\rho_c} K_n^{\alpha\rho_c}. \end{aligned}$$

We can then rewrite the optimization problem as

$$\max_{K_n} R^{1-\rho} P^{\rho} \left(\frac{\alpha_l}{\alpha_k}\frac{r}{w+p_e}\right)^{\rho\alpha_l} (\phi\Delta_n^C)^{\rho} K_n^{\alpha\rho} - \left(\frac{\alpha_l}{\alpha_k} r + r\right) K_n,$$

where we represent total variable costs as $(w+p_e)L_n = \left(\frac{\alpha_l}{\alpha_k} r\right) K_n$. Note that this optimization problem is exactly the same as our in baseline model except for the new definition of Δ_n^C. The within-conglomerate revenue share is now $\left(\frac{\delta^{i-1}}{\Delta_n^C}\right)^{\frac{\rho_c}{1-\alpha\rho_c}}$, which equals the share in the case of perfect substitution when $\rho_c = 1$. Similarly, the results of Proposition 2 continue to hold with the new Δ_n^C.

Regulated Conglomerates

Given ϕ and k_i^*, we can write $q(\phi)^{\rho_c}$ as

$$\phi^{\rho_c} \sum_i (\delta^{(i-1)\rho_c})(\delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}} k_1^*)^{\alpha_k\rho_c} l_i^{\alpha_l\rho_c} \equiv (\phi (k_1^*)^{\alpha_k})^{\rho_c} \sum_i (\delta^{(i-1)\rho_c})^{\frac{(1-\alpha_l\rho_c)}{1-\alpha\rho_c}} l_i^{\alpha_l\rho_c}.$$

We can similarly define the short-run response as

$$\max_{\{l_i\}_{i=1}^n} R^{1-\rho} P^{\rho} \left[\phi^* \left(\sum_i \delta^{\frac{(i-1)\rho_c(1-\alpha_l\rho_c)}{1-\alpha\rho_c}} l_i^{\alpha_l\rho_c}\right)^{1/\rho_c}\right]^{\rho} - (w+p_e)\sum_i l_i.$$

The first-order condition in this case becomes

$$R^{1-\rho} P^{\rho} \rho \left[\phi^* \left(\sum_i \delta^{\frac{(i-1)\rho_c(1-\alpha_l\rho_c)}{1-\alpha\rho_c}} l_i^{\alpha_l\rho_c}\right)^{1/\rho_c}\right]^{\rho-1} \times \phi^* \delta^{\frac{(i-1)\rho_c(1-\alpha_l\rho_c)}{1-\alpha\rho_c}} \alpha_l\rho_c (l_i)^{\alpha_l\rho_c - 1} = w + p_e + \lambda(\phi) I[i=1].$$

The within-conglomerate input allocation is now

$$\frac{l_j}{l_2} = \delta^{\frac{(j-2)\rho_c}{1-\alpha\rho_c}}, j > 2 \quad \text{and} \quad \frac{l_i}{l_1} = \delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}} \left[1 + \frac{\lambda(\phi)}{w+p_e}\right]^{\frac{1}{1-\alpha_l\rho_c}}.$$

The conglomerate composite output is then

$$q(\phi,n) = q_1(\phi,n)\left[1+\sum_{i>1}\delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}}\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l\rho_c}{1-\alpha_l\rho_c}}\right]^{1/\rho_c}.$$

Recalling that the original pre-regulation optimal composite output is

$$q^*(\phi,n) = q_1^*(\phi,n)\left[\sum_{i}^{n}\delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}}\right]^{1/\rho_c} \equiv q_1^*(\phi,n)(\Delta_n^C)^{\frac{1}{1-\alpha\rho_c}}$$

and using the fact that $q_1(\phi,n) = q_1^*(\phi,n)\xi^{\alpha_l}$, we obtain

$$q(\phi,n) = q^*(\phi,n)\underbrace{\left[\xi^{\alpha_l}\frac{\left(1+((\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}}-1)\left[1+\frac{\lambda(\phi)}{w+p_e}\right]^{\frac{\alpha_l\rho_c}{1-\alpha_l\rho_c}}\right)^{1/\rho_c}}{(\Delta_n^C)^{\frac{1}{1-\alpha\rho_c}}}\right]}_{\xi_{q,n}^C(\phi)}.$$

Using similar arguments from our baseline model, we have that the shadow cost of the regulation takes the form

$$1+\frac{\lambda(\phi)}{w+p_e} = (\xi)^{\alpha_l\rho_c-1}\left(\frac{P}{P^*}\right)^{\rho}(\xi_{q,n}^C)^{\rho-1}.$$

As in previous cases, the shadow cost depends only on n. Unregulated conglomerates are not affected by imperfect substitution within regulated conglomerates. We thus still have that

$$l_1 = l_1^*\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}$$

and

$$q(\phi,n) = q^*(\phi,n)\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}.$$

Product Market Equilibrium

We now again define the industry price index in terms of quantity of both regulated and unregulated conglomerates. The equilibrium price is defined by

$$P = \left[\int_{\phi_1} p(\phi,n)^{\frac{\rho}{\rho-1}}\frac{M}{1-G(\phi_1)}dG\right]^{\frac{\rho-1}{\rho}} \quad \text{where} \quad p(\phi,n)^{\frac{\rho_c}{\rho_c-1}} = \sum_{i=1}^{n} p(\phi,i)^{\frac{\rho_c}{\rho_c-1}}.$$

Substituting the residual demand curve $p(\phi,i) = R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-\rho_c}q(\phi,i)^{\rho_c-1}$, we have the composite conglomerate price index

$$p(\phi,n) = R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-\rho_c}\left(\sum_i q(\phi,i)^{\rho_c}\right)^{\frac{\rho_c-1}{\rho_c}} \equiv R^{1-\rho}P^{\rho}q(\phi,n)^{\rho-1}.$$

Substituting this expression into the aggregate price index P, we have

$$P^{\frac{\rho}{\rho-1}} = \int_{\phi_1} (R^{1-\rho}P^{\rho})^{\frac{\rho}{\rho-1}} q(\phi,n)^{\rho} \frac{M}{1-G(\phi_1)} dG,$$

which is equivalent to $P^{-\rho} = R^{-\rho}\int_{\phi_1} q(\phi,n)^{\rho}\frac{M}{1-G(\phi_1)}dG$. That is, the equation describing the change in equilibrium prices, Equation J.7, continues to hold with the new definition of Δ_n^C and $\xi_{q,n}^C$.

Aggregate Energy Use

Consider first the regulated conglomerate. Using the fact that $e^*(\phi,n) = e_1^*(\phi,n)\sum_i \delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}} = e_1^*(\phi,n)(\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}}$, we have that

$$\begin{aligned} e(\phi,n) &= e_1(\phi,n)\left[1+\sum_{i>1}\delta^{\frac{(i-1)\rho_c}{1-\alpha\rho_c}}\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{1}{1-\alpha_l\rho_c}}\right] \\ &= e^*(\phi,n)\underbrace{\left[\xi\frac{\left(1+((\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}}-1)\left[1+\frac{\lambda_n}{w+p_e}\right]^{\frac{1}{1-\alpha_l\rho_c}}\right)}{(\Delta_n^C)^{\frac{\rho_c}{1-\alpha\rho_c}}}\right]}_{\xi_{e,n}^C}. \end{aligned}$$

The rest of the aggregation results then similarly carry through with $\xi_{e,n}^C$ defined as above.

M.4 Economy with No Conglomerate Spillovers

Initial Equilibrium

We outline a standard model with no conglomerate or associated related-firm spillovers.

Consider a firm with productivity ϕ.

$$\pi(\phi) = \max_{l,k}\left\{R^{1-\rho}P^{\rho}[\phi k^{\alpha_k} l^{\alpha_l}]^{\rho} - (w+p_e)l - rk\right\}.$$

The first-order conditions for l and k imply that $l = \frac{\alpha_l}{\alpha_k}\frac{r}{(w+p_e)}k$.

Substituting this expression, we can write the profit maximization problem as

$$\pi(\phi) = \max_k\left\{R^{1-\rho}P^{\rho}\left[\phi k^{\alpha}\left(\frac{\alpha_l}{\alpha_k}\frac{r}{(w+p_e)}\right)^{\alpha_l}\right]^{\rho} - \left(\frac{\alpha}{\alpha_k}r\right)k\right\}.$$

Defining the same constant $C_\pi = (1-\alpha\rho)\left[\left(\frac{\rho\alpha_l}{w+p_e}\right)^{\alpha_l\rho}\left(\frac{\rho\alpha_k}{r}\right)^{\alpha_k\rho}\right]^{\frac{1}{1-\alpha\rho}}$ as our baseline case, we have that the optimal capital k and the firm profit are

$$k^* = \frac{R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi}{(1-\alpha\rho)}\frac{\rho\alpha_k}{r}(\phi)^{\frac{\rho}{1-\alpha\rho}} \quad \text{and} \quad \pi(\phi) = R^{\frac{1-\rho}{1-\alpha\rho}}P^{\frac{\rho}{1-\alpha\rho}}C_\pi(\phi)^{\frac{\rho}{1-\alpha\rho}}.$$

We now compute the price level. Define the zero-profit cutoff $\pi(\bar{\phi}) = f$; the price level is given by

$$P^{-\rho} = R^{-\rho} \int_{\bar{\phi}} q(\phi)^{\rho} \frac{g(\phi) M}{1 - G(\bar{\phi})} d\phi.$$

We define the optimal output $q^*(\phi)$ as

$$q^*(\phi) \;\; = \;\; \phi \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w + p_e)} \right)^{\alpha_l} k^{*\alpha} = (\phi)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + p_e} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q}.$$

The price level in the absence of the regulation is then

$$\begin{aligned}
P^{*-\rho} &= R^{-\rho} \int_{\bar{\phi}} \left((\phi)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} C_Q \right)^{\rho} \frac{g(\phi) M}{1 - G(\bar{\phi})} d\phi \\
P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} &= \int_{\bar{\phi}} (\phi)^{\frac{\rho}{1-\alpha\rho}} \frac{g(\phi) M}{1 - G(\bar{\phi})} d\phi. \\
P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1 - G(\bar{\phi})}{M} &= \underbrace{\int_{\bar{\phi}} (\phi)^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi}_{=\bar{\pi}}.
\end{aligned}$$

Given the assumption that ϕ follows a log-normal distribution, we can express

$$\bar{\pi} = \exp\left\{ \frac{\tilde{\sigma}^2}{2} \right\} [1 - \Phi(\tilde{\bar{\phi}} - \tilde{\sigma})],$$

where $\tilde{\sigma} = \frac{\rho}{1-\alpha\rho} \sigma_{\phi}$ and $\tilde{\bar{\phi}} = \frac{\rho}{1-\alpha\rho} \frac{\ln(\bar{\phi})}{\tilde{\sigma}}$.

The Top 1,000 program limits energy use at the largest firm e_1 to a fraction $\xi < 1$ of the energy use in the unregulated case e_1^*; recall that we use starred variables to denote the optimal choices in the unregulated case. We assume that the number of firms n and the capital allocations $\{k_i^*\}_{i=1}^n$ are quasi-fixed but that the conglomerate can adjust $\{l_i\}_{i=1}^n$.

Regulated Firms

We can write the profit maximization problem as

$$\max_{l} \left\{ R^{1-\rho} P^{\rho} \left[\phi^* l_i^{\alpha_l} \right]^{\rho} - (w + p_e) l - r k^* \right\} \;\; \text{subject to} \;\; l \leq \xi l^*,$$

where $\phi^* = \phi (k^*)^{\alpha_k}$. The first-order conditions for l is then

$$\frac{\partial \pi}{\partial l} = \underbrace{R^{1-\rho} P^{\rho}}_{\text{Market Demand}} \underbrace{\rho [\phi^* l^{\alpha_l}]^{\rho - 1}}_{\text{Residual Revenue}} \underbrace{\phi^* \alpha_l (l)^{\alpha_l - 1}}_{\text{Marginal Product}} = w + p_e + \underbrace{\lambda(\phi)}_{\substack{\text{Shadow Cost} \\ \text{of Regulation}}}.$$

Similarly to in the baseline case, we maintain the assumption that k^* is fixed and $l = \xi l^*$; we then have that

$$q(\phi) = \phi(k^*)^{\alpha_k}(l)^{\alpha_l} = \phi(k^*)^{\alpha_k}(l^*)^{\alpha_l}\xi^{\alpha_l} = q^*(\phi)\xi^{\alpha_l}.$$

To obtain the shadow cost of regulation, we can rewrite the first-order condition for l in terms of the capital and labor choices in the unregulated case:

$$\underbrace{R^{1-\rho}P^{*\rho}\rho[\phi(k_i^*)^{\alpha_k}(l^*)^{\alpha_l}]^{\rho-1} \times \phi\alpha_l(k^*)^{\alpha_k}(l^*)^{\alpha_l-1}}_{\text{FOC Unregulated Case}}$$
$$\times(\xi)^{\alpha_l-1}\left(\frac{P}{P^*}\right)^{\rho}(\xi^{\alpha_l})^{\rho-1} = w + p_e + \lambda(\phi),$$

where $\frac{P}{P^*}$ is the equilibrium change in the industry-level price. Using the fact that the first-order condition in the unregulated case equals $w + p_e$, we have

$$(\xi)^{\rho\alpha_l-1}\left(\frac{P}{P^*}\right)^{\rho} \quad = \quad 1 + \frac{\lambda(\phi)}{w + p_e}.$$

This expression shows that the shadow cost does not depend on ϕ but depends on the equilibrium price changes.

Unregulated Firms

It is straightforward to show that, for unregulated firms,

$$l = l^*\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}}$$

and

$$q(\phi) = q^*(\phi)\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}.$$

Regulated Product Market Equilibrium

The price level under the regulation is

$$\begin{aligned} P^{-\rho} \quad &= \quad R^{-\rho}\int_{\bar{\phi}}^{\tilde{\phi}}\left(\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}(\phi)^{\frac{1}{1-\alpha\rho}}\,R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}\,P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\bar{\phi})}d\phi \\ &+ \quad R^{-\rho}\int_{\tilde{\phi}}\left(\xi^{\alpha_l}\,(\phi)^{\frac{1}{1-\alpha\rho}}\,R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}}\,P^{*\frac{\rho\alpha}{1-\alpha\rho}}C_Q\right)^{\rho}\frac{g(\phi)M}{1-G(\bar{\phi})}d\phi. \end{aligned}$$

Separate the regulated and unregulated regime into $\tilde{\pi}_1$ and $\tilde{\pi}_2$ as follows:

$$\tilde{\pi}_1 = \int\limits_{\bar{\phi}}^{\tilde{\phi}}(\phi)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi \quad \text{and} \quad \tilde{\pi}_2 = \int\limits_{\tilde{\phi}}(\phi)^{\frac{\rho}{1-\alpha\rho}}g(\phi)d\phi.$$

We can then manipulate the expression for the equilibrium price as follows:

$$P^{-\rho} P^{*\frac{-\rho^2\alpha}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1-G(\bar{\phi})}{M} \quad = \quad \left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} \tilde{\pi}_1 + \xi^{\alpha_l\rho}\tilde{\pi}_2.$$

Note that this equation holds in the case without the regulation if we set $\xi = 1$ and $P = P^*$. Let $s_{\tilde{\phi}} = \frac{\tilde{\pi}_2}{\tilde{\pi}_1+\tilde{\pi}_2}$ be the output share of the Top 1,000 conglomerates prior to the regulation. We obtain

$$\left(\frac{P}{P^*}\right)^{-\rho} \quad = \quad (1-s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\xi^{\alpha_l\rho}.$$

We can similarly derive the change in aggregate energy use as

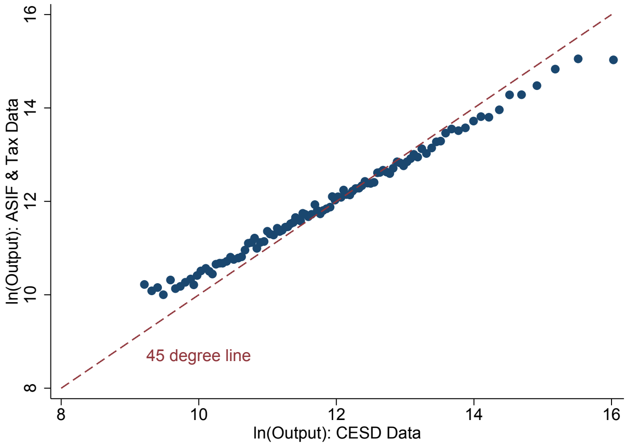
$$\frac{E}{E^*} = (1-s_{\tilde{\phi}})\left(\frac{P}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}}\xi.$$

These results show that the model without conglomerate spillovers is a special case of our model. In particular, the results of Propositions 4 and 5 continue to hold by our setting $\Delta_n = 1$.

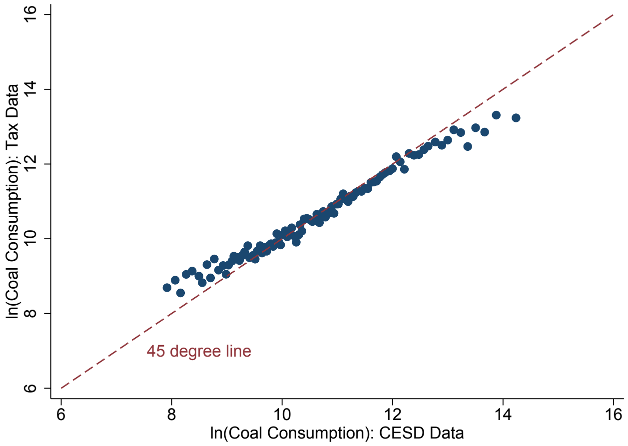
Appendix Figures

Figure A.1: Data Comparison: ASIF, CESD, Tax Data

A. Comparison of Output Data, 2001–2010

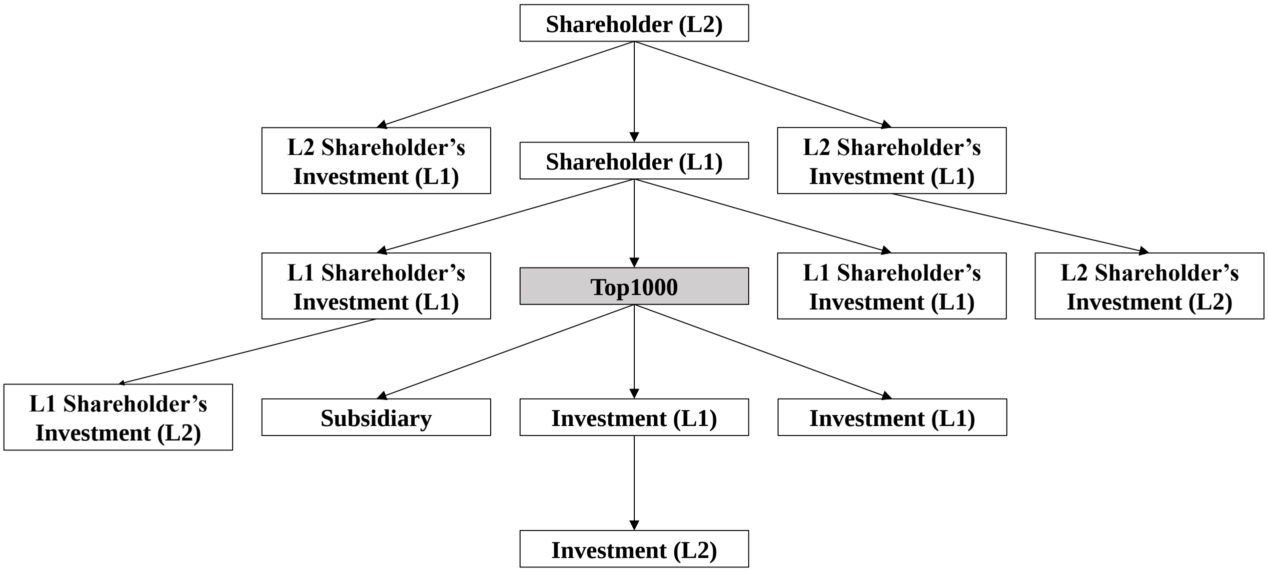


B. Comparison of Coal Consumption, 2007–2010



Notes: Authors' calculations using data from the ASIF, CESD and ATS. This figure shows the comparison of output and energy consumption data between the ASIF, ATS and CESD. Panel A shows that the output data from the ASIF and tax data are highly correlated with those from the CESD. This panel uses data from the ASIF for 2001–2006 and the ATS for 2007–2010. Panel B shows that the coal consumption figures from the tax data almost mirror those from the CESD (tax data are available only after 2007).

Figure A.2: Types of Related Parties

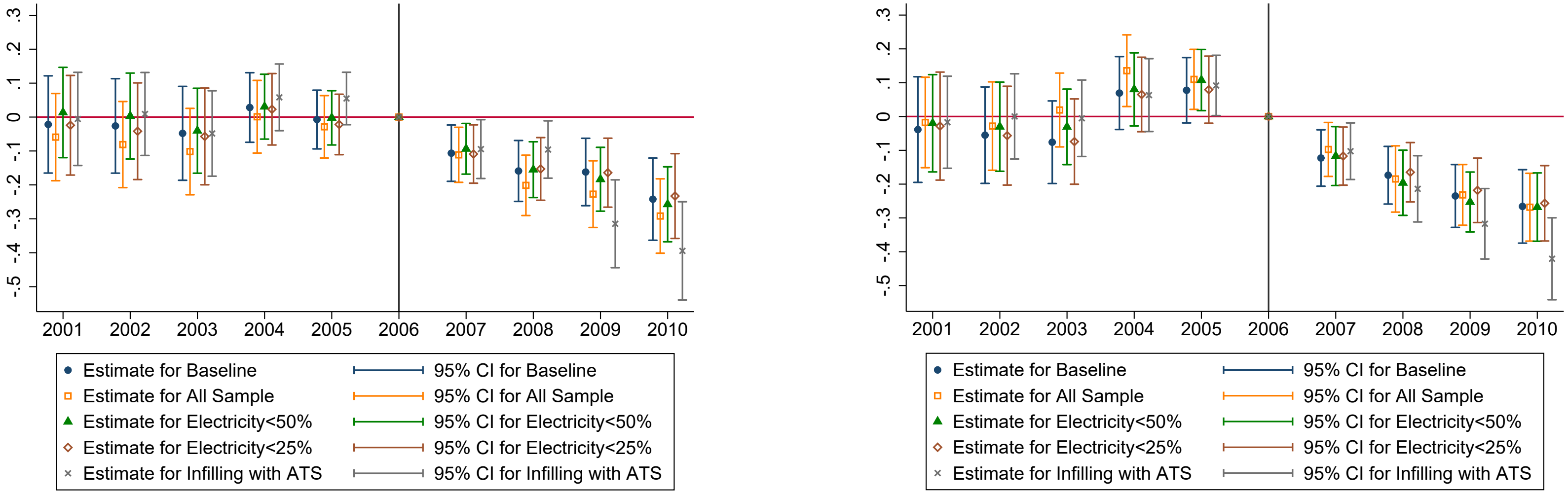


Notes: This figure depicts all possible types of firm relations within 2 levels of ownership. See Section 1.2 for the definition of each related firm type, and see Figure 2 for examples.

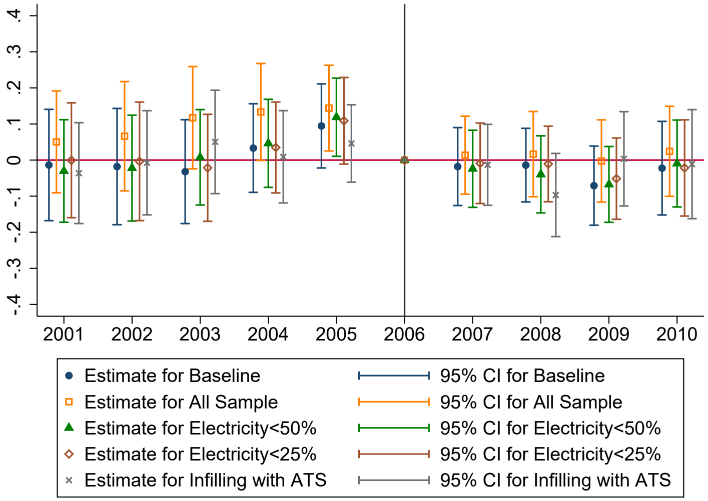
Figure A.3: Effects of the Top 1,000 Program on Regulated Firms: Robustness to Different Samples

A. Energy Use

B. Output



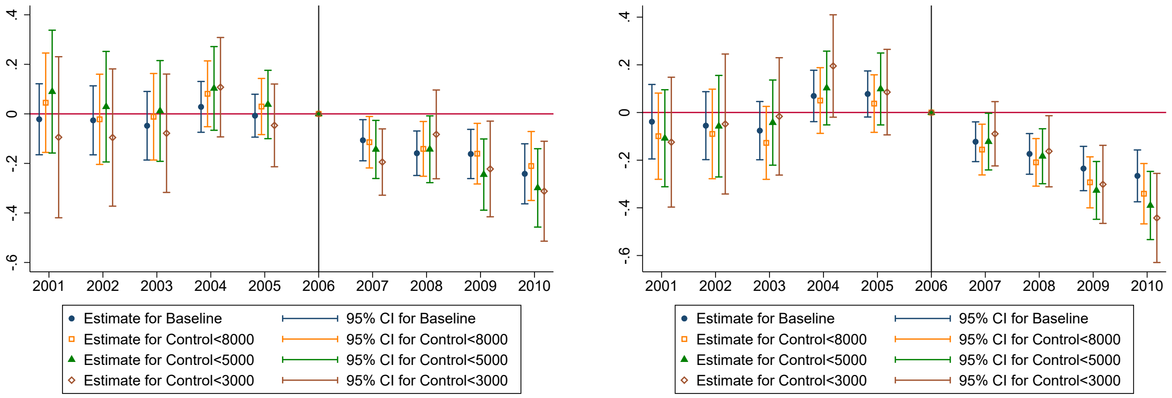
C. Energy Efficiency



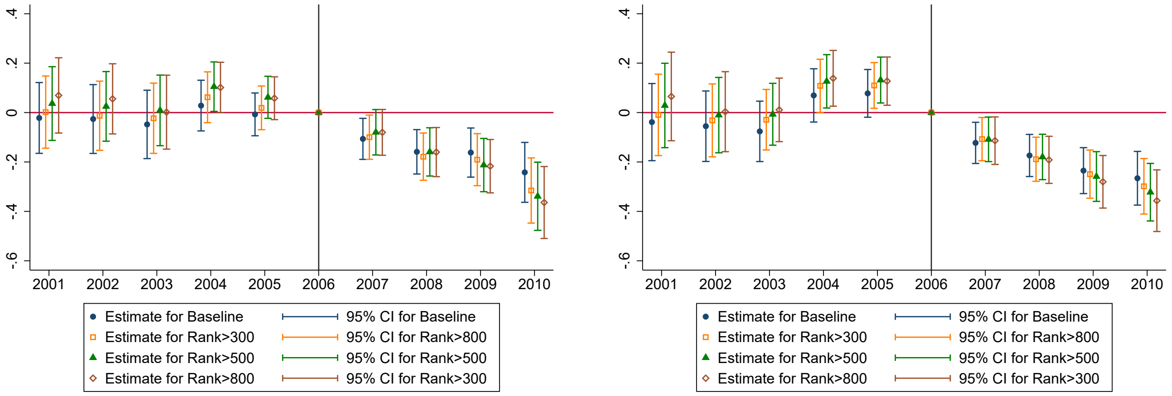
Notes: Authors' calculations using data from the CESD and ATS. This figure shows the coefficients for event studies of log firm energy consumption, output and energy efficiency for regulated firms under varying data restrictions. This figure corresponds to Panels A, C and E in Figure 4 with our baseline results (firms in industries with electricity consumption accounting for less than 30%) shown by blue lines, the results for firms in all industries by yellow lines, the results for firms in industries with electricity consumption accounting for less than 50% by green lines, the results for firms in industries with electricity consumption accounting for less than 25% by brown lines, and the results when we fill in the missing data with the ATS by the gray lines. The point estimates are shown in Table A.4. Standard errors are clustered at the firm level.

Figure A.4: Effects of the Top 1,000 Program on Regulated Firms: Robustness to Trimming the Sample by Top 10,000 Rank

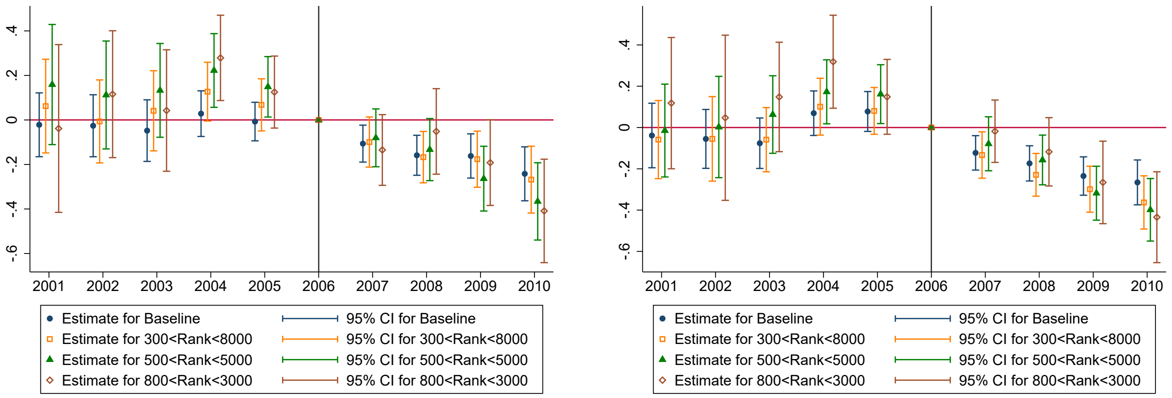
A. Energy: Trimming Controls **B. Output: Trimming Controls**



C. Energy: Trimming Treated **D. Output: Trimming Treated**

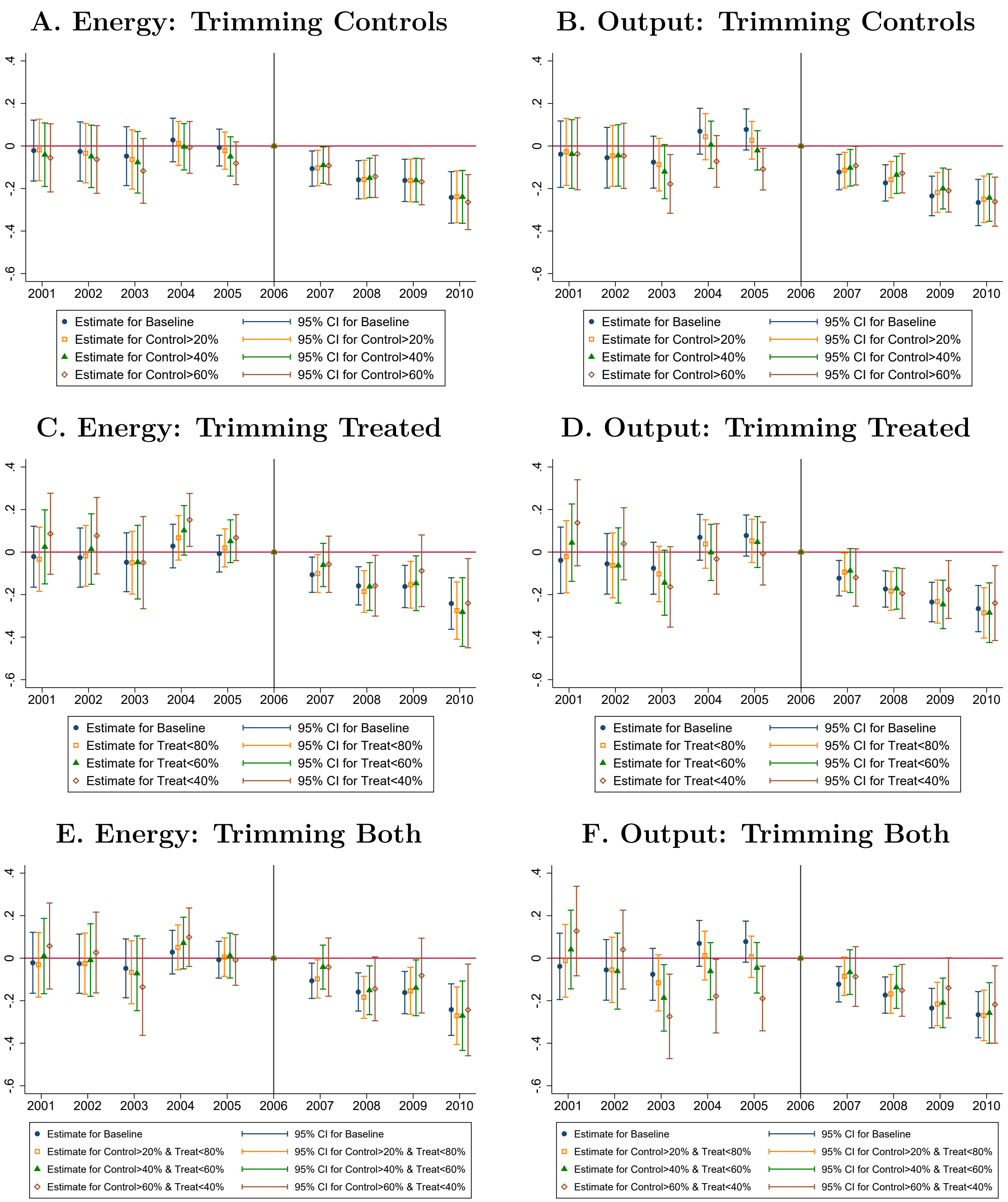


E. Energy: Trimming Both **F. Output: Trimming Both**



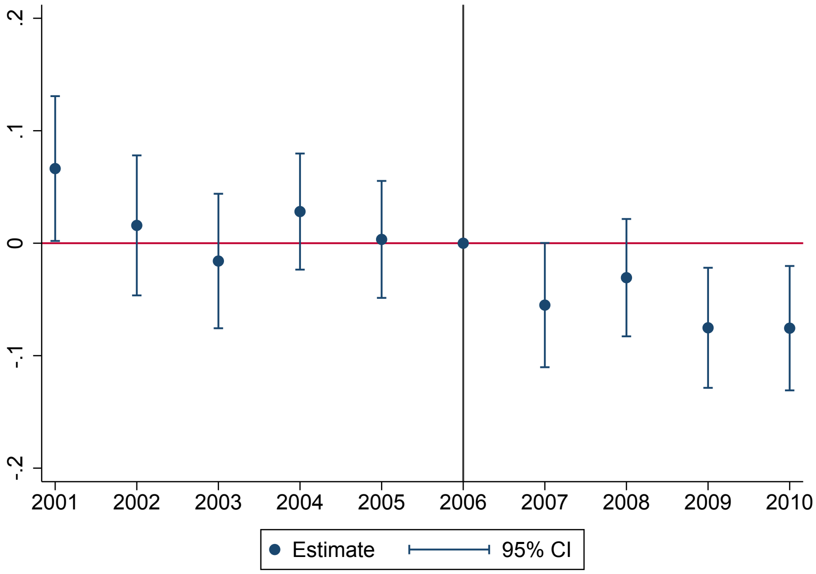
Notes: Authors' calculations using data from the CESD. This figure corresponds to Panels A and C in Figure 4 and uses firm rank in the Top 10,000 program to trim the sample. Our baseline results are shown by the blue lines; the other lines show the effects of trimming the sample so that treated and control firms have more comparable ranks. Panels A and B trim the sample of control firms to exclude firms with ranks below 8,000, 5,000, and 3,000. Panels C and D trim the sample of treated firms to include those with ranks greater than 300, 500, and 800. Panels E and F impose these rank restrictions on both control and treated firms. The point estimates are shown in Table A.5. Standard errors are clustered at the firm level.

Figure A.5: **Effects of the Top 1,000 Program on Regulated Firms: Robustness to Trimming the Sample Using an Estimated Propensity of Treatment**



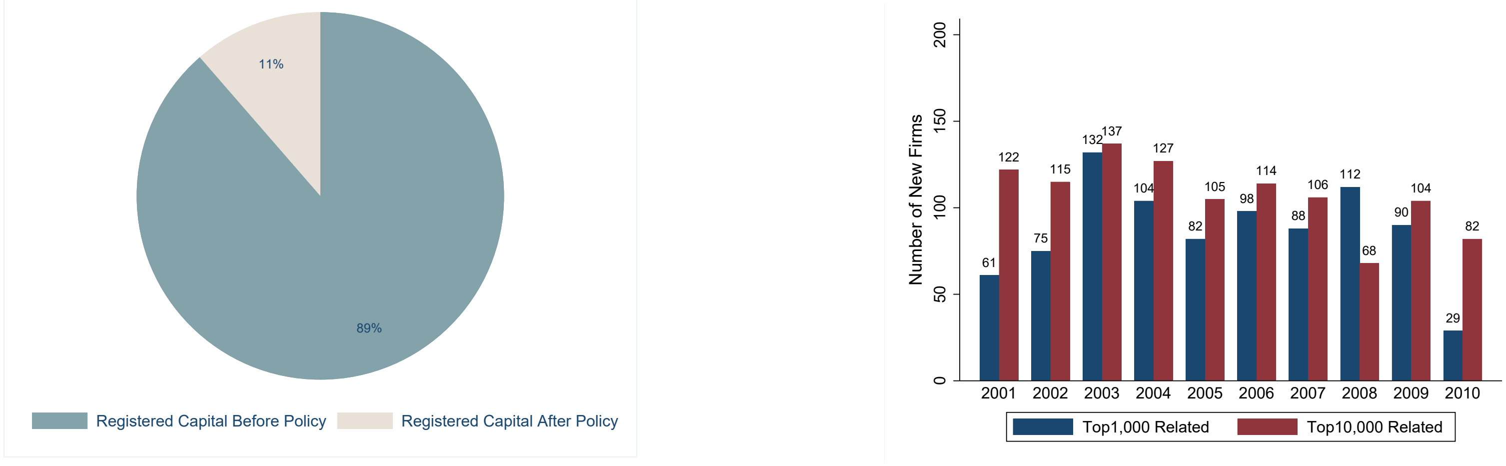
Notes: Authors' calculations using data from the CESD. This figure corresponds to Panels A and C in Figure 4 and uses firms' estimated propensity of being treated to trim the estimation sample. The propensity score is a logit model of the firm's likelihood of being in the Top 1,000 program that includes mean firm size (output) in years 2003–2005 and industry and province fixed effects and allows for heterogeneous effects of firm size by industry. Our baseline results are shown by the blue lines; the other lines show the effects of trimming the sample. Panels A and B trim the sample of control firms to exclude firms with propensity scores below the 20^{th}, 40^{th}, and 60^{th} percentiles of the distribution of propensity scores among control firms. Panels C and D trim the sample of treated firms to exclude firms with propensity scores above the 80^{th}, 60^{th}, and 40^{th} percentiles of the distribution of propensity scores among treated firms. Panels E and F impose these restrictions on both control and treated firms. The point estimates are shown in Table A.6. Standard errors are clustered at the firm level.

Figure A.6: Effects of the Policy on Regulated Firms' Investment



Notes: Authors' calculations using data from the ASIF and ATS. This figure shows estimates of Equation 1 where the dependent variable is firm investment choice, defined as whether a firm invests. See Section 1 for more information about the data-generating procedure. This figure shows that regulated firms were less likely to invest than similar control firms (Top 10,000 firms not related to Top 1,000 firms) after the regulation. The point estimates are displayed in Table A.12. See Section 2 for additional discussion. Standard errors are clustered at the firm level.

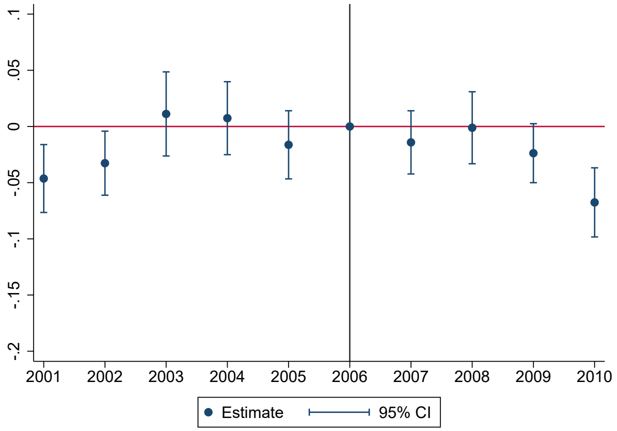
Figure A.7: Lack of Entry Effect



A. Relative Size of New Affiliates

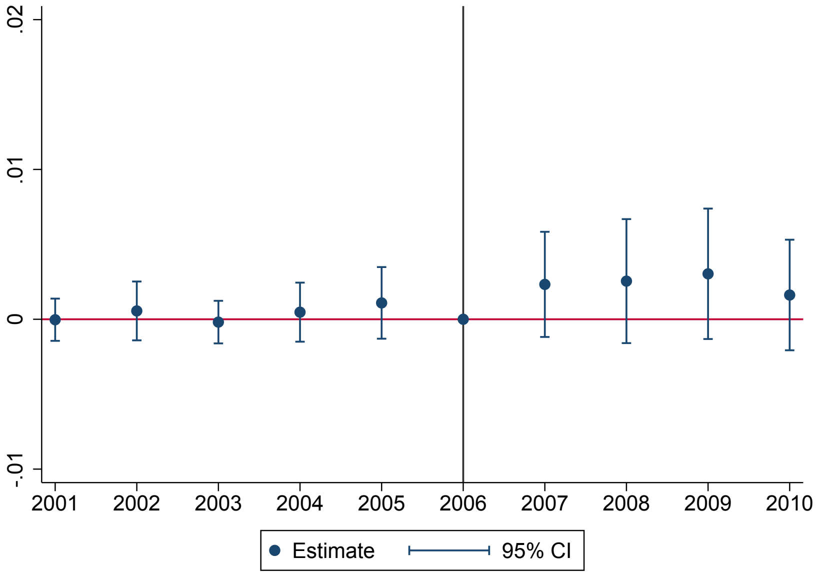
B. New Affiliates by Year

C. Entry Rate Event Study



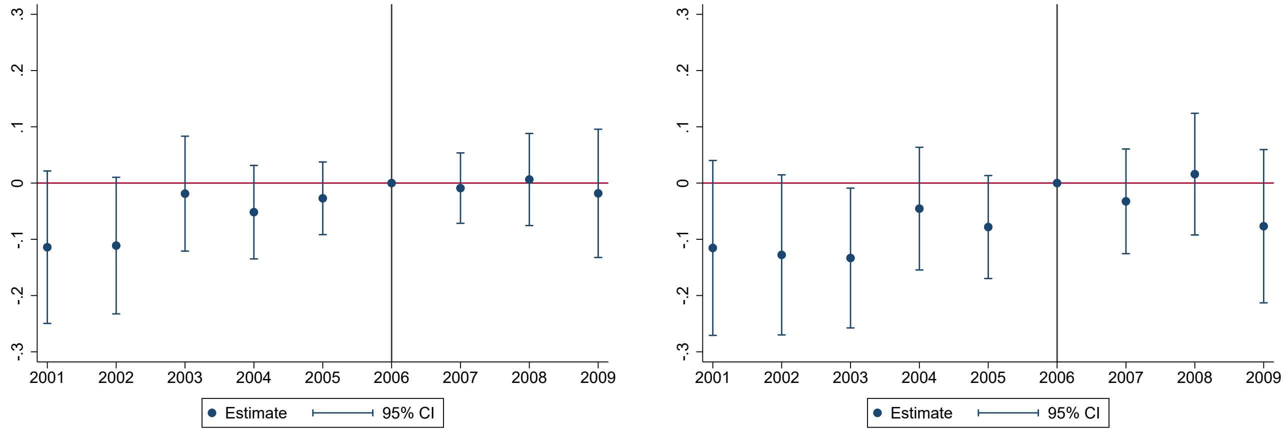
Notes: Authors' calculations using data from the CARD. Panel A shows that related firms of Top 1,000 firms that entered after the program started correspond to only 11% of the overall related registered capital. Panel B compares the number of new affiliates of Top 1,000 and Top 10,000 firms. Panel C estimates the effects of the Top 1,000 program on the entry rate, defined as the number of new affiliates relative to the size of the conglomerate prior to the regulation. Both Panels B and C show that regulated conglomerates did not add more affiliates than nonregulated conglomerates after the regulation was put in place.

Figure A.8: Innovation in Top 1,000 Firms: Energy-Saving Patent Applications



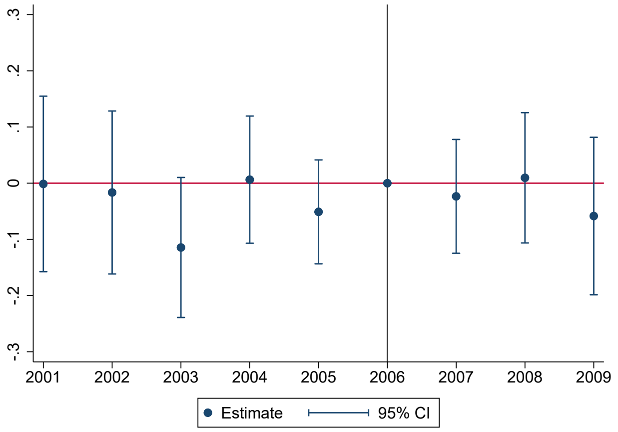
Notes: Authors' calculations using data from Incopat. This figure shows estimates of Equation 1 where the dependent variable is log firm patent applications. No significant effects on patent applications can be found for regulated firms relative to similar control firms (Top 10,000 firms not related to Top 1,000 firms). Standard errors are clustered at the firm level.

Figure A.9: Substitution Effects of the Program on Regulated Firms



A.Capital–Labor

B.Energy–Labor

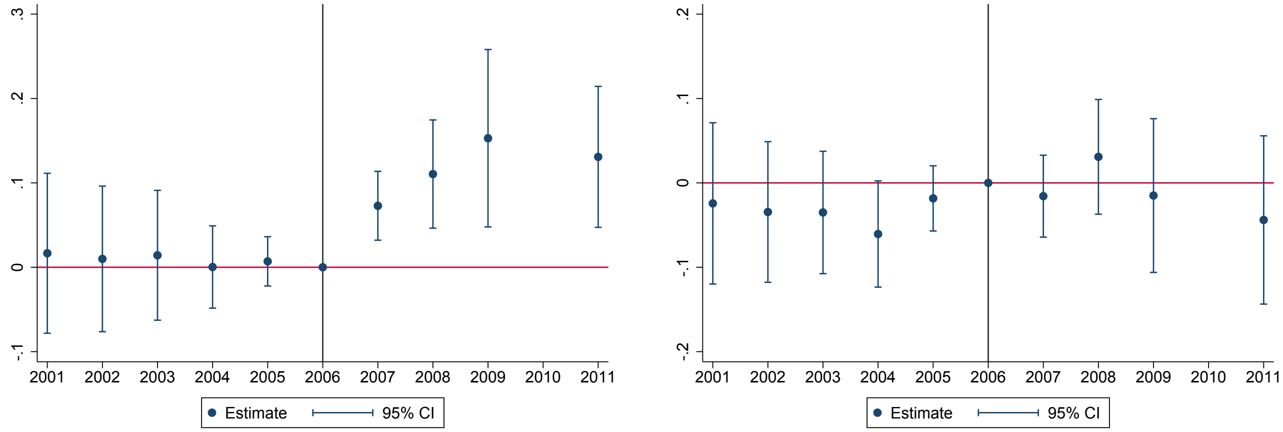


C.Energy–Capital

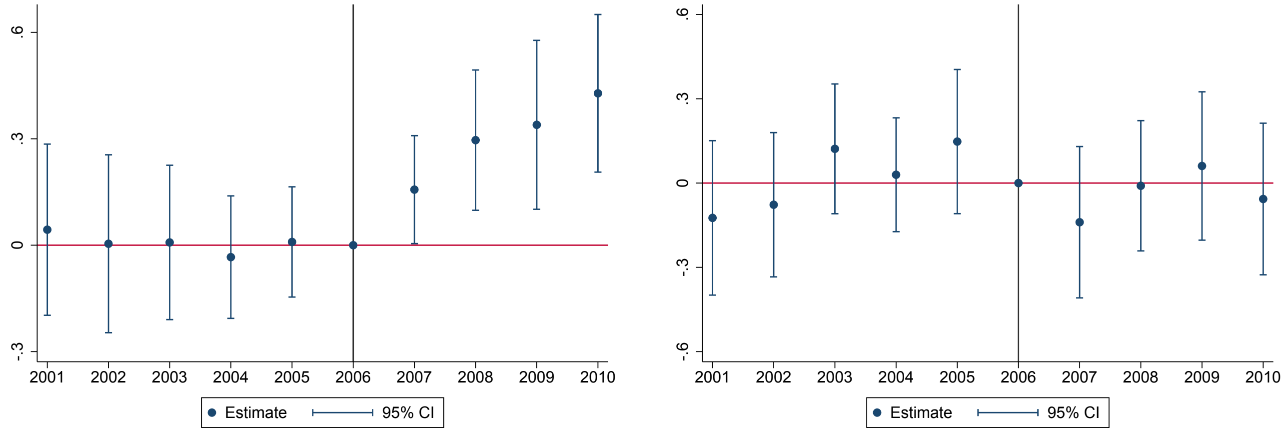
Notes: Authors' calculations using data from the CESD, ASIF and ATS. This table shows estimates of Equation 1 where the dependent variables are the log capital–labor ratio in Panel A, log energy–labor ratio in Panel B and log energy–capital ratio in Panel C. This figure shows no significant substitution among different inputs of regulated firms after they are regulated. The point estimates are displayed in Table A.13. Standard errors are clustered at the firm level.

Figure A.10: Robustness of Spillovers to Related Firms: Entropy Matching

A. Output **B. Placebo**

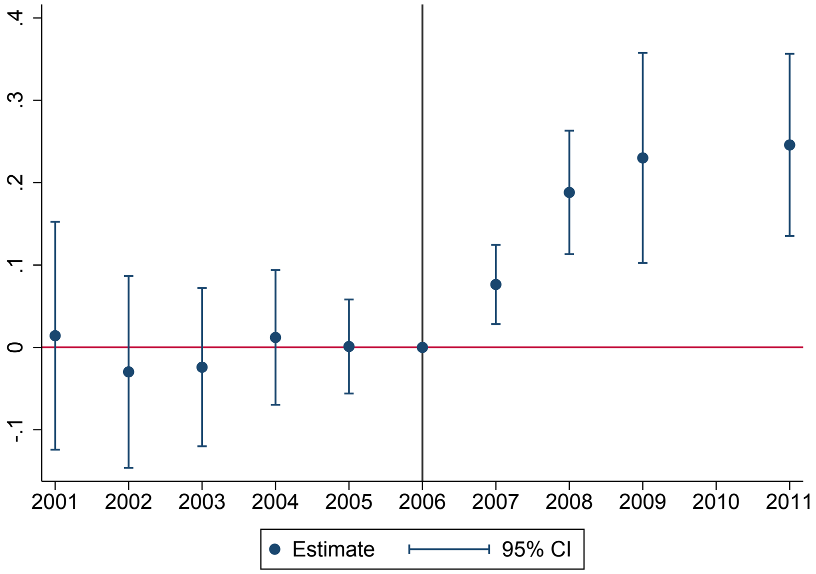


C. Energy Use **D. Energy Efficiency**



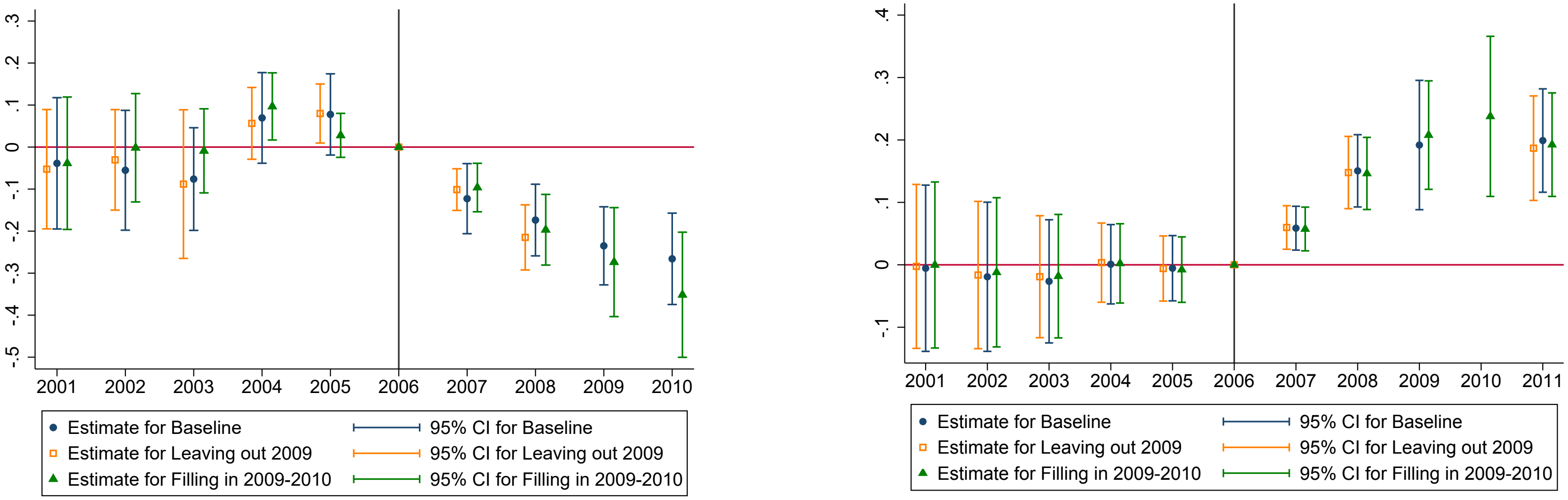
Notes: Authors' calculations using data from the ASIF and CESD. This figure shows the coefficients for event studies of log output, log energy consumption and log energy efficiency for firms related to regulated firms and event studies of log firm output for placebo firms. This figure corresponds to Figure 5 but deploys the additional matching method of entropy matching. The point estimates are displayed in Tables A.15 and A.16. Standard errors are clustered at the firm level.

Figure A.11: Robustness to Dropping Electric Power Generation and Supply



Notes: Authors' calculations using data from the ASIF. This figure shows the coefficients for an event study of log firm output for firms related to regulated firms, where related firms are restricted to those in industries other than electric power generation and supply. This figure corresponds to Panel A of Figure 5 but drops all observations in the electric power generation and supply industry. The point estimates are shown in Table A.18. Standard errors are clustered at the firm level.

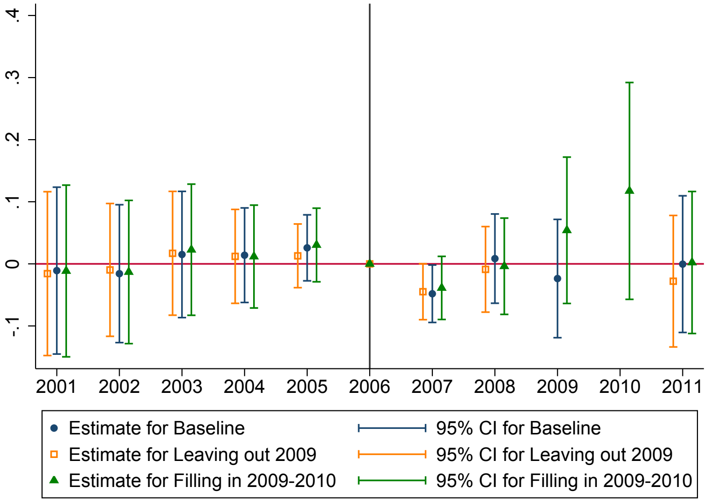
Figure A.12: Data Quality Robustness



A. Regulated Firms

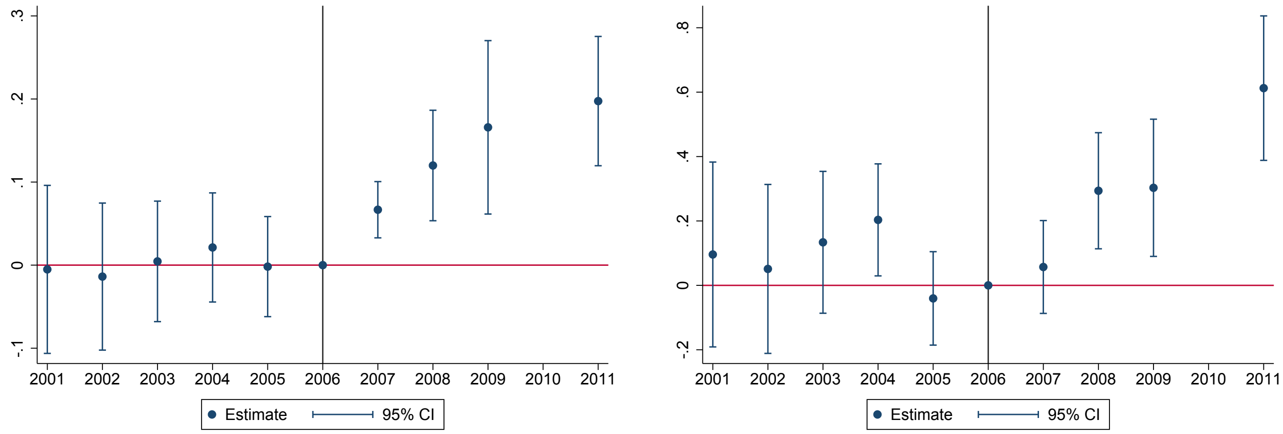
B. Related Firms

C. Placebo on Related Firms

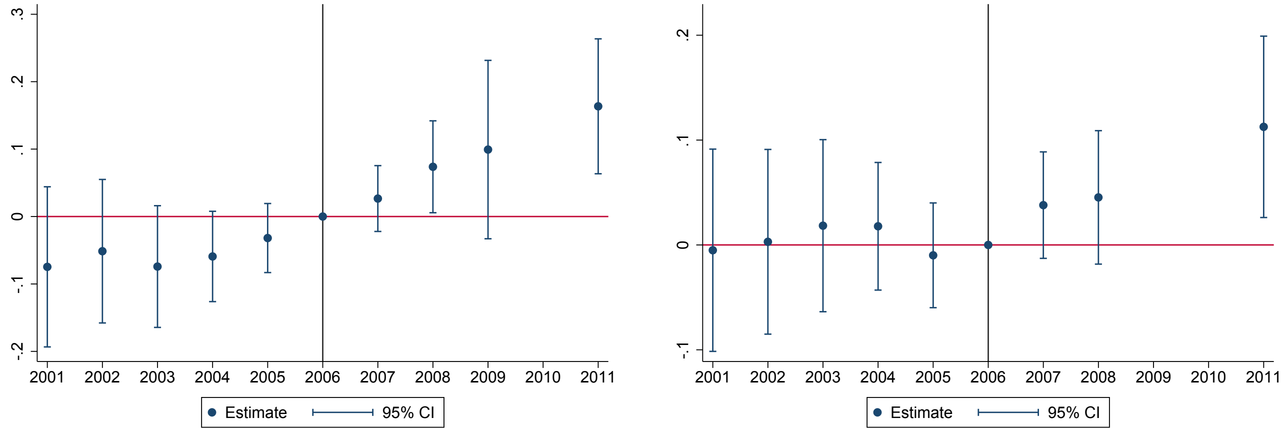


Notes: Panel A corresponds to Panel C in Figure 4, and Panels B and C correspond to Panels A and B in Figure 5. These figures drop data for 2009 from the ASIF (yellow lines) and fill in the 2009–2010 data with tax data (green lines). To mitigate the impact of outliers, firms with extreme output values and extreme output differences between various datasets (outside the 95% confidence interval) are excluded from Panel A. Standard errors are clustered at the firm level.

Figure A.13: Additional Spillover Effects of the Program



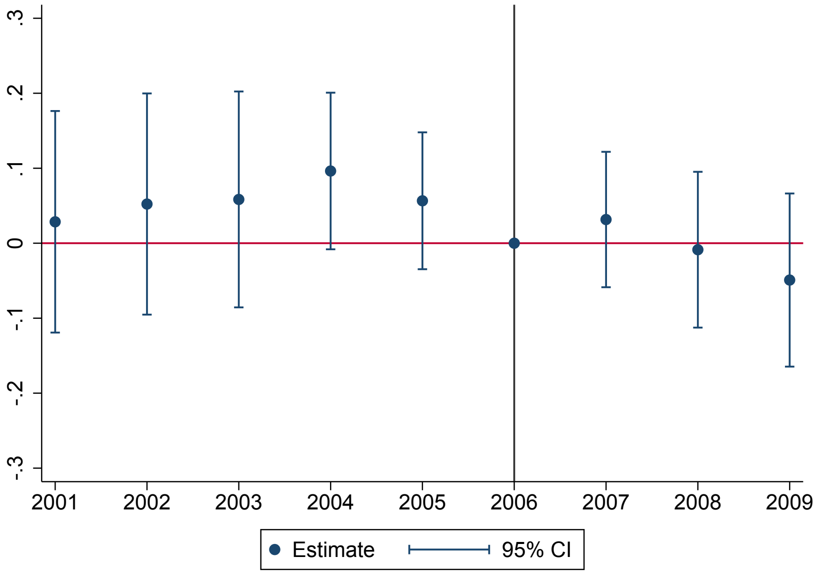
A. Sales **B. Profit**



C. Capital **D. Labor**

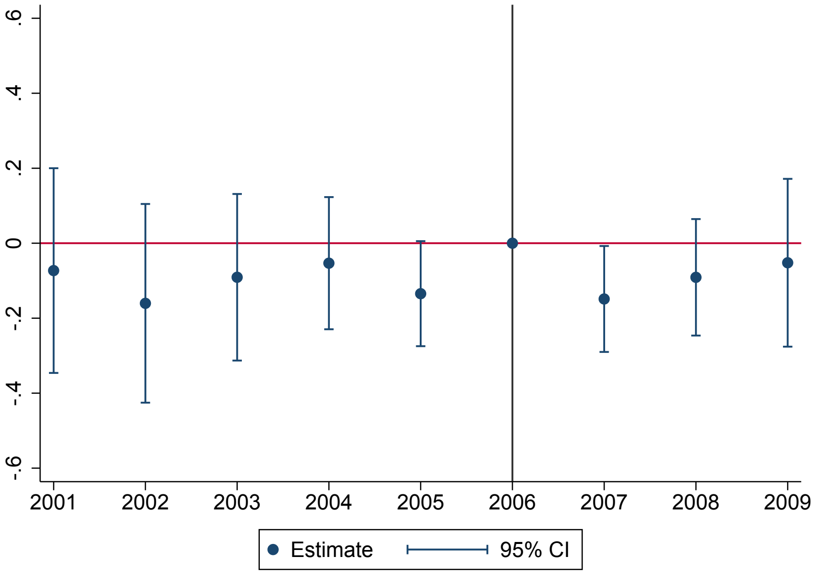
Notes: Authors' calculations using data from the ASIF. This figure shows the coefficients for event studies of log firm sales, profit, capital and labor for firms related to regulated firms. It shows that the Top 1,000 Energy Conservation Program had a persistent effect on the production and performance of related firms. The point estimates are displayed in Table A.21. Standard errors are clustered at the firm level.

Figure A.14: Robustness of Effects on the Energy Efficiency of Regulated Firms



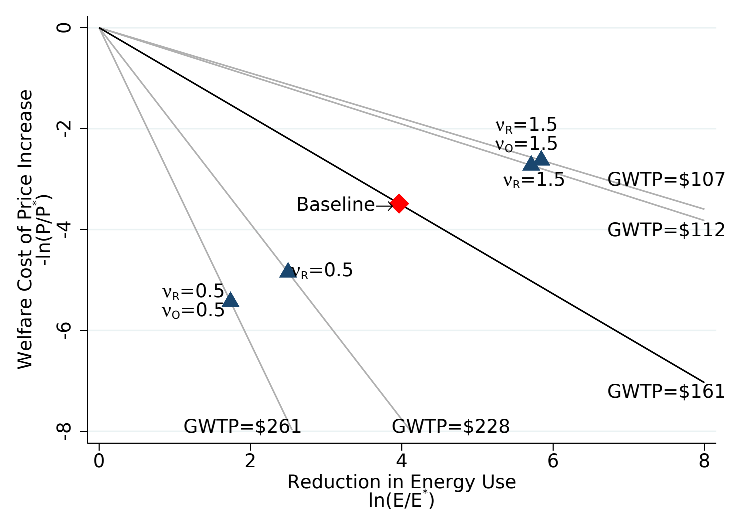
Notes: Authors' calculations using data from the CESD and ASIF. This figure shows the coefficients for an event study of log firm energy efficiency in regulated firms, where energy efficiency is defined as the inverse of the energy share in variable input costs. Variable input is calculated from the ASIF in terms of sales cost. This figure corresponds to Panel E of Figure 4 but with an alternative definition of energy efficiency. The point estimates are shown in Table A.25. Standard errors are clustered at the firm level.

Figure A.15: Robustness of Effects on the Energy Efficiency of Related Firms



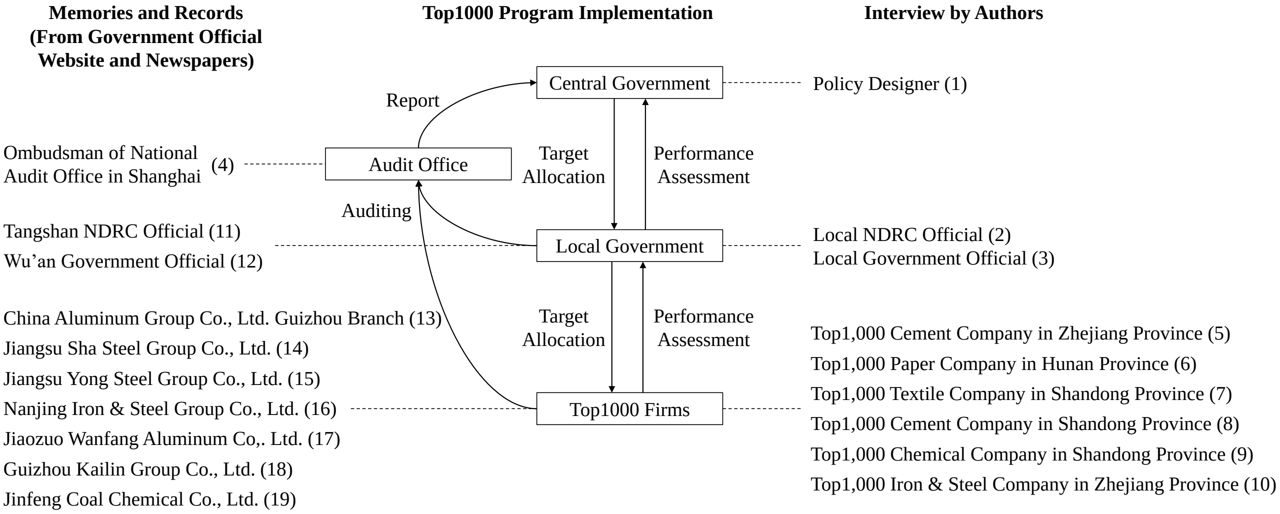
Notes: Authors' calculations using data from the CESD and ASIF. This figure shows the coefficients for an event study of log firm energy efficiency in firms related to regulated firms, where energy efficiency is defined as the inverse of the energy share in variable input costs. Variable input is calculated from the ASIF in terms of sales cost. This figure corresponds to Panel D of Figure 5 but with an alternative definition of energy efficiency. The point estimates are shown in Table A.26. Standard errors are clustered at the firm level.

Figure A.16: Welfare Effects: Additional Robustness to Differences in Energy Efficiency



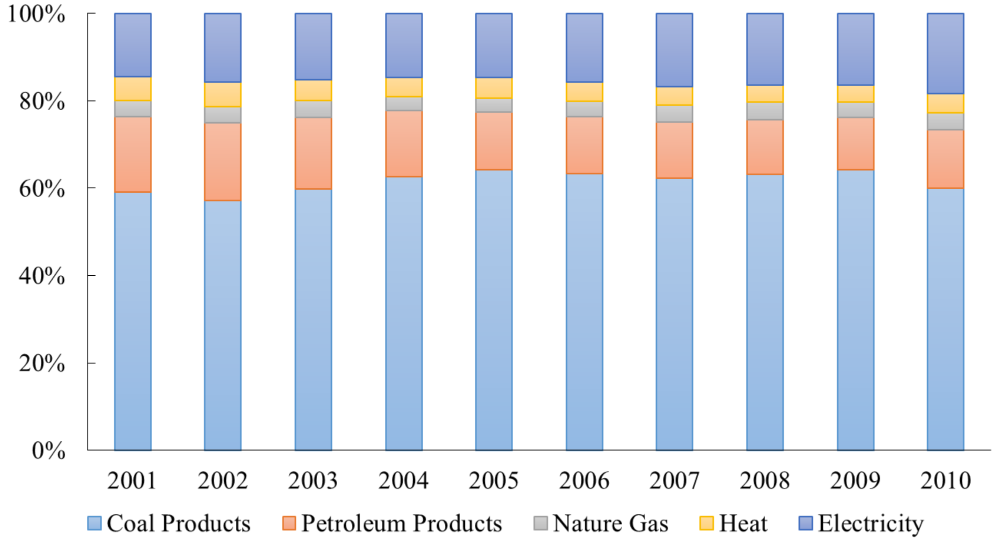
Notes: This figure shows the robustness of the welfare effects of the Top 1,000 program. The black lines plot indifference curves for the baseline value of GWTP=$161, and the light gray lines plot values of the GWTP according to different extensions. As discussed in Section 7.2, we explore the robustness of our results to our allowing for preexisting differences in energy efficiency. $\nu_R > 1$ denotes that related firms are more energy efficient than Top 1,000 firms, while $\nu_O < 1$ denotes that unregulated and unrelated firms are less efficient than Top 1,000 firms and vice versa. See Section 7.2 for details. While Panel B of Figure 10 allows differences in energy efficiency of up to 20%, this figure further allows differences in energy efficiency of up to 50%. Across these values, the GWTP that rationalizes the Top 1,000 program lies between $107 and $261.

Figure A.17: Interview Record



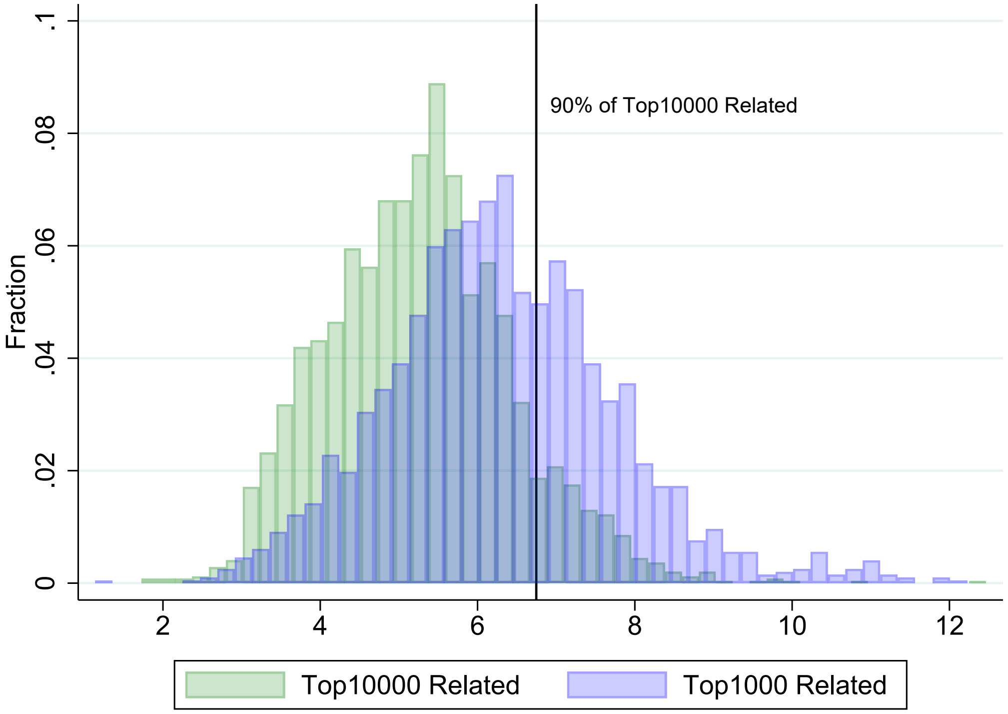
Notes: Sources (1), (2)–(3) and (5)–(10) are interviews conducted by the authors with the main policy designer of the Top 1,000 Energy Conservation Program, two local government officials, and executives from six Top 1,000 firms. Source (4) is an official report written by the ombudsman of the National Office in Shanghai (Xu, 2010), which can be found on the Official Website of the National Audit Office of the People's Republic of China. Sources (11)–(19) are interviews extracted from newspapers during the 11^{th} Five-Year Plan with two local government officials and executives from seven Top 1,000 firms.

Figure A.18: Energy Mix of Nine Energy-Intensive Industries



Notes: This figure uses data from the China Energy Statistical Yearbook to plot the aggregate energy mix of firms in the nine energy-intensive industries regulated by the Top 1,000 program. Since Top 1,000 firms account for nearly 60% of the energy consumption of these nine energy-intensive industries, significant changes in their energy mix would also be visible in the aggregate quantities. However, this figure shows that the energy mix for firms in these industries was stable during this decade.

Figure A.19: Size Distribution of Firms Related to Top 10,000 Firms



Notes: This figure uses ASIF data to plot the size distribution of firms related to Top 1,000 firms and Top 10,000 firms. Size is calculated as the log industrial output (in million RMB) of each firm. Firms related to Top 10,000 firms are much smaller than those related to Top 1,000 firms.

Appendix Tables

Table A.1: Waterfall Table

Data Cleaning Process	Firm Number
Top1000 (2004)	1008
Top1000 (2004) & CESD (2001–2010)	802
Drop Firms That Do Not Report Energy Data after 2006	687
Drop Firms That Do Not Report Coal Data Annually	527
Drop Extreme Energy Values	520
Drop Firms in 4-Digit Industries with Electricity>50%	479
Drop Firms in 4-Digit Industries with Electricity>30%	427
Drop Firms in 4-Digit Industries with Electricity>25%	420

Notes: Authors' calculations using data from the CESD. This table reports the number of Top 1,000 firms in the data according to different sample restrictions.

Table A.2: Dataset Matching

Datasets	Top 1,000		Top 10,000	
	Number	Ratio	Number	Ratio
List	1008	-	14641	-
ASIF	1001	99.31%	14227	97.17%
CESD	802	79.56%	10662	72.82%
ASIF & CESD	795	78.87%	9433	64.43%

Notes: Authors' calculations using data from the ASIF and CESD. This table shows the result of dataset matching. Over 99% of the Top 1,000 firms and over 97% of the Top 10,000 firms can be found in the ASIF, and nearly 80% of the Top 1,000 firms and over 70% of the Top 10,000 firms can be found in the CESD. See Section 1.2 for additional discussion.

Table A.3: Policy Compliance

Type	Orig. List	Evaluation			
Year		2007	2008	2009	2010
Firm Number	1008	953	922	901	881
Noncompliant Firms	-	74	36	28	15
Noncompliant Firms in Estimation Sample	-	23	20	12	8

Notes: Authors' calculations using data from the NDRC. This table shows the compliance of Top 1,000 firms during the 11FYP. The first row shows the number of Top 1,000 firms evaluated by the government in each year, the second row shows the number of noncompliant firms, and the last row shows the number of noncompliant firms in our estimation sample. See Section 1.1 for additional discussion.

Table A.4: Robustness to Different Samples

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)	Top1000
Panel A: CESD Data				
(1) All Sample	-0.168***	-0.237***	-0.071	687
	(0.046)	(0.038)	(0.047)	
(2) Drop Missing Coal Data	-0.173***	-0.235***	-0.061	527
	(0.042)	(0.040)	(0.044)	
(3) Drop Extreme Values	-0.172***	-0.236***	-0.063	520
	(0.042)	(0.040)	(0.044)	
(4) Industry(Electricity<50%)	-0.170***	-0.233***	-0.064	479
	(0.043)	(0.042)	(0.045)	
(5) Industry(Electricity<30%)	-0.156***	-0.204***	-0.049	427
	(0.047)	(0.042)	(0.046)	
(6) Industry(Electricity<25%)	-0.146***	-0.195***	-0.048	420
	(0.048)	(0.043)	(0.047)	
Panel B: CESD and ATS Data				
Infilling Missing Coal Data	-0.236***	-0.288***	-0.043	602
(From Row (2) in Panel A)	(0.048)	(0.040)	(0.047)	
Firm FE	Y	Y	Y	
Industry × Year FE	Y	Y	Y	
Province × Year FE	Y	Y	Y	

Notes: Authors' calculations using data from the CESD and ATS. This table reports the estimates from regressions of log firm energy consumption, output, and energy efficiency on regulated firms interacted with an indicator for years after 2006 under varying data restrictions. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 3 with various samples. Row (1) includes all firms with energy data after 2006. Row (2) drops firms with missing coal data. Row (3) restricts to firms without outlier energy values. Rows (4), (5), and (6) show the results after we drop industries with electricity consumption accounting for more than 50%, 30%, and 25%, respectively—our baseline is row (5). Panel B shows the results when we use ATS data to fill in missing energy use data. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Effects of the Top 1,000 Program on Regulated Firms: Robustness to Trimming the Sample by Top 10,000 Rank

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Baseline	-0.156***	-0.204***	-0.049
	(0.047)	(0.042)	(0.046)
Rank < 8000	-0.178***	-0.230***	-0.054
	(0.055)	(0.047)	(0.058)
Rank < 5000	-0.246***	-0.276***	-0.029
	(0.066)	(0.054)	(0.069)
Rank < 3000	-0.186***	-0.288***	-0.104
	(0.069)	(0.065)	(0.072)
Rank > 300	-0.202***	-0.241***	-0.039
	(0.049)	(0.042)	(0.048)
Rank > 500	-0.232***	-0.263***	-0.033
	(0.050)	(0.043)	(0.049)
Rank > 800	-0.243***	-0.289***	-0.048
	(0.050)	(0.045)	(0.047)
300 < Rank < 8000	-0.226***	-0.272***	-0.048
	(0.057)	(0.048)	(0.060)
500 < Rank < 5000	-0.328***	-0.314***	0.014
	(0.069)	(0.055)	(0.072)
800 < Rank < 3000	-0.293***	-0.337***	-0.049
	(0.079)	(0.073)	(0.085)
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table reports the estimates from regressions of log firm energy consumption, output and energy efficiency on regulated firms interacted with an indicator for years after 2006 under varying data restrictions. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 3 and use the rank in the Top 10,000 program to trim the sample. The Top 10,000 rank is generated with the energy savings targets of each Top 10,000 firm provided by the NDRC. Row (1) is our baseline result. Rows (2)–(4) restrict to firms ranking in the top 8000, 5000, and 3000, respectively. Rows (5)–(7) restrict to firms ranking below 300, 500, and 800, respectively. Rows (8)–(10) further restrict to firms ranking within 300–8000, 500–5000, and 800–3000, respectively. This table corresponds to a pooled version of the regression displayed in Figure A.4 and shows that our baseline results are robust to trimming the sample to include firms with more comparable ranks. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Effects of the Top 1,000 Program on Regulated Firms: Robustness to Trimming the Sample by Estimated Propensity of Treatment

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Baseline	-0.156***	-0.204***	-0.049
	(0.047)	(0.042)	(0.046)
Control > 20%	-0.146***	-0.175***	-0.031
	(0.047)	(0.042)	(0.046)
Control > 40%	-0.126***	-0.139***	-0.016
	(0.048)	(0.042)	(0.048)
Control > 60%	-0.117**	-0.102**	0.014
	(0.051)	(0.044)	(0.051)
Treat < 80%	-0.179***	-0.188***	-0.012
	(0.051)	(0.047)	(0.050)
Treat < 60%	-0.186***	-0.179***	0.004
	(0.060)	(0.054)	(0.059)
Treat < 40%	-0.186**	-0.165**	0.016
	(0.074)	(0.066)	(0.073)
Control > 20% & Treat < 80%	-0.168***	-0.159***	0.005
	(0.051)	(0.047)	(0.051)
Control > 40% & Treat < 60%	-0.152**	-0.112**	0.034
	(0.062)	(0.055)	(0.061)
Control > 60% & Treat < 40%	-0.129*	-0.047	0.077
	(0.076)	(0.066)	(0.077)
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table reports the estimates from regressions of log firm energy consumption, output and energy efficiency on regulated firms interacted with an indicator for years after 2006 under varying data restrictions. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 3 and use firms' estimated propensity of being treated to trim the estimation sample. The propensity score is a logit model of the likelihood of a firm's being in the Top 1,000 program that includes mean firm size (output) in 2003–2005 and industry and province fixed effects and allows for heterogeneous effects of firm size by industry. Row (1) is our baseline result. Rows (2)–(4) drop control firms below the 20^{th}, 40^{th}, and 60^{th} percentiles of likelihood of being treated, respectively. Rows (5)–(7) drop treated firms that are above the 20^{th}, 40^{th}, and 60^{th} percentiles of likelihood of being treated, respectively. Rows (8)–(10) impose these restrictions on both control and treated firms. The last row of the table shows that the effect on output is no longer significant when we rely on the 159 treated firms that meet this condition (relative to the 427 firms in our baseline analysis). This table corresponds to a pooled version of the regression displayed in Figure A.5 and shows that our baseline results are generally robust to trimming the sample to eliminate treated firms with a very high likelihood of being treated or control firms with a very low likelihood of being treated. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Robustness to Entry and Exit: Regulated Firm Response

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat × Post	-0.141***	-0.170***	-0.029
	(0.049)	(0.044)	(0.048)
Observations	18,506	18,385	18,385
R^2	0.886	0.890	0.831
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table reports the estimates from regressions of log firm energy consumption, output and energy efficiency on regulated firms interacted with an indicator for years after 2006 when we exclude firm entry and exit. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 3 with the exclusion of firms that enter the CESD data after 2006 or exit before 2010. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Approximately Balanced Panel: Regulated Firm Response

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat × Post	-0.153**	-0.145**	0.006
	(0.069)	(0.057)	(0.065)
Observations	10,336	10,265	10,265
R^2	0.892	0.910	0.835
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table reports the estimates from regressions of log firm energy consumption, output and energy efficiency on regulated firms interacted with an indicator for years after 2006 with an approximately balanced panel. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 3 with the exclusion of firms with missing data for two years or more. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Robustness to Consideration of Concurrent Policies

A. Effects of Concurrent Policies on Top 1,000 Firms

	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Monitor × Post	-0.003	0.023	0.019
	(0.097)	(0.089)	(0.106)
Observations	3,358	3,322	3,322
R^2	0.865	0.899	0.812
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

B. Robustness to Effects on Top 1,000 Firms

	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat × Post	-0.186**	-0.234***	-0.046
	(0.075)	(0.080)	(0.083)
Observations	20,655	20,511	20,511
R^2	0.864	0.858	0.847
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table shows that our estimated effects of the Top 1,000 program are robust to consideration of a concurrent policy—the National Specially Monitored Firms (NSMF) program. Panel A estimates a difference-in-differences model to show the effects of the NSMF program within the Top 1,000 firm sample. We can see that the NSMF program had little effect on the energy consumption, output and energy efficiency of Top 1,000 firms. Panel B estimates the same regression as in Table 3 while excluding all treated firms included under both polices. It shows that taking the NSMF program into account does not affect our main results. See Section D for both a detailed description of the NSMF program and additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Heterogeneous Effects on Regulated Firms by Industry

Variables	ln(Output)				
	Baseline	Drop Power	Processing	Materials	Mining
Treat × Post	-0.204***	-0.206***	-0.219***	-0.153***	-0.251
	(0.042)	(0.042)	(0.073)	(0.053)	(0.275)
Observations	22,991	21,748	5,440	12,662	545
R^2	0.889	0.887	0.893	0.865	0.863
Firm FE	Y	Y	Y	Y	Y
Industry × Year FE	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table shows estimates of Equation 2 by industry where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variable is log firm output, corresponding to Panel B of Table 3. Column (1) is our baseline result. Column (2) shows the results after we drop the production and supply of electric power and heating power industry. Column (3) shows the results for processing industries, including smelting and pressing of ferrous metals, smelting and pressing of nonferrous metals, processing of petroleum, coking, and processing of nuclear fuel. Column (4) shows the result for material industries, including the manufacture of raw chemical materials and chemical products and the manufacture of nonmetallic mineral products. Column (5) shows the result for mining industries, including mining and washing of coal and extraction of petroleum and natural gas. This table shows no significant differences in regulated firm response among different industries. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Heterogeneous Effects on Regulated Firms by Ownership

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat × Post	-0.135***	-0.167***	-0.034
	(0.051)	(0.047)	(0.049)
Treat × SOE × Post	0.041	-0.111	-0.141
	(0.115)	(0.104)	(0.120)
Observations	21,924	21,778	21,778
R^2	0.892	0.889	0.841
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y
SOE × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD. This table shows estimates of Equation 2 by industry where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006. This table explores whether there are heterogeneous effects on firms that are state owned. This table shows no significant differences in the response of regulated firms along ownership margins. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Effects of Policy on Regulated Firms' Investment

Variables	If Firm Invest			
Treat × Post	-0.056***	-0.070***	-0.070***	-0.071***
	(0.013)	(0.014)	(0.014)	(0.014)
Observations	47,231	47,211	47,208	45,675
R^2	0.192	0.201	0.209	0.214
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using ASIF and ATS. This table shows estimates of Equation 2 where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variable is firm investment choice. Investment choice is defined as whether a firm invests. See Section 1 for more information about the data-generating procedure. This table corresponds to a pooled version of the regression displayed in Figure A.6. It shows that regulated firms decreased their possibility of investment by 5.6%–7.1% relative to that of similar control firms (Top 10,000 firms not related to Top 1,000 firms) after the policy implementation. See Section 2 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Effects of the Top 1,000 Program on Input Substitution in Regulated Firms

Variables	ln(Capital/Labor)	ln(Energy/Labor)	ln(Energy/Capital)
Treat × Post	0.034	0.041	0.007
	(0.035)	(0.044)	(0.047)
Observations	13,600	13,600	13,600
R^2	0.736	0.812	0.808
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Authors' calculations using data from the CESD, ASIF and ATS. This table shows estimates of Equation 2, where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variables are log capital labor ratio, log energy labor ratio and log energy capital ratio. This table corresponds to a pooled version of the regression displayed in Figure A.9 and shows no significant substitution among different inputs of regulated firms after they become regulated. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Heterogeneous Effects on Regulated Firms by Reported Compliance Status

Variables	ln(Energy Use)			
Compliance × Post	-0.138***	-0.168***	-0.169***	-0.141***
	(0.042)	(0.046)	(0.047)	(0.048)
Noncompliance × Post	0.395	0.338	0.316	0.333
	(0.347)	(0.366)	(0.359)	(0.346)
Observations	23,607	23,602	23,151	20,571
R^2	0.887	0.890	0.892	0.898
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using the CESD. This table shows estimates of heterogeneous effects within regulated firms by government-reported compliance status. Compliance is an indicator for regulated firms that are reported to have achieved their energy savings targets under the 11FYP, while Noncompliance is an indicator for regulated firms whose energy savings targets were not met. This table shows that the compliant regulated firms decreased their energy consumption by 13%–17% after the policy implementation while the noncompliant regulated firms, in contrast, show a 31%–40% increase in their energy consumption. However, due to the limited size of the sample of noncompliant firms (only 15 firms had not achieved their energy savings targets at the end of the 11FYP), the results for Noncompliance show a large standard deviation. This result implies that the official compliance status can be verified to some extent by firm energy consumption when the energy quota is used. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Robustness of Spillovers to Related Firms on Output: Entropy Matching

A. Output

Variables	ln(Output)			
Related × Post	0.143***	0.139***	0.106***	0.133***
	(0.039)	(0.038)	(0.037)	(0.036)
Observations	119,064	119,064	119,064	116,064
R^2	0.874	0.881	0.890	0.896

B. Placebo Test on Output

Variables	ln(Output)			
Related × Post	0.016	0.019	0.020	0.036
	(0.041)	(0.040)	(0.039)	(0.038)
Observations	150,997	150,997	150,997	147,431
R^2	0.909	0.914	0.922	0.929
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 2 with an alternative matching method, where Related × Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006 in Panel A and an indicator for related firms in the same 2-digit industry (but not the same 4-digit industry) interacted with an indicator for years after 2006 in Panel B. This table corresponds to Table 4 but deploys the additional matching method of entropy matching. Panels A and B correspond to a pooled version of the regression displayed in Panels A and B of Figure A.10. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: Robustness of Spillovers to Related Firms on Energy Use and Energy Efficiency: Entropy Matching

A. Energy Use

Variables	ln(Energy Use)			
Related × Post	0.247^{**}	0.248^{***}	0.289^{***}	0.263^{**}
	(0.097)	(0.095)	(0.090)	(0.104)
Observations	20,254	20,254	20,101	14,507
R^2	0.855	0.858	0.871	0.879

B. Energy Efficiency

Variables	ln(Energy Efficiency)			
Related × Post	-0.021	-0.019	-0.056	-0.039
	(0.099)	(0.098)	(0.103)	(0.104)
Observations	20,122	20,122	19,971	14,424
R^2	0.822	0.826	0.839	0.845
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the CESD and ASIF. This table shows estimates of Equation 2 with an alternative matching method, where Related × Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006 and the dependent variables are log firm energy consumption in Panel A and log firm energy efficiency in Panel B. This table corresponds to Panels C and D of Table 4 but deploys the additional matching method of entropy matching and corresponds to a pooled version of the regression displayed in Panels C and D of Figure A.10. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Spillover Effects on the Output of Related Firms: Robustness to Different Definitions of Related Parties

Variables	ln(Output)				
	Baseline	Drop Changes	6 Level, 20%	2 Level, 20%	2 Level, 51%
Related × Post	0.127***	0.125***	0.133***	0.124***	0.154***
	(0.035)	(0.036)	(0.033)	(0.035)	(0.039)
Observations	17,905	17,030	20,036	18,185	14,589
R^2	0.889	0.889	0.889	0.888	0.892
Firm FE	Y	Y	Y	Y	Y
Industry × Year FE	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y
Firm-Level Controls	Y	Y	Y	Y	Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 2 where Related × Post is an indicator for related firms interacted with an indicator for years after 2006 and the dependent variable is log firm output. This table corresponds to Table 4 but uses different definitions for related parties. Column (1) is our baseline result. Column (2) shows the results after we drop all related firms with shareholding changes after the policy implementation (which account for 3.89% of total related firms). It shows that our results are robust to dropping related firms with shareholding changes. Column (3) shows the results for related firms within six levels of shareholder links and with an ownership requirement of at least 20%. It shows that, under this broader definition, related firms increased output by 13.3% after the policy implementation; this means that the conglomerate shift accounts for 46.0% ($\approx 2.80 \times 17.9\% \times 13.3\%/14.5\%$) of the output decline in regulated firms. Column (4) shows the results for related firms within two levels of shareholder links and with an ownership requirement of at least 20%. It shows that, under this definition, related firms increased output by 12.4% after the policy implementation; this means that the conglomerate shift accounts for 41% ($\approx 2.49 \times 19.3\% \times 12.4\%/14.5\%$) of the output decline in regulated firms. Column (5) shows the results for related firms within two levels of shareholder links and with an ownership requirement of more than 50%. It shows that, under this narrower definition, related firms increased output by 15.4% after the policy implementation; this means that the conglomerate shift accounts for 41% ($\approx 1.95 \times 19.9\% \times 15.4\%/14.5\%$) of the output decline in regulated firms. See Section 3 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.18: Robustness to Dropping Electric Power Generation and Supply

Variables	ln(Output)			
Related*Post	0.180***	0.173***	0.163***	0.167***
	(0.043)	(0.041)	(0.039)	(0.037)
Observations	11,232	11,229	11,222	10,982
R^2	0.850	0.862	0.872	0.883
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 2 where Related × Post is an indicator for related firms in industries other than electric power generation and supply interacted with an indicator for years after 2006 and the dependent variable is log firm output. This table corresponds to Panel A of Table 4 but with the exclusion of all firms in the electric power generation and supply industry and corresponds to a pooled version of the regression displayed in Figure A.11. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.19: Data Quality Robustness

A. Leave out 2009 in ASIF Data

Variables	ln(Output)			
Treat × Post	-0.054	-0.198***	-0.166***	-0.112***
	(0.035)	(0.036)	(0.039)	(0.038)
Observations	10,028	10,023	9,860	9,678
R^2	0.955	0.959	0.961	0.963
Related × Post	0.131***	0.125***	0.102***	0.111***
	(0.034)	(0.033)	(0.033)	(0.031)
Observations	16,454	16,452	16,450	15,970
R^2	0.890	0.898	0.905	0.913
Placebo × Post	-0.033	-0.031	-0.025	-0.019
	(0.039)	(0.037)	(0.038)	(0.037)
Observations	7,971	7,970	7,955	7,809
R^2	0.908	0.914	0.921	0.927

B. Fill in 2009–2010 with Tax Data

Variables	ln(Output)			
Treat × Post	-0.235***	-0.262***	-0.237***	-0.201***
	(0.047)	(0.047)	(0.048)	(0.048)
Observations	11,805	11,803	11,586	11,102
R^2	0.921	0.925	0.929	0.932
Related × Post	0.159***	0.156***	0.137***	0.135***
	(0.037)	(0.037)	(0.037)	(0.035)
Observations	19,293	19,289	19,287	18,735
R^2	0.869	0.876	0.885	0.893
Placebo × Post	0.000	-0.003	-0.005	-0.008
	(0.041)	(0.040)	(0.040)	(0.039)
Observations	9,287	9,286	9,266	9,102
R^2	0.874	0.880	0.889	0.897
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using the ASIF and ATS. This table shows estimates of Equation 2 with different datasets. Treat is an indicator for regulated firms. The reported effects on regulated firms correspond to the results in Panel A in Table 3. Related is an indicator for related firms in the same 4-digit industry, and Placebo is an indicator for related firms in the same 2-digit industry but not the same 4-digit industry. The reported effects on related firms correspond to the results in Panel A in Table 4, and the placebo results correspond to the results in Panel B in Table 4. In this table, Panel A uses ASIF data with exclusion of the 2009 data, and Panel B uses tax data to fill in the ASIF data for years 2009–2010. The estimates in this table also correspond to a pooled version of the regressions displayed in Figure A.12. To mitigate the impact of outliers, firms with extreme output values and extreme output differences between various datasets (outside the 95% confidence interval) are excluded from the regulated firms regression. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.20: Heterogeneous Spillover Effects on Output by Related Firm Size

Variables	ln(Output)			
Related × Post(0%-30%)	0.104*	0.109**	0.048	0.078
	(0.054)	(0.052)	(0.052)	(0.051)
Related × Post(30%-60%)	0.130***	0.123***	0.096**	0.111**
	(0.047)	(0.045)	(0.046)	(0.043)
Related × Post(60%-100%)	0.164***	0.156***	0.161***	0.161***
	(0.045)	(0.044)	(0.043)	(0.040)
Observations	17,691	17,691	17,689	17,212
R^2	0.892	0.900	0.907	0.915
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the ASIF. This table corresponds to Panel A of Table 4 and shows estimates of heterogeneous spillover effects by terciles of firm size. We can see that the related-firm spillovers are greater for larger related firms. See Section 3 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.21: Additional Spillover Effects of the Program

Variables	ln(Sale)	ln(Profit)	ln(Capital)	ln(Labor)
Related × Post	0.115***	0.190***	0.114***	0.063**
	(0.033)	(0.055)	(0.040)	(0.026)
Observations	17,867	13,147	17,901	15,966
R^2	0.893	0.826	0.904	0.897
Firm FE	Y	Y	Y	Y
Industry × Year FE	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y
Firm-Level Controls	Y	Y	Y	Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 2 where Related × Post is an indicator for related firms interacted with an indicator for years after 2006 and the dependent variables are log firm sales, profit, capital and labor. The estimates in this table correspond to a pooled version of the regression displayed in Figure A.13. They show that related firms in the same 4-digit industries increased sales by 11.5%, profit by 19.0%, capital by 11.4%, and labor by 6.3% after the policy implementation. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.22: Spillover Effects on the Output of Related Firms: Heterogeneous Effects by Industry

Variables	ln(Output)				
	Baseline	Drop Power	Processing	Materials	Mining
Related × Post	0.118***	0.164***	0.105	0.180***	0.197**
	(0.037)	(0.040)	(0.096)	(0.056)	(0.080)
Observations	18,418	11,152	2,586	5,566	2,641
R^2	0.881	0.872	0.871	0.846	0.883
Firm FE	Y	Y	Y	Y	Y
Industry × Year FE	Y	Y	Y	Y	Y
Province × Year FE	Y	Y	Y	Y	Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 2 by industry where Related × Post is an indicator for related firms interacted with an indicator for years after 2006 and the dependent variable is log firm output, corresponding to Panel A of Table 4. Column (1) is our baseline result. Column (2) shows the results after we drop the production and supply of electric power and heat power industry. Column (3) shows the result for processing industries, including smelting and pressing of ferrous metals, smelting and pressing of nonferrous metals, processing of petroleum, coking, and processing of nuclear fuel. Column (4) shows the result for material industries, including manufacture of raw chemical materials and chemical products and manufacture of nonmetallic mineral products. Column (5) shows the results for mining industries, including mining and washing of coal and extraction of petroleum and natural gas. This table shows that we do not find significant differences in the response of related firms across different industries. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.23: Market Spillovers: Heterogeneous Effects across 4-Digit Industries with Top 1,000 Firms

Variables	ln(Output)			
	All Sample		Energy-Intensive Industries	
Spillover × Post	0.087^{***}	0.078^{***}	0.082^{***}	0.084^{***}
× $\mathbb{I}$[Top 1,000 Firm in 4-Digit Ind.]	(0.022)	(0.020)	(0.021)	(0.026)
Spillover × Post	0.037	0.040	0.040	0.046
× $\mathbb{I}$[No Top 1,000 Firm in 4-Digit Ind.]	(0.039)	(0.031)	(0.056)	(0.057)
Observations	2,557,940	2,557,940	843,313	843,313
R^2	0.840	0.856	0.831	0.848
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry-Level Controls	Y	Y	Y	Y
Firm-Level Controls		Y		Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 3 where Spillover × Post is an indicator for industry-level exposure to the Top 1,000 program interacted with an indicator for years after 2006 and the dependent variable is log firm output. This table corresponds Panel A of Table 5 but the specification allows for heterogeneous effects across industries with and without Top 1,000 firms. This table shows that market spillover effects are more pronounced in 4-digits industries with Top 1,000 firms; we do not estimate statistically significant effects in 4-digit industries without Top 1,000 firms. See Section 3 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.24: Robustness to Entry: Market Spillovers

Variables	ln(Output)			
	All Sample		Energy-intensive Industries	
Spillover × Post	0.083***	0.075***	0.092***	0.092***
	(0.022)	(0.020)	(0.021)	(0.025)
Observations	2,129,911	2,129,911	716,518	716,518
R^2	0.847	0.862	0.837	0.853
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry-Level Controls	Y	Y	Y	Y
Firm-Level Controls		Y		Y

Notes: Authors' calculations using data from the ASIF. This table shows estimates of Equation 3 where Spillover × Post is an indicator for industry-level exposure to the Top 1,000 program interacted with an indicator for years after 2006 and the dependent variable is log firm output. This table corresponds Panel A of Table 5 with the exclusion of firms that enter the ASIF data after 2006. It shows that average market-level spillovers led to a 7.5%–9.2% increase in output for unregulated existing firms. See Section 3 for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.25: Robustness of Effects on the Energy Efficiency of Regulated Firms

Variables	ln(Variable Input/Energy)			
Treat × Post	-0.020	-0.082*	-0.053	-0.029
	(0.043)	(0.047)	(0.048)	(0.048)
Observations	17,096	17,091	16,824	16,452
R^2	0.865	0.868	0.872	0.873
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the CESD and ASIF. This table shows estimates of Equation 2 where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variable is log energy efficiency. Energy efficiency in this table is defined as the inverse of the energy share in variable input costs, and variable input is calculated from the ASIF in terms of sales cost. This table corresponds to Panel C of Table 3 but with an alternative definition of energy efficiency and corresponds to a pooled version of the regression displayed in Figure A.14. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.26: Robustness of Effects on the Energy Efficiency of Related Firms

Variables	ln(Variable Input/Energy)			
Related × Post	-0.026	-0.033	-0.026	-0.008
	(0.078)	(0.080)	(0.088)	(0.088)
Observations	2,503	2,497	2,449	2,424
R^2	0.904	0.907	0.917	0.918
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Authors' calculations using data from the CESD and ASIF. This table shows estimates of Equation 2 where Related × Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006 and the dependent variable is log energy efficiency. Energy efficiency in this table is defined as the inverse of the energy share in variable input costs, and variable input is calculated from the ASIF in terms of sales cost. This table corresponds to Panel D of Table 4 but with an alternative definition of energy efficiency and corresponds to a pooled version of the regression displayed in Figure A.15. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.27: Structural Estimation: Robustness to Alternative Specifications

	Data	Baseline	Low ρ	High ρ	Low α	High α	Imperfect Substitution
1. Fixed Values							
Elasticity of substitution ρ		0.750	0.700	0.900	0.750	0.750	0.750
Within-conglomerate Elasticity of substitution ρ_c							0.900
Return to scale α		0.900	0.900	0.900	0.850	0.950	0.900
2. Method of Moments							
Efficiency depreciation δ		0.900	0.900	0.900	0.853	0.949	0.800
		(0.003)	(0.003)	(0.003)	(0.005)	(0.001)	(0.007)
Dispersion of ln-ability σ_m		1.239	1.500	0.579	1.359	1.063	1.271
		(0.055)	(0.060)	(0.045)	(0.087)	(0.053)	(0.097)
Survival threshold ϕ_1		0.609	0.461	0.792	0.579	0.435	0.985
		(0.166)	(0.134)	(0.193)	(0.273)	(0.118)	(0.535)
3. Policy Parameters							
Policy threshold $\tilde{\phi}$		9.289	14.093	2.658	10.915	6.050	10.977
4. Moments							
Share of firms $< 1M$	0.336	0.347	0.350	0.350	0.347	0.351	0.344
Share of firms $5-20M$	0.105	0.155	0.155	0.160	0.157	0.156	0.156
Share of firms $20-100M$	0.071	0.072	0.072	0.076	0.073	0.074	0.071
Share of firms $100M+$	0.025	0.026	0.026	0.027	0.027	0.027	0.025
Share of output $5-20M$	0.051	0.072	0.072	0.074	0.073	0.073	0.072
Share of output $20-100M$	0.146	0.144	0.144	0.152	0.147	0.148	0.142
Share of output $100M+$	0.722	0.733	0.733	0.723	0.729	0.729	0.734
Relative output 1st–2nd	0.289	0.348	0.350	0.347	0.347	0.350	0.347
Relative output 2nd–3rd	0.203	0.121	0.122	0.120	0.120	0.122	0.121

Notes: This table summarizes the parameters that we set or estimate. Panel 1 lists the various parameter values that we calibrate. Across all cases, we set α_l to match the cost share of variable inputs given α. Panel 2 reports the estimated parameter moments with standard errors in parentheses. See Section 5.1 for the detailed estimation procedure. Panel 3 reports the policy threshold $\tilde{\phi}$. This threshold is selected to match the share of energy use by regulated firms, which itself depends on the parameter values. Panel 4 reports the data moments and the moments predicted by the model parameters. Section 7.3 discusses the results when we vary α or ρ, and Section 7.4 discusses the results when we set $\rho_c = 0.9$.

Table A.28: **Heterogeneous Spillover Effects by Local Pollution and Density**

Variable	High to Low	Horizontal	Low to High
Related × Post	0.115^{**}	0.101	0.224^{***}
	(0.047)	(0.072)	(0.073)
Observations	10,256	3,740	3,457
R^2	0.895	0.883	0.897
Output Share	57.9%	13.9%	28.2%
Aggregate Effect	6.7%	1.4%	6.3%
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y
Firm-Level Controls	Y	Y	Y

Notes: Authors' calculations using data from the ASIF. This table show estimates of Equation 2 where Related × Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006 and the dependent variable is log firm output. The estimates correspond to Panel A of Table 4 but divide related firms into three groups according to their pollution exposure. Column (1) includes related firms whose city-level pollution exposure is less than (by more than 10%) that of their corresponding Top 1,000 firms; Column (2) includes related firms whose city-level pollution exposure is similar to (within a 10% difference) that of their corresponding Top 1,000 firms; and Column (3) includes related firms whose city-level pollution exposure is more than (by more than 10%) that of their corresponding Top 1,000 firms. City-level pollution exposure is defined by city-level so_2 density × city population. This table shows that related firms in places with lower pollution exposure increased output by 11.5% and related firms in places with similar pollution exposure increased output by 10.1% while related firms in places with more pollution exposure increased output by 22.4% after the policy implementation. However, considering that a higher share of the related output (57.9%) was concentrated in less exposed areas, we see similar output increases in more and less exposed areas. See Section H for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.29: Heterogeneous Spillover Effects by Location

Variable	East&Mid to West	Within Region	West to East&Mid
Related × Post	0.117	0.164^{***}	0.177
	(0.129)	(0.041)	(0.275)
Observations	1,133	9,230	456
R^2	0.900	0.887	0.907
Output Share	8.8%	83.4%	7.8%
Aggregate Effect	1.03%	13.7%	1.38%
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y
Firm-Level Controls	Y	Y	Y

Notes: Authors' calculations using data from the ASIF. This table show estimates of Equation 2 where Related × Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006 and the dependent variable is log firm output. The estimates correspond to Panel A of Table 4 but divide related firms into three groups according to their location. Column (1) includes related firms located in the Western regions while their corresponding Top 1,000 firms are located in the Eastern or Mid-China regions (if a firm is related to multiple Top 1,000 firms, then we require that over 30% of the corresponding Top 1,000 firms are in different regions from the related firm). Column (2) includes related firms located in the same region with their corresponding Top 1,000 firms (if a firm is related to multiple Top 1,000 firms, then we require that over 70% of its corresponding Top 1,000 firms are in same region as the related firm). Column (3) includes related firms located in the Eastern or mid-China regions while their corresponding Top 1,000 firms are located in the Western regions (if a firm is related to multiple Top 1,000 firms, then we require that over 30% of the corresponding Top 1,000 firms are in different regions from the related firm). In addition, we leave out all firms in the electric power generation and supply industry. This table shows that related firms located in less developed places than the corresponding regulated firms increased output by 11.7%, that related firms in places similar to the corresponding regulated firms increased output by 16.4%, and that related firms in more developed places than the corresponding regulated firms increased output by 17.7% after the policy implementation. However, considering that most of the related output (83.4%) was concentrated in the same region as the regulated firms, within-region spillover is the main contributor of related-firm spillovers, and the cross-sectional effects mostly offset each other. Standard errors clustered at the firm level are shown in parentheses with p-values below.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.30: Heterogeneous Effects on Regulated Firms by Electricity Consumption

Variables	ln(Energy)	ln(Output)	ln(Efficiency)
Treat × Post × Electricity-Intensive	0.044	-0.029	-0.075
	(0.081)	(0.074)	(0.081)
Treat × Post	-0.169**	-0.154***	0.014
	(0.068)	(0.058)	(0.068)
Observations	22,623	22,471	22,471
R^2	0.890	0.886	0.841
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Author's calculations using data from the CESD and ATS. This table shows estimates of heterogeneous effects within regulated firms by firm electricity consumption. Electricity-intensive is an indicator for regulated firms above the median electricity consumption in 2007 (the first year for which we have electricity consumption data in ATS). This table shows no significant effect heterogeneity between regulated firms that use more electricity and those that use less electricity. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.31: Effects of the Program on Regulated Firms' Energy Mix

Variables	Coal Share	Petroleum Share	Gas Share
Treat × Post	-0.005	0.001	0.003
	(0.007)	(0.004)	(0.006)
Observations	23,151	23,151	23,151
R^2	0.679	0.636	0.667
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Notes: Author's calculations using data from the CESD. This table shows estimates of Equation 2 where Treat × Post is an indicator for regulated firms interacted with an indicator for years after 2006 and the dependent variables are firm coal consumption as a proportion of firm total energy consumption, firm petroleum consumption as a proportion of firm total energy consumption, and firm gas consumption as a proportion of firm total energy consumption (all calculated in standard coal equivalent). Note that the CESD does not include firm electricity consumption data, and thus, the firm total energy consumption in this analysis is a sum of firm coal, petroleum and gas consumption. This table shows that the energy mix of regulated firms did not change after policy. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.32: Heterogeneous Effects on Regulated Firms by Firm Size and Energy Consumption

A. Heterogeneous Effects on Output by Firm Size

Variables	ln(Output)			
Treat × Post × Large	-0.031	-0.074	-0.053	-0.078
	(0.076)	(0.071)	(0.071)	(0.072)
Treat × Post	-0.071	-0.180***	-0.167***	-0.098
	(0.064)	(0.059)	(0.059)	(0.060)
Observations	23,079	23,074	22,648	20,147
R^2	0.879	0.884	0.886	0.891

B. Heterogeneous Effects on Output by Energy Consumption

Variables	ln(Output)			
Treat × Post × High Energy Consumption	0.021	-0.026	-0.007	-0.017
	(0.076)	(0.071)	(0.071)	(0.072)
Treat × Post	-0.102*	-0.211***	-0.197***	-0.131**
	(0.058)	(0.056)	(0.056)	(0.057)
Observations	23,119	23,114	22,683	20,168
R^2	0.879	0.885	0.887	0.891

C. Heterogeneous Effects on Energy Efficiency by Energy Consumption

Variables	ln(Energy Efficiency)			
Treat × Post × High Energy Consumption	0.064	0.037	0.037	0.028
	(0.079)	(0.078)	(0.080)	(0.080)
Treat × Post	-0.003	-0.091	-0.069	-0.029
	(0.056)	(0.056)	(0.057)	(0.059)
Observations	23,119	23,114	22,683	20,168
R^2	0.837	0.840	0.842	0.848
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

Notes: Author's calculations using data from the CESD. This table shows estimates of heterogeneous effects within the regulated firms by firm size and energy consumption. Panel A shows output estimates based on firm size of the regulated firms, where *Large* is an indicator for regulated firms above the median output size within each 4-digit industry in the selection year (2004). Panels B and C show output and energy efficiency estimates based on firm energy consumption of the regulated firms, where *High Energy Consumption* is an indicator for regulated firms with above-median energy consumption within each 4-digit industry in the selection year (2004). This table shows no significant effect heterogeneity among the different regulated firms, which indicates that there was probably little reallocation within treated firms. This result is consistent with the program design whereby the energy savings target were allocated according to each firm's total energy consumption before the policy, thus leading all regulated firms to tend to have a similar response. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.33: Estimates of Total Pollution Damages

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Source/ Method	Total Pollution Damages (Bn \$)	Year	CO_2 Emissions (Bn tn)		Damages Per Ton	
			All Sources	Coal Only	All Sources	Coal Only
Mohan et al. (2020) Gross External Damage	\$107.99	2006	6.38	4.91	\$16.93	\$22.00
World Bank (2007) Willingness to Pay	\$62.79	2003	4.45	3.32	\$13.83	\$18.91
World Bank (2007) Adjusted Human Capital	\$19	2003	4.45	3.32	\$4.23	\$5.72
Ito and Zhang (2020*b*) Willingness to Pay	\$55.8	2013	9.80	7.49	\$5.69	\$7.45

Notes: This table uses estimates of total pollution damages from a range of studies (column (1)) and estimates of CO_2 emissions from Ritchie, Roser and Rosado (2020) (column (4)) to compute average damages per ton of emissions from all sources (column (6)). Column (7) computes this average damage under the assumption that all emissions derive from coal usage. See Appendix G for additional details.

Table A.34: Summary Statistics for Top 10,000 & Top 10,000 Related Firms

	Top 10,000		Top 10,000 Related	
	Number	Output	Number	Output
Panel A. All Firms Matched in ASIF				
All Industries	14300	2201.53	9787	885.63
Energy-Intensive Industries	10962	1861.45	5383	925.67
Panel B. Exclude Top 1,000 and Top 1,000 Related				
All Industries	12524	1588.03	8522	694.00
Energy-Intensive Industries	9222	970.27	4152	545.35

Notes: This table reports summary statistics for Top 10,000 firms and Top 10,000 related firms in the same 4-digit industries. Panel A shows firm counts and mean output from the ASIF for Top 10,000 and Top 10,000 related firms, while Panel B excludes Top 1,000 firms and their related firms. Output value is in million RMB. This table shows that Top 10,000 firms have fewer related firms than Top 1,000 firms. On average, a Top 10,000 firm in an energy-intensive industry has 0.5 related firms in the same 4-digit industry, and the size of the related firm is approximately 50% of that of the Top 10,000 firm.

Table A.35: Direct and Spillover Effects of the Top 10,000 Program

A. Regulated Firms

Variables	ln(Output)			
Treat × Post	-0.158***	-0.156***	-0.163***	-0.130***
	(0.007)	(0.007)	(0.007)	(0.007)
Observations	319,346	319,346	319,346	310,966
R^2	0.879	0.880	0.883	0.895

B. Related Firms

Variables	ln(Output)			
Related(0%-90%) × Post	0.001	-0.001	0.007	0.015
	(0.021)	(0.021)	(0.023)	(0.021)
Related(90%-100%) × Post	0.067	0.083*	0.107**	0.115**
	(0.049)	(0.050)	(0.052)	(0.050)
Observations	13,397	13,397	13,397	13,082
R^2	0.873	0.875	0.879	0.892
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry × Year FE		Y	Y	Y
Province × Year FE			Y	Y
Firm-Level Controls				Y

C. Market Spillover

Variables	ln(Output)			
	All Sample		Energy-Intensive Industries	
Incremental Spillover × Post	0.045***	0.043***	0.019	0.018
	(0.015)	(0.014)	(0.016)	(0.014)
Observations	949,277	949,277	280,734	280,734
R^2	0.867	0.878	0.861	0.873
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry-level Controls	Y	Y	Y	Y
Firm-Level Controls		Y		Y

Notes: Authors' calculations using data from the ASIF. This table shows effects of the Top 10,000 Program. Panel A corresponds to Panel B in Table 3, while Treat×Post is an indicator for Top 10,000 firms interacted with an indicator for years after 2011. Panel B corresponds to Panel A in Table 4, while Related here is an indicator for Top 10,000 related firms, and 0%-90% (90%-100%) refers to percentile ranges of firm size. Panel C corresponds to Panel A of Table 5 while Incremental Spillover here refers to the industry-level exposure to the Top 10,000 program minus the Top 1,000 exposure. In this table, we show that, when the size-dependent policy expands, the regulated firms still experience significant negative effects but the related-firm spillover and market spillover effects are significantly reduced. Standard errors clustered at the firm level are shown in parentheses with p-values below. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.