

Regulating Conglomerates: Evidence from an Energy Conservation Program in China

Online Appendix

Qiaoyi Chen Zhao Chen Zhikuo Liu
Fudan University Fudan University Fudan University

Juan Carlos Suárez Serrato Daniel Yi Xu
Stanford University GSB Duke University
& NBER & NBER

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 the data construction in more detail. Appendix D discusses the robustness of the effects of the program on regulated firms and 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 discusses the robustness of our spillover analyses. Appendix G uses our estimated model to compute the shadow costs of the program and relates them to the aggregate effects of the program. Appendix H shows that our welfare analysis is not confounded by reallocation across regulated firms of different size. Appendix I describes how we can incorporate the health benefits from pollution reduction in our policy analysis. Appendix J explores the robustness of our model results to allowing for endogenous energy efficiency, heterogeneous energy efficiency, alternative parameterizations, or alternative patterns of substitution within a conglomerate, or in a hypothetical scenario without conglomerates. Appendix K shows that the program did not lead to a significant geographic reallocation of economic activity. Appendix L estimates the effects of the Top 10,000 program. Appendix M provides a detailed derivation of our baseline model. Appendix N describes how we model alternative regulations. Appendix O examines the effects of the program when we shut down market- and conglomerate-level spillovers. Finally, Appendix P details the derivations of the model extensions.

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.

1. Policy Goals

A key focus of the 11th Five-Year Plan was to significantly improve energy efficiency. Specifically, the goal of the plan was to reduce the ratio of aggregate energy consumption 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 tons of coal equivalent by 2011 (NDRC, 2006).³¹ Table A.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 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.³² According to these official documents, pollution reduction and the reallocation of economic activity across different regions were not primary goals of the program.³³ While Wang et al. (2008) briefly mention environmental benefits from the program, the focus is on the potential benefits from reducing energy use.

³¹In addition to the Top 1,000 program, the 11th Five-Year Plan included government support to encourage energy savings for all firms. Zhou et al. (2010), Price et al. (2011), and Zhao et al. (2014) discuss other forms of government support.

³²Specifically, 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.

³³We also document empirically that the program did not lead to disproportionate production increases in more polluted or more populated areas (see Table A.2) and that it did not lead to significant reallocation of economic activity across different geographic regions (see Table A.3).

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, to decrease energy consumption, or to 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, which may include 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 tons in carbon dioxide emissions. Because the program lasted 5 years, the implied average yearly reduction would be 48.08 million tons of carbon emissions. Below, we compare the Top 1,000 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 tons over 14 years, for an average yearly reduction of 10.7 million tons.
- 2) The “2030 Climate and Energy Framework” from the European Union (2014–2030) implies a carbon emission reduction of 859.44 million tons over 17 years.³⁴ The average yearly reduction for this program is 50.56 million tons.
- 3) The “Law on Energy Transition for Green Growth (LTECV)” from France (2015–present) implies a carbon emission reduction of 135.71 million tons over 16 years.³⁵ The average yearly reduction for this program is 8.48 million tons.
- 4) The “Agreement with Industry on CO₂ Emissions Cuts” from Germany (2000–2012) implied a carbon emission reduction of 224.27 million tons over

³⁴The program’s objectives are to reduce the 2014 level of emissions of 3537.80 million tons to 60% of the 1990 level of 4463.94 million tons, which implies a reduction of 859.44 million tons.

³⁵The program’s objectives are to reduce the 2015 level of emissions of 433.22 million tons to 60% of the 1990 level of 495.85 million tons, which implies a reduction of 135.71 million tons.

13 years.³⁶ The average yearly reduction for this program is 17.25 million tons.

- 5) The “Climate VISION (Voluntary Innovative Sector Initiatives: Opportunities Now)” from the United States (2003–2012) implied a carbon emission reduction of 500 million tons over 10 years for an average yearly reduction of 50 million tons.

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.

2. Intended Program Implementation

As described by Price et al. (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 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.³⁷

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

³⁶The program’s objectives are to reduce the 2000 level of emissions of 957.53 million tons to 65% of the 1990 level of 1128.09 million tons, which implies a reduction of 224.27 million tons.

³⁷Zhao et al. (2016) discuss the limitations of the metrics used to measure energy savings.

our narrative and interview evidence below, the original energy saving formulas in CSBTS (1991) were not used by local governments to assess firm performance.³⁸

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 Hainan Province (2006) and of Chongqing City (2007), as well as 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 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' electricity use by relying on the "live energy consumption monitoring system" 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 that 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.

In practice, local officials had strong incentives to implement the program, because 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. As we confirm in our interviews with executives of Top 1,000 firms below, the "live energy consumption monitoring system" allowed the governments both to track energy usage and to credibly threaten to shut down production at plants that exceeded their energy

³⁸For 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.

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).

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 (which 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.4 displays our interview list and how each stakeholder is connected to the program.

INTERVIEW WITH POLICY DESIGNER

The designer of the Top 1,000 Energy Conservation Program—number (1) in Figure A.4—wrote many reports and books about the energy saving target allocation and assessment of the Top 1,000 Program during the 11th Five-Year Plan (see, e.g., Wang et al., 2008; Price et al., 2008, 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 [National

Development Reform Commission] 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.

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 (NDRC) in Municipality A (listed as number (2) in Figure A.4). 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 sets 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 11th 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 11th 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.4). 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 11th 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 governments 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 savings. 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.

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.4) to corroborate the interviews with local officials.

First, the auditor's report confirms that it was not feasible to rely on the 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.

INTERVIEWS WITH TOP 1,000 FIRMS

We asked the same set of questions to each of the Top 1,000 firm representatives interviewed. 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 the government assess your energy saving performance?
- 4) Was your production affected by the program?
- 5) Did you make energy-saving investments during the 11th 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 11th 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 11th 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 11th 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 11th Five-Year Plan, but the energy saving policy during the 11th 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 a chemical engineer in the company who is familiar with the energy conservation policy during the 12th Five-Year Plan. (The target under the “Top 10,000 Energy Conservation and Carbon Reduction Program” during the 12th Five-Year Plan, which asked energy-intensive firms to improve their energy efficiency and save a total of 250 mcte, is very similar to that under the “Top 1,000 Energy Conservation Program” during the 11th Five-Year Plan.) 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 12th 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 scraps, 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, and improve energy efficiency and self-generated electricity.

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 11th 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 is 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 stops 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.

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 et al., 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 11th Five-Year Plan, local government officials had strong incentives to meet energy savings targets in their own jurisdictions. 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 et al. (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.³⁹

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.

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.

³⁹In Table A.4, 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.5, 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.

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.

Another attempt at evaluating the effects of the Top 1,000 program, Ke et al. (2012), uses time-series data on energy consumption and value added for 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), because Ke et al. (2012) rely only on time-series data, it is not possible to separate the effects of the program from underlying 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 enable 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 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” (Porter and van der Linde (1995)), which posits that firms respond to regulation through innovation that leads to improvements in firm performance. To evaluate this hypothesis, Shen et al. (2015) use a difference-in-differences design and ASIF data from between 2005 and 2007 to study the impact of the Top 1,000 program on firms’ export volume. They estimate a positive treatment effect of 10.3% on the export volume of 240 Top 1,000 firms relative to that of unregulated firms in the same industry. 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 et al. (2021) use the same data and time window but focus on chemical firms and find that the program decreased firm annual total factor productivity (TFP) growth. While Ai et al. (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 et al., 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, in Figure A.5 we use Incopat (2001-2010) data to show 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 I.B, 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.

A potential concern is that firms may manipulate reported data as a way to meet compliance requirements. Indeed, Karplus et al. (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.6, we show that firms report similar levels of 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.

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 are 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 permits 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 can take a number of steps if it discovers that a firm has falsified data, including:

- 1) Instructing the firm to immediately stop falsifying data;
- 2) Requiring the firm to re-monitor 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, taking 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.

1. Data Merge and Summary Statistics

Table A.6 shows the results of our data construction. We merge the lists of regulated firms using both the firm name and unique legal identifier. Because 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.

Panel A of Table A.7 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 because Top 1,000 firms are more likely to be in energy-intensive heavy industries. As we show in the paper, our empirical analyses are robust to controlling for these firm-level characteristics.

To compute an overall measure of firm-level energy use, we combine data on the three primary sources of energy: coal, oil and gas. We use a factor of 0.9kgce/kg to convert coal to standard coal, and a factor of 1.4286kgce/kg to convert oil to standard coal. To account for the fact that different industries use different types of gas, we use three industry-specific conversion factors. For industries that mainly rely on natural gas, we use a factor of 1.33kgce/cu.m to convert gas to standard coal. For industries that mainly rely on other types of gas, we use a conversion rate of 0.5kgce/cu.m, which corresponds to the average conversion

rate among all other types of gas that listed in the China Energy Statistical Yearbook. Finally, we use a zero conversion rate for industries that mainly use gas generated as a byproduct of their production process (coking, iron, steel, and ferroalloy smelting)(NBS, 2001-2010). Note that firms that use gas as their primary energy source are quite scarce in China, with natural gas contributing less than 4% to the total energy consumption in energy-intensive sectors. Our results are also robust to focusing solely on coal and oil data.

Panel B of Table A.7 shows that we can identify 46,178 related parties of Top 1,000 firms in the CARD data.⁴⁰ Because 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.⁴¹ Merging the CARD data with the ASIF and the CESD alleviates the potential concern that some of the related firms may be holding companies (and thus not engaged in production), ensuring that our results are driven by real economic activity in industrial firms. In our baseline regressions, we also require related firms to be in the same 4-digit industry as a related Top 1,000 firm.⁴² Our main sample of related firms includes 2,466 industrial firms. Because 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.

Panel B of Table A.7 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 only 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.⁴³ Moreover, it is important to note that regulated firms could not escape the regulation by splitting into smaller firms. Because local policymakers face regional energy use targets, they have strong incentives to ensure that any firm that was initially regulated meets its energy target.⁴⁴ If firms split, the energy use targets would accompany the firms

⁴⁰Figure A.7 depicts all the possible links that we consider.

⁴¹As 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.

⁴²Omitting firms in unrelated industries is unlikely to affect our results because Top 1,000 firms would not be able to shift production to service firms or very small firms.

⁴³Using 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.

⁴⁴In 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 K, we also show that the program did not significantly alter the allocation of production across cities with different levels of emissions and population density.

after any such separation.

Table A.8 shows the effects of the sample restrictions on the number of Top 1,000 firms in the data. As we show in Figure A.8 and Table A.9, our results are robust to different sample construction decisions.

D. ROBUSTNESS OF EFFECTS ON REGULATED FIRMS

This appendix explores the robustness of the results in Section II. We first consider general robustness checks in Appendix D.1 and then consider the role of other environmental policies in Appendix D.2.

1. Robustness to Sample Construction, Additional Outcomes, and Heterogeneous Effects

We now explore the robustness of the effects of the Top 1,000 program on regulated firms. First, as we discuss in Section I, our analysis sample excludes some firms based on missing data and firms in industries that rely primarily on electricity. Table A.9 and Figure A.8 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. 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.10–A.11 and Figures A.9–A.10 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.

Second, our results are robust to constructing the sample of firms in our regression in different ways and to the exclusion of new firms and those that exit during our sample period. Table A.12 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.13 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.

Third, an additional 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 Figure 3 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.

We also explore the effects of the program on other outcomes. In Table A.14 and Figure A.11, we show that regulated firms experienced a decline in their probability of investing after the regulation was enacted. Similarly, Figure A.12 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.5 shows no increase in the filing of patents related to energy efficiency in regulated firms.

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.15 shows similar effects of the program across different industry groups, and Table A.16 shows that SOEs do not have statistically distinct responses to the program.

2. *Robustness to Controlling for Competing Policies*

This appendix explores the robustness of our results to accounting for competing policies. 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. The policy applied to firms with chemical oxygen demand (COD), NH_4^+ , SO_2 , smoke, and industrial dust emissions above a certain threshold in 2005. Half of the Top 1,000 firms and 14% of the Top 10,000 firms were included on the NSMF list.

We 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.17 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 policies and 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.

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.⁴⁵

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 (NBS, 2001-2010). 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.13 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 mix⁴⁶

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 experienced smaller declines in output and larger declines in our coal-based measure of energy use. However, as Table A.18 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. We also proxy for the switching cost to electricity by assessing each firm's change in electricity consumption from 2007 to 2010, relative to its industry average.⁴⁷ We then estimate the heterogeneous effect of the program using an indicator for whether the firm was above the median in terms of electricity use change. Table A.19 reports no significant difference in output decline based on this measure of heterogeneity. The non-electricity energy use decline

⁴⁵Note 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.

⁴⁶A stable aggregate energy mix may also align with a scenario in which the Top 1,000 companies transition from coal to electricity, while unregulated firms exhibit a contrasting pattern. This transition could potentially be facilitated by factors such as a reduction of input prices. While our aggregate data cannot definitively rule out this possibility, the estimated program impact on energy efficiency should be interpreted as an upper bound. To see this, recall our energy efficiency measure is given by the ratio of total sales to coal usage; a measure that could also reflect increases in electricity use would therefore show smaller gains in energy efficiency.

⁴⁷One potential limitation of this approach is the possibility that the classification itself might be subject to influence by the Top 1,000 program, a phenomenon sometimes referred to as the "bad controls problem" in the literature.

is somewhat smaller for firms that experience larger increases in electricity use, although the point estimate is not statistically significant. These results are inconsistent with the hypothesis that potential substitution from coal to electricity is a central margin of substitution in this time period.

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.20 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. While a drawback of our setting is the lack of firm-level data on electricity use prior to the Top 1,000 program, we use both industry-level data and firm-level changes in electricity use after the Top 1,000 program to show that substitution from coal to electricity was not a central margin of substitution in this time period.

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. This appendix explores these assumptions empirically by estimating the effects of the program on firms' capital-labor, energy-labor, and energy-capital ratios. Table A.21 and Figure A.14 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 1), 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. In practice, the evolution of a firm's capital stock over time is shaped by its decisions to cease investments and allow its existing capital to depreciate. As we report in Table A.14, firms regulated by the Top

1,000 program were less likely to have positive investment following the onset of the program. To account for this possibility, Appendix M.3 develops an extension of our model that allows firms to adjust their capital in response to the program. As we discuss in Section J.3, this alternative assumption leads to a comparable albeit slightly larger estimate of the government willingness to pay (GWTP) (see Panel C of Figure A.2).

F. ROBUSTNESS OF SPILLOVER ANALYSES

This appendix describes robustness checks related to the spillover effects of the Top 1,000 program. We first study the robustness of our results to the set of control firms we use. In Section III, 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. We obtain similar results when we use the entropy balancing method of Hainmueller (2012) to find controls for related firms (see Figure A.15 and Tables A.22–A.23).

Second, 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 conglomerates after the program.

Third, these results are robust to alternative definitions of ownership networks. Table A.24 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 the power generation industry (see Table A.25 and Figure A.16).

Finally, we assuage concerns that our results may be affected by data quality issues by showing similar effects when we rely on tax data to measure the output of related firms (see Figure A.17 and Table A.26).⁴⁸

In Section III, we compare the direct effects of the Top 1,000 program on regulated firms to the spillover effects on related firms (see footnote 13). We estimate that the output increase of related firms was commensurable with 41% of the 14.5% decrease in the output of regulated firms. We consider the sensitivity of this estimate to the measurement of business networks in three ways. First, using results from the specifications in columns (3) of Tables 1 and 2, we obtain an updated estimate of 27%. (*i.e.*, $27\% \approx 2.45 \times 19.3\% \times 11.8\% / 20.4\%$). Second, if we supposed that regulated firms have an average of 3 related firms, we find that spillovers would account for 51% of the output decline in regulated firms.

⁴⁸We also find positive spillover effects on other measures of economic activity. Table A.27 shows estimates of positive spillover effects on sales, profits, capital and labor (see Figure A.18 for the corresponding event studies).

Finally, we consider the role of related firms that are not in the ASIF, which are not present in our analysis. 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.

We also explore the potential for heterogeneity in the spillovers across related firms. Table A.28 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.29 shows no significant differences in how related firms in different industries responded to the program.

As with the spillovers to related firms, market spillovers are concentrated in the same 4-digit industries as Top 1,000 firms; see Table A.30. We also confirm that the market spillover results are not driven by firm entry. Specifically, in Table A.31, we report similar estimates of market-level spillovers when we restrict the sample to firms in operation prior to 2006.

G. SHADOW COST OF THE POLICY

In Section IV, we show that the shadow cost is the fundamental building block of our model outcomes, including the aggregate and welfare effects of the program. This appendix quantifies this shadow cost. To compute the shadow cost of the policy, we solve for the regulated equilibrium as in Proposition 5. Panel A of Figure A.3 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. Because 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.⁴⁹

A somewhat surprising feature of Panel A of Figure A.3 is that the shadow cost appears to be constant with respect to conglomerate productivity ϕ . Panel B of Figure A.3 zooms in to show the different shadow costs for regulated firms (i.e., $\phi > \tilde{\phi}$). As we discuss in Section IV, 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

⁴⁹In 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%.

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 (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:

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

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 -34.3% estimate in the first column of Panel B of Table 3 and our value of $\alpha_l = 0.80$, we estimate a shadow cost of 8.95% .⁵⁰ 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 A.3 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 A.3 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.⁵¹ These calculations showcase the importance of accounting for both market and

⁵⁰This 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 A.3, there is little variation in $\lambda(\phi)$.

⁵¹This 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% .

conglomerate spillovers in the measurement of the shadow cost of the regulation.

We also use our model to compute the shadow cost under the alternative policies we consider in Section VI.B. Panel A of Figure A.3 shows that a conglomerate-level regulation has a lower shadow cost of 4.9% (instead of 8.7%). This figure also shows that the policymaker could obtain the same energy reduction as that under the Top 1,000 program by taxing inputs at 4.9%, which is very similar to the shadow cost of the conglomerate-level regulation.⁵²

H. 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 further 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 could 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:

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

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:

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

Using Δ to denote percentage changes, we have that:

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

⁵²As 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. Because 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% ($\approx \frac{4.9\%}{15\%}$).

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 total output of Top 1,000 firms.

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

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

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

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

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 (i.e., whether the Top 1,000 program had heterogeneous impacts). Table A.32 estimates the effects of the Top 1,000 program on regulated firms, as in Table 1, while allowing 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 1. We also obtain average effects that are very similar to those in Table 1 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 1 reports a value of 0.204. Our derivation above shows that the program's aggregate effects on energy use can be affected by the presence of heterogeneity in its firm-level output and energy efficiency effects 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 thus show that the program did not lead to significant reallocation of

economic activity and energy use across regulated firms, which suggests 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.⁵³

I. INCORPORATING POLLUTION DAMAGES

Our analysis of the welfare effects of the program captures 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 positive global externalities associated with reducing the amount of CO₂ in the atmosphere. If this were the only benefit, our estimates would correspond directly to the social cost of carbon (SCC). However, a reduction in energy use also leads to reductions in local pollution and its associated damages—specifically, health damages.

This section details how we incorporate pollution damages into the welfare effects of the policy. As we discuss in Section VI.A, 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— κ_0 —and the local health costs of associated pollution— κ_1 . These parameters are related to values of the SCC and to estimates of pollution damages as follows:

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

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 local pollution damage per ton of carbon.

To implement these expressions, we rely on four estimates of pollution damages. First, we use the estimates of gross external damage (GED) from Mohan et al. (2020), which derive the values of both social cost of carbon and local damage from pollution 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 thus equaled \$337.67 billion. Because Mohan et al. (2020) rely on a \$36 estimate of the SCC, the damages from China’s

⁵³One 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.

6.38 billion tons of CO₂ emissions in 2006 are \$229.68 billion. Thus, the estimate from Mohan et al. (2020) implies total pollution damages of \$107.99 billion (= \$337.67 – \$229.68).⁵⁴

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 RMB in 2003. Converting to US dollars with an exchange rate of 8.28 RMB per dollar yields a value of total health damages of \$62.79 billion.

As a third estimate of pollution damages, we rely on an evaluation from the World Bank (2007) that uses an adjusted human capital approach. The World Bank (2007) reports an estimate of total health costs from air pollution of 157.3 billion RMB in 2003. Converting to dollars with an exchange rate of 8.28 RMB 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 demonstrate reduces local life expectancy. Ito and Zhang (2020) use this discontinuity and the purchase of air purifiers to estimate households’ willingness to pay to reduce their exposure to air particulates (PM₁₀). They estimate that, across China, the total willingness to pay to lower PM₁₀ exposure by 1 μ g/m³ is \$0.45 billion USD. On the basis of the 124 μ g/m³ national average of PM₁₀ cited in Ito and Zhang (2020), total pollution-related health damages in China in 2013 according to this measure were \$55.8 billion USD.

Table A.33 collects these estimates of total pollution damages in China in column (2). Column (4) lists estimates of total CO₂ emissions in China from Ritchie et al. (2020). In column (6), we report the ratio of total damages from column (2) to the total CO₂ emissions from column (4). Across these four methodologies, we estimate that reducing CO₂ 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, because 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 CO₂ emissions from coal in column (5) (Ritchie et al., 2020). Dividing the total pollution damages in column (2) by the CO₂ emissions from coal in column (5) gives an upper bound of the

⁵⁴As Muller et al. (2011) note, the external damages from pollution may be higher or lower depending on where it is emitted. As we show in Appendix K, 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.

local pollution damages associated with reducing the energy use of Top 1,000 firms of \$5–\$22 per ton of carbon. It is noteworthy that these upper-bound estimates are not much larger than the estimates in column (6). This is a consequence of the fact that approximately 75% of Chinese CO₂ emissions are due to coal, so attributing all 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 Top 1,000 policy is justified if the total benefits from energy reduction are greater than \$161. Subtracting the \$4–\$17 estimates of the local health benefits from pollution reduction in column (6) of Table A.33 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 contrasts 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 that the US had a GDP of \$13.8 trillion and emitted 6.05 billion tones of CO₂ at the time, we compute that reducing one ton of CO₂ has a GED of \$114. Subtracting the authors’ assumed SCC of \$36 implies a \$78 value of health benefits per ton of CO₂. 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.

J. MODEL RESULTS: EXTENSIONS AND ROBUSTNESS

This appendix explores four extensions of our model results. 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 for imperfect substitution between products produced by firms in a given conglomerate.

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 II–III. 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 P.1 provides additional details of this model extension.

We assume that a 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 regulation does not impact the choice of ν_i for unregulated firms. We use these results and the fact that $\nu_i = \frac{l_i}{e_i}$ to restate the conglomerate problem as

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

where we substitute the regulatory constraint into the cost of energy efficiency and abstract from the cost of the regulated energy, $p_e \xi e_1^*$, because 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

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

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 regulated firm improves by $\frac{\nu_1}{\nu^*}$ relative to that in the unregulated case.⁵⁵ 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 because Propositions 1–2 continue to describe the unregulated equilibrium and because 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. Because we estimate statistically insignificant effects of the program on the revenue-to-energy (i.e., energy efficiency) ratio of regulated firms, we calibrate γ using the upper bound of the 95% confidence interval.⁵⁶ In Appendix P.1, we show that when $\gamma = 7.4$, the improvement in the revenue-to-energy ratio of regulated firms in our model equals 11.6%, which matches the 95% confidence interval of column (1) in Panel C of Table 1.⁵⁷ 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 regulated firm. In this case, the shadow cost is 6.5% instead of the 8.7% we find 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 A.2 plots the aggregate effects of the regulation at different values of γ . This model nests our baseline model under the assumption that $\gamma \rightarrow \infty$. 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 regulation 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 regulation reduces aggregate energy use by 5.6% for a price increase of 2.8%. This calculation also yields a bound on the GWTP: as we show in Panel A of Figure A.2, the

⁵⁵The 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 $\lambda(\phi) = [\nu_1 c'(\nu_1) + c(\nu_1)] - (\nu^* c'(\nu^*) + c(\nu^*))$.

⁵⁶We also find that the Top 1,000 program did not lead to improvements in the variable-costs-to-energy ratio. Figure A.19 and Table A.34 show a null effect on this measure of energy efficiency for regulated firms, and Figure A.20 and Table A.35 reveal the same result for related firms.

⁵⁷Alternatively, our model implies a value of $\gamma = 12.8$ if we target the 99% confidence interval of column (4) of Panel C of Table 1.

program raises welfare at this value of γ as long as the $\text{GWTP} \geq \$114$.⁵⁸

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.⁵⁹

2. *Heterogeneous Energy Efficiency*

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

One possibility is that the government selected the firms for regulation because those firms are particularly 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 regulation may lead to smaller reductions in energy use if regulated firms are more energy efficient than other firms.

Appendix P.2 generalizes our analysis to allow the energy efficiency of related firms to differ from that of regulated 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, as 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 regulated firms by a factor of ν_O . These differences impact both the response of related and unrelated firms to the regulation as well as overall energy use. To explore the sensitivity of our results to baseline 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 regulated firms (leading to negative leakage; e.g., Baylis et al., 2014).⁶⁰

⁵⁸The 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.

⁵⁹The 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.

⁶⁰Unfortunately, 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

Panel B of Figure A.2 shows the effect of heterogeneity in energy efficiency on our welfare calculations. We first assume that firms related to regulated 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 unregulated, unrelated 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 within-conglomerate energy leakage is an important contributor to the overall energy effects of the program. If we alternatively assume that regulated firms are particularly energy inefficient, i.e., set $\nu_R = \nu_O = 1.2$, the program can be rationalized with a GWTP value as low as \$134. To the best of our knowledge, regulated firms are not particularly inefficient relative to their related firms. Nevertheless, this calculation provides interesting bounds for the welfare effects of the program.

3. Robustness to Alternative Parameters and Timing Assumptions

The structural model in Section IV uses 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. We reestimate the structural parameters for different values of the calibrated parameters following the same procedure as in Section IV. Finally, we solve for the regulated equilibrium that is implied by every pair of values of α and ρ .⁶¹

Panel C of Figure A.2 shows how varying these parameters affects the estimated aggregate effects of the regulation 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 A.2 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 A.2 by the blue triangles. In this case, we find that lower values of α lead to both larger reductions in

competing firms in the market. Using the CESD, we observe in Panel A of Table A.7 that regulated 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, in Figure A.21, we report the SCC that rationalizes the policy when energy efficiencies differ by up to 50%.

⁶¹The value of α_I 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.36 reports the estimated model parameters across different specifications.

energy use and larger increases in price. 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% ($\alpha = 0.85$). These values imply annual aggregate energy savings of between 37 and 57 million tce. Because the parameter α governs the extent of energy leakage to related and unrelated 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 indicate 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 the different parameterizations, the implied GWTP that rationalizes the policy varies only between \$156 and \$167.

Our baseline model assumes that conglomerates were not able to respond to the regulation by adjusting their capital. While this is a convenient assumption in the short run, as we discuss in Appendix E.2, it is possible that some conglomerates may have partially responded to the regulation by adjusting their capital. Perhaps more importantly, it is likely that conglomerates would be able to adjust their capital over time. We extend our model to allow for this possibility in Appendix M.3. Because the model is estimated based on pre-reform data, this extension does not require us to reestimate the model. The X-mark in Panel C of Figure A.2 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 as well as assumptions about the timing of capital adjustment do not significantly affect our quantitative assessment of the fundamental trade-off faced by the government.

4. Imperfect Substitution within Conglomerates

The previous section shows that the welfare effects of the regulation 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 $q(\omega) = (\sum_i q(\omega, i)^{\rho_c})^{1/\rho_c}$, 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 P.3, many of

the results of our baseline model extend to this case after slight modifications.⁶² 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 $\left(\frac{\delta^{i-1}}{\Delta_n^C} \right)^{\frac{\rho_c}{1-\alpha\rho_c}}$. Because $\rho_c < 1$, the within-conglomerate distribution of output in Panel A of Figure 2 implies a lower value of δ . Intuitively, because firms related to regulated 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 approximately 11%.

The black square in Panel C of Figure A.2 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 both larger price effects and larger energy savings. As the black square in Panel C of Figure A.2 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; conglomerates do not jointly optimize production decisions across their related firms. The red cross in Panel C of Figure A.2 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.⁶³

K. 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

⁶²Indeed, we obtain our baseline when $\rho_c = 1$. Table A.36 reports the estimated model under this extension.

⁶³Appendix P.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.

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.⁶⁴ 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:

$$\begin{aligned} \frac{d \ln W}{1 - \kappa} &= -\ln\left(\frac{P}{P^*}\right) - \frac{\kappa_0}{1 - \kappa} \ln\left(\frac{E}{E^*}\right) - \sum_n \frac{\lambda_n}{1 - \kappa} \ln\left(\frac{E_n}{E_n^*}\right) \\ \text{(K1)} \quad &= -\ln\left(\frac{P}{P^*}\right) - \frac{\kappa}{1 - \kappa} \ln\left(\frac{E}{E^*}\right) - \sum_n \frac{\lambda_n}{1 - \kappa} \left[\ln\left(\frac{E_n}{E_n^*}\right) - \ln\left(\frac{E}{E^*}\right) \right], \end{aligned}$$

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

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, nor were they 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.

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

One important concern is that the spillovers we identify may have shifted the location of production and of related emissions to more populated areas or areas with higher levels of preexisting industrial emissions. To address this concern, we

⁶⁴For 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 K1.

first use city-level data on sulfur dioxide emissions and population density to generate the following measure of pre-policy 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.2 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 the production of related firms 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 preexisting levels of industrial emissions.

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

During the time period of the Top 1,000 program, the Chinese government was also interested in reallocating economic activity to less developed areas in the Western regions of China. This appendix section 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. In our sample, 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.3 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.

L. 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 produc-

tion, energy use, and potential relationships 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 the initiative through the Top 10,000 program, which was put in place as part of the 12th Five-Year Plan starting in 2012. The Top 10,000 program expanded the scope of regulation by targeting firms that consumed energy in excess of 10,000 tons of coal equivalent in 2010. This requirement meant that about 15,000 additional firms were 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. Both of these constraints arise because the CESD data is not available for this period.

Panel A of Table A.37 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 Top 10,000 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.37 shows that our data cover 12,524 Top 10,000 firms not previously regulated or related and 8,522 firms related to these new firms.

Relative to Top 1,000 firms, Top 10,000 firms have fewer related firms; these related 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.22 plots the size distribution of firms related to Top 1,000 and Top 10,000 firms. This figure shows that the 90th 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.38 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 firms experienced increases in output as a result of the regulation. As in our analyses of within-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.38 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: as Table A.28 shows, we estimate larger spillover effects for larger related firms.⁶⁵

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.38 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 because 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.38 show that unregulated, 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 3).

Because 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. This suggests that the magnitude of within-conglomerate and market-level spillover effects depends on the policy coverage.

M. MODEL DERIVATION DETAILS

This appendix provides detailed model derivations.

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

⁶⁵An important consideration for our analysis in Section II 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 because 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.

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 because

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

Using these results, we can write the profit maximization problem in terms of $K_n = \sum_{i=1}^n k_i$. 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

$$\begin{aligned}
 q^*(\phi, n) &= \Delta_n^{\frac{1}{1-\alpha}} q_1^*(\phi, n) \\
 &= \Delta_n^{\frac{1}{1-\alpha}} \phi \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w+p_e)} \right)^{\alpha_l} k_1^{*\alpha} \\
 \text{(M1)} \quad &= (\phi \Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{-\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w+p_e} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q}.
 \end{aligned}$$

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^* .

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

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

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

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

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, because 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

$$\begin{aligned}
 (M2) \quad q(\phi, n) &= q^*(\phi, n) \xi^{\alpha_l} \frac{\left[1 + \sum_{i>1}^n \delta^{\frac{i-1}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right]}{\sum_i^n \delta^{\frac{i-1}{1-\alpha}}} \\
 &= q^*(\phi, n) \underbrace{\left[\xi^{\alpha_l} \frac{1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right]}_{=\xi_{q,n}},
 \end{aligned}$$

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:

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

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

$$\begin{aligned}
 (M3) \quad & (\xi)^{\alpha_l-1} \left(\frac{P}{P^*} \right)^\rho \left[\xi^{\alpha_l} \frac{1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w+p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right]^{\rho-1} \\
 & = 1 + \frac{\lambda(\phi)}{w + p_e}.
 \end{aligned}$$

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

$$\begin{aligned}
 (M4) \quad & \left[1 + \frac{\lambda_n}{w + p_e} \right]^{\frac{1}{1-\rho}} \left[1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + p_e} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right] \\
 & = \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\rho}} \xi^{\frac{-(1-\alpha_l\rho)}{1-\rho}} \Delta_n^{\frac{1}{1-\alpha}}.
 \end{aligned}$$

This expression does not have a general closed-form solution. Consider, however, the special case where $\frac{1}{1-\rho} = \frac{\alpha_l}{1-\alpha_l}$.⁶⁶ 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 M4. 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 price adjustment in the product market. The first-order conditions for l_i ($1 \leq i \leq n$) are then

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

Because 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

$$(M5) \quad l_1 = l_1^* \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}}$$

⁶⁶While this is a knife-edge case, it holds in the empirically relevant case of $\rho = .75$ and $\alpha_l = 0.8$, so that $\frac{1}{1-\rho} = \frac{\alpha_l}{1-\alpha_l} = 4$.

and both

$$(M6) \quad \begin{aligned} q_1(\phi, n) &= q_1^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho}{1-\alpha_l \rho}} \quad \text{and} \\ q(\phi, n) &= q^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho}{1-\alpha_l \rho}}. \end{aligned}$$

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 M2 for regulated firms and Equation M6 for unregulated firms, the price level under the regulation is then

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

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 $s_{\tilde{\phi}} = \frac{\Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha \rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \Delta_n^{\frac{\rho}{1-\alpha \rho}} \pi_n}{\sum_n \Delta_n^{\frac{\rho}{1-\alpha \rho}} \pi_n}$ be the output share of the regulated

conglomerates prior to the regulation. Additionally, we introduce the notation:

$$\begin{aligned} \mathbb{E}_e \left[x_n \mid \phi > \tilde{\phi} \right] &= x_{\tilde{n}} \left(\frac{\Delta_n^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2}{\Delta_n^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n} \right) \\ &+ \sum_{n=\tilde{n}+1} x_n \left(\frac{\Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n}{\Delta_n^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n} \right) \end{aligned}$$

to denote the average of a size-dependent variable x_n conditional on being part of the regulation 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

$$(M7) \quad \left(\frac{P}{P^*} \right)^{-\rho} = (1 - s_{\tilde{\phi}}) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho^2}{1-\alpha_l \rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{q,n}^\rho \mid \phi > \tilde{\phi} \right].$$

Equations M4 and M7 jointly determine the shadow costs of the regulation $\{\lambda_n\}_{n \geq \tilde{n}}$ and the increase in the price level $\frac{P}{P^*}$.

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

$$\begin{aligned} e^*(\phi, n) &= \left(\frac{\alpha_l}{\alpha_k} \frac{r}{(w + p_e)} \right) K_n^* \\ (M8) \quad &= (\phi \Delta_n)^{\frac{\rho}{1-\alpha\rho}} R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \underbrace{\rho^{\frac{1}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + p_e} \right)^{1-\alpha_k \rho} \left(\frac{\alpha_k}{r} \right)^{\alpha_k \rho} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_E}. \end{aligned}$$

The total energy use prior to the regulation is then

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

REGULATED CONGLOMERATES

We now characterize the change in energy use for regulated conglomerates. Because $e_i = l_i$, 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

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

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

UNREGULATED CONGLOMERATES

Proposition 1 and Equation M5 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

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

This implies that

$$\begin{aligned}
\frac{E}{E^*} &= \frac{\left(\sum_{n=1}^{\tilde{n}-1} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n + \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1 \right) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}} + \xi_{e,\tilde{n}} \Delta_{\tilde{n}}^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \xi_{e,n} \Delta_n^{\frac{\rho}{1-\alpha\rho}} \pi_n}{\sum_n (\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \pi_n} \\
(M9) \quad &= (1 - s_{\tilde{\phi}}) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{e,n} \mid \phi > \tilde{\phi} \right].
\end{aligned}$$

Equation M9—along with the equilibrium price increase and shadow costs determined by Equations M4 and M7—allows us to compute the effect of the regulation on welfare.

SOLVING THE NEW EQUILIBRIUM WITH THE REDUCED-FORM ESTIMATES

In Section VI, we present the full solution to the model using the derivations above. We now show 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 substitute these numbers into 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.

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

This appendix section extends our baseline model to allow conglomerates to adjust both capital and labor allocations.

REGULATED CONGLOMERATES

The profit maximization problem is

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

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

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

The first-order conditions for capital 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

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

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 regulated 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 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

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

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

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

$$\begin{aligned} &\underbrace{R^{1-\rho} P^{*\rho} \rho (q^*(\phi, n))^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l-1}}_{\text{FOC Unregulated Case}} \\ &\times (\xi)^{\alpha-1} \left(1 + \frac{\lambda(\phi)}{w + p_e} \right)^{\alpha_k} \left(\frac{P}{P^*} \right)^{\rho} [\xi_{q,n}(\phi)]^{\rho-1} = w + p_e + \lambda(\phi). \end{aligned}$$

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

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

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

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

N. ALTERNATIVE REGULATIONS

1. 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 M9 imply that

$$(N1) \quad \frac{E^C}{E^*} = (1 - s_{\tilde{\phi}}) \left(\frac{P^C}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[\xi_{e,n} \mid \phi > \tilde{\phi} \right],$$

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

Because 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 M7 imply that

$$(N2) \quad \left(\frac{P^C}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) \left(\frac{P^C}{P^*}\right)^{\frac{\alpha_l \rho^2}{1 - \alpha_l \rho}} + s_{\tilde{\phi}} \mathbb{E}_e \left[(\xi_{e,n})^{\alpha_l \rho} | \phi > \tilde{\phi} \right].$$

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

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

2. 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

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

Because 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 M9 imply that

$$(N3) \quad \frac{E^\tau}{E^*} = (1 - s_{\tilde{\phi}}) \left(\frac{P^\tau}{P^*}\right)^{\frac{\rho}{1-\alpha_l\rho}} + s_{\tilde{\phi}} \xi_\tau.$$

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

$$(N4) \quad \left(\frac{P^\tau}{P^*}\right)^{-\rho} = (1 - s_{\tilde{\phi}}) \left(\frac{P^\tau}{P^*}\right)^{\frac{\alpha_l\rho^2}{1-\alpha_l\rho}} + s_{\tilde{\phi}} (\xi_\tau)^{\alpha_l\rho}.$$

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:

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

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

O. INSPECTING THE EFFECT MECHANISMS OF THE TOP 1,000 REGULATION

1. 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

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

The new price is given by

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

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

2. 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

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

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

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

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

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

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

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

3. Shutting down Both Market and Conglomerate Leakage

In this case, the formulas for 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

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

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.

P. MODEL EXTENSIONS

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

$$(P1) \quad \underbrace{R^{1-\rho}(P^*)^\rho \rho \left[\phi^* \sum_{i=1}^n \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} (l_i^*)^{\alpha_l} \right]^{\rho-1} \phi^* \alpha_l (l_1^*)^{\alpha_l-1} \left(\frac{P}{P^*} \right)^\rho}_{\text{FOC Unregulated}}$$

$$\left[\underbrace{\left(\frac{\xi \nu_1}{\nu^*} \right)^{\alpha_l} \frac{\left[1 + \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^\gamma} \right]^{\frac{\alpha_l}{1-\alpha_l}} \sum_{j>1} \delta^{\frac{j-1}{1-\alpha}} \right]}{\sum_i \delta^{\frac{j-1}{1-\alpha}}}}_{\xi_{q,n}^\nu} \right]^{\rho-1} \times \left(\frac{\xi \nu_1}{\nu^*} \right)^{\alpha_l-1} = (w + (\nu^*)^\gamma) \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^\gamma} \right].$$

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

$$(P2) \quad \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^\gamma} \right]^{\frac{1}{1-\rho}} \left[1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{w + (\nu^*)^\gamma} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right]$$

$$= \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\rho}} \left(\frac{\xi \nu_1}{\nu^*} \right)^{\frac{-(1-\alpha_l)\rho}{1-\rho}} \Delta_n^{\frac{1}{1-\alpha}}.$$

This equation has one difference from Equation M4. Because 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 . In this equation, as with Equation M4, 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 improve energy efficiency by the same amount 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 $e(\phi, n) = \frac{l(\phi, n)}{\nu^*}$ because these firms do not change their investment in energy efficiency. However, for

regulated conglomerates, we have

$$\begin{aligned}
e(\phi, n) &= \frac{l_1}{\nu_{1,n}} + \frac{1}{\nu^*} \sum_{i>1} l_i = \frac{l_1}{\nu_{1,n}} \left(1 + \frac{\nu_{1,n}}{\nu^*} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + (\nu^*)^\gamma} \right]^{\frac{1}{1-\alpha_l}} \right) \\
&= e_1(\phi, n) \left(1 + \frac{\nu_{1,n}}{\nu^*} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + (\nu^*)^\gamma} \right]^{\frac{1}{1-\alpha_l}} \right) \\
&= e_1^*(\phi, n) \xi \left(1 + \frac{\nu_{1,n}}{\nu^*} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + (\nu^*)^\gamma} \right]^{\frac{1}{1-\alpha_l}} \right) \\
&= e^*(\phi, n) \underbrace{\left[\xi \frac{1 + \frac{\nu_{1,n}}{\nu^*} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda_n}{w + (\nu^*)^\gamma} \right]^{\frac{1}{1-\alpha_l}}}{\Delta_n^{\frac{1}{1-\alpha}}} \right]}_{\equiv \xi_{e,n}^\nu}.
\end{aligned}$$

The term $\xi_{e,n}^\nu$ incorporates the insight that the shifting of production to related firms now leads to a larger increase in total energy use because these firms do not improve their energy efficiency, i.e., $\frac{\nu_{1,n}}{\nu^*} > 1$. Using this term in Equation M9 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,

$$\frac{R^{1-\rho} P^\rho q(\phi, n)^{\rho-1} q_1(\phi, n)}{e_1(\phi, n)} = \frac{R^{1-\rho} P^{*\rho} (q(\phi, n)^*)^{\rho-1} q_1^*(\phi, n)}{e_1^*(\phi, n)} \times \frac{\left(\frac{P}{P^*}\right)^\rho (\xi_{q,n}^\nu)^{\rho-1} \left(\xi \frac{\nu_{1,n}}{\nu^*}\right)^{\alpha_l}}{\xi}.$$

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 P1 as follows:

$$\frac{\left(\frac{P}{P^*}\right)^\rho (\xi_{q,n}^\nu)^{\rho-1} \left(\xi \frac{\nu_{1,n}}{\nu^*}\right)^{\alpha_l}}{\xi} = \left(\frac{\nu_{1,n}}{\nu^*}\right) \left[1 + s_e \left(\left(\frac{\nu_{1,n}}{\nu^*}\right)^\gamma - 1 \right) \right].$$

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 $\ln \left[1 + \frac{\lambda_n}{w+(\nu^*)^\gamma} \right]$.

Because 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

$$\frac{e_1(\phi, n)}{e_1^*(\phi, n)} = \frac{e(\phi, n)}{e^*(\phi, n)} = \left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l \rho}} \quad \text{and} \quad \frac{q_1(\phi, n)}{q_1^*(\phi, n)} = \frac{q(\phi, n)}{q^*(\phi, n)} = \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho^2}{1-\alpha_l \rho}}.$$

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

$$(P3) \quad \beta^{EE} = \mathbb{E} \left[\ln \left(\frac{\nu_{1,n}}{\nu^*} \right) \middle| \phi > \tilde{\phi} \right] + \mathbb{E} \left[\ln \left[1 + s_e \left(\left(\frac{\nu_{1,n}}{\nu^*} \right)^\gamma - 1 \right) \right] \middle| \phi > \tilde{\phi} \right].$$

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 P2 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 P3.

We then choose the value of γ that is consistent with our empirical estimates of β^{EE} . Because 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} .

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

$$\begin{aligned} q^*(\phi, n) &= (\phi \Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_O}} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}} \\ &= (\phi \Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \rho^{\frac{\alpha}{1-\alpha\rho}} \underbrace{\left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_1}} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q} (d_O)^{\frac{-\alpha_l}{1-\alpha\rho}}, \end{aligned}$$

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, because $l_1 = \nu_O e_1$ 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

$$(P4) \quad q_1(\phi, n) = q_1^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}} \quad \text{and} \quad q(\phi, n) = q^*(\phi, n) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l\rho}{1-\alpha_l\rho}}.$$

REGULATED CONGLOMERATES

Because $\nu_1 \neq \nu_R$, 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

$$(P5) \quad q^*(\phi, n) = \left(\phi \hat{\Delta}_n \right)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} \underbrace{\rho^{\frac{\alpha}{1-\alpha\rho}} \left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_1}} \right)^{\alpha_l} \left(\frac{\alpha_k}{r} \right)^{\alpha_k} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_Q}.$$

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 because 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

$$\max_{\{l_i\}_1^n} \left\{ R^{1-\rho} P^\rho \left[\phi^* \sum_{i=1}^n \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} d_i^{\frac{-\alpha_l \alpha_k}{1-\alpha}} \right]^\rho - \left(w + \frac{p_e}{\nu_1} \right) \sum_{i=1}^n d_i l_i - r \sum_{i=1}^n k_i^* \right\} \text{ subject to } l_1 \leq \xi l_1^*,$$

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

$$\begin{aligned} & \underbrace{R^{1-\rho} P^\rho}_{\text{Market Demand}} \underbrace{\rho \left[\phi^* \sum_{i=1}^n \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} l_i^{\alpha_l} d_i^{\frac{-\alpha_l \alpha_k}{1-\alpha}} \right]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi^* \delta^{\frac{(i-1)(1-\alpha_l)}{1-\alpha}} \alpha_l (l_i)^{\alpha_l-1} d_i^{\frac{-\alpha_l \alpha_k}{1-\alpha}}}_{\text{Marginal Product}} \\ &= \left(w + \frac{p_e}{\nu_1} \right) d_i + \underbrace{\lambda(\phi) \mathbb{I}[i=1]}_{\substack{\text{Shadow Cost} \\ \text{of Regulation}}} . \end{aligned}$$

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

$$\begin{aligned} & \phi(k_1^*)^{\alpha_k} l_1^{\alpha_l} \left[1 + \sum_{i>1} \delta^{\frac{i-1}{1-\alpha}} d_i^{\frac{-\alpha_l}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1} \right)} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right] \\ &= \phi(k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l} \xi^{\alpha_l} \left[1 + (\hat{\Delta}_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1} \right)} \right]^{\frac{\alpha_l}{1-\alpha_l}} \right] . \end{aligned}$$

The first-order condition for the regulated firm is then

$$\begin{aligned}
& \underbrace{R^{1-\rho} P^{*\rho} \rho \left[\phi \sum_{i=1}^n \delta^{\frac{i-1}{1-\alpha}} (k_i^*)^{\alpha_k} (l_i^*)^{\alpha_l} \right]^{\rho-1} \times \phi \alpha_l (k_1^*)^{\alpha_k} (l_1^*)^{\alpha_l-1}}_{\text{FOC Unregulated Case}} \\
& \times (\xi)^{\alpha_l-1} \left(\frac{P}{P^*} \right)^\rho \left[\underbrace{\xi^{\alpha_l} \frac{1 + (\hat{\Delta}_n^{\frac{1}{1-\alpha}} - 1) \left[1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1} \right)} \right]^{\frac{\alpha_l}{1-\alpha_l}}}}_{\hat{\Delta}_n^{\frac{1}{1-\alpha}} \equiv \xi_{q,n}^d} \right]^{\rho-1} = \left(w + \frac{p_e}{\nu_1} \right) + \lambda(\phi),
\end{aligned}$$

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

$$(\text{P6}) \quad 1 + \frac{\lambda(\phi)}{\left(w + \frac{p_e}{\nu_1} \right)} = (\xi)^{\alpha_l-1} \left(\frac{P}{P^*} \right)^\rho (\xi_{q,n}^d)^{(\rho-1)}.$$

PRODUCT MARKET EQUILIBRIUM

The price level absent the regulation is then

$$\begin{aligned}
P^{*- \rho} &= R^{-\rho} \int_{\phi_1}^{\tilde{\phi}} \left((\phi \Delta_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} C_Q (d_O)^{\frac{-\alpha_l}{1-\alpha\rho}} \right)^\rho \frac{g(\phi)M}{1-G(\phi_1)} d\phi \\
&+ R^{-\rho} \int_{\tilde{\phi}_1}^{\tilde{\phi}} \left((\phi \hat{\Delta}_n)^{\frac{1}{1-\alpha\rho}} R^{\frac{(1-\rho)\alpha}{1-\alpha\rho}} P^{*\frac{\rho\alpha}{1-\alpha\rho}} C_Q \right)^\rho \frac{g(\phi)M}{1-G(\phi_1)} d\phi \\
P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1-G(\phi_1)}{M} &= \int_{\phi_1}^{\tilde{\phi}} (\phi \Delta_n d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi + \int_{\tilde{\phi}}^{\tilde{\phi}} (\phi \hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}} g(\phi) d\phi \\
P^{*\frac{-\rho}{1-\alpha\rho}} R^{\frac{(1-\alpha)\rho}{1-\alpha\rho}} C_Q^{-\rho} \frac{1-G(\phi_1)}{M} &= \sum_{n=1}^{\tilde{n}-1} (\Delta_n d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}} \pi_n + (\Delta_{\tilde{n}} d_O^{-\alpha_l})^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1 \\
&+ (\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} (\hat{\Delta}_n)^{\frac{\rho}{1-\alpha\rho}} \pi_n.
\end{aligned}$$

After the regulation, the equilibrium is then

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

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

Comparing the regulated and unregulated equilibrium conditions, we then have

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

$$\text{where } s_{\tilde{\phi}}^d = \frac{(\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha \rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} (\hat{\Delta}_n)^{\frac{\rho}{1-\alpha \rho}} \pi_n}{(d_O)^{\frac{-\alpha_l \rho}{1-\alpha \rho}} \left[\sum_{n=1}^{\tilde{n}-1} (\Delta_n)^{\frac{\rho}{1-\alpha \rho}} \pi_n + (\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha \rho}} \tilde{\pi}_1 \right] + (\hat{\Delta}_{\tilde{n}})^{\frac{\rho}{1-\alpha \rho}} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} (\hat{\Delta}_n)^{\frac{\rho}{1-\alpha \rho}} \pi_n} \text{ and}$$

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

$$\frac{l(\phi, n)}{l^*(\phi, n)} = \xi \frac{\left[1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \left[1 + \frac{\lambda(\phi)}{(w + \frac{p_e}{\nu_1})} \right]^{\frac{1}{1-\alpha_l}} \right]}{1 + (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}}}.$$

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 regu-

lation, 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

$$\begin{aligned} l^*(\phi, n) &= l_1^* \left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \right) \\ l^*(\phi, n) &= \left(\phi \hat{\Delta}_n \right)^{\frac{\rho}{1-\alpha\rho}} R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \rho^{\frac{1}{1-\alpha\rho}} \underbrace{\left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_1}} \right)^{1-\alpha_k\rho} \left(\frac{\alpha_k}{r} \right)^{\alpha_k\rho} \right]^{\frac{1}{1-\alpha\rho}}}_{=C_E} \end{aligned}$$

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. Because 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

$$\begin{aligned} e^*(\phi, n) &= \frac{l^*(\phi, n)}{\nu_O} \\ &= \frac{1}{\nu_O} (\phi \Delta_n)^{\frac{\rho}{1-\alpha \rho}} R^{\frac{1-\rho}{1-\alpha \rho}} P^{*\frac{\rho}{1-\alpha \rho}} \rho^{\frac{1}{1-\alpha \rho}} \underbrace{\left[\left(\frac{\alpha_l}{w + \frac{p_e}{\nu_1}} \right)^{1-\alpha_k \rho} \left(\frac{\alpha_k}{r} \right)^{\alpha_k \rho} \right]^{\frac{1}{1-\alpha \rho}}}_{=C_E} (d_O)^{\frac{-(1-\alpha_k \rho)}{1-\alpha \rho}}. \end{aligned}$$

The total energy use prior to the regulation is then

$$\begin{aligned} E^* &= \int_{\phi_1} e^*(\phi, n) \frac{g(\phi)M}{1-G(\phi_1)} d\phi \\ &= R^{\frac{1-\rho}{1-\alpha \rho}} P^{*\frac{\rho}{1-\alpha \rho}} \frac{C_E M}{1-G(\phi_1)} \left[\frac{(d_O)^{\frac{-(1-\alpha_k \rho)}{1-\alpha \rho}}}{\nu_O} \left(\sum_{n=1}^{\tilde{n}-1} (\Delta_n)^{\frac{\rho}{1-\alpha \rho}} \pi_n + (\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha \rho}} \tilde{\pi}_1 \right) \right. \\ &\quad + \left(\hat{\Delta}_{\tilde{n}} \right)^{\frac{\rho}{1-\alpha \rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) \right)} \tilde{\pi}_2 \\ &\quad \left. + \sum_{n=\tilde{n}+1} \left(\hat{\Delta}_n \right)^{\frac{\rho}{1-\alpha \rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \right)} \pi_n \right]. \end{aligned}$$

The total energy use after the regulation is then

$$\begin{aligned}
E &= \int_{\phi_1} e(\phi, n) \frac{g(\phi)M}{1-G(\phi_1)} d\phi \\
&= R^{\frac{1-\rho}{1-\alpha\rho}} P^{*\frac{\rho}{1-\alpha\rho}} \frac{C_E M}{1-G(\phi_1)} \left[\left(\frac{P}{P^*} \right)^{\frac{\rho}{1-\alpha_l\rho}} \frac{(d_O)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}}{\nu_O} \left(\sum_{n=1}^{\tilde{n}-1} (\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \pi_n + (\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1 \right) \right. \\
&\quad + \xi_{e,\tilde{n}}^d \left(\hat{\Delta}_{\tilde{n}} \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) \right)} \tilde{\pi}_2 \\
&\quad \left. + \sum_{n=\tilde{n}+1} \xi_{e,n}^d \left(\hat{\Delta}_n \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \right)} \pi_n \right].
\end{aligned}$$

The change in aggregate energy use is then

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

where

$$\frac{1 - s_{\tilde{\phi}}^e}{s_{\tilde{\phi}}^e} = \frac{\frac{(d_O)^{\frac{-(1-\alpha_k\rho)}{1-\alpha\rho}}}{\nu_O} \left(\sum_{n=1}^{\tilde{n}-1} (\Delta_n)^{\frac{\rho}{1-\alpha\rho}} \pi_n + (\Delta_{\tilde{n}})^{\frac{\rho}{1-\alpha\rho}} \tilde{\pi}_1 \right)}{\left(\hat{\Delta}_{\tilde{n}} \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) \right)} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \left(\hat{\Delta}_n \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \right)} \pi_n}$$

and where we evaluate $\mathbb{E}_e$ using the conditional probabilities

$$\begin{aligned}
&\left(\hat{\Delta}_{n'} \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{n'}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{n'}^{\frac{1}{1-\alpha}} - 1) \right)} \pi_{n'} \\
\Pr \left[n = n' \mid \phi > \tilde{\phi} \right] &= \frac{\left(\hat{\Delta}_{n'} \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{n'}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{n'}^{\frac{1}{1-\alpha}} - 1) \right)} \pi_{n'}}{\left(\hat{\Delta}_{\tilde{n}} \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_{\tilde{n}}^{\frac{1}{1-\alpha}} - 1) \right)} \tilde{\pi}_2 + \sum_{n=\tilde{n}+1} \left(\hat{\Delta}_n \right)^{\frac{\rho}{1-\alpha\rho}} \frac{\frac{1}{\nu_1} \left(1 + \frac{\nu_1}{\nu_R} (\Delta_n^{\frac{1}{1-\alpha}} - 1) d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} \right)}{\left(1 + d_R^{\frac{-(1-\alpha_k)}{1-\alpha}} (\Delta_n^{\frac{1}{1-\alpha}} - 1) \right)} \pi_n}.
\end{aligned}$$

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

$$q(\omega) = \left(\sum_i q(\omega, i)^{\rho_c} \right)^{1/\rho_c}, \text{ where } 0 < \rho < \rho_c < 1.$$

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

$$\begin{aligned} \max_{\{l_i\}_{i=1}^n, \{k_i\}_{i=1}^n} & R^{1-\rho} P^\rho q(\phi)^{\rho-\rho_c} \sum_i q(\phi, i)^{\rho_c} - (w + p_e) \sum_i l_i - r \sum_i k_i \\ = & \max_{\{l_i\}_{i=1}^n, \{k_i\}_{i=1}^n} R^{1-\rho} P^\rho \left[\sum_i q(\phi, i)^{\rho_c} \right]^{\rho/\rho_c} - (w + p_e) \sum_i l_i - r \sum_i k_i, \end{aligned}$$

where $q(\phi, i) = \phi \delta^{i-1} k_i^{\alpha_k} l_i^{\alpha_l}$. The first-order conditions for $l_1, \dots, l_n$ imply that $\frac{l(\phi, i)}{l(\phi, 1)} = \left(\frac{q(\phi, i)}{q(\phi, 1)} \right)^{\rho_c}$. Because 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

$$\begin{aligned} R^{1-\rho} P^\rho \rho \left[\phi^* \left(\sum_i \delta^{\frac{(i-1)\rho_c(1-\alpha_l \rho_c)}{1-\alpha \rho_c}} l_i^{\alpha_l \rho_c} \right)^{1/\rho_c} \right]^{\rho-1} &\times \phi^* \delta^{\frac{(i-1)\rho_c(1-\alpha_l \rho_c)}{1-\alpha \rho_c}} \alpha_l \rho_c (l_i)^{\alpha_l \rho_c - 1} \\ &= w + p_e + \lambda(\phi) I[i = 1]. \end{aligned}$$

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

As in other model extensions, we 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 M7, 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.

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 base-

line 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

$$\bar{\pi} = \exp \left\{ \frac{\tilde{\sigma}^2}{2} \right\} [1 - \Phi(\tilde{\phi} - \tilde{\sigma})],$$

where $\tilde{\sigma} = \frac{\rho}{1-\alpha\rho} \sigma_\phi$ and $\tilde{\phi} = \frac{\rho}{1-\alpha\rho} \frac{\ln(\bar{\phi})}{\tilde{\sigma}}$.

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

$$\max_l \{ R^{1-\rho} P^\rho [\phi^* l_i^{\alpha_l}]^\rho - (w + p_e)l - rk^* \} \quad \text{subject to } l \leq \xi l^*,$$

where $\phi^* = \phi(k^*)^{\alpha_k}$. The first-order conditions for l is then

$$\frac{\partial \pi}{\partial l} = \underbrace{R^{1-\rho} P^\rho}_{\text{Market Demand}} \underbrace{\rho [\phi^* l^{\alpha_l}]^{\rho-1}}_{\text{Residual Revenue}} \underbrace{\phi^* \alpha_l (l)^{\alpha_l-1}}_{\text{Marginal Product}} = w + p_e + \underbrace{\lambda(\phi)}_{\text{Shadow Cost of Regulation}}.$$

As 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

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

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

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

Separate the regulated and unregulated regime into $\tilde{\pi}_1$ and $\tilde{\pi}_2$ as follows:

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

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

$$P^{-\rho} P^{*\frac{-\rho^2 \alpha}{1 - \alpha \rho}} R^{\frac{(1 - \alpha)\rho}{1 - \alpha \rho}} C_Q^{-\rho} \frac{1 - G(\bar{\phi})}{M} = \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho^2}{1 - \alpha_l \rho}} \tilde{\pi}_1 + \xi^{\alpha_l \rho} \tilde{\pi}_2.$$

Note that this equation holds in the case without the regulation if we set $\xi = 1$ and $P = P^*$. Let $s_{\bar{\phi}} = \frac{\tilde{\pi}_2}{\tilde{\pi}_1 + \tilde{\pi}_2}$ be the output share of the Top 1,000 conglomerates prior to the regulation. We obtain

$$\left(\frac{P}{P^*} \right)^{-\rho} = (1 - s_{\bar{\phi}}) \left(\frac{P}{P^*} \right)^{\frac{\alpha_l \rho^2}{1 - \alpha_l \rho}} + s_{\bar{\phi}} \xi^{\alpha_l \rho}.$$

We can similarly derive the change in aggregate energy use as

$$\frac{E}{E^*} = (1 - s_{\bar{\phi}}) \left(\frac{P}{P^*} \right)^{\frac{\rho}{1 - \alpha_l \rho}} + s_{\bar{\phi}} \xi.$$

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

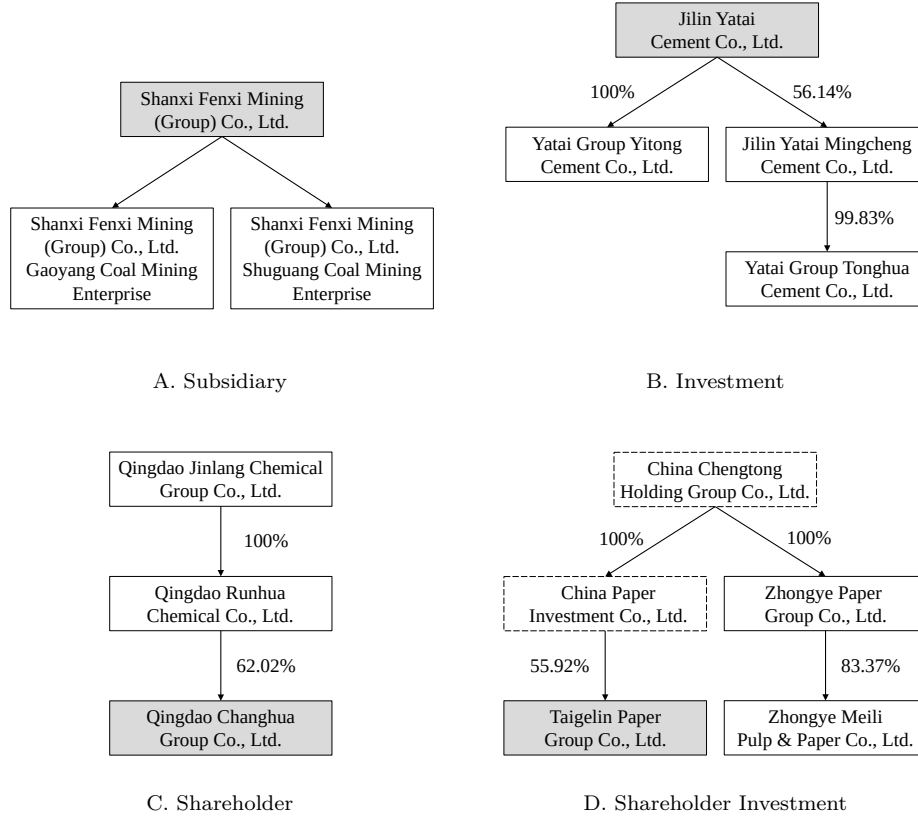
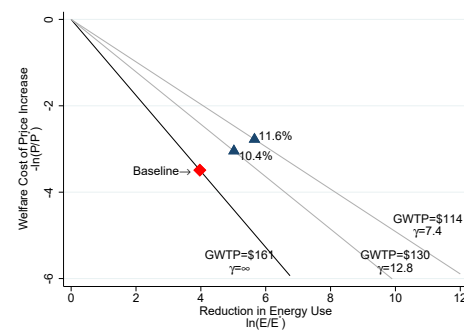


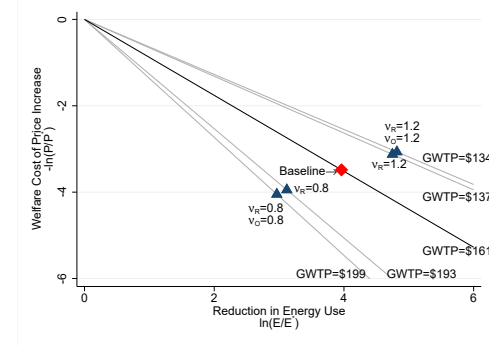
FIGURE A.1. EXAMPLES OF FIRM RELATIONS

Note: 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 I.C for the definition of related firms.

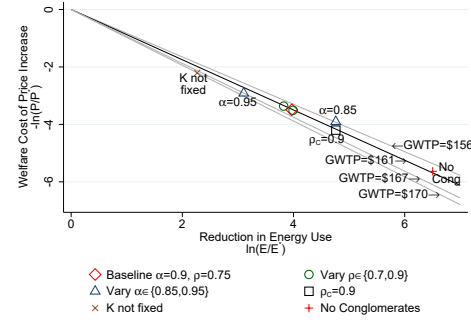
Source: Authors' calculations using data from CARD.



A. Endogenous Energy Efficiency



B. Heterogeneous Energy Efficiency

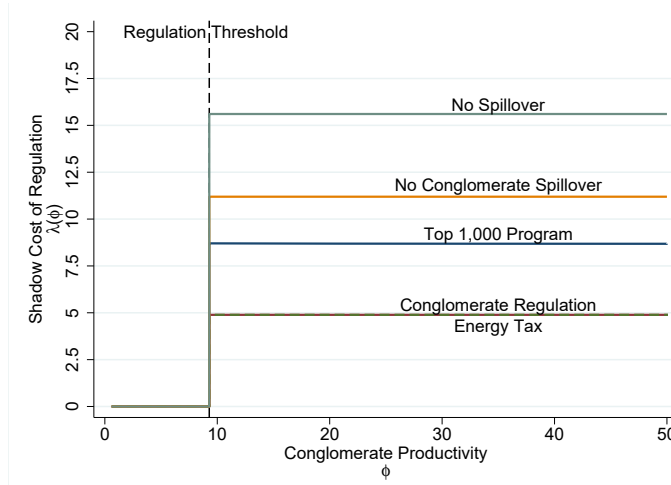


C. Alternative Parameters

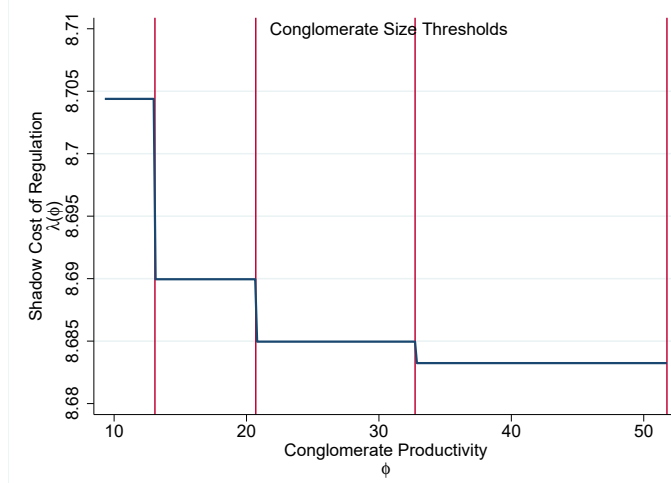
FIGURE A.2. WELFARE EFFECTS: ROBUSTNESS

Note: 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 J.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 J.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 J.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 J.4 for details. Across these wide-ranging assumptions, the GWTP that rationalizes the Top 1,000 program lies between \$114 and \$199.

Source: Author's calculations using the model developed in Section IV.



A. Shadow Costs of Alternative Regulations



B. Size Distortions in Top 1,000 Program

FIGURE A.3. MODEL-BASED ESTIMATES OF THE SHADOW COSTS OF REGULATION

Note: 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 N and O 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 VI.A and VI.B for additional discussion.

Source: Author's calculations using the model developed in Section IV.

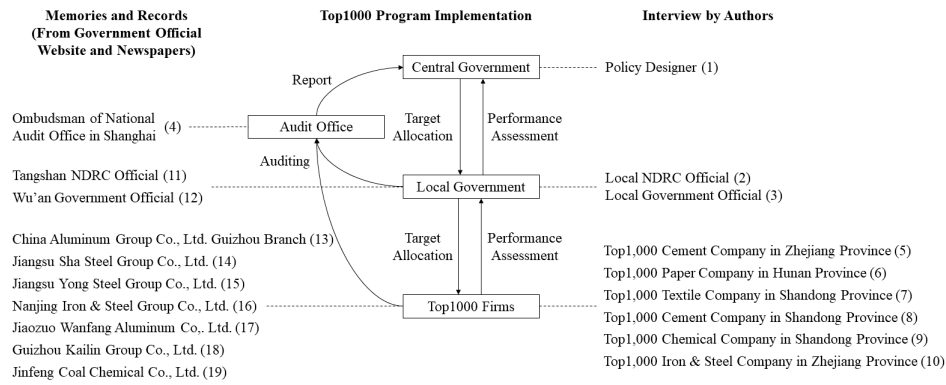


FIGURE A.4. INTERVIEW RECORD

Note: 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 11th Five-Year Plan with two local government officials and executives from seven Top 1,000 firms.

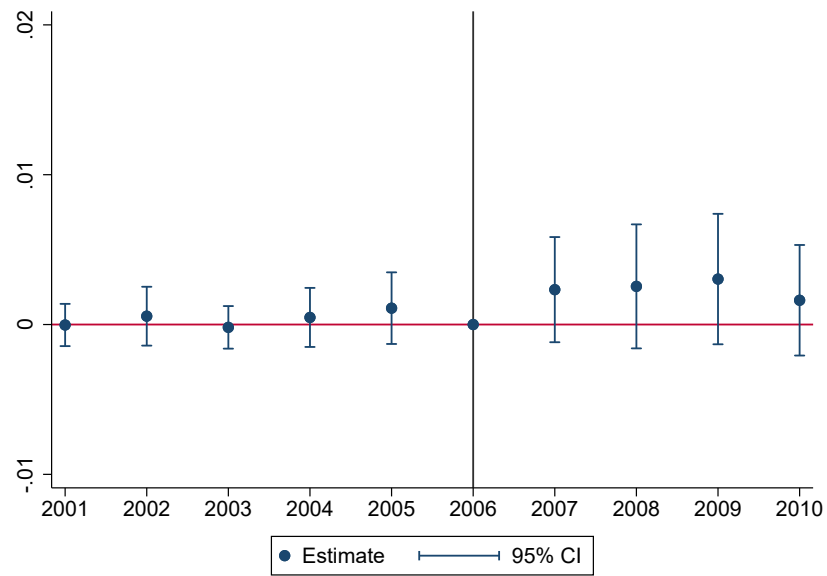
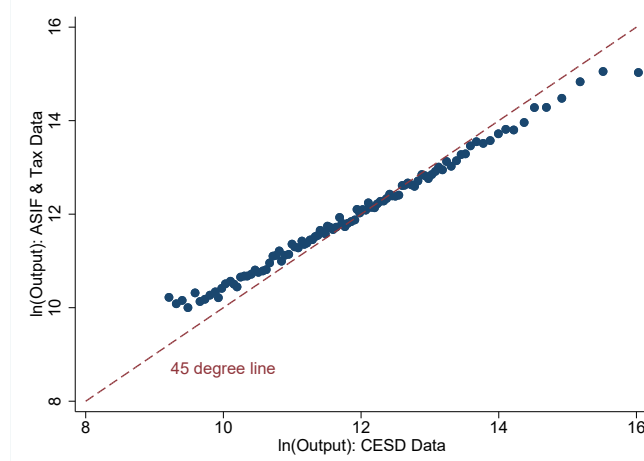


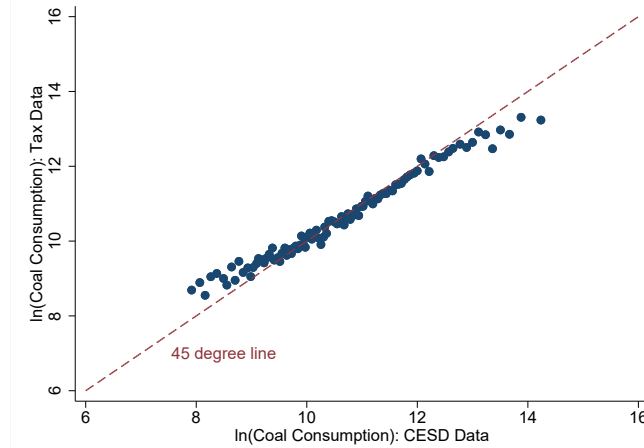
FIGURE A.5. THE EFFECT OF THE TOP 1,000 PROGRAM ON ENERGY-SAVING PATENT APPLICATIONS

Note: This figure shows estimates of Equation 1 where the dependent variable is log firm patent applications in energy saving. The Top 1,000 program appears to have no significant effects on patent applications 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.

Source: Authors' calculations using data from IncoPat.



A. Comparison of Output Data, 2001–2010



B. Comparison of Coal Consumption, 2007–2010

FIGURE A.6. DATA COMPARISON: ASIF, CESD, TAX DATA

Note: 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 ATS 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 ATS data almost mirror those from the CESD (ATS data are available only after 2007).

Source: Authors' calculations using data from the ASIF, CESD and ATS.

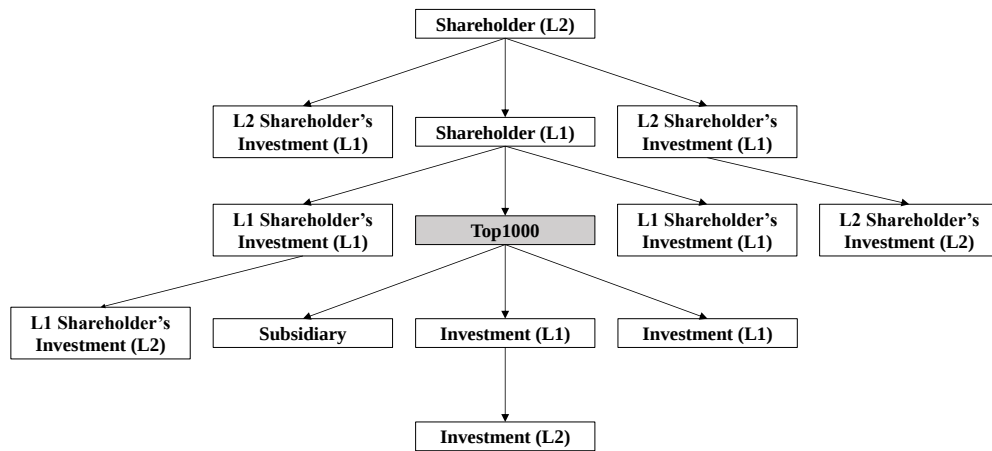
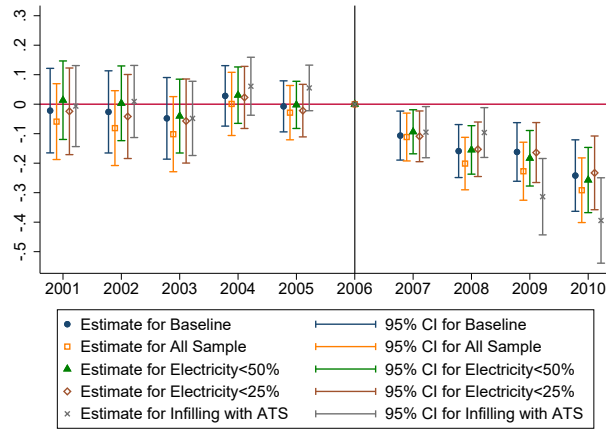
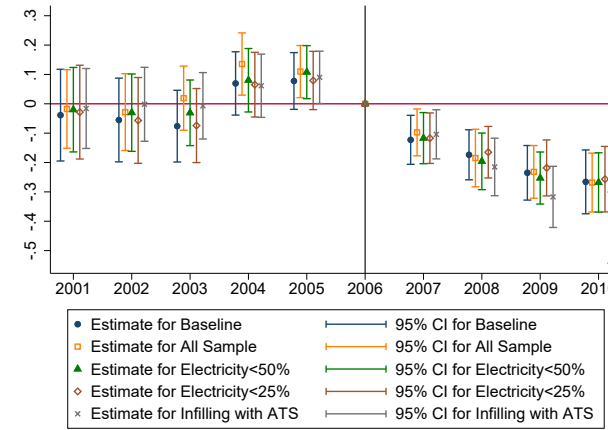


FIGURE A.7. TYPES OF RELATED PARTIES

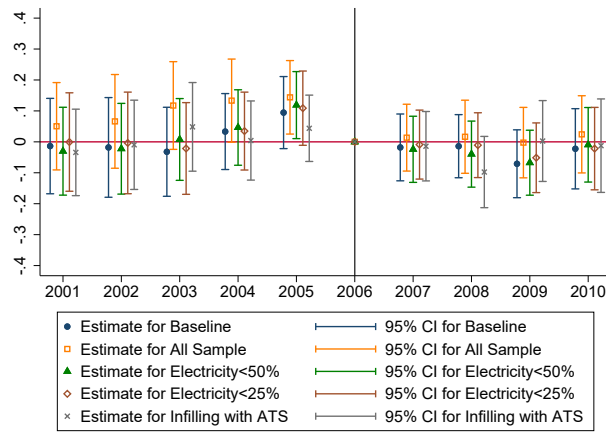
Note: This figure depicts all possible types of firm relations within 2 levels of ownership. See Section I.B for the definition of each related firm type, and see Figure A.1 for examples.



A. Energy Use



B. Output

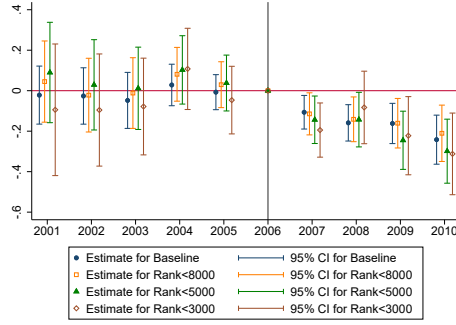


C. Energy Efficiency

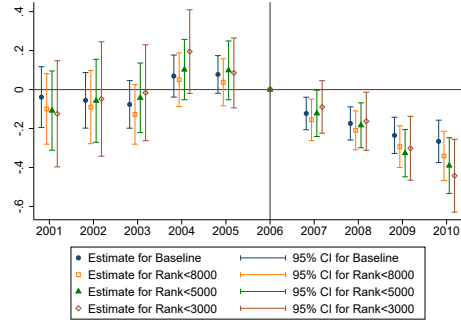
FIGURE A.8. EFFECTS OF THE TOP 1,000 PROGRAM ON REGULATED FIRMS: ROBUSTNESS TO DIFFERENT SAMPLES

Note: 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 3 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.9. Standard errors are clustered at the firm level.

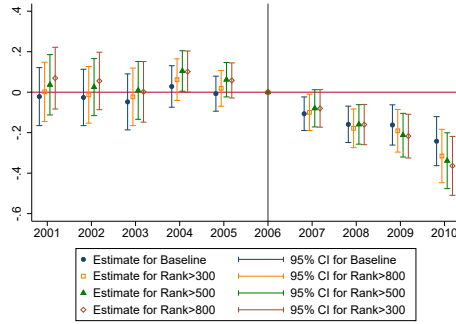
Source: Authors' calculations using data from the CESD and ATS.



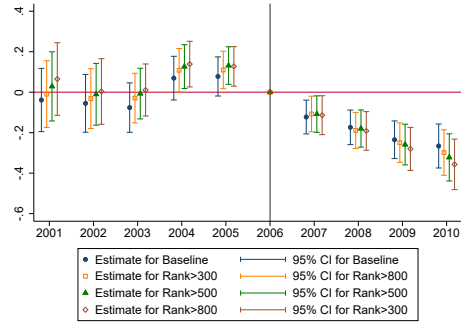
A. Energy: Trimming Controls



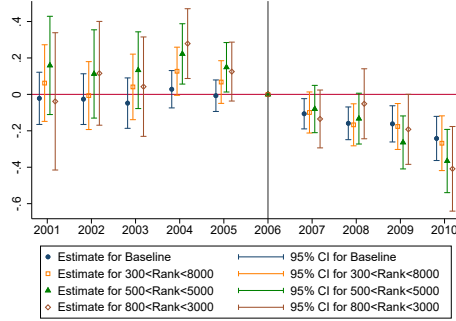
B. Output: Trimming Controls



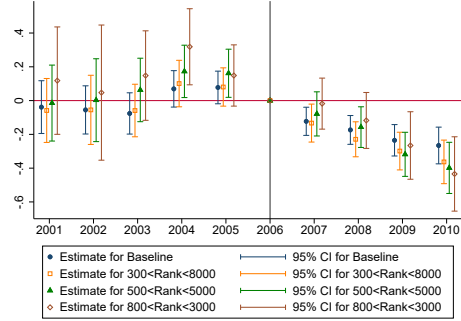
C. Energy: Trimming Treated



D. Output: Trimming Treated



E. Energy: Trimming Both



F. Output: Trimming Both

FIGURE A.9. EFFECTS OF THE TOP 1,000 PROGRAM ON REGULATED FIRMS: ROBUSTNESS TO TRIMMING THE SAMPLE BY TOP 10,000 RANK

Note: This figure corresponds to Panels A and C in Figure 3 and uses firm rank in the Top 10,000 program to trim the sample. Our baseline results are shown in blue; 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.10. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the CESD.

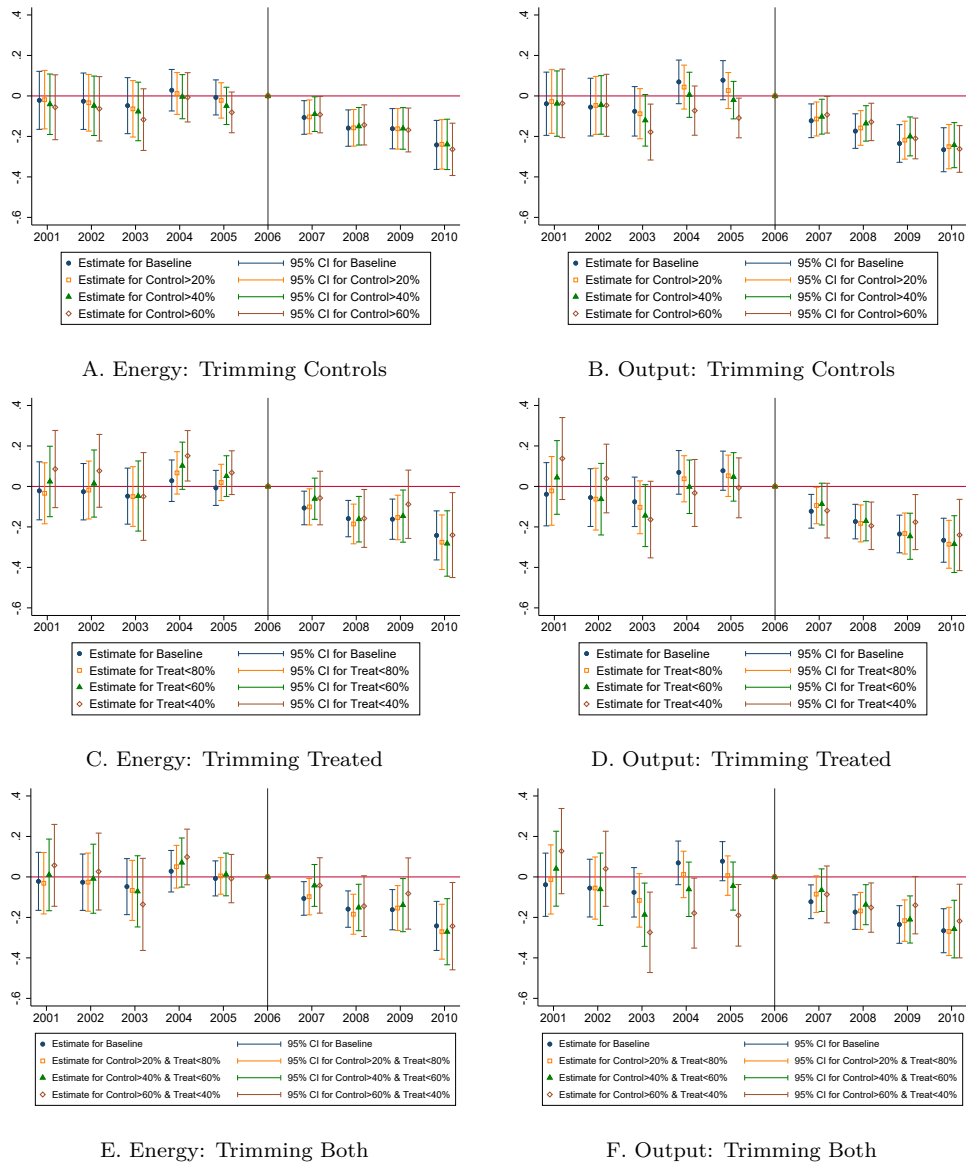


FIGURE A.10. EFFECTS OF THE TOP 1,000 PROGRAM ON REGULATED FIRMS: ROBUSTNESS TO TRIMMING THE SAMPLE USING AN ESTIMATED PROPENSITY OF TREATMENT

Note: This figure corresponds to Panels A and C in Figure 3 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 in blue; 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 20th, 40th, and 60th 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 80th, 60th, and 40th 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.11. Standard errors are clustered at the firm level.

Source: Authors' calculations using data from the 90ESD.

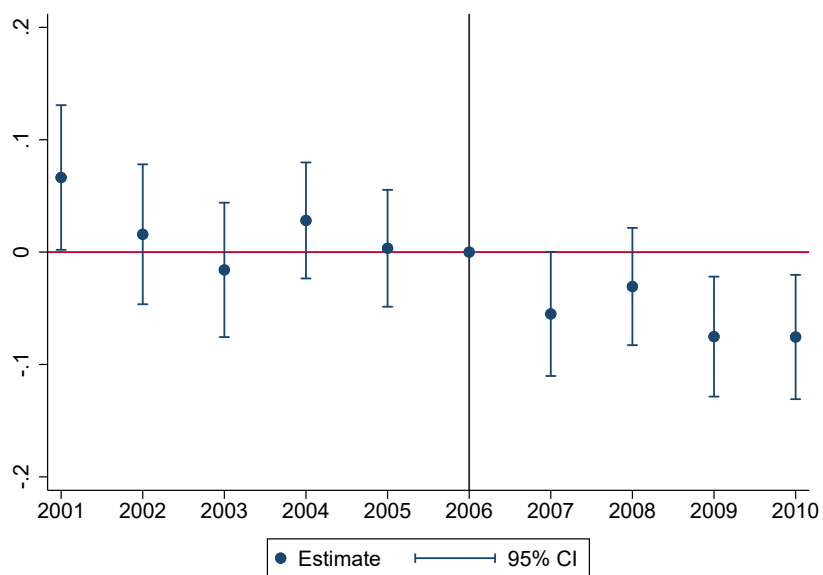
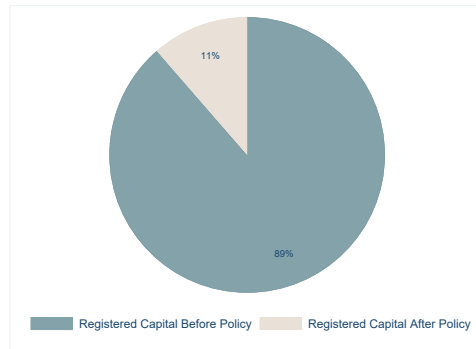
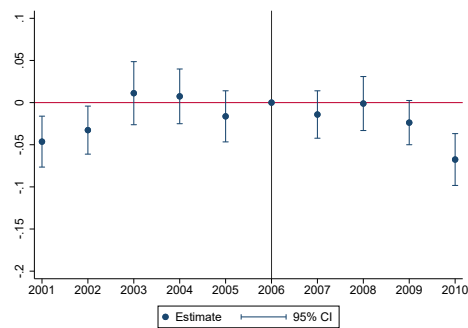


FIGURE A.11. EFFECTS OF THE POLICY ON REGULATED FIRMS' INVESTMENT

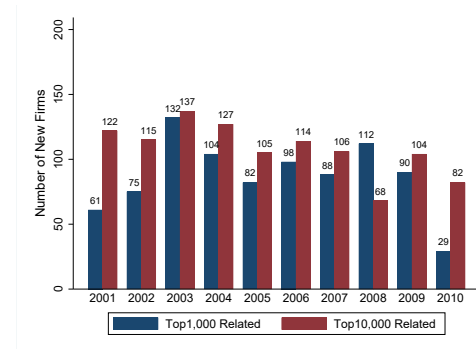
Note: This figure shows estimates of Equation 1 where the dependent variable is an indicator for whether a firm invests. See Section I for more information about the data-generating procedure. This figure shows that regulated firms were less likely to invest than similar control firms (Top 10,000 firms not related to Top 1,000 firms) after the regulation. The point estimates are displayed in Table A.14. See Section II for additional discussion. Standard errors are clustered at the firm level.
Note: Authors' calculations using data from the ASIF and ATS.



A. Relative Size of New Affiliates



C. Entry Rate Event Study



B. New Affiliates by Year

FIGURE A.12. LACK OF ENTRY EFFECT

Note: Panel A shows that firms related to 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.

Source: Authors' calculations using data from the CARD.

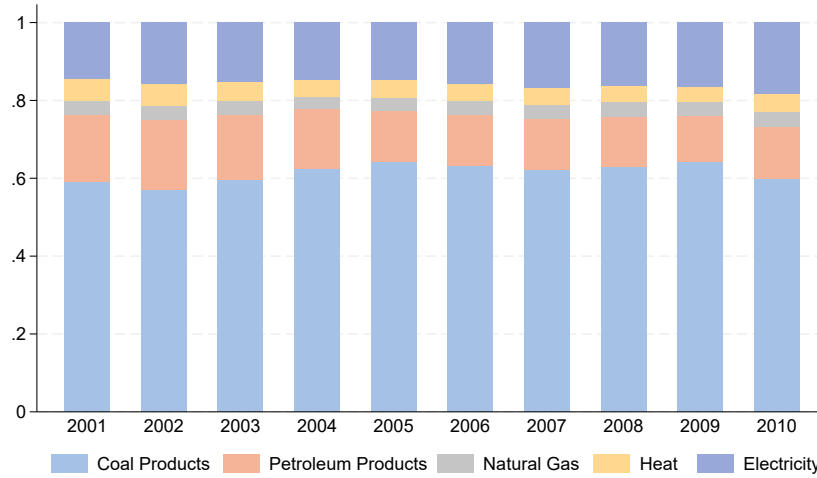


FIGURE A.13. ENERGY MIX OF NINE ENERGY-INTENSIVE INDUSTRIES

Note: 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. Because 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.

Source: Author's calculations from the China Energy Statistical Yearbook

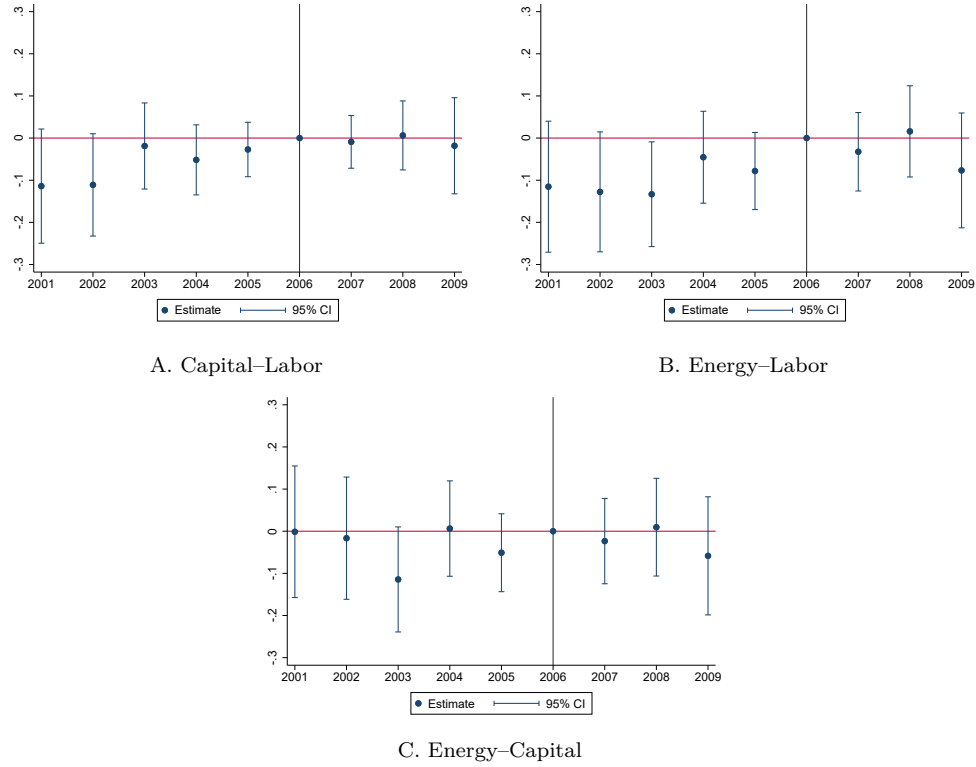


FIGURE A.14. SUBSTITUTION EFFECTS OF THE PROGRAM ON REGULATED FIRMS

Note: 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.21. Standard errors are clustered at the firm level.

Source: Authors' calculations using data from the CESD, ASIF and ATS.

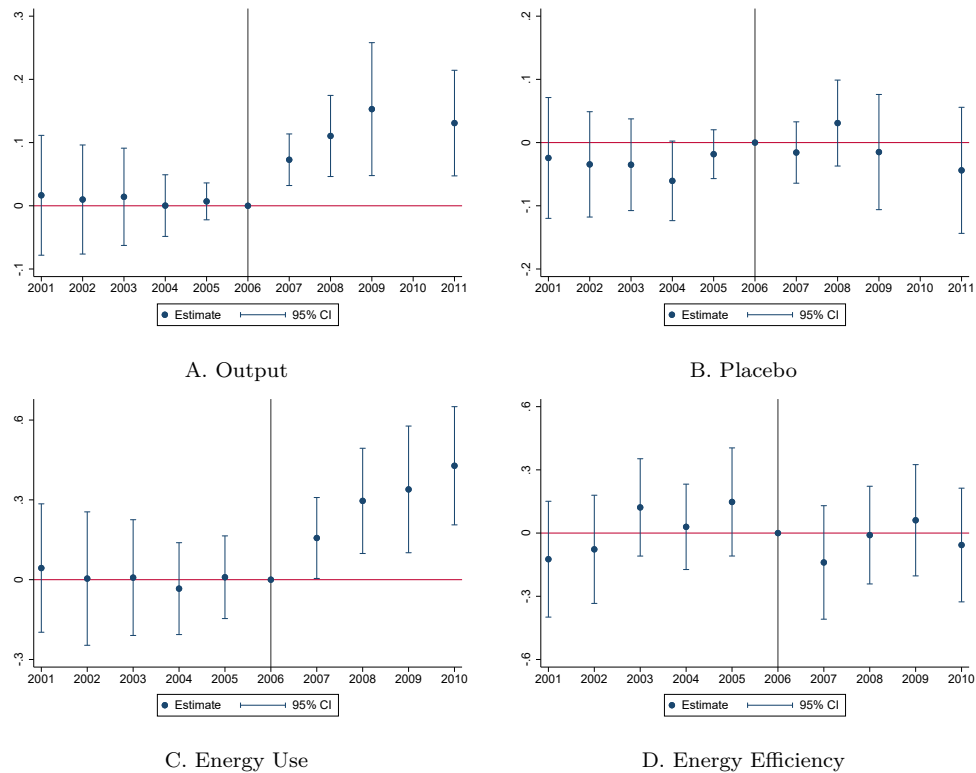


FIGURE A.15. ROBUSTNESS OF SPILLOVERS TO RELATED FIRMS: ENTROPY MATCHING

Note: This figure shows the coefficients for event studies of log output, log energy consumption and log energy efficiency for firms related to regulated firms as well as an event study of log firm output for placebo firms. This figure corresponds to Figure 4 but uses the method of entropy matching. The point estimates are displayed in Tables A.22 and A.23. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the ASIF and CESD.

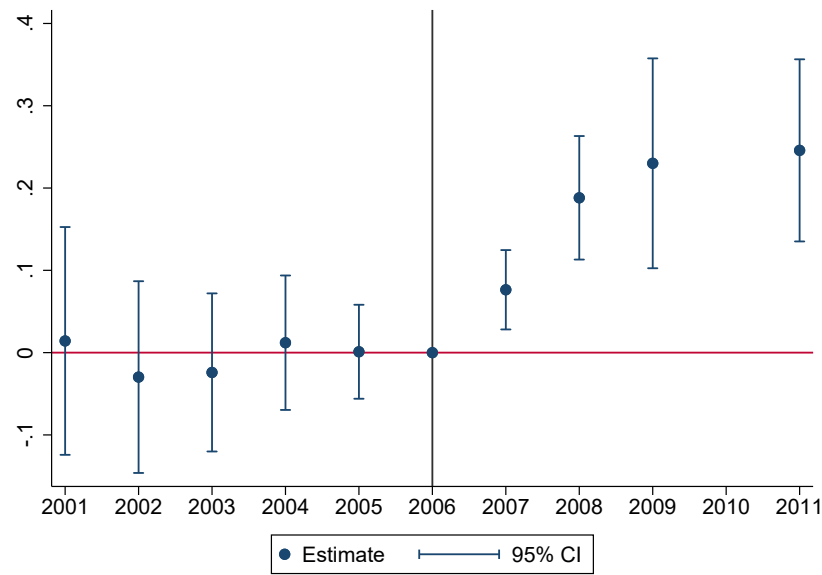
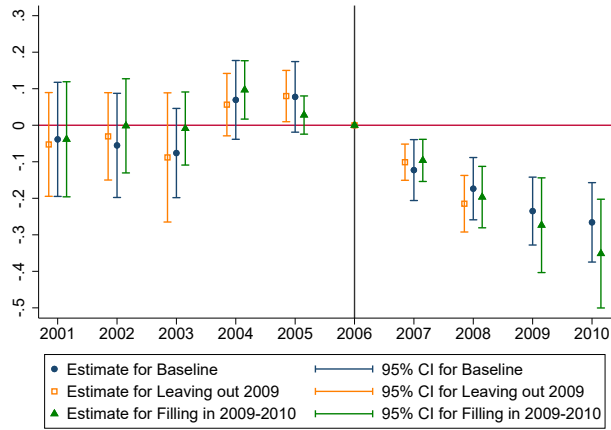
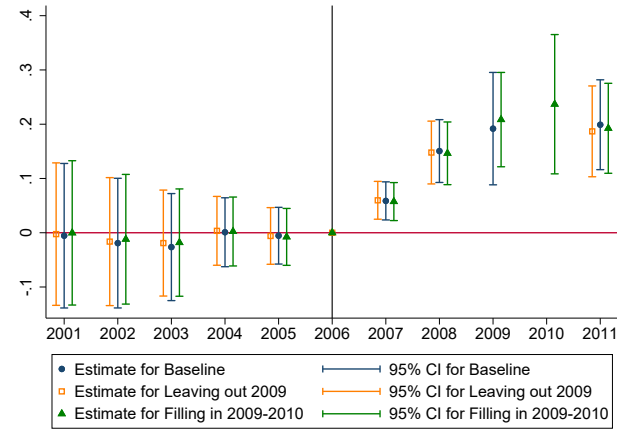


FIGURE A.16. ROBUSTNESS TO DROPPING ELECTRIC POWER GENERATION AND SUPPLY

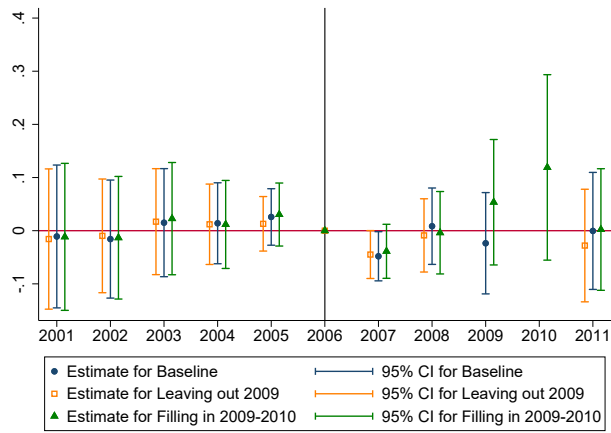
Note: This figure shows the coefficients for an event study of log firm output for firms related to regulated firms, where related firms are restricted to those in industries other than electric power generation and supply. This figure corresponds to Panel A of Figure 4 but drops all observations in the electric power generation and supply industry. The point estimates are shown in Table A.25. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the ASIF.



A. Regulated Firms



B. Related Firms



C. Placebo on Related Firms

FIGURE A.17. DATA QUALITY ROBUSTNESS

Note: Panel A corresponds to Panel C in Figure 3, and Panels B and C correspond to Panels A and B in Figure 4. These figures drop data for 2009 from the ASIF (yellow lines) and fill in the 2009–2010 data with ATS 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.

Source: Authors' calculations using data from the ASIF and ATS.

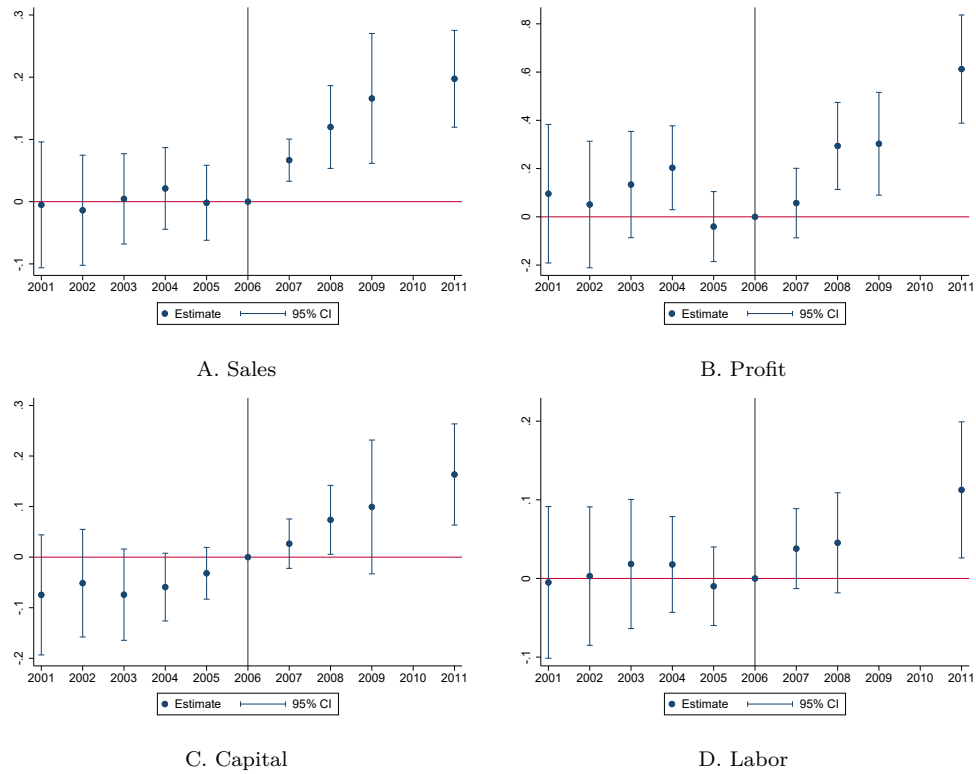


FIGURE A.18. ADDITIONAL SPILLOVER EFFECTS OF THE PROGRAM

Note: 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.27. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the ASIF.

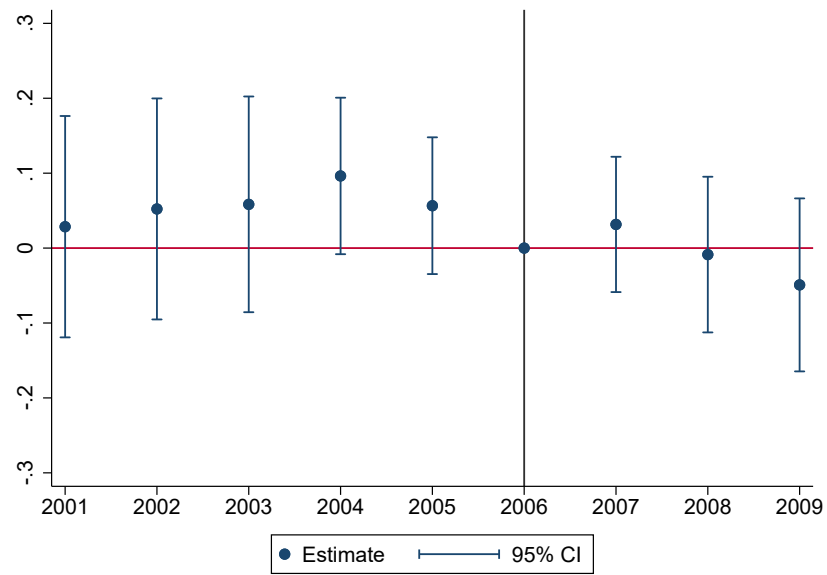


FIGURE A.19. ROBUSTNESS OF EFFECTS ON THE ENERGY EFFICIENCY OF REGULATED FIRMS

Note: 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 cost of sales. This figure corresponds to Panel E of Figure 3 but with an alternative definition of energy efficiency. The point estimates are shown in Table A.34. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the CESD and ASIF.

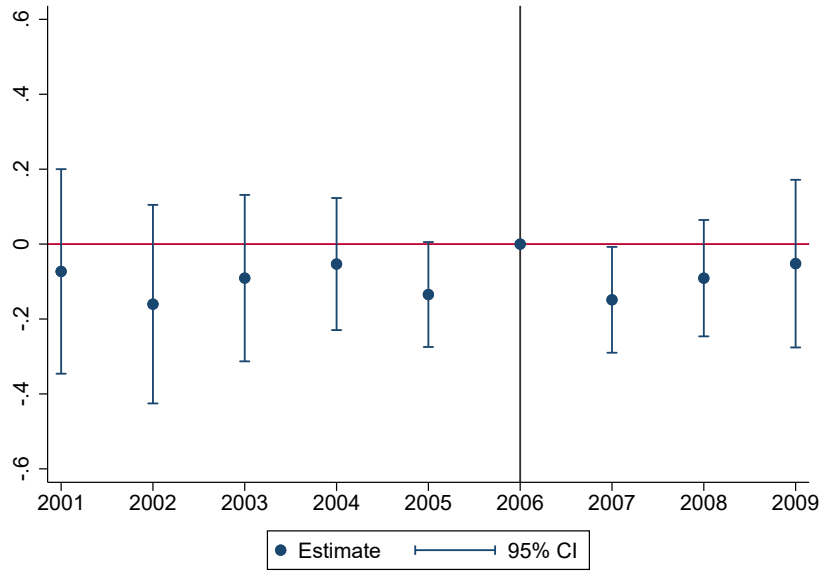


FIGURE A.20. ROBUSTNESS OF EFFECTS ON THE ENERGY EFFICIENCY OF RELATED FIRMS

Note: 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 4 but with an alternative definition of energy efficiency. The point estimates are shown in Table A.35. Standard errors are clustered at the firm level.
Source: Authors' calculations using data from the CESD and ASIF.

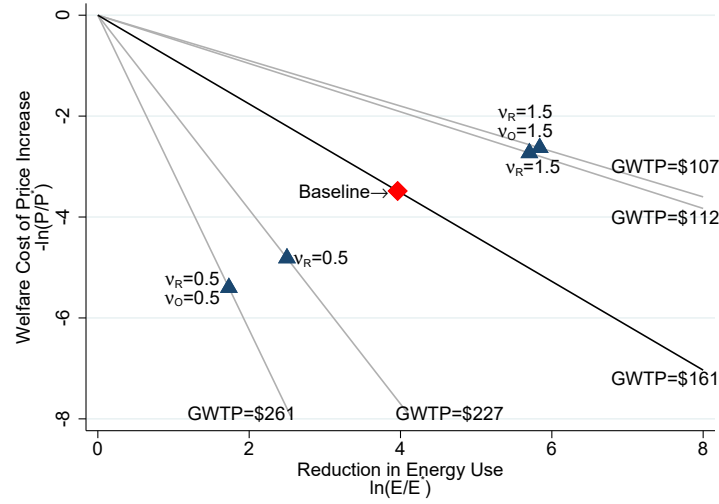


FIGURE A.21. WELFARE EFFECTS: ADDITIONAL ROBUSTNESS TO DIFFERENCES IN ENERGY EFFICIENCY

Note: 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 $\text{GWTP}=\$161$, and the light gray lines plot values of the GWTP according to different extensions. As discussed in Section J.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 J.2 for details. While Panel B of Figure A.2 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.

Source: Author's calculations using the model developed in Section IV.

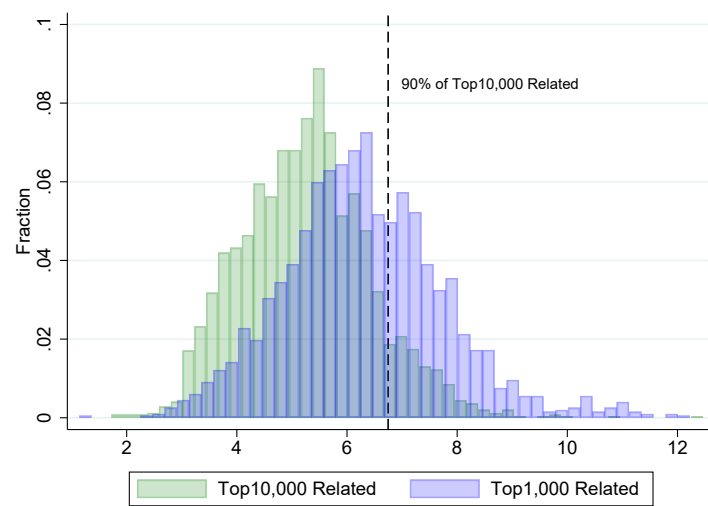


FIGURE A.22. SIZE DISTRIBUTION OF FIRMS RELATED TO TOP 10,000 FIRMS

Note: 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.
Source: Author's calculations using data from ASIF.

Appendix Tables

TABLE A.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

Note: 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 1,008 Top 1,000 firms are included in this report.
Source: Author's calculations using data from the NDRC and NBS.

TABLE A.2—HETEROGENEOUS SPILLOVER EFFECTS BY LOCAL POLLUTION AND DENSITY

Variables	High to Low	Horizontal	Low to High
Related $\times$ Post	0.115 (0.047) [0.015]	0.101 (0.072) [0.159]	0.224 (0.073) [0.002]
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	6.64%	1.40%	6.31%
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y
Firm-Level Controls	Y	Y	Y

Note: This table show estimates of Equation 2 where Related $\times$ 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 2 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 that of their corresponding Top 1,000 firms (by more than 10%); column (2) includes related firms whose city-level pollution exposure is similar to that of their corresponding Top 1,000 firms (within a 10% difference); and column (3) includes related firms whose city-level pollution exposure is more than that of their corresponding Top 1,000 firms (by more than 10%). City-level pollution exposure is defined by city-level so_2 density $\times$ city population. This table shows that, after the implementation of the policy, related firms increased output by 11.5% in places with lower pollution exposure, by 10.1% in places with similar pollution exposure, and by 22.4% in places with more pollution exposure. 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 K for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.3—HETEROGENEOUS SPILLOVER EFFECTS BY LOCATION

Variables	East&Mid to West	Within Region	West to East&Mid
Related $\times$ Post	0.117 (0.129) [0.367]	0.164 (0.041) [0.000]	0.177 (0.275) [0.523]
Observations	1,133	9,230	456
R^2	0.900	0.887	0.907
Output Share	8.70%	83.40%	7.80%
Aggregate Effect	1.02%	13.70%	1.38%
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y
Firm-Level Controls	Y	Y	Y

Note: This table shows estimates of Equation 2 where Related $\times$ 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 2 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 in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.4—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	

Note: 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 I.A for additional discussion.

Source: Authors' calculations using data from the NDRC.

TABLE A.5—HETEROGENEOUS EFFECTS ON REGULATED FIRMS BY REPORTED COMPLIANCE STATUS

Variables	ln(Energy Use)			
Compliance $\times$ Post	-0.138 (0.042) [0.001]	-0.168 (0.046) [0.000]	-0.169 (0.047) [0.000]	-0.141 (0.048) [0.003]
Non-Compliance $\times$ Post	0.395 (0.347) [0.255]	0.338 (0.366) [0.355]	0.316 (0.359) [0.379]	0.333 (0.346) [0.337]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: 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 was implemented, while the noncompliant regulated firms show a 31%–40% increase in their energy consumption. However, because only 15 firms had not achieved their energy savings targets at the end of the 11FYP, the results for Noncompliance have large standard errors. This result suggests that the official compliance status can be verified to some extent by firm energy consumption. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using the ASIF and CESD data.

TABLE A.6— 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%

Note: 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 I.B for additional discussion.

Source: Authors' calculations using data from the ASIF and CESD.

TABLE A.7—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

Note: 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 I.B 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 I.C for additional discussion.

Source: Author's calculations using ASIF, CESD, and CARD data.

TABLE A.8—WATERFALL TABLE

Data Cleaning Process	Firm Number
Top 1,000 (2004)	1008
Top 1,000 (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

Note: This table reports the number of Top 1,000 firms in the data according to different sample restrictions.

Source: Authors' calculations using data from the CESD.

TABLE A.9—ROBUSTNESS TO DIFFERENT SAMPLES

	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)	Top 1,000
Panel A: CESD Data				
All Sample	-0.168 (0.046) [0.000]	-0.237 (0.038) [0.000]	-0.071 (0.047) [0.133]	687
Drop Missing Coal Data	-0.173 (0.042) [0.000]	-0.235 (0.040) [0.000]	-0.061 (0.044) [0.164]	527
Drop Extreme Values	-0.172 (0.042) [0.000]	-0.236 (0.040) [0.000]	-0.063 (0.044) [0.145]	520
Industry(Electricity<50%)	-0.170 (0.043) [0.000]	-0.233 (0.042) [0.000]	-0.064 (0.045) [0.154]	479
Industry(Electricity<30%)	-0.156 (0.047) [0.001]	-0.204 (0.042) [0.000]	-0.049 (0.046) [0.290]	427
Industry(Electricity<25%)	-0.146 (0.048) [0.002]	-0.195 (0.043) [0.000]	-0.048 (0.047) [0.305]	420
Panel B: CESD and ATS Data				
Infilling Missing Coal Data	-0.236 (0.048) [0.000]	-0.288 (0.040) [0.000]	-0.043 (0.047) [0.365]	602
Firm FE	Y	Y	Y	
Industry $\times$ Year FE	Y	Y	Y	
Province $\times$ Year FE	Y	Y	Y	

Note: 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 1 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 in brackets below.

Source: Authors' calculations using data from the CESD and ATS.

TABLE A.10—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 (0.047) [0.001]	-0.204 (0.042) [0.000]	-0.049 (0.046) [0.290]
Rank < 8000	-0.178 (0.055) [0.001]	-0.230 (0.047) [0.000]	-0.054 (0.058) [0.351]
Rank < 5000	-0.246 (0.066) [0.000]	-0.276 (0.054) [0.000]	-0.029 (0.069) [0.677]
Rank < 3000	-0.186 (0.069) [0.007]	-0.288 (0.065) [0.000]	-0.104 (0.072) [0.148]
Rank > 300	-0.202 (0.049) [0.000]	-0.241 (0.042) [0.000]	-0.039 (0.048) [0.413]
Rank > 500	-0.232 (0.050) [0.000]	-0.263 (0.043) [0.000]	-0.033 (0.049) [0.500]
Rank > 800	-0.243 (0.050) [0.000]	-0.289 (0.045) [0.000]	-0.048 (0.047) [0.308]
300 < Rank < 8000	-0.226 (0.057) [0.000]	-0.272 (0.048) [0.000]	-0.048 (0.060) [0.424]
500 < Rank < 5000	-0.328 (0.069) [0.000]	-0.314 (0.055) [0.000]	0.014 (0.072) [0.845]
800 < Rank < 3000	-0.293 (0.079) [0.000]	-0.337 (0.073) [0.000]	-0.049 (0.085) [0.564]
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: 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 1 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.9 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 in brackets below.

Source: Authors' calculations using data from the CESD.

TABLE A.11—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 (0.047) [0.001]	-0.204 (0.042) [0.000]	-0.049 (0.046) [0.290]
Control > 20%	-0.146 (0.047) [0.002]	-0.175 (0.042) [0.000]	-0.031 (0.046) [0.503]
Control > 40%	-0.126 (0.048) [0.009]	-0.139 (0.042) [0.001]	-0.016 (0.048) [0.740]
Control > 60%	-0.117 (0.051) [0.023]	-0.102 (0.044) [0.020]	0.014 (0.051) [0.779]
Treat < 80%	-0.179 (0.051) [0.000]	-0.188 (0.047) [0.000]	-0.012 (0.050) [0.816]
Treat < 60%	-0.186 (0.060) [0.002]	-0.179 (0.054) [0.001]	0.004 (0.059) [0.949]
Treat < 40%	-0.186 (0.074) [0.012]	-0.165 (0.066) [0.013]	0.016 (0.073) [0.831]
Control > 20% & Treat < 80%	-0.168 (0.051) [0.001]	-0.159 (0.047) [0.001]	0.005 (0.051) [0.927]
Control > 40% & Treat < 60%	-0.152 (0.062) [0.014]	-0.112 (0.055) [0.041]	0.034 (0.061) [0.576]
Control > 60% & Treat < 40%	-0.129 (0.076) [0.090]	-0.047 (0.066) [0.476]	0.077 (0.077) [0.316]
Firm FE	Y	Y	Y
Industry × Year FE	Y	Y	Y
Province × Year FE	Y	Y	Y

Note: 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 1 and use firms' estimated propensity of being treated to trim the estimation sample. The propensity score comes from 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, 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 20th, 40th, and 60th percentiles of likelihood of being treated, respectively. Rows (5)–(7) drop treated firms that are above the 20th, 40th, and 60th 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.10 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 in brackets below.

Source: Authors' calculations using data from the CESD.

TABLE A.12—ROBUSTNESS TO ENTRY AND EXIT: REGULATED FIRM RESPONSE

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat $\times$ Post	-0.141 (0.049) [0.004]	-0.170 (0.044) [0.000]	-0.029 (0.048) [0.549]
Observations	18,506	18,385	18,385
R^2	0.886	0.890	0.831
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: This table reports the difference-in-differences estimates from regressions of log firm energy consumption, output and energy efficiency on regulated firms when we exclude firm entry and exit. The estimates shown in this table correspond to Equation 2 and the baseline estimates in Table 1 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 in brackets below.

Note: Authors' calculations using data from the CESD.

TABLE A.13—REGULATED FIRM RESPONSE USING AN APPROXIMATELY BALANCED PANEL

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat $\times$ Post	-0.153 (0.069) [0.026]	-0.145 (0.057) [0.012]	0.006 (0.065) [0.923]
Observations	10,336	10,265	10,265
R^2	0.892	0.910	0.835
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: 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 1 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 in brackets below.

Source: Authors' calculations using data from the CESD.

TABLE A.14—EFFECTS OF POLICY ON REGULATED FIRMS' INVESTMENT

Variables	If Firm Invest			
Treat $\times$ Post	-0.056 (0.013) [0.000]	-0.070 (0.014) [0.000]	-0.070 (0.014) [0.000]	-0.071 (0.014) [0.000]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table shows estimates of Equation 2 where Treat $\times$ 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 I for more information about the data-generating procedure. This table corresponds to a pooled version of the regression displayed in Figure A.11. It shows that regulated firms decreased their likelihood 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 II for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using ASIF and ATS.

TABLE A.15—HETEROGENEOUS EFFECTS ON REGULATED FIRMS BY INDUSTRY

Variables	ln(Output)				
	Baseline	Drop Power	Processing	Materials	Mining
Treat $\times$ Post	-0.204 (0.042) [0.000]	-0.206 (0.042) [0.000]	-0.219 (0.073) [0.003]	-0.153 (0.053) [0.004]	-0.251 (0.275) [0.363]
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 $\times$ Year FE	Y	Y	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y	Y	Y

Note: This table shows estimates of Equation 2 by industry where Treat $\times$ 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 1. 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 in brackets below.

Source: Authors' calculations using data from the CESD.

TABLE A.16—HETEROGENEOUS EFFECTS ON REGULATED FIRMS BY OWNERSHIP

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat $\times$ Post	-0.135 (0.051) [0.008]	-0.167 (0.047) [0.000]	-0.034 (0.049) [0.497]
Treat $\times$ SOE $\times$ Post	0.041 (0.115) [0.723]	-0.111 (0.104) [0.288]	-0.141 (0.120) [0.239]
Observations	21,924	21,778	21,778
R^2	0.892	0.889	0.841
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y
SOE $\times$ Year FE	Y	Y	Y

Note: Authors' calculations using data from the CESD. This table shows estimates of Equation 2 by firm ownership type where Treat $\times$ 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 in brackets below.

Source: Authors' calculations using data from the CESD and ASIF.

TABLE A.17—ROBUSTNESS TO CONSIDERATION OF CONCURRENT POLICIES

A. Effects of Concurrent Policies on Top 1,000 Firms

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Monitor $\times$ Post	-0.003 (0.097) [0.978]	0.023 (0.089) [0.799]	0.019 (0.106) [0.855]
Observations	3,358	3,322	3,322
R^2	0.865	0.899	0.812
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

B. Robustness to Effects on Top 1000 Firms

Variables	ln(Energy Use)	ln(Output)	ln(Energy Efficiency)
Treat $\times$ Post	-0.186 (0.075) [0.013]	-0.234 (0.080) [0.003]	-0.046 (0.083) [0.581]
Observations	20,655	20,511	20,511
R^2	0.864	0.858	0.847
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: 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 1 while excluding all treated firms included under both policies. It shows that taking the NSMF program into account does not affect our main results. See Section D.2 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 in brackets below.

Source: Authors' calculations using data from the CESD.

TABLE A.18—HETEROGENEOUS EFFECTS ON REGULATED FIRMS BY ELECTRICITY CONSUMPTION

Variables	ln(Energy)	ln(Output)	ln(Efficiency)
Treat $\times$ Post $\times$ Electricity-Intensive	0.044 (0.081) [0.582]	-0.029 (0.074) [0.697]	-0.075 (0.081) [0.358]
Treat $\times$ Post	-0.169 (0.068) [0.012]	-0.154 (0.058) [0.008]	0.014 (0.068) [0.838]
Observations	22,623	22,471	22,471
R^2	0.890	0.886	0.841
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: 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 heterogeneity in effects 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 in brackets below.

Source: Author's calculations using data from the CESD and ATS.

TABLE A.19—HETEROGENEOUS EFFECTS ON REGULATED FIRMS BY CHANGE IN ELECTRICITY USE

Variables	ln(Energy)	ln(Output)	ln(Efficiency)
Treat $\times$ Post $\times$ Electricity-Increase	0.070 (0.134) [0.601]	0.034 (0.139) [0.808]	-0.032 (0.153) [0.834]
Treat $\times$ Post	-0.126 (0.056) [0.025]	-0.172 (0.051) [0.001]	-0.050 (0.056) [0.368]
Observations	21,801	21,656	21,656
R^2	0.884	0.879	0.842
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: This table shows estimates of heterogeneous effects within regulated firms by changes in firm electricity consumption. Electricity-Increase is an indicator that equals one if the firms' change in electricity use between 2007 and 2010 is above the industry's average. This table shows no significant effect heterogeneity between regulated firms with larger increases in electricity use during the post-period. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Author's calculations using data from the CESD and ATS.

TABLE A.20—EFFECTS OF THE PROGRAM ON REGULATED FIRMS' ENERGY MIX

Variables	Coal Share	Petroleum Share	Gas Share
Treat $\times$ Post	-0.005 (0.007) [0.494]	0.001 (0.004) [0.706]	0.003 (0.006) [0.582]
Observations	23,151	23,151	23,151
R^2	0.679	0.636	0.667
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: This table shows estimates of Equation 2 where Treat $\times$ Post is an indicator for regulated firms interacted with an indicator for years after 2006. 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 in brackets below.

Source: Author's calculations using data from the CESD.

TABLE A.21—EFFECTS OF THE TOP 1,000 PROGRAM ON INPUT SUBSTITUTION IN REGULATED FIRMS

Variables	ln(Capital/Labor)	ln(Energy/Labor)	ln(Energy/Capital)
Treat $\times$ Post	0.034 (0.035) [0.339]	0.041 (0.044) [0.354]	0.007 (0.047) [0.882]
Observations	13,600	13,600	13,600
R^2	0.736	0.812	0.808
Firm FE	Y	Y	Y
Industry $\times$ Year FE	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y

Note: This table shows estimates of Equation 2, where Treat $\times$ 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.14 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 in brackets below.

Source: Authors' calculations using data from the CESD, ASIF and ATS.

TABLE A.22—ROBUSTNESS OF SPILLOVERS TO RELATED FIRMS ON OUTPUT USING ENTROPY MATCHING

A. Output				
Variables	ln(Output)			
Related $\times$ Post	0.143 (0.039) [0.000]	0.139 (0.038) [0.000]	0.106 (0.037) [0.004]	0.133 (0.036) [0.000]
Observations	119,064	119,064	119,064	116,064
R^2	0.874	0.881	0.890	0.896
B. Placebo on Output				
Variables	ln(Output)			
Related $\times$ Post	0.016 (0.041) [0.694]	0.019 (0.040) [0.630]	0.020 (0.039) [0.612]	0.036 (0.038) [0.342]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table shows estimates of Equation 2 with an alternative matching method, where Related $\times$ 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 2 but uses the 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.15. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.23—ROBUSTNESS OF SPILLOVERS TO RELATED FIRMS ON ENERGY USE AND ENERGY EFFICIENCY:
ENTROPY MATCHING

A. Energy Use				
Variables	ln(Energy Use)			
Related $\times$ Post	0.247 (0.097) [0.011]	0.248 (0.095) [0.009]	0.289 (0.090) [0.001]	0.263 (0.104) [0.011]
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 $\times$ Post	-0.021 (0.099) [0.836]	-0.019 (0.098) [0.844]	-0.056 (0.103) [0.584]	-0.039 (0.104) [0.706]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table shows estimates of Equation 2 with an alternative matching method, where Related $\times$ Post is an indicator for related firms in the same 4-digit industry interacted with an indicator for years after 2006. 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 2 but uses the method of entropy matching and corresponds to a pooled version of the regression displayed in Panels C and D of Figure A.15. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the CESD and ASIF.

TABLE A.24—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 $\times$ Post	0.127 (0.035) [0.000]	0.125 (0.036) [0.001]	0.133 (0.033) [0.000]	0.124 (0.035) [0.000]	0.154 (0.039) [0.000]
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 $\times$ Year FE	Y	Y	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y	Y	Y
Firm-Level Controls	Y	Y	Y	Y	Y

Note: This table shows estimates of Equation 2 where Related $\times$ 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 2 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 with up to 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 with up to 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 implementation of the policy; 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 with up to 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 III for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.25—ROBUSTNESS TO DROPPING ELECTRIC POWER GENERATION AND SUPPLY

Variables	ln(Output)			
Related $\times$ Post	0.180 (0.043) [0.000]	0.173 (0.041) [0.000]	0.163 (0.039) [0.000]	0.167 (0.037) [0.000]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table shows estimates of Equation 2 where Related $\times$ 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 2 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.16. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.26—DATA QUALITY ROBUSTNESS

A. Leave out 2009 in ASIF Data				
Variables	ln(Output)			
Treat $\times$ Post	-0.054 (0.035) [0.123]	-0.198 (0.036) [0.000]	-0.166 (0.039) [0.000]	-0.112 (0.038) [0.003]
Observations	10,028	10,023	9,860	9,678
R^2	0.955	0.959	0.961	0.963
Related $\times$ Post	0.131 (0.034) [0.000]	0.125 (0.033) [0.000]	0.102 (0.033) [0.002]	0.111 (0.031) [0.000]
Observations	16,454	16,452	16,450	15,970
R^2	0.890	0.898	0.905	0.913
Placebo $\times$ Post	-0.033 (0.039) [0.392]	-0.031 (0.037) [0.402]	-0.025 (0.038) [0.505]	-0.019 (0.037) [0.599]
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 $\times$ Post	-0.235 (0.047) [0.000]	-0.262 (0.047) [0.000]	-0.237 (0.048) [0.000]	-0.201 (0.048) [0.000]
Observations	11,805	11,803	11,586	11,102
R^2	0.921	0.925	0.929	0.932
Related $\times$ Post	0.159 (0.037) [0.000]	0.156 (0.037) [0.000]	0.136 (0.037) [0.000]	0.135 (0.035) [0.000]
Observations	19,291	19,287	19,285	18,733
R^2	0.869	0.876	0.885	0.893
Placebo $\times$ Post	0.000 (0.041) [0.990]	-0.003 (0.040) [0.949]	-0.005 (0.040) [0.906]	-0.007 (0.039) [0.851]
Observations	9,290	9,289	9,269	9,105
R^2	0.875	0.880	0.889	0.897
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: 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 1. 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 2, and the placebo results correspond to the results in Panel B in Table 2. 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.17. 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 in brackets below.

Source: Authors' calculations using the ASIF and ATS.

TABLE A.27—ADDITIONAL SPILLOVER EFFECTS OF THE PROGRAM

Variables	ln(Sale)	ln(Profit)	ln(Capital)	ln(Labor)
Related $\times$ Post	0.115 (0.033) [0.001]	0.190 (0.055) [0.001]	0.114 (0.040) [0.004]	0.063 (0.026) [0.016]
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 $\times$ Year FE	Y	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y	Y
Firm-Level Controls	Y	Y	Y	Y

Note: This table shows estimates of Equation 2 where Related $\times$ 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.18. 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 in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.28—HETEROGENEOUS SPILLOVER EFFECTS ON OUTPUT BY RELATED FIRM SIZE

Variables	ln(Output)			
Related $\times$ Post(0%-30%)	0.104 (0.054) [0.053]	0.109 (0.052) [0.037]	0.048 (0.052) [0.359]	0.078 (0.051) [0.123]
Related $\times$ Post(30%-60%)	0.130 (0.047) [0.005]	0.123 (0.045) [0.006]	0.096 (0.046) [0.035]	0.111 (0.043) [0.011]
Related $\times$ Post(60%-100%)	0.164 (0.045) [0.000]	0.156 (0.044) [0.000]	0.161 (0.043) [0.000]	0.161 (0.040) [0.000]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table corresponds to Panel A of Table 2 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 III for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.29—SPILLOVER EFFECTS ON THE OUTPUT OF RELATED FIRMS: HETEROGENEOUS EFFECTS BY INDUSTRY

Variables	ln(Output)				
	Baseline	Drop Power	Processing	Materials	Mining
Related $\times$ Post	0.118 (0.037) [0.001]	0.164 (0.040) [0.000]	0.105 (0.096) [0.275]	0.180 (0.056) [0.001]	0.197 (0.080) [0.015]
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 $\times$ Year FE	Y	Y	Y	Y	Y
Province $\times$ Year FE	Y	Y	Y	Y	Y

Note: This table shows estimates of Equation 2 by industry where Related $\times$ 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 2. 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, which include the smelting and pressing of ferrous metals, the smelting and pressing of nonferrous metals, the processing of petroleum, coking, and the processing of nuclear fuel. Column (4) shows the result for material industries, which include the manufacture of raw chemical materials and chemical products as well as the manufacture of nonmetallic mineral products. Column (5) shows the results for mining industries, which include the mining and washing of coal as well as the 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 in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.30—MARKET SPILLOVERS: HETEROGENEOUS EFFECTS ACROSS 4-DIGIT INDUSTRIES WITH TOP 1,000 FIRMS

Variables	ln(Output)			
	All Sample	Energy-Intensive Industries		
Spillover $\times$ Post $\times$ $\mathbb{I}[\text{Top 1,000 Firm in 4-Digit Ind}]$	0.087 (0.022) [0.000]	0.078 (0.020) [0.000]	0.082 (0.021) [0.003]	0.084 (0.026) [0.010]
Spillover $\times$ Post $\times$ $\mathbb{I}[\text{No Top 1,000 Firm in 4-Digit Ind.}]$	0.037 (0.039) [0.354]	0.040 (0.031) [0.207]	0.040 (0.056) [0.494]	0.046 (0.057) [0.435]
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

Note: This table shows estimates of Equation 3 where Spillover $\times$ 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 3 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 III for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

TABLE A.31—ROBUSTNESS TO ENTRY: MARKET SPILLOVERS

Variables	ln(Output)			
	All Sample	Energy-Intensive Industries		
Spillover $\times$ Post	0.083 (0.022) [0.001]	0.075 (0.020) [0.000]	0.092 (0.021) [0.002]	0.092 (0.025) [0.005]
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

Note: This table shows estimates of Equation 3 where Spillover $\times$ 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 to Panel A of Table 3 but excludes 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 III for additional discussion. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

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 (0.076) [0.682]	-0.074 (0.071) [0.299]	-0.053 (0.071) [0.458]	-0.078 (0.072) [0.276]	
Treat × Post	-0.071 (0.064) [0.266]	-0.180 (0.059) [0.002]	-0.167 (0.059) [0.005]	-0.098 (0.060) [0.104]	
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 (0.076) [0.782]	-0.026 (0.071) [0.708]	-0.007 (0.071) [0.922]	-0.017 (0.072) [0.815]	
Treat × Post	-0.102 (0.058) [0.080]	-0.211 (0.056) [0.000]	-0.197 (0.056) [0.000]	-0.131 (0.057) [0.021]	
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 (0.079) [0.419]	0.037 (0.078) [0.636]	0.037 (0.080) [0.645]	0.028 (0.080) [0.724]	
Treat × Post	-0.003 (0.056) [0.952]	-0.091 (0.056) [0.109]	-0.069 (0.057) [0.228]	-0.029 (0.059) [0.627]	
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	

Note: 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 targets were allocated according to each firm's total energy consumption before the policy, which resulted in all regulated firms having a similar response. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Author's calculations using data from the CESD.

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) All Sources	Coal Only	Damages Per Ton 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 (2020b) Willingness to Pay	\$55.8	2013	9.80	7.49	\$5.69	\$7.45

Note: This table uses estimates of total pollution damages from a range of studies (column (1)) and estimates of CO₂ emissions from Ritchie et al. (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 I for additional details.

Source: Author's calculations using data from Mohan et al. (2020), World Bank (2007), Ito and Zhang (2020), and Ritchie et al. (2020).

TABLE A.34—ROBUSTNESS OF EFFECTS ON THE ENERGY EFFICIENCY OF REGULATED FIRMS

Variables	ln(Variable Input/Energy)			
Treat × Post	-0.020 (0.043) [0.641]	-0.082 (0.047) [0.078]	-0.053 (0.048) [0.273]	-0.029 (0.048) [0.546]
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

Note: 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 1 but with an alternative definition of energy efficiency. It also corresponds to a pooled version of the regression displayed in Figure A.19. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the CESD and ASIF.

TABLE A.35—ROBUSTNESS OF EFFECTS ON THE ENERGY EFFICIENCY OF RELATED FIRMS

Variables	ln(Variable Input/Energy)			
Related $\times$ Post	-0.026 (0.078) [0.741]	-0.033 (0.080) [0.680]	-0.026 (0.088) [0.770]	-0.008 (0.088) [0.923]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

Note: This table shows estimates of Equation 2 where Related $\times$ 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 2 but with an alternative definition of energy efficiency. It also corresponds to a pooled version of the regression displayed in Figure A.20. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the CESD and ASIF.

TABLE A.36—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 (0.003)	0.900 (0.003)	0.900 (0.003)	0.853 (0.005)	0.949 (0.001)	0.800 (0.007)
Dispersion of ln-ability σ_m		1.239 (0.055)	1.500 (0.060)	0.579 (0.045)	1.359 (0.087)	1.063 (0.053)	1.271 (0.097)
Survival threshold ϕ_1		0.609 (0.166)	0.461 (0.134)	0.792 (0.193)	0.579 (0.273)	0.435 (0.118)	0.985 (0.535)
3. Policy Parameters							
Policy threshold ϕ		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

Note: 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 V.A for the detailed estimation procedure. Panel 3 reports the policy threshold ϕ . 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 J.3 discusses the results when we vary α or ρ , and Section J.4 discusses the results when we set $\rho_c = 0.9$.

Source: Author's calculations from ASIF, CARD, and 2004 Economic Census data.

TABLE A.37—SUMMARY STATISTICS FOR TOP 10,000 & TOP 10,000 RELATED FIRMS

	Top 10,000		Top 10,000 Related	
	Number	Output	Number	Output
<u>Panel A. All Firms Matched in ASIF</u>				
All Industries	14300	2201.53	9787	885.63
Energy-Intensive Industries	10962	1861.45	5383	925.67
<u>Panel B. Exclude Top 1,000 and Top 1,000 Related</u>				
All Industries	12524	1588.03	8522	694.00
Energy-Intensive Industries	9222	970.27	4152	545.35

Note: 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.

Source: Author's calculations using data from the ASIF.

TABLE A.38—DIRECT AND SPILLOVER EFFECTS OF THE TOP 10,000 PROGRAM

A. Regulated Firms

Variables	ln(Output)			
Treat $\times$ Post	-0.158 (0.007) [0.000]	-0.156 (0.007) [0.000]	-0.163 (0.007) [0.000]	-0.130 (0.007) [0.000]
Observations	319,346	319,346	319,346	310,966
R^2	0.879	0.880	0.883	0.895
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Industry $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

B. Related Firms

Variables	ln(Output)			
Related(0%-90%) $\times$ Post	0.001 (0.021) [0.956]	-0.001 (0.021) [0.981]	0.007 (0.023) [0.771]	0.015 (0.021) [0.477]
Related(90%-100%) $\times$ Post	0.067 (0.049) [0.168]	0.083 (0.050) [0.095]	0.107 (0.052) [0.041]	0.115 (0.050) [0.021]
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 $\times$ Year FE		Y	Y	Y
Province $\times$ Year FE			Y	Y
Firm-Level Controls				Y

C. Market Spillover

Variables	ln(Output)			
	All Sample	Energy-Intensive Industries		
Incremental Spillover $\times$ Post	0.045 (0.015) [0.004]	0.043 (0.014) [0.004]	0.019 (0.016) [0.255]	0.018 (0.014) [0.255]
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

Note: This table shows effects of the Top 10,000 Program. Panel A corresponds to Panel B in Table 1, but where Treat $\times$ 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 2, but where Related is an indicator for Top 10,000 related firms, and 0%-90% and 90%-100% refer to percentile ranges of firm size. Panel C corresponds to Panel A of Table 3, but where Incremental Spillover refers to the industry-level exposure to the Top 10,000 program minus the Top 1,000 exposure. When the size-dependent policy expands, the regulated firms still experience significant negative effects. However, compared to the Top 1,000 program, the related-firm and market spillover effects are significantly reduced. Standard errors clustered at the firm level are shown in parentheses with p-values in brackets below.

Source: Authors' calculations using data from the ASIF.

REFERENCES

- Ai, Hongshan, Yang Hu, and Ke Li**, “Impacts of Environmental Regulation on Firm Productivity: Evidence from China’s Top 1000 Energy-Consuming Enterprises Program,” *Applied Economics*, 2021, 53 (7), 830–844.
- Baylis, Kathy, Don Fullerton, and Daniel H. Karney**, “Negative Leakage,” *Journal of the Association of Environmental and Resource Economists*, 2014, 1 (1/2), 51–73.
- Cao, Kaihu**, “Power Limit Affect Electrolytic Aluminum Industry, Jiaozuo Wanfeng Aluminum Co., Ltd. Reduced Production by 1/3,” *China Business News*, 2010.
- CCCPC**, “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),” Technical Report, Central Committee of the Communist Party of China October 2005.
- Chen, Qiaoyi, Zhao Chen, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu**, “Industrial Energy Regulation: The Role of Business Conglomerates in China,” *AEA Papers and Proceedings*, May 2021, 111, 396–400.
- CSBTS**, “Method of Calculating Energy Saved in Industrial Enterprise (企业节能量计算方法),” Technical Report, China State Bureau of Technical Supervision November 1991.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan**, “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India,” *The Quarterly Journal of Economics*, 2013, 128 (4), 1499–1545.
- , —, —, and —, “The Value of Regulatory Discretion: Estimates from Environmental Inspections in India,” *Econometrica*, 2018, 86 (6), 2123–2160.
- Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou**, “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*, 2017, 114 (39), 10384–10389.
- Filippini, Massimo, Thomas Geissmann, Valerie J Karplus, and Da Zhang**, “The Productivity Impacts of Energy Efficiency Programs in Developing Countries: Evidence from Iron and Steel Firms in China,” *China Economic Review*, 2020, 59, 101364.

- Hainmueller, Jens**, “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies,” *Political Analysis*, 2012, 20 (1), 25–46.
- IEA**, “IEA Policies Database,” Technical Report, International Energy Agency 2023. Accessible at <https://www.iea.org/policies?sector%5B0%5D=Industry&topic%5B0%5D=Energy%20Efficiency> (Accessed: April 12, 2023).
- Incopat**, “Global Patent Database,” Patent Application and Grant Dataset <https://www.incopat.com/> 2001-2010.
- Ito, Koichiro and Shuang Zhang**, “Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China,” *Journal of Political Economy*, 2020, 128 (5), 1627–1672.
- Karplus, Valerie J, Xingyao Shen, and Da Zhang**, “Herding Cats: Firm Non-Compliance in China’s Industrial Energy Efficiency Program,” *The Energy Journal*, 2020, 41 (4).
- Ke, Jing, Lynn Price, Stephanie Ohshita, David Fridley, Nina Zheng Khanna, Nan Zhou, and Mark Levine**, “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*, 2012, 50, 562 – 569.
- Kostka, Genia and William Hobbs**, “Local Energy Efficiency Policy Implementation in China: Bridging the Gap between National Priorities and Local Interests,” *The China Quarterly*, 2012, 211, 765–785.
- Li, Bin**, “Tangshan: The Energy Eaving Dilemma of a Steel City,” *China Business*, 2010.
- Li, Ruoxin and Xiaohui Li**, “Energy Saving Sprint Set Production Limit, Energy Intensive Firms All Suffered,” *China Securities Journal*, 2010.
- Ma, Liang and Jiaqi Liang**, “The Effects of Firm Ownership and Affiliation on Government’s Target Setting on Energy Conservation in China,” *Journal of Cleaner Production*, 2018, 199, 459–465.
- Mohan, Aniruddh, Nicholas Z Muller, Akshay Thyagarajan, Randall V Martin, Melanie S Hammer, and Aaron van Donkelaar**, “The Growth of Nations Revisited: Global Environmental Accounting from 1998 to 2018,” Working Paper 27398, National Bureau of Economic Research June 2020.
- Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus**, “Environmental Accounting for Pollution in the United States Economy,” *American Economic Review*, August 2011, 101 (5), 1649–75.

- NBS, “Chinese Energy Statistical Yearbook,” Statistical Yearbook 2001-2010.
- NDRC, “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),” Technical Report, National Development and Reform Commission April 2006.
- NPC, “Environmental Protection Law of the People’s Republic of China,” Technical Report, National People’s Congress 1989. Accessible at <https://www.cecc.gov/resources/legal-provisions/prc-environmental-protection-law> (Accessed: May 3, 2023).
- of Chongqing City, The People’s Government**, “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),” Technical Report, The People’s Government of Chongqing City October 2007.
- of Hainan Province, The People’s Government**, “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),” Technical Report, The People’s Government of Hainan Province December 2006.
- Porter, Michael E. and Claas van der Linde**, “Green and Competitive,” *Harvard Business Review*, Sept-October 1995, pp. 120–134.
- Price, Lynn, Mark D Levine, Nan Zhou, David Fridley, Nathaniel Aden, Hongyou Lu, Michael McNeil, Nina Zheng, Yining Qin, and Ping Yowargana**, “Assessment of China’s Energy-Saving and Emission-Reduction Accomplishments and Opportunities during the 11th Five Year Plan,” *Energy policy*, 2011, 39 (4), 2165–2178.
- , **Xuejun Wang, and Jiang Yun**, “China’s Top-1000 Energy-Consuming Enterprises Program: Reducing Energy Consumption of the 1000 Largest Industrial Enterprises in China,” Technical Report, Lawrence Berkeley National Laboratory 6 2008.
- , —, **and** —, “The Challenge of Reducing Energy Consumption of the Top-1000 Largest Industrial Enterprises in China,” *Energy Policy*, 2010, 38 (11), 6485–6498.
- Ritchie, Hannah, Max Roser, and Pablo Rosado**, “China: CO2 Country Profile,” May 2020.
- Shen, Meng, Yanping Zeng, and Ruxiao Qu**, “Environmental Regulation and Exportation: Micro Evidence from China’s Top-1000 Policy,” *Journal of International Trade (In Chinese)*, 2015, 8, 43–50.

- Wang, Jie**, “Luo Bingsheng, Vice Executive President of CISA Oppose the ”One Size Fits All” Policy,” *21st Century Business Herald*, 2010.
- Wang, Qinyong**, “China Aluminum Group Co., Ltd. Guizhou Branch Faces the Largest Power Limit Crisis,” *China Nonferrous Metals News*, 2008.
- Wang, Xuejun, Zhihua Fu, Wei Zhang, and Youbin Ye**, “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),” in “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),” China Sustainable Energy Project, January 2008.
- Wang, Yufei, Huimin Li, Qijiao Song, and Ye Qi**, “The Consequence of Energy Policies in China: A Case Study of the Iron and Steel Sector,” *Resources, Conservation and Recycling*, 2017, 117, 66–73.
- World Bank**, “4,” in “Cost of Pollution in China : Economic Estimates of Physical Damages (English),” World Bank Group, 2007.
- Xiao, Yuxian, Haitao Yin, and Jon J Moon**, “Coping with Externally Imposed Energy Constraints: Competitiveness and Operational Impact of China’s Top-1000 Energy-Consuming Enterprises Program,” *The Energy Journal*, 2023, 44 (2).
- Xu, Min**, “From Dilemma to Path – A New Perspective on the Energy Conservation Supervision and Audit (从困境到路径——节能监管审计新视角),” Technical Report, National Audit Office of the People’s Republic of China October 2010.
- Yang, Mian, Dalia Patino-Echeverri, Fuxia Yang, and Eric Williams**, “Industrial Energy Efficiency in China: Achievements, Challenges and Opportunities,” *Energy Strategy Reviews*, 2015, 6, 20–29.
- Yu, Hong**, “Fertilizer Firm Show a Wry Face to the Power Limit,” *China Chemical Industry News*, 2011.
- Zhao, Xiaofan, Huimin Li, Liang Wu, and Ye Qi**, “Implementation of Energy-Saving Policies in China: How Local Governments Assisted Industrial Enterprises in Achieving Energy-Saving Targets,” *Energy Policy*, 2014, 66, 170–184.
- , —, —, and —, “Enterprise-Level Amount of Energy Saved Targets in China: Weaknesses and a Way Forward,” *Journal of Cleaner Production*, 2016, 129, 75–87.
- Zhou, Nan, Mark D Levine, and Lynn Price**, “Overview of Current Energy-Efficiency Policies in China,” *Energy policy*, 2010, 38 (11), 6439–6452.

Zhu, Junming, Limin Niu, Matthias Ruth, and Lei Shi, “Technological Change and Energy Efficiency in Large Chinese Firms,” *Ecological Economics*, 2018, *150*, 241–250.