Ray of Hope? China and the rise of Solar Energy

Preliminary: Do not circulate without authors' consent

Ignacio Banares-Sanchez[*], Robin Burgess[†], David Laszlo[‡],
Pol Simpson[§], John Van Reenen[¶], and Yifan Wang[‖]

July 18, 2023

Abstract

The rapid decline in the global cost of solar panels coincided with China's growing market dominance in solar photovoltaics (PV) from the early 2000s. We evaluate the effectiveness of local, city-level policy efforts to encourage the growth of the solar industry in China. We develop novel measures covering all policies since their inception based on text analysis; construct original data on patenting, production and trade; and implement a synthetic-difference-in-differences approach. We show that city-level subsidies for solar production cause large increases in the production of solar panels and - with a lag - also raise innovation and the productivity of solar panel manufacturers relative to firms in matched control cities that did not implement such policies. Cities which combined production subsidies with R&D support showcased even larger positive effects on solar innovation and production. We also document positive impacts on solar firm numbers, revenues, and exports from these production and innovation subsidies. In contrast, demand/installation subsidies targeted at increasing city-level generation of solar electricity have no significant impact on city-level solar innovation and production as additional demand can be met with solar panels produced in other Chinese cities. These results are consistent with the predictions of a model in which electricity is supplied locally using components (e.g. solar panels) sourced from heterogeneous manufacturers across China, who endogenously choose whether to pay the fixed costs of innovating and/or exporting. Taken together, these results suggest that industrial policy can foster innovation and growth in clean energy. The fact that we observe this in an industry that is displacing dirty energy generation worldwide magnifies the importance of our findings.

JEL classification: L5, L52, O31, H25, L25, N5
Keywords: Solar, Energy Transition, Renewable Energy, Green Energy Subsidies, Innovation, Climate Change, Industrial Policy, China.

[*] LSE, i.banares-sanchez@lse.ac.uk
[†] LSE, r.burgess@lse.ac.uk
[‡] LSE, d.laszlo@lse.ac.uk
[§] LSE, p.simpson1@lse.ac.uk
[¶] LSE and MIT, j.vanreenen@lse.ac.uk
[‖] LSE, y.wang390@lse.ac.uk

Acknowledgements: We would like to thank participants in the Berkeley Energy Camp, Chicago EEE and IO seminars, LSE Environment Week and POID, Munich CESIFO Environmental Economics Conference and Oslo for helpful comments. Jiahua Liu and Bo Li have given both great comments and generous help with data. We would also like to thank the IGC, STICERD and the ESRC/UKRI through POID for helping to fund this work.

1 Introduction

Of all the changes required to halt climate change, none is more important than the energy transition. This is because energy (electricity, heat and transport) accounts for 73.2% of global greenhouse gas emissions.[1] Recent improvements in renewable energy technologies have made solar and wind cost-competitive with fossil fuel technologies in many parts of the world. Between 1990 and 2019, the average annual growth rate of solar energy supply exceeded that of any other energy source.[2] The rise of solar offers a "ray of hope" that we may be able to curb emissions without large-scale reductions in energy usage.

This begs the question of what underpins the dramatic cost reductions in solar energy that are driving global diffusion. Understanding this is central to ensuring that the transition to clean energy continues and may also yield insights into how to encourage other clean sources of energy (such as wind, tidal and hydrogen).

It is striking that the fastest reductions in solar photovoltaic (PV) costs since the mid-1970s have occurred in the last decade and a half, coinciding with the take-off of the solar industry in China (Figure 1). We observe that between 2004 and 2013, Chinese solar firms increased their annual production by 76% per year, and by 2016, China's dominance of global solar manufacturing had become all-encompassing. The country produced 52% of polysilicon, 81% of silicon wafer, 59% of silicon cell, and had 70% of crystalline module capacity worldwide (Ball, Reicher, Sun, & Pollock, 2017). This production was not solely low-value. Chinese firms have been innovating extensively in solar technologies and processes: the data showcases a 23% increase per year in aggregate solar patenting between 2004 and 2019.

This impressive increase in industrial activity was accompanied by the implementation of a series of major pro-solar policies by local governments in China, including production, innovation, and installation subsidies. In this paper, we assess empirically the contribution of such place-based industrial policies to the development of the solar industry in China. To do so, we exploit variation in the implementation of solar policies across city-regions. Subsidies to solar manufacturing and generation were managed and allocated by local governments. As a result, the timing, size, and targeting of policy support varied significantly depending on the

[1] Climate Watch, The World Resources Institute (2020)

[2] IEA (2021), Renewables Information: Overview. Solar PV grew at an average of 36% annually, followed by Wind (22.6%), Biogases (11.31%), Solar Thermal (10.52%), and Liquid biofuels (9.58%). The rest of renewables (Municipal waste, Geothermal, Hydro, Tide, wave & ocean, and Solid biofuels) grew at a rate lower than 5%.

city-region. To account for potential non-random implementation of policies, we use a synthetic difference-in-differences approach (Arkhangelsky, Athey, Hirshberg, Imbens, & Wager, 2021).

We find that Chinese cities that introduced local solar policies enjoyed positive and long-lasting benefits (up to at least 13 years after treatment). Notably, we estimate that the number of patents filed by solar manufacturers in treated cities increases by over 50% per year in the long run. We find similarly sizeable impacts on the number of solar manufacturers in treated cities, their total revenue, and total solar panel production. The magnitudes are small and insignificant for demand subsidies, but large for production subsidies, especially when combined with innovation subsidies.

To perform this analysis, we construct a novel longitudinal database covering city-level solar policies and city-level solar industry innovation, production and exporting outcomes. To measure policy support, we use a comprehensive data set of China's legal information (the PKULaw database), which includes all laws, regulations, and any related legal information implemented by the central and local governments since 1949. We build on recent attempts to use this or similar data sets to generate micro-level quantitative measures of industrial policy in China (e.g. Chen and Xie (2019), Wang and Yang (2021)) by using text analysis to identify all regulations that pertain to solar photovoltaics and classify these by type (e.g., subsidies) and target (installation, production, innovation).

To estimate the effectiveness of local solar subsidies, we gather a variety of city-level solar industry outcome data from a wide range of sources. We identify solar manufacturers in China using an industry directory (ENF) which covers the near-universe of solar-related companies worldwide from 2004-2021 and contains detailed company location information. From 2004-2013 we capture production and capacity information for solar manufacturers using market research reports undertaken by ENF, which include measures of solar module (and cell) production and capacity in Mega Watts per hour (MWh).

To complement the production data, we are able to obtain revenue (and other financial information such as assets, employees and cost of goods sold) of solar firms over the whole 2004-2019 period using company accounts data drawn from a variety of sources such as BVD Orbis, ASIE, and the National Firm Registry (a Census). To study innovation, we obtain patenting activity for our sample of solar manufacturers from the State Intellectual Property Office (SIPO) and PATSTAT. We classify these patents in several ways (IPC codes, SIPO classification, and text analysis) to identify solar patents and their innovative nature. We then aggregate this firm-level information to the city level using the location of each firm's headquarters as recorded in the ENF database. Most firms in the ENF data set operate exclusively in one city.

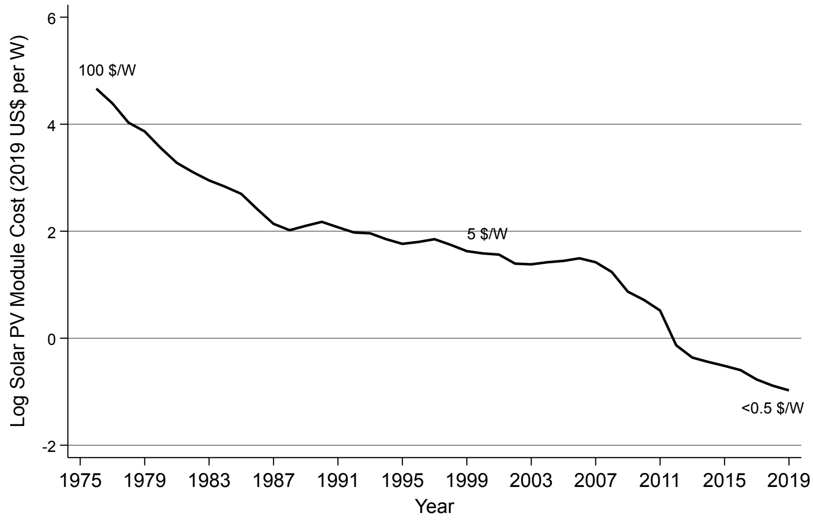
Our paper speaks to a long literature that attempts to explain the channels through which industrial policies may facilitate development (Murphy, Shleifer, and Vishny (1989), Aghion et al. (2015), Kalouptsidi (2018), Criscuolo, Martin, Overman, and Van Reenen (2019), Buera, Hopenhayn, Shin, and Trachter (2021), Lane (in press), Buera, Moll, and Shin (2013), Itskhoki and Moll (2019), Choi and Levchenko (2021)). Whilst much of this work is theoretical, there is a small, recent empirical literature that attempts to overcome challenging measurement and endogeneity problems to establish the causal effect of industrial policies. For a review on the empirical evidence on industrial policy, see Lane (2020).

In this context, our primary contribution is to show that subsidies to production lead to increases in innovation as measured by patenting activity. Previous work has documented the link between production subsidies and sustained firm growth (Manelici & Pantea, 2021), or between technology adoption subsidies, technological upgrading, and long-term outcomes (Choi & Shim, n.d.). Our main result showcases a previously unexplored link - that between production subsidies and sustained innovative activity, which is consistent with theories of learning by doing, whereby current production, enabled by policy support, affects future productivity and hence innovative activity.

We are able to draw this link as a result of our novel approach to the measurement of industrial policy. Recent work by Juhász, Lane, Oehlsen, and Pérez (2022) has used text analysis to identify policies which likely constitute industrial policy. We extend this approach to detect specific subsidies and their targets (demand, production, innovation). As a result, we are able to study how different forms of industrial policy have varying effects on a range of industrial outcomes and, in doing so provide suggestive evidence of the underlying mechanisms that generate these effects.

The structure of the paper is as follows. Section 2 provides background information on the evolution of China's solar industrial policy and our approach towards measuring it. Section 3 details the rest of the data, Section 4 gives the basics of our model and Section 5 our empirical strategy. The main results are in Section 6, some extensions in Section 7, and Section 8 concludes. Online Appendices give more details of Institutional Background (A), Data (B), Theory (C) and Further Results (D).

Figure 1: Global average price of solar PV modules (in 2019 US$) per Watt



Source: LaFond et al. (2017) & IRENA Database

2 Institutional Background: China's Industrial Policy Towards Solar PV

We now give a brief history of the intervention of the Chinese state in the solar industry back to the early 2000s, borrowing extensively from the excellent account in Ball et al. (2017) (more details are in Appendix A). First, we describe the increasing prominence of solar in the Chinese government's Five-Year Plans, which outline national economic priorities and sectoral industrial policies. Second, we describe the decentralised policy-making which led to considerable local heterogeneity in policy support towards the solar industry. Finally, we discuss the challenges with measuring industrial policy and our novel measurement approach.

2.1 Solar PV in the Government's Five-Year Plans

The Chinese government outlines its vision for sectoral industrial policies in its Five-Year Plans. These plans reflect the priorities of the central government, and provide guidance for policy-makers at all levels of government. The solar industry has occupied an increasingly prominent position in the Five-Year Plans, beginning with the Tenth Five-Year Plan (2001-2005), where it received only a brief mention[3]. By contrast, the Eleventh Five-Year Plan (2006-

[3] "Actively developing new and renewable energy sources such as wind power, solar power, and geothermal energy. Promoting energy conservation and comprehensive utilization technologies."

2010) featured an increased emphasis on R&D and mentioned funding for solar manufacturing and innovation for the first time. This was further built on in the Twelfth (2011-2015) and Thirteenth (2016-2020) Five-Year Plans –the later of which was accompanied by a specific Solar Energy Development Plan issued by China's National Energy Administration.

Figure 2 traces the evolution of the Chinese solar industry over this time period, using our data on the universe of solar manufacturers in China (see Section 3). Panel A shows that solar patents went through a revolution, rising from a few hundred in 2004 to over 10,000 in 2020. Panel D shows that there was near zero panel production capacity in 2004, but this rose to about 70,000 MWh by 2013. Similarly, Panel C shows that revenues of solar firms rose from close to zero to over 700 billion Yuan by 2019 spread across over 1,500 firms (Panel B).[4]

2.2 Policy Support toward Solar Manufacturing

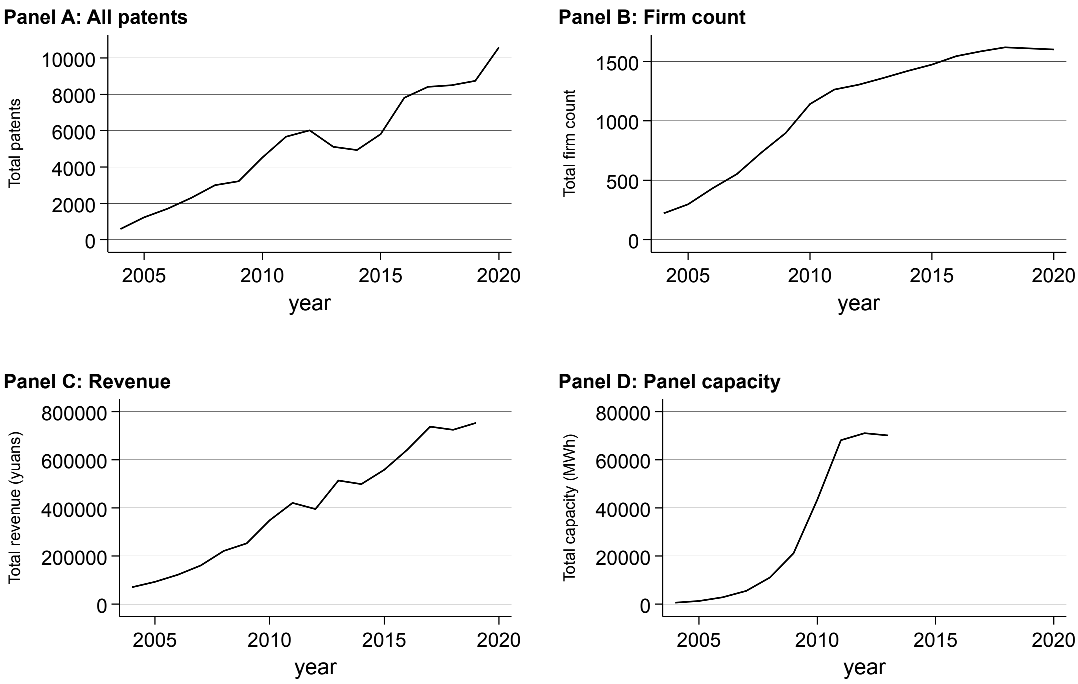
Whilst the Five-Year Plans provide national policy guidance, the power to implement industrial policy is dispersed across different levels of government in China, resulting in considerable uncertainty about the extent and nature of policy support for the industry.

Qualitative research by Ball et al. (2017) based on interviews with government officials, business leaders, managers and academics, provides some clarity. Their work suggests that subsidies to solar manufacturing were managed and allocated by local governments, following the national guidance embedded in the Five-Year Plans. Since at least 2006, many city bureaucrats have competed to build up solar manufacturers, offering tax incentives, discounts for land acquisitions, and cash investments. Bai, Hsieh, and Song (2020) give a rich description of the local policy landscape in China. City bureaucrats have strong administrative competence and compete aggressively on offering special deals to private firms.

Ball et al. (2017) estimate a lower bound of around $300 million spending in solar R&D subsidies by national and local governments between 2001 and 2015, compared to $1.44 billion of private R&D solar spending. They also find evidence of considerable regional heterogeneity in solar R&D funding.

[4] It might seem surprising that the number and revenues of solar firms can be non-zero in 2004 when production capacity is zero. This is mainly because because some solar firms are multi-product (there are also some firms who produce non-solar PV such as cells, but this is a small number). Hence, they may earn revenues on non-Solar products and services. To check this does not change the results, we set to zero firm counts and revenues if solar production capacity was zero. The results were essentially unaltered. Since we can only do this exercise pre-2014 due to the ENF data constraint, we prefer to use a consistent approach in our baseline and not impose this additional restriction.

Figure 2: Chinese solar manufacturers activity over time



Notes: Time series for total number of patents filed by solar firms at the SIPO; firm count obtained from the Chinese firm registration platform; revenue obtained from Orbis and panel capacity, obtained from ENF Market Research reports. The sample is our near-universe of solar panel manufacturers in China, obtained from ENF's register.

2.3 Measuring Solar Industrial Policy

This still leaves open the question of how we can identify and measure the dispersed implementation of industrial policies - particularly when there may be no one organisation who has oversight of the full range of implemented policies. Given these, and other challenges, some researchers have relied on model-based approaches to detect industrial policy subsidies (Kalouptsidi, 2018).

In this paper, we follow an approach based on text analysis of policy documents, similar to that of Juhász et al. (2022)[5]. We extract data on industrial policy towards solar manufacturing, innovation, and installation from PKULaw's Laws & Regulations dataset[6]. The Laws & Regulations database is a comprehensive and reliable source of China's legal information, including all laws, regulations, and any related legal information implemented by the central

[5]We manually inspect and classify all solar policies, while Juhász et al. (2022) use an automated classification algorithm

[6]https://www.pkulaw.com/law/

and local governments since 1949. We obtain data disaggregated by industry and gather all regulations pertaining to solar photovoltaics. The first sub-national solar policy we identified was in 2006.

The dataset contains information on the title, validity, administrative level, department, release date, and implementation date of each policy. We additionally scrape the original policy documents, which contain the text of each regulation or announcement. We manually inspect the full text of each policy and classify them into types. We focus on subsidy policies where there is direct financial support. We further disaggregate subsidy policies according to whether they target demand (solar installation), production and/or innovation.

Table 1 illustrates the criteria we follow to classify policy documents and provides examples of key text extracts that guided the classification. The table shows that there have been 78 subsidy policies in total (sometimes a policy is a bundle of demand and supply subsidy policies which is why the sum of the disaggregated policy numbers exceeds 78). City-level demand subsidies are the most common - there are 61 of these since 2006. There are 27 production subsidy policies, 12 of which also contain innovation subsidies. We did not find any standalone city-level innovation policies. So when we compare across policies we are implictly comparing standalone demand or production policies to a bundle of a production and innovation policy. We return to this when interpreting the empirical results.

Figure 3 presents the time series of city subsidy policies. As we are unable to accurately measure the end date of policies, this captures the cumulative number of cities which have at some point implemented a solar industrial policy. The total policy line (hollow circles) shows that there was a steady increase in the number of cities using solar policies since 2006. In 2010 only 10 cities had solar policies, but this had reached 18 by 2013. There was then almost a doubling to 32 in 2014, driven by an increase in demand policies (hollow diamonds). The number of cities with policies levels off after 2017, finishing at 43 (out of 358 cities across China) by 2022.

Interestingly, the solid dots show that early policies were production subsidies (usually bundled with innovation policies) - the first demand policy only began in 2010. By the end of our sample period 19 cities had production subsidies and 10 also had innovation subsidies. By contrast, a full 30 cities had demand subsidies.

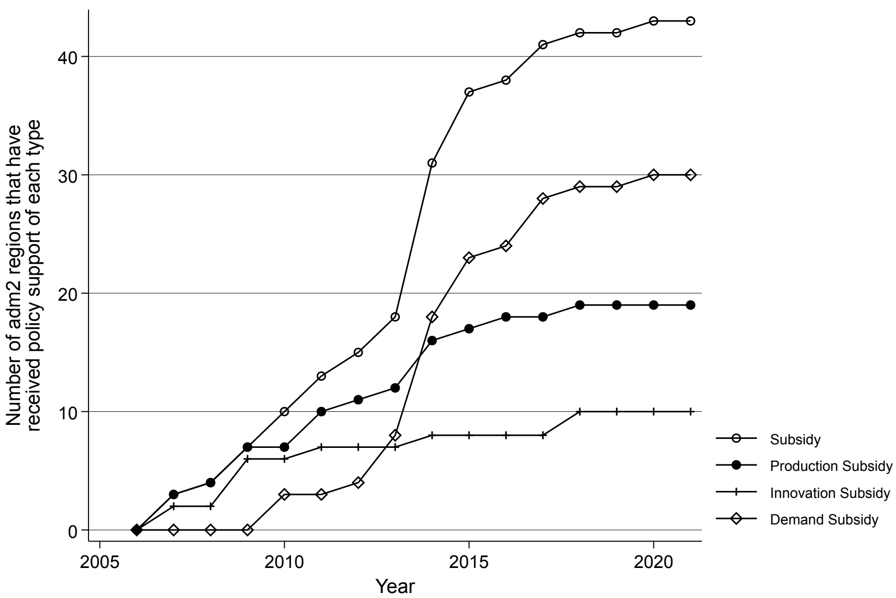
Note that in our PKULaw dataset we cannot accurately identify the end date of policies. Therefore, the main signal that we exploit in the data is the staggered adoption of policies over time (and a comparison with those cities who do not). Despite this caveat, most cities report that

Table 1: City-level solar policies

Type of policy	*Number*	*key feature*	*Example*
Subsidy	78	Policy text contains precise information on the size of the subsidy	
1. Production subsidy	27	Subsidises solar production	*"A new solar production line built in Hefei will be subsidized by 12% (2018)"*
2. Innovation subsidy	12	Subsidises solar innovation	*"Firms will be awarded 10,000 RMB if they earn provincial level R&D center certification (Guilin, 2011)"*
3. Demand subsidy	61	Subsidises the installation of solar panels	*"1 RMB per watt for the electricity generated by solar projects installed in Beijing (2010)"*

Note: All policies are at the city (admin2 region) level over the 2006-2022 period. There are 358 cities. 43 cities are treated by some subsidy by the end of our sample.

Figure 3: Number of cities treated with Solar supply & demand subsidy policies



Note: All policies are at the admin2 region level. There are in total 358 admin2 regions in China (we remove Taiwan, Hongkong and Macao from the analysis). The time series for 'Subsidy' includes any of demand, production, or innovation subsidies.

their policies were still in place when this information was last updated. Moreover, cities often implement multiple policies, so its unlikely a city completely abandons policy support towards solar in our data. Hence, our treatment of policy support as an absorbing state.

Our approach complements recent attempts to provide micro-level quantitative measures of industrial policy in China. Chen and Xie (2019) use the Chinese Law and Regulation Database, which is a subset of our PKULaw dataset, to provide a micro-level measure of the number of industrial policies at the Chinese prefectural-city level. We extend these recent approaches by using the full universe of laws and regulations targeting the solar industry and by analyzing policy documents to identify targeting (production, innovation, or installation) as well as spatial heterogeneity. In addition, our manual text classification allows us to carefully distinguish financial support in the form of subsidies from solar industry announcements, records, or other type of policies that do not include explicit support to manufacturers.

3 Data

This section provides an overview of our data set on the Chinese solar industry. More details are in Appendix B.

We gather a variety of firm-level outcome data that we aggregate at the city-level and combine with our policy support indicators. First, we construct a sample of solar manufacturers over time using the historical directories of solar panel producers from ENF Solar Industry Directory, available from 2010 until 2021 (henceforth, *ENF register*). The ENF Solar Industry Directory is a register of 50,800 worldwide photovoltaic companies. Because it is the leading solar website, most companies self-register on ENF's platform. ENF additionally reviews daily news regarding the solar industry, as well as available lists of key solar exhibitions, to incorporate the remaining new solar companies. It also relies on government organisations and a variety of web-searching techniques to complete the full list of firms. To detect firm exit, ENF uses automatic scanning of company updates, which triggers careful checks from ENF database experts to update manufacturers' information, or to report firms as ceasing their activities. Hence, ENF is able to reasonably capture a snapshot of all solar panel manufacturers each year.

We obtain our first measure of panel and cell production from the last edition of ENF's Chinese Cell & Panel Manufacturers Report. This dataset (henceforth, *ENF production*) allows us to measure, for each firm, their production and capacity figures (in MWh) for both solar panels and solar cells across the 2004-2013 period. We match *ENF production* and *ENF register* for their overlapping period based on firm name and extensive contact details information (address, phone, website, fax, and email). Together, we are left with a sample of 1,718 Chinese solar panel manufacturers, operating at some point between 2004 and 2021, which includes production and capacity data for each manufacturer during the 2004-2013 period. Both ENF datasets contain detailed address information, which allows us to geo-locate all firms through the Google and Baidu APIs, and assign them their corresponding city. In order to expand the time horizon of our analysis and estimate long-run effects on production beyond 2013, we use Bureau Van Dijk's Orbis dataset, which gives us rich financial data, including total assets, revenue, employees, and cost of goods sold, throughout the 2004-2019 period. We use the comprehensive firm contact information included in both the Orbis and ENF register datasets to merge the two datasets, and obtain Orbis variables for our sample of solar manufacturers. We aggregate all production, capacity, and revenue figures from ENF cell and panel manufacturers at the city-level.

We validate the entry and exit information on ENF register using the Chinese firm registration information, which we access through the Qichacha platform (https://www.qcc.com/). The Qichacha platform gathers detailed firm-level information, spanning from registration to exit, and is updated periodically following government requirements. It collects this information from multiple data sources, but mostly relies on government's official sources, which include the National Enterprise Credit Information Publicity System, the China Court Judgment Documents Network, and the China Enforcement Information Disclosure Network. This allows us to obtain the number of operating solar manufacturers in each city and year.

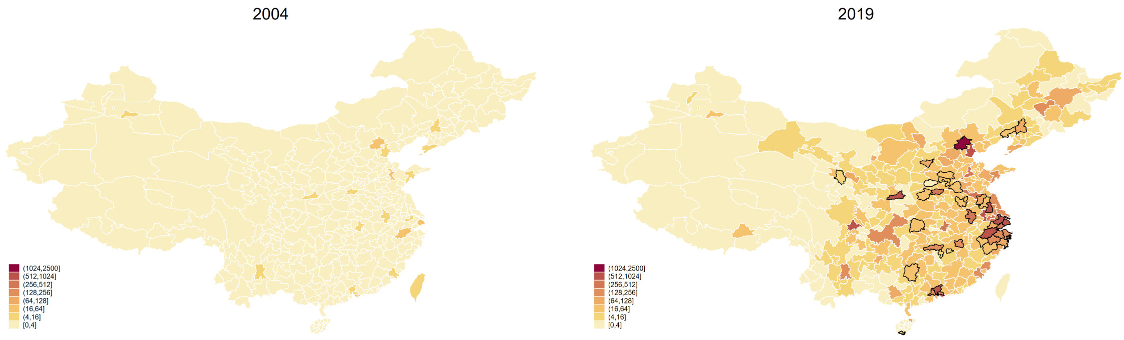
The Qichacha platform also contains detailed intellectual property information from the former State Intellectual Property Office (SIPO). We extract, for each ENF manufacturer, the name, patent ID, type, application date, publication date, and assignee, of the patents it has filed. We then use the SIPO patent ID to extract IPC codes and patent abstracts from the PATSTAT database. To understand the nature of the underlying innovation, we classify the patents filed by our sample of manufacturers into several categories. First, we rely on the SIPO classification of patents into Invention, Utility Model and Design patents[7]. Invention patents

[7] An invention patent refers to a new technical solution or improvement for a product or method. Unlike utility model patents, which are restricted to products, invention patents can apply to both products and methods. The protection period for invention patents is the longest in the domestic patent classification, up to 20 years. A utility model patent involves a new technical and practical solution regarding the shape, structure or combination of a product or products. The protection period for utility model patents is of 10 years. Design Patents

have longer protection periods, require paying higher filing costs, and involve a more cumbersome administrative process. They are therefore generally of higher quality and a more innovative nature[8]. Second, using IPC codes, we further classify invention and utility model patents into solar and non-solar patents[9]. Again, we aggregate the patents filed by our sample of ENF manufacturers to the city-level.

The maps in Figure 4 illustrate the spatial variation in solar patents and subsidies across Chinese cities, which we exploit in our empirical analysis. What is striking is that we capture the *whole* of the development of the solar industry in China from beginning to the present. In 2004 (map on the left), the starting year of our analysis, there was very little innovation in solar and no policy support. On the other hand, by 2019 (map on the right), we observe a total of 43 cities whose solar industry has been subsidised, and crucially, innovative activity skyrocketed across the country. 82 of the 358 city-regions in China had some patenting by solar firms in 2019, compared to only 25 in 2004.

Figure 4: Number of solar patents in each city and subsidy policy



Note: Each white-bordered region represents an admn2 level city region. Black circled cities are treated by any subsidy policy. We use a heat map scale, where cities colored in a stronger red are filing more solar patents

Finally, in order to explore if policy support encouraged learning by doing, we also use text mining techniques to classify invention and utility model patents based on the text in patent abstracts. We classify patents into learning-by-doing patents (those that include a productivity-increasing process innovation) and non-learning-by-doing patents (which often reflect either the invention of new products or basic science research around the chemistry and material

cover product improvements of an aesthetic nature, which are suitable for industrial application. Broadly, all original designs around a product's appearance could apply for a design patent. The protection period of design patents is of 15 years.

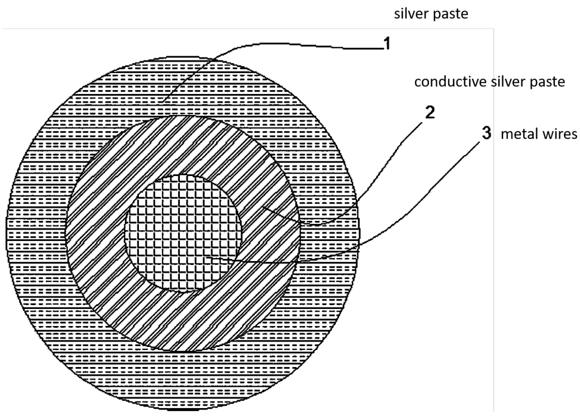
[8]The firms in our solar manufacturers dataset file mostly invention and utility model patents. Only 10% are design patents

[9]We follow the categorisation into solar developed in Shubbak (2019).

science that informs the production of wafers and cells). To do so, we build on the work of Liu (2023), who classified a sample of Chinese solar patents into process and product related innovations by hand-reading the full patent text. Leveraging his work as a training dataset, we classify the remaining patents in our much larger sample using machine-learning techniques.

Figure 5 shows an example of a learning-by-doing patent. The patent abstract alludes to reducing production costs, minimizing production errors, and being suitable for mass production compared to prior art. In the data Appendix B, we provide additional learning-by-doing and non-learning-by-doing patent examples and their abstracts.

Figure 5: Learning by doing patent example



Patent Abstract: "The present invention discloses a grid line structure for a solar cell, which comprises metal wires, conductive silver paste and silver paste. The grid line is woven from metal wires, with a layer of silver paste applied to the metal wires and then a layer of silver paste, which ensures excellent adhesion between the silver paste and the metal wires and ensures good ohmic contact between the sub-grid line and the silicon wafer. The silver paste used for the main grid line does not contain glass material, which ensures good adhesion between the main grid line and the silicon wafer and reduces the recombination of minority carriers under the main grid line. *Compared with the prior art, the present invention greatly reduces the amount of silver paste used, thus saving more expensive silver paste, effectively reducing production costs, and ensuring excellent aspect ratios of the grid lines, eliminating the possibility of broken lines and false prints, thereby improving the photovoltaic conversion efficiency of the solar cell, and being suitable for large-scale industrial production*"

4 Model

To provide an intuition for the impact of place-based industrial policies on the evolution of the solar industry, we develop a model of electricity demand, production of power-plant components (such as solar panels), exporting, and innovation in China. The model builds on Bustos (2011) and Shapiro and Walker (2018). More details and derivations can be found in Appendix C.

The model features multiple regions which we index by d (for "destination") when referring to the region that is generating or consuming electricity and o when referring to the region that is producing the components of power plants. Each region d has a representative consumer who demands electricity services e^d. To satisfy this demand, a grid-planner in each region builds and runs power plants of different types s, combining their output to supply final electricity services. To simplify, we assume that there is no trade in electricity across regions. To build power plants, the grid planner purchases (differentiated) sector-specific power plant manufactured components (e.g. solar panels) from producers in all origin regions o. Manufacturers of power plant components such as solar PV producers have heterogeneous productivity and make decisions about entry, exit, production, and international exporting. They also have the opportunity to innovate or 'upgrade their technology' –increase their productivity for a fixed cost.

By using a heterogeneous firm framework, we are able to get predictions for the impact of subsidies on a range of industry outcomes, including firm numbers, output, and exports and innovation. By incorporating internal economic geography (a key feature that distinguishes our model from approaches such as Bustos (2011)) we can generate separate predictions for the impact of subsidies targeting demand (installation of solar panels or use of solar panels to generate electricity) as opposed to the supply side (production or innovation). For example, increased use of solar due to demand subsidies in one city can be met by increased supply from other cities (subject to transport costs). Since there are no 'local content' requirements, this mutes the effect of demand subsidies on the growth of solar production in the same city.

We make several key simplifications in the model. First, electricity services are only used for final consumption, not as an input to production. Second, we abstract from the production, trade and consumption of goods and services other than electricity. Third, we take as exogenous the demand function for electricity in each city. We therefore abstract from the response of population or industrial production to changes in energy prices across space. Fourth, power plants are not durable. We make these simplification to focus on the problem of a solar panel producer, since this is the centre of our empirical analysis and where we have the richest data.

4.1 The Grid Planner Problem

In each region d, there is a representative consumer that obtains utility from electricity services e. We abstract from consumption of other goods and services for simplicity, as our focus in on the production of varieties of power plant components.

$$U_d = u\left(e_d\right) \tag{1}$$

This representative consumer provides L_d effective units of labour inelastically. L_d reflects both the population of region d and the human capital of the local population. We consider L_d to be, for now, exogenous to the model - implying that we abstract from the migratory response of labour to changes in electricity prices.

Electricity services in region d are the combined electrical output of many power plants of different energy sectors which are built and operated by a regional 'grid-planner'. Given that we focus on subsidies to the solar industry which do not apply to other electricity sectors, we assume these sectors to be solar s and non-solar s'. The electrical output of solar ($e_{d,s}$) and non-solar ($e_{d,s'}$) plants are combined using a CES technology.

$$e_d = \left(\kappa_{d,s'} e_{d,s'}^{\rho} + \kappa_{d,s} e_{d,s}^{\rho}\right)^{1/\rho} \tag{2}$$

This assumption captures in a reduced-form-way the differential timing and flexibility of generation from different electricity sources. For example, solar is not a perfect substitute for coal because it only provides electricity whilst the sun is up. Coal on the other hand, can produce electricity all day and therefore meet nighttime demand. $\kappa_{d,\text{sector}}$ allows for the overall productivity of different electricity sources to differ across regions, capturing differences in solar potential or mineral resources across regions. Note that the grid-planner can only use power plants located within their own region d and cannot trade electricity across regions.

To generate electricity in each sector $e_{d,s}$, the 'grid planner' in region d combines sector-specific intermediate manufactured inputs using a CES aggregator.

$$e_{d,s} = \left(\sum_{o} \int_{\omega \in \Omega_{o,s}} q_{od,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega\right)^{\frac{\sigma_s}{\sigma_s - 1}} \tag{3}$$

Here, $q_{od,s}(\omega)$ denotes the quantity of the variety ω of goods in sector s from region o used by the grid-planner in region d. Note that each variety ω is produced in a single region only and by a single firm. The elasticity of substitution across varieties is represented by the sector-specific parameter σ_s, whereas σ represents the elasticity of substitution across different energy sectors.

We conceive this CES aggregate of intermediate inputs as a 'power plant'. For example, in order to install a solar power system providing solar electricity services, the planner combines different varieties of solar cells or panels, a racking system, and the necessary associated power inverters, charge controllers, batteries, and wiring as needed. These intermediate inputs can be sourced from all regions of the country. Despite our grid-planner 'building' power plants, we abstract from their durable nature.

The grid-planner chooses the overall quantity of electricity services to produce, the mix of solar and non-solar electricity, and the mix of intermediate inputs for each energy sector in order to maximise their profits taking as given the price of final electricity in their region (p_d) and the price of all intermediate inputs. In practice, since utility depends only on electricity services (and is strictly increasing in e_d), and as production of final electricity services is constant returns to scale, we can equivalently solve this as a problem in which the grid-planner supplies as much electricity as possible in the minimal cost way given the income of the representative household I_d.

With this framing, the grid-planner problem can be solved in two stages. First, the grid-planner chooses the mix of intermediates (e.g. varieties of solar panel) in order to minimise the cost of generating a given level of electricity in sector s.

$$\min_{q_{o,d}(\omega)} \left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{od,s}(\omega) p_{od,s}(\omega) \right)$$
$$\text{s.t.} \quad \left(\sum_o \int_{\omega \in \Omega_{od,s}} q_{od,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{\sigma_s}{\sigma_s - 1}} = e_{d,s}$$

This yields the price (and the optimal intermediate inputs mix) of generating one unit of electricity using solar and non-solar electricity in region d: $P_{d,s}$ and $P_{d,s'}$.

Given these price indices, the grid-planner chooses a mix of solar and non-solar eletricity generation to maximise electricity output.

$$\max_{e_{d,s}, e_{d,s'}} \left(\kappa_{d,s'} e_{d,s'}^{\rho} + \kappa_{d,s} e_{d,s}^{\rho} \right)^{1/\rho}$$
$$\text{s.t. } P_{d,s} e_{d,s} + P_{d,s'} e_{d,s'} = I_d$$

Solving this nested CES problem yields the following demand for each intermediate input variety:

$$q_{od,s}(\omega) = \left(\frac{p_{od,s}(\omega)}{P_{d,s}} \right)^{-\sigma_s} \left(\frac{\kappa_{d,s}}{P_{d,s}} \right)^{\sigma} \frac{I_d}{\kappa_{d,s'}^{\sigma} P_{d,s'}^{1-\sigma} + \kappa_{d,s}^{\sigma} P_{d,s}^{1-\sigma}} \tag{4}$$

4.1.1 Demand-side industrial policy

We model *demand* subsidies targeting sector s as a policy which increases the productivity of solar as a means of producing electricity - that is, a factor θ_d which multiplies $\kappa_{s,d}$ in the grid-planner problem. This captures in a reduced form way policies such as feed-in tariffs which guarantee a higher per-unit electricity price for solar power plant operators when selling their electricity to the grid. Demand for each variety becomes:

$$q_{od,s}(\omega) = \left(\frac{p_{od,s}(\omega)}{P_{d,s}}\right)^{-\sigma_s} \left(\frac{\theta_d \kappa_{d,s}}{P_{d,s}}\right)^{\sigma} \frac{I_d}{\kappa^{\sigma}_{d,s'} P^{1-\sigma}_{d,s'} + \theta_d \kappa^{\sigma}_{d,s} P^{1-\sigma}_{d,s}} \tag{5}$$

A solar demand subsidy in location d will lead to a shift in the composition of electricity production towards solar. This, in turn, will result in an increase in the demand for solar intermediate inputs in location d. Assuming that all prices are fixed (i.e. considering just the partial equilibrium response) this will lead to the same proportionate increase in region d demand for all solar intermediate varieties. Demand subsidies implemented in region d therefore influence demand for varieties produced in all regions of China, indirectly impacting firm decisions.

4.2 The Manufacturer Problem

Intermediate inputs for each sector are produced by firms in different regions of China. Each region o has a continuum of potential manufacturing firms i in each sector s, which operate under monopolistic competition.

4.2.1 Manufacturing Technology

Firm i, who produces intermediate goods for electricity sector s (e.g. solar panels for the solar electricity sector), uses effective units of labour $L_{o,s,i}$, with unit cost $w_{o,s}$. Firm subscripts i are dropped from now onwards for notational simplicity. To operate, a firm must pay a sunk cost $w_{o,s} f^e_{o,s}$, which we express in terms of effective units of labour[10]. Upon paying the entry cost, the firm draws an initial level of productivity φ, from a Pareto productivity distribution, whose cumulative distribution function is:

$$G\left(\varphi; b_{o,s}\right) = 1 - \left(\frac{\varphi}{b_{o,s}}\right)^{-\theta_s}$$

[10]This sunk cost could be understood as the cost incurred in initial product definition and development.

Each firm in operation produces a differentiated variety of sector-specific intermediate good. We therefore equivalently index firms by either their productivity φ or the variety they produce ω. To produce $q_{o,s}(\varphi)$ units of a variety, a firm requires an amount of effective labor

$$l_{o,s} = f_{o,s} + \frac{q_{o,s}}{\varphi}$$

where $f_{o,s}$ is the fixed cost of production, expressed in terms of effective units of labour, and $\frac{1}{\varphi}$ is the marginal cost of production.

4.2.2 Innovation

After observing its initial productivity φ, a firm can choose to upgrade its technology (innovate), which increases the fixed cost of production: $\eta_{o,s} f_{o,s}$, with $\eta_{o,s} > 1$, but reduces its marginal cost, now: $\frac{1}{\xi_{o,s}\varphi}$, with $\xi_{o,s} > 1$

4.2.3 Exporting

In addition to innovation, the firm also can choose which regions to sell to. The d regions consiste of many Chinese second administrative level regions ('cities') and one foreign region which we index by $\tilde{d}$.

We assume there are no fixed costs of trade within China. On the other hand, a firm must pay an international exporting fixed cost $w_{o,s} f^x_{o,\tilde{d},s}$ if it wants to sell to the foreign market.

Trade (intra-China and international) is subject to iceberg trade costs such that in order for $q_{od,s}(\varphi)$ to arrive to destination d, a firm in o needs to produce $\tau_{od,s} q_{od,s}(\varphi)$ units of the variety, with $\tau_{od,s} \geq 1$. Trade costs are normalised, such that they are equal to 1 if and only if $d = o$.

4.2.4 Supply-side industrial policy

Firms are directly or indirectly subject to three different types of subsidies which are set by local government: production, innovation and demand.

We model *production* subsidies targeting sector s as a reduction in input costs for targeted firms, whose marginal cost becomes $\frac{1-s_{o,s}}{\xi_{o,s}\varphi}$. This captures the sort of production subsidy given as an example in Table 1, where the entire run of a product is subsidised by a given percentage.

We model the *innovation* subsidy targeting sector s as a reduction in the fixed cost involved in technological upgrading. The fixed costs for an innovator facing a subsidy are $\phi_{o,s}\eta_{o,s} f_{o,s}$, with $\phi_{o,s} < 1$ (and $\eta_{o,s} > 1$). This corresponds to the example innovation subsidy in Table 1, where firms are given a fixed payment for incurring a fixed cost (establishing an R&D centre).

4.2.5 Firm profits

We now derive an expression for firm profits after paying the entry cost and drawing the productivity. This comes from combining the firm technology and industrial policy as described above. Since profits depend on the exporting and innovation decision of the firm, their profits here are a maximum over three alternative profits - if they neither innovate or export, if they export but don't innovate, and if they both export and innovate. We ignore the case in which the firm innovates but does not export as this appears not to be an empirically relevant case.

$$
\begin{aligned}
\pi_{o,s}(\varphi) = \max\Bigg\{ & \sum_{d \neq \tilde{d}} \left\{ p_{od,s}(\varphi) q_{od,s}(\varphi) - w_{o,s} \frac{\tau_{od,s}(1-s_{o,s}) q_{od,s}(\varphi)}{\varphi} \right\} - w_{o,s} f_{o,s}, \\
& \sum_{d} \left\{ p_{od,s}(\varphi) q_{od,s}(\varphi) - w_{o,s} \frac{\tau_{od,s}(1-s_{o,s}) q_{od,s}(\varphi)}{\varphi} - \mathbb{1}[d = \tilde{d}] \left(w_{o,s} f^{x}_{o,d',s} \right) \right\} - w_{o,s} f_{o,s}, \\
& \sum_{d} \left\{ p_{od,s}(\varphi) q_{od,s}(\varphi) - w_{o,s} \frac{\tau_{od,s}(1-s_{o,s}) q_{od,s}(\varphi)}{\xi_{o,s}\varphi} - \mathbb{1}[d = \tilde{d}] \left(w_{o,s} f^{x}_{o,d',s} \right) \right\} - w_{o,s} \phi_{o,s} \eta_{o,s} f_{o,s} \Bigg\}
\end{aligned}
$$

In each case, profits are given by revenues in each destination minus the costs of production, which consist of marginal costs, the fixed cost of exporting and the fixed cost of production. Marginal costs depend on productivity, the production subsidy, trade costs, productivity and innovation decisions. The fixed cost of production depends on the innovation subsidy and the innovation decision.

4.2.6 Firms optimal choices

Firms maximise profits by making decisions about price, which regions to export to, whether to innovate, whether to exit and whether to enter in the first place. To solve, we can break this problem down into stages.

Given the firm's choice over innovation and exporting, we can solve for the firms optimal price by taking the FOC of firm profits with respect to $p_{od,s}(\varphi)$, and replacing the optimal $q_{od,s}(\omega)$ above, we obtain the usual result that manufacturers price as a constant markup over marginal costs. For example, if the firm is exporting and innovating price would be given by:

$$
p_{od,s}(\varphi) = \frac{\sigma_s}{\sigma_s - 1} \frac{w_{o,s} \tau_{od,s}(1 - s_{o,s})}{\xi_{o,s}\varphi} \tag{6}
$$

Replacing the optimal pricing and demand functions in the expression for firm profits, we can obtain the potential value functions for each technology and exporting choice (full details in Appendix C). Optimal profits are therefore:

$$\Pi_{o,s}(\varphi) = \max\left\{ \sum_{d\neq\tilde{d}} \left\{ \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}\left(1-s_{o,s}\right)}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s}f_{o,s}, \right.$$

$$\sum_{d} \left\{ \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}\left(1-s_{o,s}\right)}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s}f^{x}_{o,\tilde{d},s} - w_{o,s}f_{o,s},$$

$$\left. \sum_{d} \left\{ \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}\left(1-s_{o,s}\right)}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s}f^{x}_{o,\tilde{d},s} - w_{o,s}\phi_{o,s}\eta_{o,s}f_{o,s} \right\}$$

Note that the demand subsidies show up in this expression through $E_{d,s}$, which depends on the demand subsidy in region d.

Given that they price optimally, firms make exporting and innovation decisions to maximise this expression. This results in a set productivity cut-offs which determine whether a firm will i) stay in the market after drawing a productivity, ii) export to the international market $\tilde{d}$, and iii) innovate.

Domestic market exit threshold: We define $\varphi^{*}_{oo,s}$ as the domestic market exit productivity threshold. This is the productivity that generates zero profits from serving the domestic market only.

$$\varphi^{*}_{oo,s} = \left(1-s_{o,s}\right) \left(\sum_{d\neq\tilde{d}} \left\{ \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{w_{o,s}f_{o,s}} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d,s}} \right)^{1-\sigma_s} \right\} \right)^{\frac{1}{1-\sigma_s}}$$

Exporting threshold: Let $\varphi^{*}_{o\tilde{d},s}$ describe the productivity level which makes a firm earn zero profits from exporting to foreign country $\tilde{d}$, and therefore indifferent between serving $\tilde{d}$ or limiting its supply to the domestic market. We also assume that the marginal exporting firm is not innovating.

$$\varphi^{*}_{o\tilde{d},s} = \frac{\tau_{o\tilde{d},s}\left(1-s_{o,s}\right)}{P_{\tilde{d},s}} \left(\frac{E_{\tilde{d},s}}{f^{x}_{o,\tilde{d},s}} \frac{(\sigma_s-1)^{\sigma_s-1}}{w_{o,s}^{\sigma_s}\sigma_s^{\sigma_s}} \right)^{\frac{1}{1-\sigma_s}}$$

Innovation threshold: Let $\varphi^{i}_{od,s}$ be the productivity level which makes a firm indifferent between upgrading its technology or not. This is given by:

$$\varphi^{i}_{od,s} = \left(1-s_{o,s}\right) \left(\sum_{d} \frac{1-\xi_{o,s}^{1-\sigma_s}}{\xi_{o,s}^{1-\sigma_s}} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{w_{o,s}\phi_{o,s}\eta_{o,s}f_{o,s}} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d,s}} \right)^{1-\sigma_s} \right)^{\frac{1}{1-\sigma_s}}$$

Next, given optimal profits they can make if they continue to operate, firms choose whether to pay the fixed cost of production or to exit. Finally, with knowledge of their optimal policies and the distribution of productivity, firms choose whether to pay a fixed cost to draw a productivity or to stay out of the market.

4.3 Industry Equilibrium

An equilibrium in this model is characterised by the following:

1. Households maximise utility
2. The grid-planner minimises cost
3. Firms maximise profits. Implying that they will price according to the pricing formula above, make exit, export and innovation decisions according to the thresholds above, and enter if the expected profits are weakly greated than the fixed cost of entry.
4. Free entry/Zero expected profits: the sunk cost of entry equals the expected profits from entry.

$$w_o f^e_{o,s} = \left(1 - G\left[\varphi^*_{oo,s}\right]\right) \mathbb{E}\left[\pi \mid \varphi > \varphi^*_{oo,s}\right] \tag{7}$$

5. Final electricity service market clears
6. Power plant component market clears
7. Labour market clears: Each region d allocates its labor supply L_d to paying all the fixed and marginal costs of production.

To determine the equilibrium price indices, number of firms, aggregate production and revenue, and mass of exporters and innovators in each region, we impose all of the above conditions with the exemption of the labour market clearing condition.

4.4 Comparative Statics

We are interested in how key features of the solar industry respond to changes in industrial policy. We consider a number of comparative static predictions of the model with respect to policy parameters. We consider three types of subsidies: production, demand, and innovation and several outcomes: production, number of solar firms, exports and innovation. We consider perturbing a single policy in a single city, holding the policies of all other cities fixed.

Full derivation of the comparative statics is still an ongoing work (work-in-progress derivations are available upon request). Here, we present our initial understanding of the direction of these comparative statics and the mechanisms through which industrial policy influences outcomes through the lens of the model.

Production Subsidies, θ:

Producers in the subsidising region will have lower marginal costs and, therefore, charge lower prices for their goods. With lower prices being charged, overall quantities increase, which we would expect to show up in the data as higher overall production in this region.

Profits increase and thus the threshold for exit is reduced. The firms that grow larger through the subsidies will be more likely to pay the fixed costs of innovating and exporting (the threshold for exporting and innovation are reduced). Hence, we would expect to see higher firm count, production, exports, and innovation in the subsidising region. Note that these effects are limited to the region in which the subsidy is implemented. More solar panels are produced in region o and shipped across China for use in electricity production in any region d.

Innovation Subsidies, θ:

The impact of innovation subsidies is more straightforward. By reducing the fixed cost of technology upgrading, innovation subsidies reduce the productivity cut-off for innovation, implying that more firms will innovate, and that, on average, less productive firms will engage in innovation.

Note that we assume that innovation subsidies are not too large, or that $\phi_{o,s}$ is not too small, so that $\varphi^*_{o\tilde{d},s} < \varphi^i_{od,s}$. That is, the marginal innovator is already exporting. Thus, the exit and international exporting thresholds remain unchanged.

The subsidizing region will also face lower effective prices for their goods as a result of more firms crossing the innovation threshold and reducing their variable costs.

Demand Subsidies, θ:

The effects of demand subsidies are more nuanced. As explained above, we model demand subsidies targeting sector s as a policy which increases the productivity of solar as a means of producing electricity. As there is no trade in electricity, this will increase demand for solar intermediate inputs produced across china in the implementing region.

Higher demand for solar intermediates will increase expected profits for firms, and therefore (holding prices and wages fixed) lead to higher production, firm count and innovation across China. The relative impact on the solar activity in implementing region d compared to other regions that did not implement the demand subsidy will depend on how much demand for intermediates in region o increases as a result of the demand subsidy in d. This will be determined by several factors.

First, trade costs. In the extreme case of infinite trade costs within China, the demand effects of the subsidy in region d will fall entirely on firms in this same region. Starting from this extreme and reducing trade costs will result in a more spatially diffuse effect of the policy. If trade costs are low (or the city where the subsidies are implemented has little comparative advantage in solar production), then there may be little or no effect on local outcomes.

Second, the initial share of region d demand for region o solar intermediates in the total production of solar intermediates in region o. Regions where d is a relatively large customer of its solar intermediate products will face a larger proportionate increase in demand for their goods, and therefore we would expect to see relatively large impacts on production, firm count and innovation.

Third, the possible presence of local content restrictions. If local content rules stipulated that increased local solar energy generation had to be met from locally produced solar panels.

In terms of all our outcomes, we expect a strict hierarchy of positive responses. For innovation, we anticipate the largest impact to come from direct innovation policies, followed by production subsidies, and the smallest impact from demand policies. For output, exports, and firm numbers, we anticipate the largest positive effects to result from production subsidies, followed by innovation subsidies, and the smallest from demand policies. Table 2 illustrates these predictions. A 'prediction' in this table represents the sign and magnitude of a potential treatment effect for each type of policy and outcome variable, demonstrating the outlined hierarchy of responses.

As described earlier, city-level innovation subsidies in China are always bundled with production subsidies. Therefore, in the last column of Table 2, we include predictions of the impact this bundle would have on solar industry outcomes. Due to their combined advantages of production and innovation subsidies, we predict that they would have the largest impacts on innovation, firm counts, revenue, production, and exports in city-level solar industries.

Upon examining our empirical results, we will find that they broadly support all of these predictions from the model.

Table 2: Summary of Model Predictions

	Demand Subsidy θ_o	*Production Subsidy* s_o	*Innovation Subsidy* ϕ_o	*Production & Innovation Subsidy* $s_o + \phi_o$
Innovation$_o$	$\approx +$	+	++	+++
Firm count$_o$	$\approx +$	++	+	+++
Panel production$_o$	$\approx +$	++	+	+++
Revenue$_o$	$\approx +$	++	+	+++
Exports$_o$	$\approx +$	++	+	+++

Notes: All outcome variables and subsidy policies are referred to the same region o. The table shows no prediction on how policies in region d affect outcomes in region o. A 'prediction' in this table represents the sign and magnitude of a potential treatment effect for each type of policy and outcome variable. That is, we are predicting the relative difference between treated and untreated regions. The last column corresponds to the type of innovation subsidies that we observe in the data, which are always implemented together with some policy support towards production. $\approx +$ indicates that we expect effects to be plausibly positive but there is some ambiguity in their sign. We rank unambiguously positive effects to provide qualitative intuition on the effectiveness of different policies on improving each outcome. The ranking, from higher to lower effects, is: $+++ > ++ > +$.

5 Empirical Strategy

Our objective is to study whether the solar industrial policy was effective in increasing solar industry activity in China. We explore the dynamics of the effects in the short and long run in cities that supported their solar industry. It is important to investigate longer-run persistence as a key aim of industrial policy is often to "kick-start" an industry (Juhász (2018), Choi and Levchenko (2021)). These potential effects are particularly relevant in our context, as concerns have been raised that subsidies may have boosted local solar manufacturing and domestic jobs in the near term but that these benefits were not sustained (Ball et al., 2017).

Our baseline results consider treatment as the first time a city implements a solar-related industrial policy. Once a city implements such a policy, it becomes an absorbing state. This modeling choice follows from our argument in Section 2, which suggests that no Chinese city has completely removed all solar subsidies once it has started a subsidy program.

There are several challenges in evaluating causal effects of solar policies. First, implementation of policies is not random. For example, local governments in areas who were hosting nascent solar industries may have been more likely to implement subsidies than those with comparative advantage in other areas. Alternatively, areas in which the solar industry was lagging behind may have used subsidies to try and get one going. We are helped here by the fact that the Chinese solar industry started prior to the first policy interventions in 2006. Having pre-2006 outcome data will help us in constructing control cities for the treatment cities.

Second, the impact of policies may vary over time. This is likely to be true for an outcome such as patenting, as even if firms direct more resources to R&D in response to the subsidy, it is likely that it takes some time until new innovations are discovered. Our approach should therefore be able to capture (and be robust to) these potential dynamic effects of policy intervention.

To help address these concerns we employ the Synthetic Difference-In-Differences (SDID) methodology proposed by Arkhangelsky et al. (2021). The approach combines a familiar difference in differences approach, with a synthetic control approach to construct a counterfactual group for each treated unit by taking a weighted average of all possible control units. Since we may be concerned that implementation of industrial policy was non-random with respect to trends in outcomes, a synthetic DID approach is well-suited to our purposes, as it allows us to match pre-treatment trends between our treated units and synthetic control, thereby establishing a more plausible counterfactual.

In our setting, cities are treated by solar policies in different years. We therefore first estimate cohort-specific ATTs by applying the SDID for a given treatment year. For a given cohort (a set of cities that implement a solar industrial policy in year t) and a given outcome of interest (e.g., patents), we select a synthetic control to make the pre-treatment trends for the treatment and control group as parallel as possible. Given the likelihood that treatment effects will be evolving over time, we construct the synthetic control using only the never-treated cities (i.e., cities that never implement a solar industrial policy in our study period). This allows us to circumvent the challenges raised by recent work on two-way fixed effects with differential treatment timing.(Callaway and Sant'Anna (2021), Sun and Abraham (2021))

Formally, we estimate the treatment effect τ by solving the minimization problem in equation 8, where Y_{it} is the outcome of interest, μ is the intercept, α_i and β_t are city fixed effects and time fixed effects, respectively, and W_{it} is a treatment dummy variable that takes the value of one for every time period post-policy (absorbing state). The weights $\hat{\omega}_i^{\text{sdid}}$ are chosen to make the pre-treatment trends in the outcome variable for treatment and control cities as parallel as possible. The weights $\hat{\lambda}_t^{\text{sdid}}$ allow us to estimate our treatment effects by placing more importance on those pre-treatment periods that better predict post-treatment outcomes.

$$\left(\hat{\tau}^{\text{SDID}}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \hat{\omega}_i^{\text{SDID}} \hat{\lambda}_t^{\text{SDID}} \right\} \quad (8)$$

After estimating cohort-specific ATTs, we then aggregate all the cohort effects using the weighted average proposed in Arkhangelsky et al. (2021), which pools effects across all cohorts and time periods and provides us with the ATTs reported in Tables 3, 4, 5, and 6. Finally, we illustrate the dynamic effects of the policy by reporting SDID event study estimates for the cities that were treated in 2007, at the beginning of The Eleventh Five-Year Plan, which are the cities for which we have a longer time span of data. The event study estimates for the 2007 cohort are in Figures 6, 7, 8, and 9.

Since our approach requires matching on pre-trends, it is worth briefly discussing the time period for which we have data available and during which solar industrial activity was taking place in China. Our policy data starts at the beginning of the Eleventh Five-Year Plan (2006-2010), when specific financial support was allocated across provinces, cities, and municipalities for the first time. The earliest policy we could therefore in principle evaluate was implemented in 2006. To do so, a minimal requirement is that we observe outcome data before 2006, which we do –our production, revenues, patents, and capacity measures all commence in 2004. In order for our SDID approach to makes sense, we would also like to observe that at least some solar industry activity was taking place before 2006 –otherwise the pre-trends that we will be matching on will all be plausibly uninformative zeros. Whilst the 2006-2010 period saw the fastest growth in the industry and decline in the price of solar modules, it was also preceded by a sizeable kickstart growth period following The Tenth Five-Year Plan (2001-2005). We therefore are able to compare outcome trends across cities where the solar industry existed even if it was at a nascent stage. The synthetic control feature of the method therefore provides us with comparable cities based on the trend exhibited, for each outcome variable, during The Tenth Five-Year Plan early industrial development.

6 Results

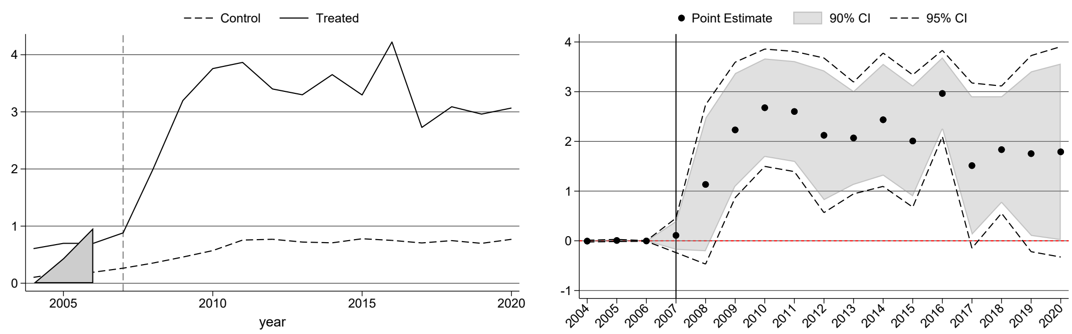
Each subsection focuses on different groups of outcomes (innovation, firm numbers, output and exports). We visualize the results in an event study focusing on the first wave of solar policies in 2007 and then show average results across all cohorts of policies in all years. More detailed cohort-specific event studies are relegated to the Appendix and referred to when necessary in the text.

6.1 Innovation

We begin with the most novel findings, focusing on the impact of subsidies on innovation using the number of patents filed by solar firms in the same city. Figure 6 depicts our estimates for the cohort of cities that were treated with any form of subsidy in 2007. The left panel displays the raw SDID results comparing patenting in treated cities to their synthetic control cities (the shaded triangles represent the pre-policy time weights). Pre-2007, there are mild trends in both treatment and control that are parallel, suggesting that SDID is doing a good job at balancing the groups. After the policy, there is a large increase in patenting in the treatment group (while there is a very small increase for the control cities), peaking about four years after the policy was introduced. Importantly, the number of patents does not revert to baseline even 13 years after the policy was introduced, which implies that firms were not simply bringing forward activity that would have occurred even in the absence of the policy.

The right panel of Figure 6 presents the results in a more classical event-study style. The dots (point estimates) capture the difference between the two lines in the left panel (treatment vs. control) and we show 90% and 95% confidence intervals (bootstrapped standard errors). There are clearly positive and highly statistically significant effects, and although the error bands get wider as we go decades out after the policy, the impact remains significant at the 10% level even in 2020, thirteen years after the introduction of the policy.

Figure 6: All Patents by Solar Firms (2007 Cohort)



Note: Estimated by Synthetic DID methods. The left-hand side graph reflects the raw trends of the control group and treated group, and the grey triangle reflects the time weight. The right-hand side graph reflects the difference between the control and treated groups with 95% and 90% confidence intervals. The outcome variable is all patents (IHS) and the treatment is any subsidy. These are estimates for the cohort of three cities treated in 2007.

Table 3 shows the aggregate ATT, which combines the treatment effects for all cohorts of policies. Column (1) reports the "any subsidy" treatment shown in 6, but for all policies in all years. The point estimate is 0.496 which is smaller than the long-run estimates of the 2007 policy cohort. Since the dependent variable is transformed through the Inverse Hyperbolic Sine (IHS), the ATT effect implies an increase of approximately 64%[11] in the number of patents in a city that introduces a solar subsidy. To put this in context, the average number of annual patents by solar firms in a city is 13.1, so this would imply an increase to 21.5, or about 8.4 extra solar patents per year.

There has been much recent discussion over the interpretation of models with the IHS transformation that we are using here (e.g. Aihounton and Hennings, 2021, Chen and Roth, 2023 and Mullahy and Norton, 2022). We also find that nontrivial magnitudes arise from alternative transformations of the patent count such as using Poisson count data models, simple levels or the log(1+Patents) transformation.

Columns (2)-(4) of Table 3 disaggregate the any solar policies in column (1) into the three alternative types of subsidy. In column (2), we observe that although the ATT of demand subsidies is positive, it is less than half the size of the first column (0.236) and is not statistically significantly different from zero. By contrast, the production subsidy in column (3) is highly significant with an ATT of 0.871, which is larger than column (1). And in the final column, the ATT for Innovation subsidies is 1.060, the largest in the Table and over twice the size of the any subsidy effect in column (1).

Table 3: All Patents

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
All patents	0.496**	0.236	0.871***	1.060***
	(0.200)	(0.275)	(0.227)	(0.367)
Observations	6,086	6,086	6,086	6,086

Notes: *0.1 ** 0.05 *** 0.01. Each observation is a city (admin2 level region) and there are 358 cities in China. 43 regions are treated by any subsidy. The time period is 2004-2020. Each column contains one Synthetic Difference In Differences (SDID) estimate of the Average Treatment of the Treated (ATT), which averages the staggered treatment effects across all cohorts (years in which there were solar policies). Column (1) has any solar policy, column (2) the demand (installation) subsidies, column (3) production subsidies and column (4) innovation subsidies. Bootstrapped standard errors below the ATT. All regressions without controls.

Taking Table 3 as a whole, we see a very striking pattern. Subsidies seem to work in in-

[11] To calculate the percentage increase, we use $e^{\beta} - 1$. However, this is only an approximate estimator to give us an intuitive sense of magnitude. This estimator is accurate only if the outcome value is relatively large. It is also important to notice that the inverse hyperbolic sine transformation is not scale invariant, which is why we have checked against other transformations such as using levels.

creasing innovation, and the magnitude of the effects are consistent with the simple theory we have laid out. The effects are largest for innovation subsidies and negligible for demand subsidies. But production subsidies also generate positive, significant and non-negligible effects on innovation over the longer-run, which is consistent with the aim of green industrial policies.

The influence of production subsidies on firm innovation occurs through two channels. First, production subsidies lower the marginal cost of production, allowing firms to incur the fixed cost associated with innovating. Second, through expanding their production, firms may engage in learning by doing. Some of these efficiency improvements are captured in patenting (which is easier to do in the Chinese patent office than in those of richer countries like the USPTO or EPO). In our extensions in Section 7, we use text-mining techniques to classify patents and replicate our results using as an outcome variable the number of patents whose abstracts reflect process efficiency improvements (which we denote as 'learning-by-doing patents'). The fact that we find positive effects for these patents is suggestive of subsidies enabling learning-by-doing and the subsequent filing of the resulting productivity improvements.

In our Section 7 extensions, we also explore whether some of the city-level estimates we document may include business stealing effects across cities, which could mute the aggregate impact of solar industrial policies. We find no evidence of negative cross-city spillovers.

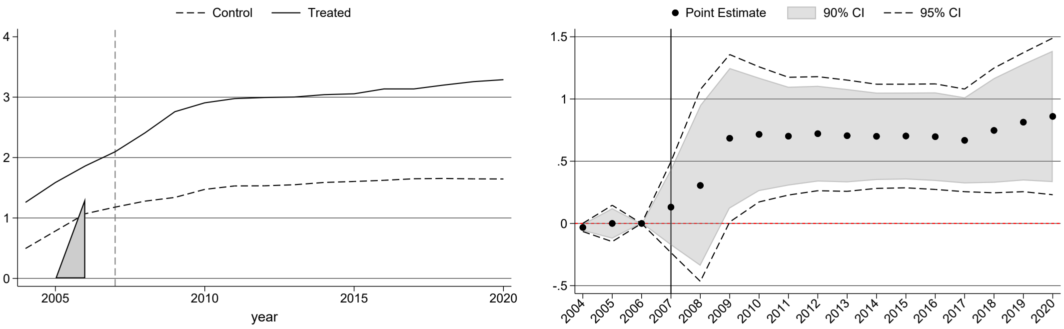
6.2 Number of Solar Firms

Our model posits that the mechanism through which production subsidies positively impact innovation is through first increasing the scale of activity of firms. We now proceed to evaluate the impact of solar subsidies on production activity at both the extensive and intensive margins.

We begin with the simplest measure, that of the number of solar firms in a city. Figure 7 replicates the analysis of Figure 6, but using IHS(number of solar firms) as the outcome. Note that the doppleganger cities used as controls can be different to those in Figure 6 as the SDID routine picks the best controls for each outcome variable separately. Similarly to the previous analysis we see parallel trends between treatment and control and a large positive and significant impact on the number if firms after a city introduces pro-solar policies. Again the effects persist for at least 13 years, and are highly significant even in the long-run.

Table 4, which parallels Table 3, presents the ATT for firm numbers. The overall effect in

Figure 7: Firm Count (2007 Cohort)



Note: Synthetic DID methods. The left-hand side graph reflects the raw trends of the control group and treated group, and the grey triangle reflects the time weight. The right-hand side graph reflects the difference between the control and treated groups with 95% and 90% confidence interval. The outcome variable is firm counts (with arcsinh transformation). These are estimates for the cohort of cities treated in 2007. There are in total 358 admin2 regions in China and 3 cities are treated in 2007.

Table 4: Firm Count

	(1)	(2)	(3)	(4)
	Any subsidy	*Demand subsidy*	*Production subsidy*	*Innovation subsidy*
Firm count	0.186***	0.060	0.288***	0.381***
	(0.064)	(0.043)	(0.090)	(0.135)
Observations	6,086	6,086	6,086	6,086

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time: 2004-2020. Each column is one sdid regression. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls and variable is transformed using IHS.

column (1) is significant, but smaller in magnitude than for patents (0.186 compared to 0.496). We see a similar qualitative pattern when looking across the policies. The demand subsidy effects are positive, but small and statistically insignificant. The production and innovation subsidies are, by contrast large and significant.

6.3 Output: Revenues and Production

To assess the intensive margin effects on production we use two different sources of data. From 2004 to 2013, we can accurately measure solar panel capacity and production in MWh using the ENF's market research reports (*ENF production* dataset)[12]. These reports are con-

[12]Capacity is defined as the maximum 12-month output that could be achieved based on the company's end of the year factory conditions. Production is defined as the likely output that will be achieved in the year based on expected orders

structed based on detailed surveys of factory conditions. They therefore represent highly accurate measures of solar activity.

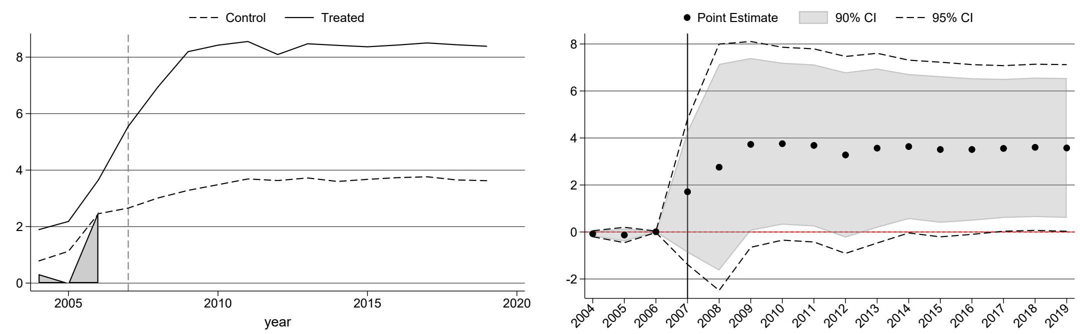
As noted in Section 3, ENF stopped collecting such detailed information after 2013. Consequently, we also examine the sales of solar firms using company accounts data. We use the the *ENF register* which contains a list of all solar firms (used in the previous subsection) and match in accounts data from a variety of sources, in particular BVD Orbis, which is reasonably comprehensive for our companies. This gives us revenue data for solar firms during the 2004-2019 period, which we aggregate to the city-by-year level.

Figure 8 displays the impact of 2007 solar policies on solar manufacturers' revenue through to 2019. The results show that solar manufacturers' total revenue (measured in RMB millions and transformed using the IHS) increases at a faster pace than that of the control group. The effects become statistically significantly different from zero after two years, at the 95% level and persists over the next decade.

Table 5 presents the pooled ATT across all city-cohorts. As before, there is a strong, positive and significant effect on revenue of 1.1 in column (1). Note that this is about five times larger than the effect on the number of firms in Table 4, which suggests that revenue per firm has increased as a result of the policy. In other words, there are not only more solar firms, they have substantially grown in size.

The difference in treatment effects with respect to the type of subsidy in Table 5 goes in the same direction as our previous results. Demand policy coefficients remain the smallest, innovation subsidies the largest and production subsidies are in between.

Figure 8: Revenue (2007 Cohort)



Note: Synthetic DID methods. The left-hand side graph reflects the raw trends of the control group and treated group. The right-hand side graph reflects the difference between the control and treated groups with 95% and 90% confidence interval. The outcome variable is firm revenue (with arcsinh transformation). These are estimates for the cohort of cities treated in 2007. There are 358 cities and 3 are treated in 2007.

Table 5: Revenue

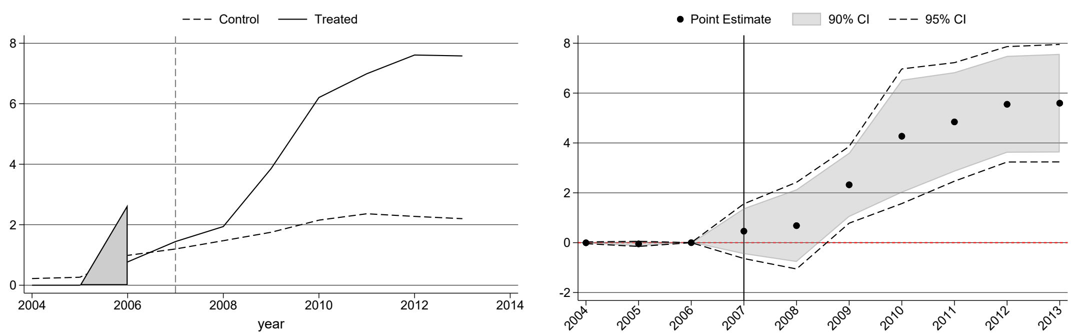
	(1) Any subsidy	(2) Demand subsidy	(3) Production subsidy	(4) Innovation subsidy
Revenue	1.100**	0.192	1.887**	2.670**
	(0.456)	(0.199)	(0.767)	(1.193)
Observations	5,728	5,728	5,728	5,728

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 42 regions are treated by any subsidy. Time: 2004-2019. Each column is one sdid regression. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

Although most of our solar firms are "pure plays", some are multi-product and produce more than just solar panels and solar cells. Hence the revenue from company accounts used in Figure 8 and Table 5 cover more than just solar activity. In addition, revenue contains a mark-up, which means we are not just measuring the quantity of solar output, but any increase in post-policy prices. We turn next, to a more precise volume measure of output in Figure 9 and Table 6 using *ENF production* data, in which we can measure directly the MWh that could be generated with the solar panel output of each manufacturer.

The data reveals that treated cities experienced a faster increase in panel capacity after policy shocks than the selected control cities did. The event-study analysis indicates that treatment effects become statistically significantly different from zero two years after policy implementation. Table 6 again shows statistically significant and positive effects of production and innovation subsidies on production capacity.

Figure 9: Panel Production Capacity (2007 Cohort)



Note: Synthetic DID methods. The left-hand side graph reflects the raw trends of the control group and treated group. The right-hand side graph reflects the difference between the control and treated groups with 95% and 90% confidence interval. The outcome variable is panel capacity MWh (with arcsinh transformation). These are estimates for the cohort of cities treated in 2007. There are in total 358 cities and 3 are treated in 2007.

According to Table 6, the impact of the ATT on production capacity is roughly twice as large

Table 6: Panel Production Capacity

	(1) Any subsidy	(2) Demand subsidy	(3) Production subsidy	(4) Innovation subsidy
Panel capacity	2.098***	0.587	2.496***	2.930***
	(0.532)	(0.467)	(0.575)	(0.773)
Observations	3,580	3,580	3,580	3,580

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 18 regions are treated by any subsidy. Time: 2004-2013. Each column is one sdid regression. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

as its effect on revenue. However, when we examine the revenue results in Table D.6 in the Appendix, limiting the revenue outcomes to end in 2013 as per the ENF production dataset, the ATT for revenue is much closer to that for panel capacity. This suggests that our revenue data is a reasonably accurate measure of production activities. The smaller overall ATT for revenue implies that some multi-product firms may be shifting output from non-solar activities to solar ones to capitalize on subsidies, or that solar companies may be using some of the solar subsidies to increase their price-cost margins.

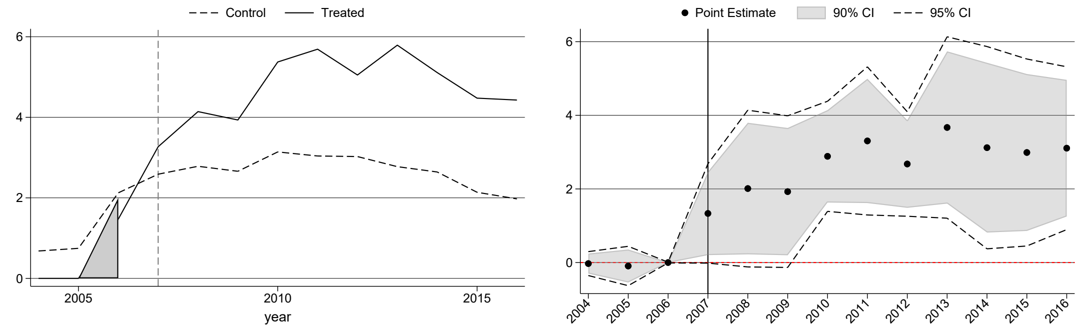
The richness of the ENF data allows us to look at many other measures of solar output. In the Appendix Figure D.1 we show results for solar PV module production (an adjusted measure of capacity based on expected orders) as well as solar cell production and capacity. The results are qualitatively similar to what we have documented in the main text.

6.4 Exports

The final set of results we consider are related to international trade. We use Chinese official customs data from 2004 to 2016 (the last year available). We match all export records with the names of ENF firms, and aggregate the data at the city level according to the location of each firm. Figure 10 illustrates the export value of solar manufacturers in treated and control cities between 2004 and 2016. As with the other outcomes, we observe an increase in total exports if a city introduces a pro-solar policy in 2007.

The results of our export analysis are presented in Table 7, which displays both export values, export volumes and unit values ("price", which we take as primarily a measure of export quality). The first two rows represent estimates for all exports from solar firms and show a familiar pattern, with positive effects in column (1) which are small and insignificant for demand subsidies in column (2) but larger and positive for production and innovation policies (last two columns). The results are weakest for the export price.

Figure 10: Export Value (2007 Cohort)



Note: Synthetic DID methods. The left-hand side graph reflects the raw trends of the control group and treated group. The right-hand side graph reflects the difference between the control and treated groups with 95% and 90% confidence interval. The outcome variable is export value (with arcsinh transformation, million dollar). These are estimates for the cohort of cities treated in 2007. There are in total 358 admin2 regions in China. 3 regions are treated in 2007.

As noted above, some of our solar firms are multi-product so will be exporting non-solar products. Since we would expect the effects of the subsidy to be larger for solar exports, we use HS code "85414020" which includes solar panels and cells only to classify exports into solar and non-solar. The last three rows of the table show the results of this analysis. Consistent with our interpretation, although the qualitative patterns are identical, the ATT effects are about twice as large when focusing on solar exports. Moreover, we now find positive effects on the quality of exports (unit values), which is suggestive of Chinese firms moving up the quality ladder (consistent with the positive effects of the subsidies on innovation). Table D.2 in the appendix also shows non-solar exports. The magnitude for the non-solar exports is much smaller than the solar exports in Table 7, which is consistent with our argument.[13]

6.5 Impact of Solar Subsidies: Summary

We have documented a set of results that appear consistent with our theoretical model. City-level solar industrial policies do appear to have important effects on the development of the solar industry. First, we find positive effects on all the solar outcomes we have examined: innovation, firm numbers, revenue, production, and exports.

Second, we find that the impact of different types of subsidies also lines up with our model

[13]There can be effects of solar subsidies on non-solar exports for a number of reasons. First, if the firm faces financial constraints, the solar subsidy can help relieve this and able greater production and exporting of all goods. Second, if there is a fixed cost to exporting, the greater size of the subsidised firm will help spread this cost over a large number of units.

Table 7: Exports

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
Export value	2.451**	0.658	3.217**	4.160**
	(1.178)	(1.130)	(1.443)	(2.143)
Export volume	2.111**	0.090	2.875**	3.826*
	(0.999)	(0.774)	(1.287)	(1.984)
Export price	0.971*	0.094	1.138	1.502
	(0.554)	(0.636)	(0.731)	(1.051)
Solar export value	4.688***	1.443	6.464***	9.360***
	(1.302)	(0.918)	(1.734)	(2.230)
Solar export volume	3.984***	0.980	5.289***	7.501***
	(1.133)	(0.688)	(1.502)	(1.953)
Solar export price	1.566**	0.197	2.107**	3.383***
	(0.630)	(0.388)	(0.868)	(0.848)

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. Time period for estimation: 2004-2016 for export value, and 2004-2015 for export volume and price. Each column is one SDID regression. The coefficient is the ATT, which averages the staggered treatment effect for all cohorts. All regressions are without controls.

predictions. Demand (installation) subsidies have positive although small and statistically insignificant effects. We interpret this as other city-regions in China being able to supply any city with solar panels in response to a policy-driven demand increase. Hence, demand policies have a muted effect on the local industry.

In contrast, production subsidies have larger and significant effects on output (higher capacity, exports, revenues and firm numbers), because they are only given to firms located in the city. Their effect on innovation is consistent with larger firms being able to cover the fixed costs of R&D and also learning-by-doing.

The largest effects come from innovation subsidies. This is unsurprising for innovation outcomes, but it may appear less obvious for variables like production, revenues, or exports. The reason, as noted above, is that in our data the cities who introduce innovation subsidies also introduce production subsidies. Hence, the innovation policy is actually a bundled production and innovation policy. The larger effects imply that this bundle is more effective than a single production subsidy, indicating an additional effect driven by the innovation subsidy.

6.6 Further Evidence on Mechanisms behind the Innovation Effects

To push our analysis of the innovation effects further, we use the richness of the information we have available on patent types.

SIPO classifies patents into three types: design, utility model and invention. Design patents are considered low value and would not receive protection in the USPTO or EPO. Hence, we would want to see that solar policies stimulated valuable innovations of the invention or utility type. Table 8 implements this test. We first reproduce the results from Table 3 in the first row. Then in the next row, we confine the results to the Design patents. As expected, the policy effects in all columns are small and statistically insignificant for these low value patents. The next row looks at the complement - invention/utility patents. Here we find the usual pattern of results with much larger and significant effects (except for demand, which remains small and insignificant).

The last two rows split the invention/utility patents into solar vs. non-solar patents using technology class codes. We find larger effects on solar patents than on non-solar patents. For example, in column (1) the ATT effect for solar is twice as large as non-solar, and the solar ATT is statistically significant, whereas the effect on the non-solar patent outcome is not.

Table 8: Results by Patent Types

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
All patents	0.496**	0.236	0.871***	1.060***
	(0.200)	(0.275)	(0.227)	(0.367)
□ Design patents	0.186	0.277	0.237	0.151
	(0.138)	(0.216)	(0.173)	(0.253)
□ Invention/utility model patents	0.529***	0.201	0.937***	1.097**
	(0.201)	(0.274)	(0.232)	(0.373)
• Solar patents	0.515***	0.189	0.857***	1.090**
	(0.168)	(0.210)	(0.216)	(0.358)
• Non-solar patents	0.247	−0.034	0.732***	0.809**
	(0.168)	(0.196)	(0.203)	(0.320)

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time period for the estimation: 2004-2020. Each coefficient represents one sdid regression. The coefficient is the ATT which averages the staggered treatment effect for all cohorts. Chinese patents can be classified as design patents, utility model patents and invention patents. Utility model and invention patents contain IPC codes and can therefore be further classified into solar patents and non-solar patents. All regressions are without controls.

These results are not only reassuring as an econometric check. They also help to rule out alternative mechanisms. Subsidies create enhanced cash flows for firms, so one concern is that the innovation effects are driven simply from relieving such financial frictions, which are believed to be large for innovation (Arrow, 1972), and a major issue in China (e.g. Song, Storesletten, and Zilibotti (2011)). The fact that we only observe effects for non-trivial, solar innovation and not for more trivial, non-solar innovation suggests financial constrains are not what primarily drives our findings.

As a further investigation of innovation mechanisms, we used the text of patent abstracts to identify those that were more closely related to learning-by-doing, as opposed to those representing new products or basic science research. We leveraged the pioneering work of Liu (2023), who classified a sample of Chinese patents into process and product-related innovations by hand-reading the full patent text. We used this as a training dataset to classify the remaining patents in our much larger sample using machine-learning techniques.

Re-running our analysis (explained in Section 7 below), we confirmed that our results were partly driven by process innovation, which is consistent with a learning-by-doing mechanism, whereby production subsidies allow firms to operate and achieve efficiency improvements, which they then file as a patent. Consequently, although fixed costs could be important, there is certainly a role for the learning-by-doing mechanism that underlies the rationale for green industrial policy.

7 Extensions and Further Robustness Tests

Our data allows us to extend the previous analysis in a number of ways. In this section, we briefly summarise some of these.

7.1 Cross-city spillovers

To what extent do the positive treatment effects we identify arise from cross-city business stealing? A plausible concern is that the introduction of a solar policy in one city may simply re-allocate activity from non-targeted to the targeted city. These still represent positive effects from the perspective of the local city. However, from the national perspective, if all subsidies do is alter the distribution of solar activity within China achieving no increase in aggregate solar activity, then these are simply beggar-thy-neighbor policies.

On the other hand, there may be positive spillovers. For example, if firms can learn from their neighbors in other cities, then policy-induced innovation or expanded production in one city may increase solar activity in a neighbor.

To investigate these potential effects we have, as usual, to identify who are the cities most "at risk" of business stealing effects. The most obvious group of cities are those who are contiguous to the cities who introduce solar policies. We set up a new set of SDID estimates which use contiguous cities as the treatment group and search for the best synthetic controls amongst the rest of the never-treated cities.

We summarize the results in Appendix Table D.3. Contrary to the business stealing concern, all the ATT effects are positive, rather than negative, rejecting business stealing. As would expect, these indirect effects are all much smaller in magnitude than the direct effects. For example, the ATT for revenue is less than half (0.5), compared to our main result (an ATT of 1.1). Moreover, the effects are insignificantly different from zero for many outcomes. For example, not only is the production capacity impact only 0.385 compared to 2.098, it is statistically insignificant.

Rather than business stealing, positive spillover effects will magnify the city-level policy impact from a national perspective.

7.2 Learning-by-doing patents

We follow standard text cleaning procedures and train a random forest algorithm on 85% of Liu (2023) data on 3,299 Chinese solar patents. We use our model to classify the universe of patents filed by ENF solar manufacturers during the full time period of our analysis into patents associated with learning-by-doing and patents that correspond to new products or basic science research.

As illustrated above, our hypothesis is that solar manufacturers will patent certain efficiency improvements that result from learning-by-doing. Thus, we can shed light on the importance of learning-by-doing in driving our results by examining treatment effects for those patents that reflect process improvements.

In Appendix Table D.5, we replicate our results using patents that have learning by doing characteristics. The results follow the same pattern observed for all patents in Table 3. Again, we find no effects for demand subsidies and a clear hierarchy of positive effects where innovation subsidies achieve the greater impact followed by production subsidies. This suggests that learning by doing may have some role in the pattern of results we observe which is something we will investigate further.

7.3 Productivity Analysis

Our accounting data from Orbis enables us to construct the inputs to production such as labor and capital. These have many of the concerns that are well discussed in the production function literature (see De Loecker and Syverson (2021), for a survey). We are fortunate, however, to have a direct quantity based measure of output from ENF which means that we can, in principle separate TFPQ from TFPR. Panel A of Table D.6 focuses on the Orbis derived measures which are available through 2019 and Panel B repeats these for the data through 2013 where we also have ENF production.

Panel A shows that the treatment effects are smaller for labor and capital than they are for revenues. For example, the ATT for production subsidies is 1.89 for revenues and only 1.4 for labor and 1.3 for capital. The ATT are also only significant at the 10% level for the inputs compared to the 5% level for output.

Since the larger impact on revenues could in principle be due to increasing prices as well as increasing quantity of output, Panel B of Table D.6 shows the analogous results over the shorter period. We see that the coefficient on Solar production is actually slightly larger than for revenues (e.g. 2.5 vs. 2.4 in column (3)), suggesting if anything, a small fall in solar prices. Consistent with the longer period results of Panel A, policy effects on inputs are smaller than for outputs, suggesting positive productivity effects.

Overall, Table D.6 suggests that solar subsidies appear to increase not only total activity in a city, but also productivity. This is independent evidence over and above the earlier results from patenting, that the policies stimulated innovation.

7.4 Other controls

Our baseline SDID analysis does not control for additional variables, since the city fixed effects should effectively absorb most of the relevant confounders. We considered specifications controlling for a number of observables, such as GDP, population and income. These are potentially "bad controls" if the policies affect the growth of the city. However, since solar is a relatively small part of the economic activity of a city they may be useful in picking up cities which are subject to unobservable shocks correlated with the introduction of solar policies that our SDID approach is not fully capturing.

Although GDP per capita tended to be positively correlated with the outcomes, its inclusion made almost no difference to the magnitude or significance of our treatment effects. Table D.7 shows the results of including such controls on all our main specifications. Although the sample is slightly smaller due to missing values on a few of the smaller cities, there is almost no discernable impact. These findings are robust to splitting up GDP from population and including other observables. All this suggests our econometric procedure is doing a good job at dealing with unobservable shocks.

8 Conclusions

In this paper, we have shown how city-level solar supply subsidies (both of production and innovation) increased innovation (as measured by patenting) and production (as measured by number of firms, revenue, panel production and exporting) in treated cities relative to those who do not implement such policies. By contrast demand side subsidies had no impact on any of these outcomes.

We interpret these effects through the lens of a model whereby production subsidies increase firm size which incentivises them to cover the fixed cost of exporting and R&D. Innovation subsidies which in Chinese cities are always layered on top of production subsidies add to this impact by directly subsidising R&D.

Demand subsidies targeted at encouraging *generation* of solar electricity do not have a sizeable effect on solar production or innovation. We argue this is because the demand stimulus can be met with production supplied from anywhere in China (solar parks and other solar generators were not required to use locally produced solar panels).

We are therefore able to document a link between government support at the early stages of an industry and persistent growth and innovation. This is the central tenet of industrial policy. The fact that we observe this in an industry that is displacing dirty energy generation worldwide magnifies the importance of our finding.

Our results indicate that city-level solar policies helped to drive up not just entry and production but also innovation and exporting. This helped to drive down solar generation costs not just in China but across the world which, in turn, may have helped to encourage global diffusion of solar energy.

What Chinese cities have achieved in the last 20 years using these policies is staggering. Our results represent a "ray of hope" for countries worldwide who are trying to balance the need for more energy to drive economic growth with the the need to drop emissions in order to avoid catastrophic climate change. There is hope not only because citizens everywhere will benefit via diffusion from cheaper solar energy built on the back of Chinese solar policies but also because what China has done can serve as a guide to what might be achieved elsewhere.

The industrial policies pursued in China, in effect, offer a route to renewable energy that is cheap enough to displace dirty sources of energy. This is a route that the US and EU (and many other countries) are also pursuing offering additional hope that we will reach net zero whilst continuing to raise living standards around the world.

References

Aghion, P., Cai, J., Dewatripont, M., Du, L., Harrison, A., & Legros, P. (2015, October). Industrial Policy and Competition. *American Economic Journal: Macroeconomics*, *7*(4), 1–32. doi: 10.1257/mac.20120103

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021, December). Synthetic Difference-in-Differences. *American Economic Review*, *111*(12), 4088–4118. doi: 10.1257/aer.20190159

Arrow, K. J. (1972). *Economic welfare and the allocation of resources for invention.* Springer.

Bai, C.-E., Hsieh, C.-T., & Song, Z. (2020). Special deals with chinese characteristics. *NBER macroeconomics annual*, *34*(1), 341–379.

Ball, J., Reicher, D., Sun, X., & Pollock, C. (2017, March). *The New Solar System: China's Evolving Solar Industry and Its Implications for Competitive Solar Power in the United States and the World* (Tech. Rep. No. 1352021). doi: 10.2172/1352021

Buera, F. J., Hopenhayn, H., Shin, Y., & Trachter, N. (2021). Big Push in Distorted Economies. *NBER Working Paper 28561.* doi: 10.3386/w28561

Buera, F. J., Moll, B., & Shin, Y. (2013, January). Well-intended policies. *Review of Economic Dynamics*, *16*(1), 216–230. doi: 10.1016/j.red.2012.10.008

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *The American Economic Review*, *101*(1), 304–340.

Callaway, B., & Sant'Anna, P. H. C. (2021, December). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. doi: 10.1016/j.jeconom.2020.12.001

Chen, J., & Xie, L. (2019, December). Industrial policy, structural transformation and economic growth: Evidence from China. *Frontiers of Business Research in China*, *13*(1).

Choi, J., & Levchenko, A. A. (2021, September). *The Long-Term Effects of Industrial Policy* (Working Paper No. 29263). National Bureau of Economic Research. doi: 10.3386/w29263

Choi, J., & Shim, Y. (n.d.). Technology Adoption and Late Industrialization. , 106.

Criscuolo, C., Martin, R., Overman, H. G., & Van Reenen, J. (2019, January). Some Causal Effects of an Industrial Policy. *American Economic Review*, *109*(1), 48–85. doi: 10.1257/aer.20160034

De Loecker, J., & Syverson, C. (2021). An industrial organization perspective on productivity. In *Handbook of industrial organization* (Vol. 4, pp. 141–223). Elsevier.

Itskhoki, O., & Moll, B. (2019). Optimal Development Policies With Financial Frictions. *Econometrica : journal of the Econometric Society*, *87*(1), 139–173. doi: 10.3982/ECTA13761

Juhász, R. (2018). Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade. *The American Economic Review*, *108*(11), 3339–3376.

Juhász, R., Lane, N., Oehlsen, E., & Pérez, V. C. (2022, August). *The Who, What, When, and How of Industrial Policy: A Text-Based Approach.* SocArXiv. doi: 10.31235/osf.io/uyxh9

Kalouptsidi, M. (2018, April). Detection and Impact of Industrial Subsidies: The Case of Chinese Shipbuilding. *The Review of Economic Studies*, *85*(2), 1111–1158. doi: 10.1093/restud/rdx050

Lane, N. (2020, June). The New Empirics of Industrial Policy. *Journal of Industry, Competition and Trade*, *20*(2), 209–234. doi: 10.1007/s10842-019-00323-2

Lane, N. (in press). Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea. *The Quarterly Journal of Economics (forthcoming).* doi: 10.31235/osf.io/6tqax

Liu, J. (2023). Quality and productivity upgrading among chinese pv panel exporter. *mimeo, Harvard.*

Manelici, I., & Pantea, S. (2021, April). Industrial policy at work: Evidence from Romania's income tax break for workers in IT. *European Economic Review*, *133*, 103674. doi: 10.1016/j.euroecorev.2021.103674

Murphy, K. M., Shleifer, A., & Vishny, R. W. (1989). Industrialization and the Big Push. *Journal of Political Economy*, *97*(5), 1003–1026.

Shapiro, J. S., & Walker, R. (2018, December). Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, *108*(12), 3814–3854. doi: 10.1257/aer.20151272

Shubbak, M. H. (2019). Advances in solar photovoltaics: Technology review and patent trends. *Renewable and Sustainable Energy Reviews*, *115*, 109383.

Song, Z., Storesletten, K., & Zilibotti, F. (2011). Growing like china. *American Economic Review*, *101*(1), 196–233.

Sun, L., & Abraham, S. (2021, December). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. doi: 10.1016/j.jeconom.2020.09.006

Wang, S., & Yang, D. Y. (2021, October). *Policy Experimentation in China: The Political Economy of Policy Learning* (Working Paper No. 29402). National Bureau of Economic Research. doi: 10.3386/w29402

ONLINE APPENDICES

A Institutional Background on China's Industrial Policy towards Solar

In this Appendix we expand on the summary in the main text regarding China's policy support towards solar manufacturing and R&D. This borrows extensively from Ball et al. (2017)'s rich account of China's photovoltaic industry. We outline several features of solar policy support in China. The 5-Year Plans provided national guideline and sectoral industrial policy focus. However, the funding and implementation of these guidelines was mostly carried out at the regional level, which generates considerable heterogeneity in policy support towards the solar industry across prefectural-cities. Finally, measuring industrial policy support to the Chinese solar industry has proven to be very challenging. We close this background section by illustrating our novel approach to obtain micro-level measures of policy support towards both solar manufacturing and R&D. Our approach is based on the analysis and classification of the policy text of the universe of laws and regulations covering the solar industry during the 2006-2022 period.

A.1 Solar PV in the Government's Five-Year Plans

The Chinese government outlines its vision for sectoral industrial policies in its five-year plans. These documents provide national guidance for scientist and inventors, being local implementation what ultimately determines the intensity of policy support. Solar became first a targeted sector in the 2001-2005 Tenth Five-Year Plan, together with other renewable energies. Subsequently, the Chinese government has promoted solar energy in its Eleventh, Twelfth, and Thirteenth Five Year Plans. These Plans have emphasised export-based manufacturing at first and guided the industry towards more sophisticated R&D spanning the whole solar value chain in later periods.

In 2001, at the start of the Tenth Five-Year Plan for New-and Renewable-Energy-Industry Development, China had no domestic solar photovoltaic industry. This plan was China's first attempt to launch renewable energy industries. With the aim of developing a solar supply chain, the State Economic and Trade Commission, encouraged the production of solar cells and modules, with specific targets to be met by the end of the Plan. While innovation appeared as a long-term objective to increase the competitiveness of the national solar industry, the Plan did not specify policy support towards R&D. The Tenth Five-Year Plan period brought considerable growth to the solar industry, exceeding government's expectations.

The Eleventh Five-Year Plan (2006-2010), for the first time, saw the solar industry as an opportunity to attain technological leadership. It emphasised the expansion of factory production and outlined strategies to increase R&D on polysilicon material and cell efficiency. It also encouraged the adoption of panels across the country. All with the broad objective of strengthening the PV manufacturing supply chain. The Eleventh Five-Year Plan included funding for R&D and manufacturing development for the first time. The solar industry witnessed exceptional growth during this period. Figure 2, which we construct with our data on the universe of solar manufacturers in China, displays a clear increase in industrial activity, along both production and patenting outcomes. On top of this, in 2006, the Chinese government kickstarted its Renewable-Energy Law to sustain and speed the incipient growth of its solar industry.

With the Twelfth Five-Year Plan (2011-2015), the government kept pushing for solar adoption, supply-chain expansion and indigenous R&D. The R&D goals gained in detail and covered all aspects of the production cycle: materials, cell, modules, auxiliary systems, and even production methods and tools. China's Thirteenth Five-Year Plan (2016-2020) again mentions solar as a sector to prioritise through industrial policy support, targeting capacity and R&D expansion, as well as industry-wide cost-reduction. Within this plan, the China's National Energy Administration issued, in December 2016, a specific Thirteenth Five Year Plan for Solar Energy Development.

A.2 Policy Support Towards Solar Manufacturing

China's national, provincial, and local governments, all provided an array of subsidies to the solar industry. However, the extent and nature of this policy support in China remains a disputed issue, which has reached international courts[14]. Ball et al. (2017)'s qualitative research, based on interviews with government officials, high-level members of the industry, manufacturing firms and academics, provides some clarity on the administrative level and characteristics of policy support and a rough estimate of its size.

Subsidies to solar manufacturing were managed and allocated by local governments, despite following the national guidance embedded in the Five-Year Plans. The timing, size, and targeting of policy support thus varied significantly depending on the city or region. Ball et al. (2017) document that local governments often engaged in competition via policy support to attract solar manufacturers to their region. Subsidies followed a similar structure to that of other sectoral industrial policies in China. They were mostly targeting manufacturing at first and included several features. Since 2006, many local governments took advantage of the national

[14] DOC investigation in the wake of the SolarWorld trade allegations

legal framework easing policy support for renewable energies and provided generous tax incentives to solar manufacturers. Many city-level governments also offered discounts for land acquisitions and cash investments for struggling solar manufacturers. Moreover, city governments hosting solar-manufacturing clusters within their administrative boundaries, offered additional mechanisms for financial assistance to resident firms. Ball et al. (2017) conjecture that this continuous and wide ecosystem of policy support may lay behind the continuous process innovation and improvements in cell-efficiency observed in China's solar manufacturing.

A.3 Policy Support Towards Solar R&D

China's national vision and ecosystem for R&D involve a variety of governmental, corporate, and academic actors, coordinated by the Plans' national guidance. At the national level, the National Development and Reform Commission (NDRC), the National Energy Administration (NEA), the Ministry of Science and Technology (MOST), the Ministry of Industry and Information Technology (MOIIT), the Ministry of Finance (MOF), and the Ministry of Education (MOE), all contribute, to varying degrees, to crafting energy-policy, and designing industrial and R&D policies targeting the solar industry.

The structure of the Chinese government policy support towards R&D is much more complex than that of policies targeting manufacturing. Government funding for solar innovation encompasses a variety of programs at a range of firms, universities and research institutions, which fund both basic and applied research. As it is the case for manufacturing subsidies, there is a lot of opaqueness around the nature and quantity of public solar R&D expenditure in China. However, Ball et al. (2017) provide some clarity on the government's solar-R&D efforts, constructed from the analysis of public information and interviews with key actors in the Chinese solar industry. In section 2.3 we explain how we provide novel city-level measurement of the government's support towards both solar manufacturing and solar R&D. The authors estimate a lower bound of $74 million spending in solar R&D during the 2000-2015 period by both the national and local governments. The period from 2005-2015 faced additional $223 millions in R&D expenditures, including both public and private investments. As it is the case with manufacturing subsidies, there is considerable regional heterogeneity in solar R&D funding, as cities and provinces support local laboratories and research centres dedicated to engineering and technology innovation. Again, retrieving quantitative data on solar R&D spending at the local level has proven to be extremely challenging in the past. In the next section, we explain the approach we use to measure industrial policy support at the city-level for both solar manufacturing and solar innovation separately.

B Data

We summarized the rich and original data we have compiled in the main text in ??. Here, we go into more details on the various datasets that we have matched and compiled.

B.1 Solar industrial policy

The main data on industrial policy towards solar manufacturing and installation comes from PKULaw's Laws & Regulations dataset. The Laws & Regulations database is a comprehensive and reliable source of China's legal information, including all laws, regulations, and any related legal information implemented by the central and local governments since 1949. We obtain data disaggregated by industry and gather all regulations pertaining to the solar photovoltaics industry, which start in 2006. The dataset contains information on the title, validity, administrative level, department, release date, and implementation date of each policy. It also includes a link to the original policy document, which contains the text of each regulation or announcement. We manually inspect the full text of each policy and classify them into subsidies, announcements, poverty alleviation policies, and records. We further classify subsidy policies according to whether they target solar installation, production, or innovation.

B.2 Solar panel and cell manufacturers register, production, and capacity data

The ENF Solar Industry Directory is a register of 50,800 worldwide photovoltaic (PV) companies. Because it is the leading solar website, most companies self-register on ENF's platform. ENF reviews daily news regarding the solar industry, as well as available lists of key solar exhibitions, to incorporate the remaining new solar companies. Additionally, ENF relies on government organizations and a variety of web-scraping techniques to complete the full list of solar companies. ENF uses automatic scanning to detect company updates, which triggers careful checks from ENF database experts to update manufacturers' information. Finally, ENF automatically scans for signs of companies ceasing their activities. Hence, ENF is able to reasonably capture a snapshot of all solar panel manufacturers each year. We obtained access to the historical directories of solar panel producers from ENF Solar Industry Directory, available from 2010 until 2021 (henceforth, "ENF register"). We also gained access to the last edition of ENF's Chinese Cell & Panel Manufacturers Report. This dataset (henceforth, "ENF production") allows us to measure, for each firm, their production and capacity figures (in MWh) for both solar panels and solar cells across the 2004-2013 period.

The ENF register and ENF production datasets overlap for the 2010-2013 period. We matched the two datasets by firm name and contact details (address, phone, website, fax, and email). We manually inspect and address remainder mismatches. We are left with a sample of 1,718 Chinese solar panel manufacturers, operating at some point between 2004 and 2021, which includes production and capacity data for each manufacturer during the 2004-2013 period.

ENF includes projections of production and capacity in 2014, but we chose not to use this. 2013 is a transition year with some actual and some projected data, so we felt comfortable with using this year. We also checked robustness of the results to ending the sample in 2012.

B.3 Firm counts, entry and exit

The Qichacha platform[15] allows us to gather detailed firm-level information, spanning from registration to exit, and updated periodically following government requirements. This includes the type of business, the identity of affiliated enterprises, a variety of judicial and legal details, company news, corporate annual reports, and our main variables of interest, firm entry and exit dates. The Qichacha platform collects this information from multiple data sources, but mostly relies on government's official sources, which include the National Enterprise Credit Information Publicity System, the China Court Judgment Documents Network, and the China Enforcement Information Disclosure Network.

To retrieve the key variables for our sample of ENF solar manufacturers, we manually search in the platform using ENF firms' Chinese names. Some of the firms included in the ENF register are based in Hongkong or Taiwan and are therefore excluded from the Qichacha platform. Our final sample of manufacturers is restricted to those with an address in mainland China.

We still face one limitation when using this approach. The 2013 and 2014 ENF solar manufacturing registers only record firms' English names, which cannot be uniquely matched to the Qichacha platform. We use Google and Baidu to obtain the corresponding Chinese name for the English-name-only firms, allowing us to further identify 462 firms (out of a total of 673 firms without Chinese name in the ENF register).

B.4 Patents and their characteristics

The Qichacha platform contains detailed intellectual property information from the State Intellectual Property Office (SIPO). This enables us to obtain, for each ENF manufacturer, the name, patent ID, type, application date, publication date, and assignee, of the patents it has

[15] https://www.qcc.com/

filed. We then use the SIPO patent ID to extract IPC codes and patent abstracts from the PATSTAT database. To understand the nature of the underlying innovation, we classify the patents filed by our sample of manufacturers into several categories. First, we rely on the SIPO classification of patents into Invention, Utility Model and Design patents. Invention patents have longer protection periods, require paying higher filing costs, and involve a more cumbersome administrative process. They are therefore patents of higher quality and a more innovative nature. The firms in our solar manufacturers dataset file mostly invention and utility model patents. Second, using IPC codes, we further classify invention and utility model patents into solar and non-solar patents. Finally, we use text mining techniques to detect learning-by-doing (LBD) patents based on the information in the patent abstracts.

B.5 Text analysis on patent abstracts

To characterize the innovative content of patents filed by our sample of solar manufacturers, we built a supervised learning model using Liu (2023)'s dataset to train our text classification procedure. This dataset contains 3,299 solar patents (according to their IPC code), manually classified by the author into *productivity-improving* or not, after careful analysis of the text of all patent abstracts. Figure B.1 display the most common words contained in the patent abstracts for productivity increasing or learning-by-doing patents[16].

We follow standard text cleaning procedures and train a random forest algorithm on 85 % of Liu (2023)'s data. The model classifies the remaining patent abstracts with an accuracy of 85-90%. We then use our model to classify the universe of patents filed by ENF solar manufacturers during the full time period of our analysis.

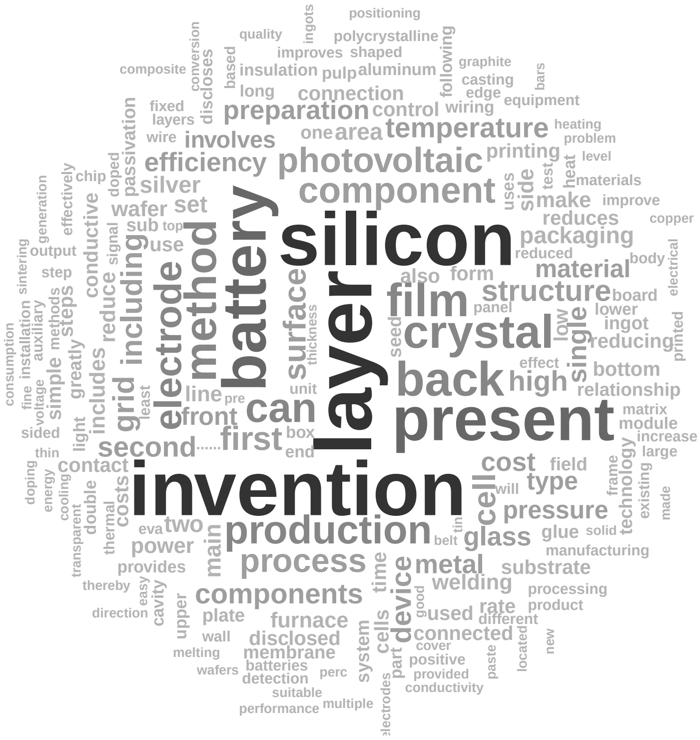
B.6 Examples of Learning by doing (LBD) patent

Figures B.2 and B.3 offer two additional instances of learning-by-doing patents. Their abstracts highlight the benefits of the current patent for production processes, product quality, and industrial development.

Figure B.4 and B.5 offer two non-learning-by-doing patents. The first non-learning-by-doing patent is a solar-related new product. This type of patent does not refer to firm productivity or process improvement. Therefore, we do not count it as a learning-by-doing patent. The second non-learning-by-doing patent improves the quality of solar cells and involves fundamental chemistry and physics. This type of patent is also not likely to reflect efficiency improvements

[16]The word 'solar' has been removed to ease visualisation.

Figure B.1: Learning-by-doing patents wordcloud

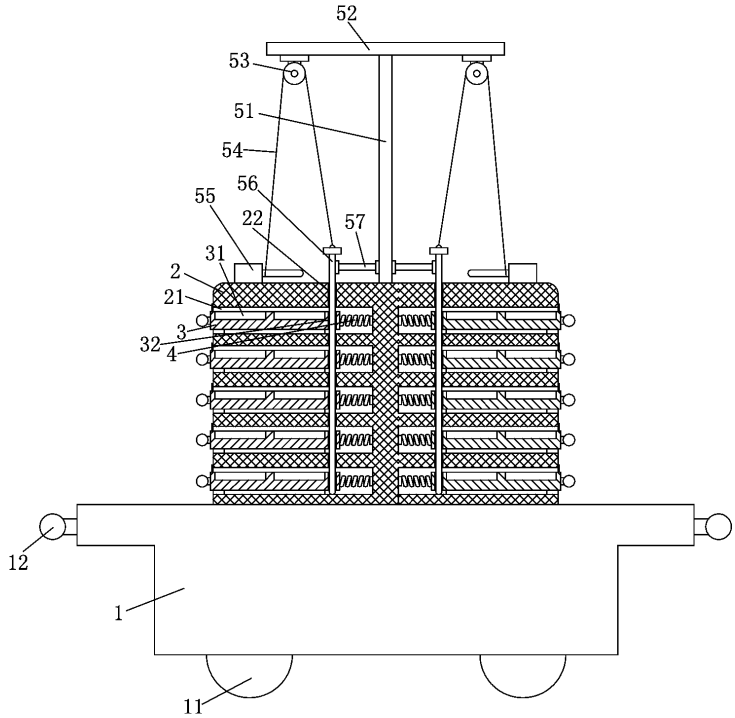


driven by increase in production.

B.7 Revenue, assets, employees, and cost of goods sold

In order to expand the time horizon of our analysis and estimate long-run effects beyond the effects on production that we calculate using ENF production data, we use Bureau Van Dijk's Orbis dataset, which gives us rich financial data, including total and tangible fixed assets, revenue, employees, and cost of goods sold, throughout the 2004-2019 period. We use the comprehensive firm contact information included in both the Orbis and ENF register datasets to merge the two datasets, and obtain Orbis variables for our sample of solar manufacturers.

Figure B.2: Learning-by-doing patent example 1



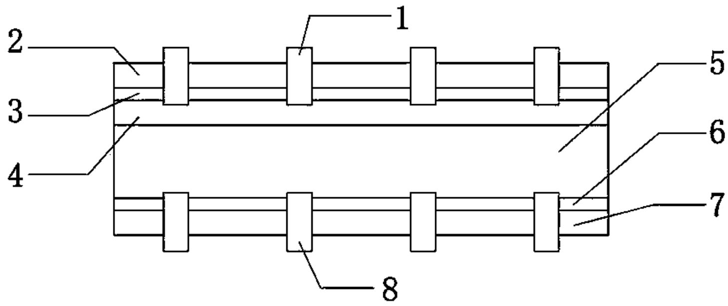
Patent Abstract: The utility model discloses a transfer assembly of a solar cell piece with a metal-stacked electrode. The assembly comprises a trolley body, a storage member arranged on the top of the trolley body, and a positioning component arranged on the storage member. A plurality of slots are opened on the storage member, and a storage plate is slidably connected in each slot. The top of the storage plate is provided with a groove, a spring is provided on the inner wall of each slot, the spring is connected to the storage plate, a first connecting hole is opened on the storage plate, and a second connecting hole penetrating all the slots is opened on the storage member. The positioning component includes a support column, a crossbar, a pulley, a rope, a motor, a limit rod, and a sliding block. *The utility model delivers the solar cell piece through the newly designed transfer assembly. The structure is simple, easy to install and transport, and will not damage the solar cell piece during transportation, reducing the defect rate and ensuring product quality.*

B.8 Validating Orbis data quality using the ASIE

We cross-validate the Orbis data making use of the Annual Survey of Industrial Enterprises (ASIE). The Annual Survey of Industrial Enterprises (ASIE), also called as Annual Survey of Industrial Firms (ASIF), is an administrative-level dataset for all large industrial[17] firms in China. The ASIE is only available between 1998 and 2013 and the sample of firms included in the survey changes over time. Before 2011, the revenue threshold for inclusion in the ASIE was 5 million RMB. After 2011, this threshold was raised to 20 million RMB. Despite these limitations, given that the Chinese government uses this dataset to construct official statistics, we use the ASIE to assess the quality of our Orbis long-run data.

[17]This includes manufacturing, mining, electricity, gas, and water firms

Figure B.3: LEARNING-BY-DOING PATENT EXAMPLE 2



Patent Abstract: The present invention discloses a new type of double-sided light-receiving solar cell, which includes a front electrode, a front anti-reflection layer, a front passivation layer, a PN junction, and a P-type silicon substrate. A back passivation layer, a back anti-reflection layer, and a back electrode are also provided on the back of the P-type silicon substrate. *The present invention reduces the preparation process of existing double-sided cells and is more conducive to industrial development.*

We match our ENF firms with the ASIE through a two-stage process. First, we search ENF firms in the Qichacha platform and retrieve their registration data, which includes the standardised official Chinese name. This name standardisation is also used in the ASIE, so we can conduct exact matching with the ASIE dataset on a second stage[18]. We are therefore able to identify ENF firms on ASIE and Orbis using two different matching procedures based on rich contact information and the standardised naming convention shared by the ASIE and the firm registration dataset. We can the compare the values registered in Orbis and ASIE for the same variable, same ENF firm, and same year. Both Orbis and ASIE include information on the value of total assets. Figure B.6 compares log(assets) in Orbis and in ASIE. Each data point represents a firm-year combination. We discard the possibility that Orbis just used ASIE data for the overlapping years by noting that there was no noticeable break in the time series of ENF firms' total yearly assets as reported in Orbis before and after ASIE became available. This is visible in Figure B.7.

B.9 Solar exports volume, value, and prices

The Chinese Customs Dataset contains information on all imports and exports between 2000 and 2016. It records all international trade transactions by Chinese firms, allowing us to observe the name of the importing/exporting firms, the value of the transaction, the quantity and price of the exchange, the product HS code, and the country of the trading partner.

[18]The ASIE and the Qichacha firm registration datasets are of administrative nature, so they share the same standard firm naming system.

We obtain exports information for our sample of ENF manufacturers following the same two-step procedure used for the ASIE data. First, we search by name in the Qichacha platform and retrieve the standardised official name for all ENF firms. This allows us to match exactly with the customs data and get information on the quantity, value and price of exports by ENF solar manufacturers.

Not all exports by ENF manufacturers are solar products. We classify exports as solar-related using the HS code "85414020" in the customs data. This code was created in 2009. Prior to that year, we use the code "854140", which is a broader code that includes LED products as well.

B.10 Data imputation for missing values in Orbis data

Broadly, there are three types of missing values for revenues in the Orbis dataset. The first type occurs when we we observe revenue data for the same firm in two different non-consecutive years, but there are missing values for the years in between. In this case, if we were not to interpolate, when aggregating at the city level, we would be assigning a value of 0 for this firm, which would create a false discontinuity in the data. Therefore we use linear imputation to fill these missing values. The second case of missing data occurs when we observe some values for revenue, but we fail to observe data for the first few years in the sample, when the firm enters the market, or the last few years, before the firm exits. In this case, we use the values we observe to replace the missing ones through extrapolation. The third case occurs when we do not observe any information for a firm. In this case, we simply drop the firm completely.

B.11 Final City (admin 2 regions) panel dataset

We construct a dataset at the "city" (second administrative level) level to run the empirical analysis, which exploits the policy variation at the city-level stemming from PKULaw data.

The ENF production dataset contains detailed address information, which allows us to geolocate all firms through the Google API, and assign them their corresponding city. We aggregate all production and capacity figures from ENF cell and panel manufacturers at the city-level. We identify, for each city, the number of ENF panel and cell manufacturers using the ENF register, ENF production, and firm registration dataset, which provides reliable firm entry and exit data. We aggregate our patent data from SIPO, revenue and assets from Orbis, as well as the total volume and total value of exports from customs data, for the same sample of ENF manufacturers, at the city level. Finally, we calculate a simple average of the price of exports at the city-level. We additionally gather annual GDP, population, number of workers, and government budget from the statistics yearbook, released by the Bureau of Statistics.

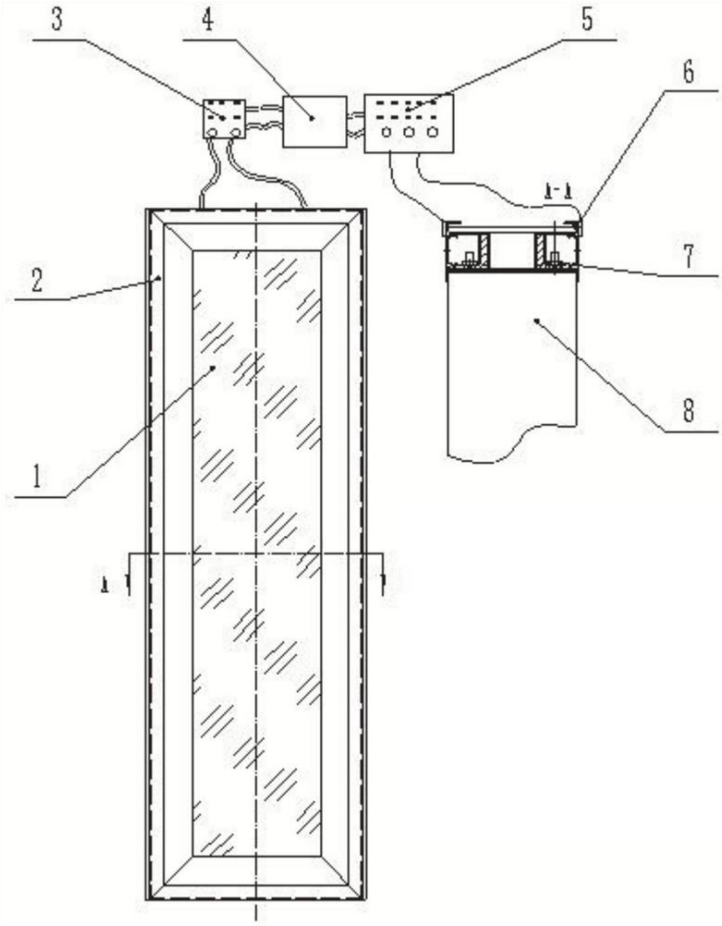
Table B.1 reports descriptive statistics for the key variables at the city-level. Over the 17 year period as a whole, the average city produced 13.1 patents by solar firms per year, a total of 79,902 over the period as a whole.

Table B.1: City-level summary statistics

	Mean	Std. Dev.	Sample Size
SIPO, 2004-2020, 358 cities:			
Total patents by solar firms	13.1	111.3	6,086
Design patents	1.2	10.4	6,086
Utility model and invention patents	11.9	102.8	6,086
Orbis and Qichacha, 358 cities:			
Total number of solar firms, 2004-2020	2.9	10.2	6,086
Total revenue of solar firms, RMB, billions, 2004-2019	0.809	5.99	5,728
ENF, 2004-2013, 358 cities:			
Total Solar Panel capacity, MWh	82.4	483.3	3,580
Total Solar Panel production, MWh	40.7	265.5	3,580
Total Solar Cell capacity, MWh	50.8	353.4	3,580
Total Solar Cell production, MWh	31.3	233.0	3,580
Total Number of Solar Panel firms	0.9	3.5	3,580
Total Number of Solar Cell firms	0.2	1.0	3,580
Customs, 358 cities:			
Total export value of solar firms, millions USD, 2004-2016	24.8	186	4,654
Total export volume of solar firms, millions, 2004-2015	3.18	43.7	4,296
Average export price of solar firms, USD, 2004-2015	9,716	480,762	4,296
Statistics Yearbook, 2004-2020, 284 cities:			
GDP, billion RMB	196.0	307.2	4,828
Population, thousand	4,453	3,176	4,828
GDP per capita, RMB	43,497	46,936	4,828

Notes: Each observation is city-year pair. There are up to 358 cities between 2004 and 2020 (6,086 observations), but differe datasets may have lower numbers of observations as noted in the table.

Figure B.4: Non-learning-by-doing patent example 1

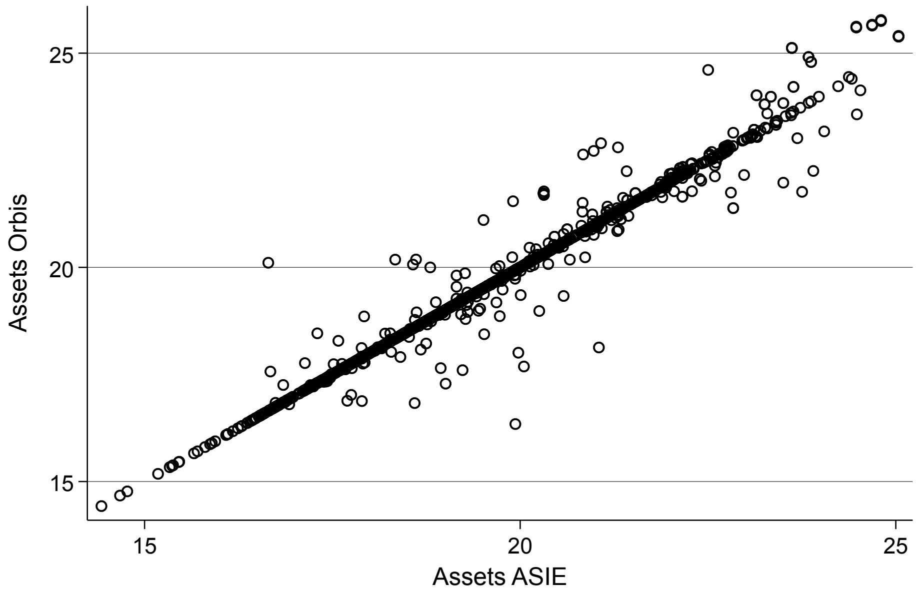


Patent Abstract: This utility model patent relates to a road cliff photovoltaic lighting device, which includes a road cliff stone or road guardrail connected to the outer surface of a photovoltaic component. The photovoltaic component is connected to the inverter and battery through a controller in sequence, and the controller is connected to the light strip. The light strip is located on one side of the road cliff stone or road guardrail facing the center of the road. By combining the photovoltaic power generation system with the road cliff or guardrail lighting, photovoltaic power generation, which serves as green energy, is closely integrated with transportation, solving the power supply and subsequent maintenance problems of traditional road lighting and reducing construction and maintenance costs. It also produces an uninterrupted power supply to indicate the road dividing lines and boundary lines, guiding the passage of vehicles and pedestrians, relieving driving fatigue and beautifying the road.

Figure B.5: NON-LEARNING-BY-DOING PATENT EXAMPLE 2

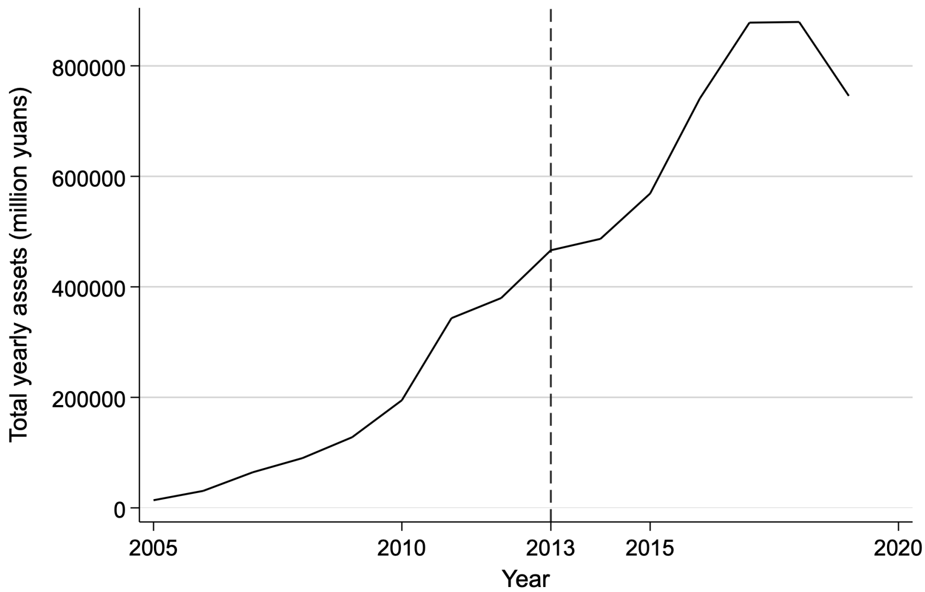
Patent Abstract: The present invention provides a carbon-doped P-type gallium phosphide material, in which carbon is used as the doping element of the P-type gallium phosphide semiconductor material. The preparation method of the material is to use metal organic chemical vapor deposition technology, introduce organic gallium source and phosphorus source into the reaction chamber, let them decompose at high temperature, and react on the surface of the substrate to produce gallium phosphide material. During the generation of gallium phosphide material, carbon impurities are introduced by inputting substances containing carbon elements, or by utilizing carbon atoms generated by the organic gallium source during thermal decomposition. In the present invention, carbon replaces Mg or Zn. Since carbon doping has a small diffusion coefficient and stable properties, highly doped GaP materials can be produced, which are characterized by high efficiency, low diffusion, and high stability.

Figure B.6: VALUE OF FIRM ASSETS IN ORBIS & ASIE



Notes: The axis is the log(assets) in the ASIE data set, and the y-axis is the log(assets) in the Orbis data set. Each point is one firm in one year. If we fit a linear line, the coefficient is 1.01, p<0.01, and R^2 = 0.9679

Figure B.7: Smoothness in Orbis total yearly non-imputed assets



Notes: The time series shows the yearly sum of Chinese solar panel manufacturers' total assets as reported in the Orbis database. The dashed vertical line at 2013 indicates the last year when the ASIE data set is publicly available.

C Theoretical Appendix

In this Appendix, we provide additional technical details to expand on the model outlined in the main body of the paper.

C.1 The Grid Planner Problem

C.1.1 Demand for Energy Sources

Each region d hosts a representative household that consumes only electricity and a grid planner, who is in charge of installing power plants to provide electricity. One way of expressing this problem, is that the grid planner chooses the electricity mix to maximise the value of the final electricity services that it produces, subject to the prices of final electricity P_d and the price of solar and non-solar electricity P_s and $P_{s'}$ respectively.

$$\max_{e_{d,s},e_{d,s'}} \left(P_d e_d - P_{d,s}e_{d,s} - P_{d,s'}e_{d,s'}\right)$$

$$\text{s.t.} \quad e_d = \left(\kappa_{d,s'}e_{d,s'}^{\rho} + \kappa_{d,s}e_{d,s}^{\rho}\right)^{1/\rho}$$

However, a number of observations allow us to re-express this problem. Because our household consumes electricity only, all of their income will go to the grid planner. The production function for electricity services is constant returns to scale. Therefore, the grid planner spends all the income they receive on the production of electricity e_d (zero profits), and the supply of electricity is perfectly elastic, so the price will be pinned down by the production side. The household has a utility function which is strictly increasing in electricity services e_d.

We can therefore rewrite our problem as though the grid-planner uses the full income of households in their area in order to maximise electricity output.

$$\max_{e_{d,s},e_{d,s'}} \left(\kappa_{d,s'}e_{d,s'}^{\rho} + \kappa_{d,s}e_{d,s}^{\rho}\right)^{1/\rho}$$

$$\text{s.t. } P_{d,s}e_{d,s} + P_{d,s'}e_{d,s'} = I_d$$

$$\mathcal{L} = \left(\kappa_{d,s'}e_{d,s'}^{\rho} + \kappa_{d,s}e_{d,s}^{\rho}\right)^{\frac{1}{\rho}} - \left(P_{d,s}e_{d,s} + P_{d,s'}e_{d,s'} - I_d\right)$$

$$\frac{\partial \mathcal{L}}{\partial e_{d,s}} = \frac{1}{\rho}\left(\kappa_{d,s'}e^{\rho}_{d,s'} + \kappa_{d,s}e^{\rho}_{d,s}\right)^{\frac{1}{\rho}}\left[\rho k_{d,s}e^{\rho-1}_{d,s}\right] - P_{d,s} = 0$$

Taking the ratio for $e_{d,s}$ and $e_{d,s'}$ FOCs:

$$\frac{\kappa_{d,s}e^{\rho-1}_{d,s}}{\kappa_{d,s'}e^{\rho-1}_{d,s'}} = \frac{P_{d,s}}{P_{d,s'}}$$

$$e_s = \left[\frac{P_{d,s}}{P_{d,s'}}\frac{\kappa_{d,s'}}{\kappa_{d,s}}\right]^{\frac{1}{\rho-1}} e_{d,s'}$$

Multiplying by $P_{d,s}$ and adding $P_{d,s'}e_{d,s'}$ on both sides of the equation, we obtain:

$$\underbrace{P_{d,s}e_{d,s} + P_{d,s'}e_{d,s'}}_{=I_d} = \left[P_{d,s}^{\frac{\rho}{\rho-1}}\left(\frac{1}{\kappa_{d,s}}\right)^{\frac{\rho}{\rho-1}} + P_{d,s'}^{\frac{\rho}{\rho-1}}\left(\frac{1}{\kappa_{d,s'}}\right)^{\frac{\rho}{\rho-1}}\right]\left(\frac{\kappa_{d,s'}}{P_{d,s'}}\right)^{\frac{1}{\rho-1}} e_{d,s'}$$

Which simplifies into:

$$I_d = \left[P^{1-\sigma}_{d,s}\kappa^{\sigma}_{d,s} + P^{1-\sigma}_{d,s'}\kappa^{\sigma}_{d,s'}\right]\left[\frac{\kappa_{d,s'}}{p_{d,s'}}\right]^{-\sigma} e_{d,s'}$$

Where $\sigma = \frac{1}{1-\rho}$

Solving for $e_{d,s'}$ and plugging back into the ratio of FOCs for each energy source, we obtain the following expression for $e^*_{d,s}$, our solar installation demand function:

$$e^*_{d,s}\left(P_{d,s}, P_{d,s'}, I_d\right) = \left(\frac{\kappa_{d,s}}{P_{d,s}}\right)^{\sigma}\frac{I_d}{\kappa^{\sigma}_{d,s'}P^{1-\sigma}_{d,s'} + \kappa^{\sigma}_{d,s}P^{1-\sigma}_{d,s}}$$

C.1.2 Demand for Energy-Sector Manufactured Inputs

In order to meet the optimal demands for energy sources e^*_s and $e^*_{s'}$, the grid-planner has to choose from the available manufactured varieties that aggregate into the final energy output (i.e. the grid planner chooses a set of solar panels to produce a solar park that meets their solar energy output requirements). The choice of manufactured inputs determines the prices P_s and $P_{s'}$. The derivations below are for a planner in any region d. For notational convenience, we omit the d spatial subscript until later in our derivations. A grid-planner can purchase manufactured varieties from any region o, which aggregates using a CES technology:

$$e_s = \left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{o,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{\sigma_s}{\sigma_s - 1}}$$

The problem of optimally delivering e_s^* is a new constrained optimisation problem, nested in the above, which we express as follows:

$$\min_{q_o(\omega)} \left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{o,s}(\omega) p_{o,s}(\omega) \right)$$

$$\text{s.t.} \quad \left(\sum_o \int_{\omega \in \Omega_{o,s}} q_{o,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{\sigma_s}{\sigma_s - 1}} = e_s^*$$

Note that the manufactured varieties in this problem are only sector s varieties (i.e. to generate solar output the planner only uses solar panels and not manufactured varieties belonging to other energy sectors). Below we detail all solution steps.

$$\mathcal{L} = -\left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{o,s}(\omega) p_{o,s}(\omega) \right) - \lambda \left[\left(\sum_o \int_{\omega \in \Omega_{o,s}} q_{o,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{\sigma_s}{\sigma_s - 1}} - e_s^* \right]$$

The FOCs (as many as solar panel varieties available across regions) are:

$$\frac{\partial \mathcal{L}}{\partial q_{o,s}(\omega)} = -p_{o,s}(\omega) - \lambda \frac{\sigma_s}{\sigma_s - 1} \left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{o,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{1}{\sigma_s - 1}} \frac{\sigma_s - 1}{\sigma_s} q_{o,s}(\omega)^{-\frac{1}{\sigma_s}} = 0$$

Which simplify as:

$$p_{o,s}(\omega) = -\lambda \left(\sum_o \int_{\omega \in \Omega_{0,s}} q_{o,s}(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega \right)^{\frac{1}{\sigma_s - 1}} q_{o,s}(\omega)^{-\frac{1}{\sigma_s}}$$

We can take the ratio of two FOC for two different varieties (within the solar sector):

$$\frac{p_{o,s}(\omega_1)}{p_{o,s}(\omega_2)} = \frac{q_{o,s}(\omega_1)^{-\frac{1}{\sigma_s}}}{q_{o,s}(\omega_2)^{-\frac{1}{\sigma_s}}}$$

This can be expressed as:

$$\frac{q_{o,s}(\omega_1)}{q_{o,s}(\omega_2)} = \left(\frac{p_{o,s}(\omega_1)}{p_{o,s}(\omega_2)} \right)^{-\sigma_s}$$

Which is equivalent to:

$$p_{o,s}(\omega_1)\, q_{o,s}(\omega_1) = \left(p_{o,s}(\omega_1)\right)^{1-\sigma_s} \left(p_{o,s}(\omega_2)\right)^{\sigma_s} q_{o,s}(\omega_2)$$

We now take the integral with respect to ω_1 and sum up across all origin regions o :

$$\underbrace{\sum_o \int_{\omega_1} p_{o,s}(\omega_1)\, q_{o,s}(\omega_1)\, d\omega_1}_{=E_s} = \sum_o \int_{\omega_1} \left(p_{o,s}(\omega_1)\right)^{1-\sigma_s} \left(p_{o,s}(\omega_2)\right)^{\sigma_s} q_{o,s}(\omega_2)\, d\omega_1$$

$$E_s = \left(p_{o,s}(\omega_2)\right)^{\sigma_s} q_{o,s}(\omega_2) \sum_o \int_{\omega_1} \left(p_{o,s}(\omega_1)\right)^{1-\sigma_s} d\omega_1$$

Where E_s is the expenditure on the solar sector (in any region d, where the spatial subscript has been omitted). We can generalise for any variety ω and express the demand function as:

$$q_{o,s}(\omega) = \frac{\left(p_{o,s}(\omega)\right)^{-\sigma_s}}{(P_s)^{1-\sigma_s}} E_s$$

Where P_s is the price index for solar that region d faces.

Replacing $E_s = e_s^* P_s$, we obtain our expression for the demand for each panel variety:

$$q_{o,s}(\omega) = \frac{\left(p_{o,s}(\omega)\right)^{-\sigma_s}}{(P_s)^{1-\sigma_s}} e_s^* P_s = \left(\frac{p_{o,s}(\omega)}{P_s}\right)^{-\sigma_s} e_s^* = \left(\frac{p_{o,s}(\omega)}{P_s}\right)^{-\sigma_s} \left(\frac{\kappa_s}{P_s}\right)^{\sigma} \frac{I}{\kappa_{s'}^{\sigma} P_{s'}^{1-\sigma} + \kappa_s^{\sigma} P_s^{1-\sigma}}$$

This applies to any region d, so we now generalise notation:

$$q_{od,s}(\omega) = \left(\frac{p_{od,s}(\omega)}{P_{d,s}}\right)^{-\sigma_s} \left(\frac{\kappa_{d,s}}{P_{d,s}}\right)^{\sigma} \frac{I_d}{\kappa_{d,s'}^{\sigma} P_{d,s'}^{1-\sigma} + \kappa_{d,s}^{\sigma} P_{d,s}^{1-\sigma}}$$

C.2 The Manufacturer Problem

Each region o has a continuum of potential manufacturing firms i in each sector s, which operate under monopolistic competition.

C.2.1 Manufacturing Technology

Firm i, who produces intermediate goods for electricity sector s (e.g. solar panels for the solar electricity sector), uses effective units of labour $L_{o,s,i}$, with unit cost $w_{o,s}$. Firm subscripts i

are dropped from now onwards for notational simplicity. To operate, a firm must pay a sunk cost $w_{o,s}f^{e}_{o,s}$, which we express in terms of effective units of labour. This sunk cost could be understood as the cost incurred in initial product definition and development. Upon paying the entry cost, the firm draws an initial level of productivity φ, from a Pareto productivity distribution, whose cumulative distribution function is:

$$G\left(\varphi; b_{o,s}\right) = 1 - \left(\frac{\varphi}{b_{o,s}}\right)^{-\theta_s}$$

That is, once a firm decides which product or variety to produce, it learns about its productivity. To produce $q_{o,s}(\varphi)$ units of a variety, a firm requires an amount of effective labor $l_{o,s} = f_{o,s} + \frac{q_{o,s}}{\varphi}$, where $f_{o,s}$ is the fixed cost of production, expressed in terms of effective units of labour, and $\frac{1}{\varphi}$ is the marginal cost of production.

Technological Upgrading/Innovation:

Upon observing its initial productivity φ, a firm can upgrade its technology (innovate), which increases the fixed cost of production: $\eta_{o,s}f_{o,s}$, with $\eta_{o,s} > 1$, but reduces its marginal cost, now: $\frac{1}{\xi_{o,s}\varphi}$, with $\xi_{o,s} > 1$

C.2.2 Firm Profits

A firm can make profits by selling the manufactured intermediate goods to grid planners in d regions. Among these regions there are Chinese second administrative level regions ('cities') and a foreign region $\tilde{d}$. We assume there are no market access fixed cost within China. On the other hand, a firm must pay an international exporting fixed cost $w_{o,s}f^{x}_{o,\tilde{d},s}$ if it wants to serve a representative foreign grid planner, whose demand function for energy sources coincides with that of each of the regional planners within China.

Trade (intra-China and international) is subject to iceberg trade costs such that in order for $q_{od,s}(\varphi)$ to arrive to destination d, a firm in o needs to produce $\tau_{od,s}q_{od,s}(\varphi)$ units of the variety, with $\tau_{od,s} \geq 1$. Trade costs are normalised, such that they are equal to 1 if and only if $d = o$. Firm profits after drawing a productivity φ are therefore:

$$\pi_{o,s}(\varphi) = \sum_{d} \left\{ p_{od,s}(\varphi)q_{od,s}(\varphi) - w_{o,s}\frac{\tau_{od,s}q_{od,s}(\varphi)}{\xi_{o,s}\varphi} \right\} - \mathbb{1}[d = \tilde{d}] \left(w_{o,s}f^{x}_{o,\tilde{d},s} \right) - w_{o,s}\eta_{o,s}f_{o,s}$$

Where $\mathbb{1}[d = \tilde{d}]$ takes the value 1 if a firm decides to serve foreign destination $\tilde{d}$ and 0 otherwise. Note that if a firm decides to export, the summation above is over Chinese regions and the foreign destination $\tilde{d}$, while if a firm chooses not to, it sums over national regions and does not pay the exporting fixed cots. Similarly, $\eta_{o,s}$ and $\xi_{o,s}$ are greater than 1 if a firm decides to innovate, and 1 otherwise.

Recall that the demand function for each manufactured variety, with our generalised notation, is:

$$q_{od,s}(\omega) = \frac{\left(p_{od,s}(\omega)\right)^{-\sigma_s}}{\left(P_{d,s}\right)^{1-\sigma_s}} E_{d,s}$$

Taking the FOC of firm profits with respect to $p_{od,s}(\varphi)$, and replacing the optimal $q_{od,s}(\omega)$ above, we obtain:

$$\frac{\partial \pi_{o,s}(\varphi)}{\partial p_{od,s}(\varphi)} = (1-\sigma_s)\frac{\left(p_{od,s}(\varphi)\right)^{-\sigma_s}}{\left(P_{d,s}\right)^{1-\sigma_s}} E_{d,s} + \frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi}\sigma_s \frac{\left(p_{od,s}(\varphi)\right)^{-\sigma_s-1}}{\left(P_{d,s}\right)^{1-\sigma_s}} E_{d,s} = 0$$

Which simplifies to:

$$p_{od,s}(\varphi) = \frac{\sigma_s}{\sigma_s - 1}\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi}$$

Thus, the optimal price is a constant markup over marginal cost, where the exporting and innovation decisions change the marginal cost of production. Substituting the optimal pricing and demand functions in the expression for firm profits, we obtain the potential value functions for each technology and exporting choice.

Domestic market only using old technology:

$$\pi_{o,s}(\varphi) = \sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f_{o,s}$$

Domestic and foreign market using old technology:

$$\pi_{o,s}(\varphi) = \sum_{d} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f^x_{o,\tilde{d},s} - w_{o,s} f_{o,s}$$

Domestic market only using new technology:

$$\pi_{o,s}(\varphi) = \sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s}\eta_{o,s} f_{o,s}$$

Domestic and foreign market using new technology:

$$\pi_{o,s}(\varphi) = \sum_{d} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f^x_{o,\tilde{d},s} - w_{o,s}\eta_{o,s} f_{o,s}$$

Assuming that the productivity thresholds governing these decisions are such that a firm does not innovate without exporting internationally first, firm optimal profits are:

$$\Pi_{o,s}(\varphi) = \max \left\{ \sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f_{o,s}, \right.$$
$$\sum_{d} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f^x_{o,\tilde{d},s} - w_{o,s} f_{o,s},$$
$$\left. \sum_{d} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} \right\} - w_{o,s} f^x_{o,\tilde{d},s} - w_{o,s}\eta_{o,s} f_{o,s} \right\}$$

C.2.3 Productivity Thresholds

We now calculate the productivity cutoffs that determine firms decisions to i) stay in the market after drawing a productivity, ii) access the international market $\tilde{d}$, and iii) innovate.

Domestic market exit threshold:

We define $\varphi^*_{oo,s}$ as the domestic market exit productivity threshold. This is the productivity that generates zero profits from serving the domestic market only.

$$\sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi^*_{oo,s}} \right)^{1-\sigma_s} \right\} - w_{o,s} f_{o,s} = 0$$

That is,

$$\varphi^*_{oo,s} = \left(\sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{w_{o,s} f_{o,s}} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d,s}} \right)^{1-\sigma_s} \right\} \right)^{\frac{1}{1-\sigma_s}}$$

Exporting threshold:

Let $\varphi^*_{o\tilde{d},s}$ describe the productivity level which makes a firm earn zero profits from exporting to foreign country $\tilde{d}$, and therefore indifferent between serving $\tilde{d}$ or limiting its supply to the domestic market. We also assume that the marginal exporting firm is using the old technology. The extra profits from serving $\tilde{d}$ are:

$$\pi_{o\tilde{d},s} = \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{o\tilde{d},s}}{\varphi} \right)^{1-\sigma_s} - w_{o,s} f^x_{o,\tilde{d},s}$$

$\varphi^*_{o\tilde{d},s}$ is the productivity level φ such that $\pi_{o\tilde{d},s}(\varphi) = 0$. That is:

$$
\begin{aligned}
\frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{o\tilde{d},s}}{\varphi^*_{o\tilde{d},s}} \right)^{1-\sigma_s} &= w_{o,s} f^x_{o,\tilde{d},s} \\
\Rightarrow \frac{(\sigma_s - 1)^{\sigma_s - 1}}{w_{o,s}^{\sigma_s}\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \left(\frac{\tau_{o\tilde{d},s}}{\varphi^*_{o\tilde{d},s}} \right)^{1-\sigma_s} &= f^x_{o,\tilde{d},s} \\
\Rightarrow \frac{1}{f^x_{o,\tilde{d},s}} \frac{(\sigma_s - 1)^{\sigma_s - 1}}{w_{o,s}^{\sigma_s}\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \tau^{1-\sigma_s}_{o\tilde{d},s} &= \left(\varphi^*_{o\tilde{d},s} \right)^{1-\sigma_s}
\end{aligned}
$$

Therefore:

$$
\varphi^*_{o\tilde{d},s} = \frac{\tau_{o\tilde{d},s}}{P_{\tilde{d},s}} \left(\frac{E_{\tilde{d},s}}{f^x_{o,\tilde{d},s}} \frac{(\sigma_s - 1)^{\sigma_s - 1}}{w_{o,s}^{\sigma_s}\sigma_s^{\sigma_s}} \right)^{\frac{1}{1-\sigma_s}}
$$

Innovation threshold:

Let $\varphi^i_{od,s}$ be the productivity level which makes a firm indifferent between upgrading its technology or not. We define profits using the old technology as low productivity profits, or $\pi_{o,s,l}$ and profits using the new technology as high productivity profits, or $\pi_{o,s,h}$. We assume that the marginal innovator is already exporting to the international market $\tilde{d}$.

$$
\pi_{o,s,l} = \sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \right\} + \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{o\tilde{d},s}}{\varphi} \right)^{1-\sigma_s} - \\
- w_{o,s} f^x_{o,\tilde{d},s}
$$

$$
\pi_{o,s,h} = \sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} \right\} + \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{\tilde{d},s}}{(P_{\tilde{d},s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{o\tilde{d},s}}{\xi_{o,s}\varphi} \right)^{1-\sigma_s} - \\
- w_{o,s} f^x_{o,\tilde{d},s} - w_{o,s}\eta_{o,s} f_{o,s}
$$

$\varphi^i_{od,s}$ therefore fulfils:

$$
\sum_d \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi^i_{od,s}} \right)^{1-\sigma_s} \right\} = \sum_d \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}\varphi^i_{od,s}} \right)^{1-\sigma_s} \right\} - w_{o,s}\eta_{o,s} f_{o,s}
$$

$$\Rightarrow \sum_d \frac{1-\xi_{o,s}^{1-\sigma_s}}{\xi_{o,s}^{1-\sigma_s}} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{(P_{d,s})^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\varphi^i_{od,s}}\right)^{1-\sigma_s} = w_{o,s}\eta_{o,s}f_{o,s}$$

$$\Rightarrow \sum_d \frac{1-\xi_{o,s}^{1-\sigma_s}}{\xi_{o,s}^{1-\sigma_s}} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{w_{o,s}\eta_{o,s}f_{o,s}} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d.s}}\right)^{1-\sigma_s} = \left(\varphi^i_{od,s}\right)^{1-\sigma_s}$$

The innovation threshold is therefore:

$$\varphi^i_{od,s} = \left(\sum_d \frac{1-\xi_{o,s}^{1-\sigma_s}}{\xi_{o,s}^{1-\sigma_s}} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{w_{o,s}\eta_{o,s}f_{o,s}} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d,s}}\right)^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}$$

Where recall, from the grid planner problem, we obtained:

$$E_{d,s} = \frac{I_d P_{d,s}^{1-\sigma}}{\left(\frac{\kappa_{d,s'}}{\kappa_{d,s}}\right)^{\sigma} P_{d,s'}^{1-\sigma} + P_{d,s}^{1-\sigma}}$$

In order to express the exporting and innovation thresholds as a function of the exit threshold, it is useful to regroup terms in each expression as follows (where we have removed the sectoral subscript s to simplify notation):

$$\varphi^*_o = \left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma w_o^\sigma}\right)^{\frac{1}{1-\sigma}} \left(\frac{1}{f_o}\right)^{\frac{1}{1-\sigma}} \left\{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}\right\}^{\frac{1}{1-\sigma}}$$

$$\varphi^*_{o\tilde{d}} = \left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma w_o^\sigma}\right)^{\frac{1}{1-\sigma}} \left(\frac{1}{f^x_{o\tilde{d}}}\right)^{\frac{1}{1-\sigma}} \left\{E_{\tilde{d}} \left(\frac{\tau_{o\tilde{d}}}{P_{\tilde{d}}}\right)^{1-\sigma}\right\}^{\frac{1}{1-\sigma}}$$

$$\varphi^i_o = \left(\frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma w_o^\sigma}\right)^{\frac{1}{1-\sigma}} \left(\frac{1-\xi_o^{1-\sigma}}{\xi_o^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} \left(\frac{1}{\eta_o f_o}\right)^{\frac{1}{1-\sigma}} \left\{\sum_d E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}\right\}^{\frac{1}{1-\sigma}}$$

We can now express the exporting and productivity thresholds as a function of the exit threshold:

$$\varphi^*_{o\tilde{d}} = \varphi^*_o \left(\frac{f_o}{f^x_{o\tilde{d}}}\right)^{\frac{1}{1-\sigma}} \left\{\frac{E_{\tilde{d}}\left(\frac{\tau_{o\tilde{d}}}{P_{\tilde{d}}}\right)^{1-\sigma}}{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}\right\}^{\frac{1}{1-\sigma}}$$

$$\varphi_o^i = \varphi_o^* \left(\frac{1-\xi_o^{1-\sigma}}{\eta_o \xi_o^{1-\sigma}} \right)^{\frac{1}{1-\sigma}} \left\{ \frac{\sum_d E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}} \right\}^{\frac{1}{1-\sigma}}$$

C.3 Industry Equilibrium

In equilibrium, grid-planners maximise utility from electricity services, manufacturers maximise profits, and labor demand equals labor supply in each region.

To determine the equilibrium price indices, number of firms, aggregate production and revenue, and mass of exporters and innovators in each region, we impose a free entry condition.

Free entry implies that the sunk entry costs equals expected profits from drawing a productivity:

$$w_o f_{o,s}^e = \left(1 - G\left[\varphi_{oo,s}^*\right]\right) \mathbb{E}\left[\pi \mid \varphi > \varphi_{oo,s}^*\right] \tag{9}$$

We can express the above condition as follows:

$$\begin{aligned} w_o f_{o,s}^e = \left(G\left[\varphi_{o\tilde{d},s}^*\right] - G\left[\varphi_{oo,s}^*\right]\right) \mathbb{E}\left[\pi_{o,s} \mid \varphi_{o\tilde{d},s}^* > \varphi > \varphi_{oo,s}^*\right] + \\ + \left(G\left[\varphi_{od,s}^i\right] - G\left[\varphi_{o\tilde{d},s}^*\right]\right) \mathbb{E}\left[\pi_{o,s} \mid \varphi_{od,s}^i > \varphi > \varphi_{o\tilde{d},s}^*\right] + \left(1 - G\left[\varphi_{od,s}^i\right]\right) \mathbb{E}\left[\pi_{o,s} \mid \varphi > \varphi_{od,s}^i\right] = \\ = \int_{\varphi_{oo,s}^*}^{\varphi_{od,s}^*} \pi_{o,s}(\varphi) g(\varphi) d\varphi + \int_{\varphi_{od,s}^*}^{\varphi_{od,s}^i} \pi_{o,s}(\varphi) g(\varphi) d\varphi + \int_{\varphi_{od,s}^i}^{\infty} \pi_{o,s}(\varphi) g(\varphi) d\varphi \end{aligned}$$

Replacing the expression for firm profits for reach range of productivities and the expression for the Pareto distribution function we obtain the following:

$$\begin{aligned} \frac{w_o f_{o,s}^e}{\theta_s b_{o,s}^{\theta_s}} = \int_{\varphi_{o,s}^*}^{\varphi_{od,s}^*} \left(\sum_{d\neq\tilde{d}} \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{\left(P_{d,s}\right)^{1-\sigma_s}} \left(w_{o,s}\tau_{od,s}\right)^{1-\sigma_s} \varphi^{\sigma_s - \theta_s - 2} \right) d\varphi - \int_{\varphi_{oo,s}^*}^{\varphi_{od,s}^i} \left(w_{o,s} f_{o,s} \varphi^{-\theta_s - 1}\right) d\varphi \\ + \int_{\varphi_{od,s}^*}^{\varphi_{od,s}^i} \left(\sum_{d} \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{\left(P_{d,s}\right)^{1-\sigma_s}} \left(w_{o,s}\tau_{od,s}\right)^{1-\sigma_s} \varphi^{\sigma_s - \theta_s - 2} \right) d\varphi - \int_{\varphi_{od,s}^*}^{\infty} \left(w_{o,s} f_{o,\tilde{d},s}^x \varphi^{-\theta_s - 1}\right) d\varphi \\ + \int_{\varphi_{od,s}^i}^{\infty} \left(\sum_{d} \frac{(\sigma_s - 1)^{\sigma_s - 1}}{\sigma_s^{\sigma_s}} \frac{E_{d,s}}{\left(P_{d,s}\right)^{1-\sigma_s}} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s}}\right)^{1-\sigma_s} \varphi^{\sigma_s - \theta_s - 2} \right) d\varphi - \int_{\varphi_{od,s}^i}^{\infty} \left(w_{o,s}\eta_{o,s} f_{o,s} \varphi^{-\theta_s - 1}\right) d\varphi \end{aligned}$$

Notice that the exit, exporting, and innovation productivity thresholds satisfy the following:

$$\left(\varphi^{*}_{oo,s}\right)^{1-\sigma_s} w_{o,s} f_{o,s} = \sum_{d\neq\tilde{d}} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} E_{d,s} \left(\frac{w_{o,s}\tau_{od,s}}{P_{d,s}}\right)^{1-\sigma_s}$$

$$\left(\varphi^{i}_{o,s}\right)^{1-\sigma_s} \frac{w_{o,s}\eta_{o,s} f_{o,s}}{1-\xi_{o,s}^{1-\sigma_s}} = \sum_{d} \frac{(\sigma_s-1)^{\sigma_s-1}}{\sigma_s^{\sigma_s}} E_{d,s} \left(\frac{w_{o,s}\tau_{od,s}}{\xi_{o,s} P_{d,s}}\right)^{1-\sigma_s}$$

$$\left(\varphi^{*}_{oo,s}\right)^{1-\sigma_s} - \frac{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}{1-\xi_{o,s}^{1-\sigma_s}} \left(\varphi^{i}_{o,s}\right)^{1-\sigma_s} = -\frac{f^{x}_{o\tilde{d},s}}{f_{o,s}} \left(\varphi^{*}_{o\tilde{d},s}\right)^{1-\sigma_s}$$

Replacing these equations into the above expression for the free entry condition, the latter simplifies to:

$$f^{e}_{o,s}\frac{\sigma_s-\theta_s-1}{\sigma_s-1} = \left(\frac{\sigma_s-\theta_s-1}{\sigma_s-1} - \eta_{o,s}\right)\left(\frac{b_{o,s}}{\varphi^{i}_{o,s}}\right)^{\theta_s} f_{o,s} - \left(\frac{b_{o,s}}{\varphi^{*}_{o\tilde{d},s}}\right)^{\theta_s} f^{x}_{o\tilde{d},s} - \left(\frac{b_{o,s}}{\varphi^{*}_{oo,s}}\right)^{\theta_s} f_{o,s} \tag{10}$$

Now, replacing the exporting and innovation thresholds with their expression as a function of the exit threshold we obtain the following expression for the exit threshold as a function of fundamentals and price indices:

$$\left(\varphi^{*}_{o,s}\right)^{\theta} = \frac{f_{o,s}}{f^{e}_{o,s}} b^{\theta}_{o,s} \frac{\sigma_s-1}{\sigma_s-\theta_s-1} \left\{\left(\frac{\sigma_s-\theta_s-1}{\sigma_s-1} - \eta_{o,s}\right)\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\theta_s}{\sigma_s-1}} \Phi_{o,s}^{-\theta_s} - \left(\frac{f^{*}_{o\tilde{d},s}}{f_{o,s}}\right)^{\frac{1-\sigma_s+\theta_s}{1-\sigma_s}} \Theta_{o,s}^{-\theta_s} - 1\right\}$$

Where

$$\Phi_{o,s} = \left\{\frac{\sum_d E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}\right\}^{\frac{1}{1-\sigma}} = \left\{\frac{\sum_d \frac{\kappa^{\sigma}_{d,s}\tau^{1-\sigma_s}_{od,s} I_d}{\kappa^{\sigma}_{d,s'}P^{1-\sigma}_{d,s'}+\kappa^{\sigma}_{d,s}P^{1-\sigma}_{d,s}}}{\sum_{d\neq\tilde{d}} \frac{\kappa^{\sigma}_{d,s}\tau^{1-\sigma_s}_{od,s} I_d}{\kappa^{\sigma}_{d,s'}P^{1-\sigma}_{d,s'}+\kappa^{\sigma}_{d,s}P^{1-\sigma}_{d,s}}}\right\}^{\frac{1}{1-\sigma}}$$

$$\Theta_{o,s} = \left\{\frac{E_{\tilde{d}} \left(\frac{\tau_{o\tilde{d}}}{P_{\tilde{d}}}\right)^{1-\sigma}}{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}\right\}^{\frac{1}{1-\sigma}} = \left\{\frac{\sum_d E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}}{\sum_{d\neq\tilde{d}} E_d \left(\frac{\tau_{od}}{P_d}\right)^{1-\sigma}} - 1\right\}^{\frac{1}{1-\sigma}} = \left(\Phi^{1-\sigma}-1\right)^{\frac{1}{1-\sigma}}$$

Replacing this expression for the exit threshold in the zero-profit condition that defines it, we get a system of equations (one for each domestic region d and sector s), which determines the price indices:

$$\sum_{d \neq \tilde{d}} \left\{ \frac{(\sigma_s - 1)^{\sigma_s - 1}}{w_{o,s} f_{o,s} \sigma_s^{\sigma_s}} \frac{\kappa_{d,s}^{\sigma} I_d}{\kappa_{d,s'}^{\sigma} P_{d,s'}^{1-\sigma} + \kappa_{d,s}^{\sigma} P_{d,s}^{1-\sigma}} \left(w_{o,s} \tau_{od,s} \right)^{1-\sigma_s} \right\} =$$

$$\left(\frac{f_{o,s}}{f_{o,s}^{e}} b_{o,s}^{\theta_s} \frac{\sigma_s - 1}{\sigma_s - \theta_s - 1} \left\{ \left(\frac{\sigma_s - \theta_s - 1}{\sigma_s - 1} - \eta_{o,s} \right) \left(\frac{1 - \xi_{o,s}^{1-\sigma_s}}{\eta_{o,s} \xi_{o,s}^{1-\sigma_s}} \right)^{\frac{\theta_s}{\sigma_s - 1}} \Phi^{-\theta_s} - \left(\frac{f_{o\tilde{d},s}^{*}}{f_{o,s}} \right)^{\frac{1-\sigma_s+\theta_s}{1-\sigma_s}} \Theta^{-\theta_s} - 1 \right\} \right)^{\frac{1-\sigma_s}{\theta_s}}$$

Note that if there are no exports the exit threshold simplifies to:

$$\left(\varphi_{oo,s}^{*} \right)^{\theta_s} = \frac{f_{o,s}}{f_{o,s}^{e}} \theta_s b_{o,s}^{\theta_s} \left(-\frac{1}{\sigma_s - \theta_s - 1} - \frac{1}{\theta_s} \right) \left(\left(\phi_{o,s} \eta_{o,s} \right)^{\frac{\theta_s + 1 - \sigma_s}{1-\sigma_s}} \left(\frac{\xi_{o,s}^{1-\sigma_s}}{1 - \xi_{o,s}^{1-\sigma_s}} \right)^{\frac{\theta_s}{1-\sigma_s}} + 1 \right)$$

C.4 Aggregate variables

In this section we derive expressions for some of our remaining aggregate (city-level) variables of interest.

C.4.1 Mass of firms

The price index in each region d for each sector s satisfies:

$$P_{d,s}^{(1-\sigma_s)} = \sum_{o \neq \tilde{d}} \int_0^{M_{od,s}} p_{od,s}(v)^{1-\sigma_s} dv \tag{11}$$

Note that foreign firms do not serve the domestic market, so the price aggregation is only over domestic suppliers. Also note that firms that are active serve every regional market within China. We can express it as follows:

$$P_{d,s}^{(1-\sigma_s)} = \sum_{o \neq \tilde{d}} M_{od,s} \int_{\varphi_{oo,s}^{*}}^{\infty} p_{od,s}(\varphi)^{1-\sigma_s} \frac{g(\varphi)}{1 - G(\varphi_{oo,s}^{*})} d\varphi = \sum_{o \neq \tilde{d}} M_{od,s} \int_{\varphi_{oo,s}^{*}}^{\infty} p_{od,s}(\varphi)^{1-\sigma_s} \theta_s (\varphi_{oo,s}^{*})^{\theta_s} \varphi^{-\theta_s - 1} d\varphi$$

$$= \sum_{o \neq \tilde{d}} M_{od,s} \theta_s (\varphi_{oo,s}^{*})^{\theta_s} \left(\int_{\varphi_{oo,s}^{*}}^{\varphi_{od,s}^{i}} p_{od,s}(\varphi)^{1-\sigma_s} \varphi^{-\theta_s - 1} d\varphi + \int_{\varphi_{od,s}^{i}}^{\infty} p_{od,s}(\varphi)^{1-\sigma_s} \varphi^{-\theta_s - 1} d\varphi \right)$$

$$= \sum_{o \neq \tilde{d}} M_{od,s} \theta_s (\varphi_{oo,s}^{*})^{\theta_s} \left(\int_{\varphi_{oo,s}^{*}}^{\varphi_{od,s}^{i}} \left(\frac{\sigma_s}{\sigma_s - 1} \frac{w_{o,s} \tau_{od,s}}{\varphi} \right)^{1-\sigma_s} \varphi^{-\theta_s - 1} d\varphi + \int_{\varphi_{od,s}^{i}}^{\infty} \left(\frac{\sigma_s}{\sigma_s - 1} \frac{w_{o,s} \tau_{od,s}}{\xi_{o,s} \varphi} \right)^{1-\sigma_s} \varphi^{-\theta_s - 1} d\varphi \right)$$

Integrating and simplifying, we obtain:

$$P_{d,s}^{(1-\sigma_s)} = \sum_{o \neq \tilde{d}} \frac{M_{od,s}\theta_s}{\sigma_s - \theta_s - 1} \left(\varphi_{oo,s}^*\right)^{\theta_s} \left(\frac{w_{o,s}\tau_{od,s}\sigma_s}{\sigma_s - 1}\right)^{1-\sigma_s} \left(\frac{\xi_{o,s}^{1-\sigma_s} - 1}{\xi_{o,s}^{1-\sigma_s}} \left(\varphi_{od,s}^i\right)^{\sigma_s - \theta_s - 1} - \left(\varphi_{oo,s}^*\right)^{\sigma_s - \theta_s - 1}\right)$$

The mass of active firms in each region is related to the mass of entrants in each region in the following way:

$$M_{od,s} = \left(1 - G\left(\varphi_{oo,s}^*\right)\right) M_{o,s}^e = \frac{\left(b_{o,s}\right)^{\theta_s}}{\left(\varphi_{oo,s}^*\right)^{\theta_s}} M_{o,s}^e \tag{12}$$

The expression for the price index therefore becomes:

$$P_{d,s}^{(1-\sigma_s)} = \sum_{o \neq \tilde{d}} \frac{b_{o,s}^{\theta_s} M_{o,s}^e \theta_s}{\sigma_s - \theta_s - 1} \left(\frac{w_{o,s}\tau_{od,s}\sigma_s}{\sigma_s - 1}\right)^{1-\sigma_s} \left(\frac{\xi_{o,s}^{1-\sigma_s} - 1}{\xi_{o,s}^{1-\sigma_s}} \left(\varphi_{od,s}^i\right)^{\sigma_s - \theta_s - 1} - \left(\varphi_{oo,s}^*\right)^{\sigma_s - \theta_s - 1}\right)$$

We thus obtain an expression relating the mass of entrants, the price index and the exit threshold:

$$P_{d,s}^{(1-\sigma_s)} = \sum_{o \neq \tilde{d}} \frac{b_{o,s}^{\theta_s} M_{o,s}^e \theta_s}{\sigma_s - \theta_s - 1} \left(\frac{w_{o,s}\tau_{od,s}\sigma_s}{\sigma_s - 1}\right)^{1-\sigma_s} (\varphi_o^*)^{\sigma_s - \theta_s - 1} \left(\Phi^{\sigma_s - \theta_s - 1} \eta_o^{\frac{\sigma_s - \theta_s - 1}{\sigma_s - 1}} \left(\frac{\xi_o^{1-\sigma_s}}{\xi_o^{1-\sigma_s} - 1}\right)^{\frac{\theta_s}{1-\theta_s}} - 1\right)$$

C.4.2 Production

We can define aggregate city-level production for city o as:

$$\begin{aligned}
Q_o = \sum_d \int_{\omega \in \Omega_{od,s}} q_{od,s}(\omega) d\omega &= \sum_d \int_0^{M_{od,s}} q_{od,s}(v) dv \\
= \sum_{d \neq \tilde{d}} M_{od,s} \int_{\varphi_{oo,s}^*}^{\infty} q_{od,s} \frac{g(\varphi)}{1 - G(\varphi_{oo,s}^*)} d\varphi &+ M_{o\tilde{d},s} \int_{\varphi_{o\tilde{d},s}^*}^{\infty} q_{o\tilde{d},s} \frac{g(\varphi)}{1 - G(\varphi_{o\tilde{d},s}^*)} d\varphi \\
= \sum_{d \neq \tilde{d}} M_{od,s} \int_{\varphi_{oo,s}^*}^{\varphi_{o,s}^i} q_{od,s}\theta_s(\varphi_{oo,s}^*)^{\theta_s} \varphi^{-\theta_s - 1} d\varphi &+ \sum_{d \neq \tilde{d}} M_{od,s} \int_{\varphi_{o,s}^i}^{\infty} q_{od,s}\theta_s(\varphi_{oo,s}^*)^{\theta_s} \varphi^{-\theta_s - 1} d\varphi \\
+ M_{o\tilde{d},s} \int_{\varphi_{o\tilde{d},s}^*}^{\varphi_{o,s}^i} q_{o\tilde{d},s}\theta_s(\varphi_{o\tilde{d},s}^*)^{\theta_s} \varphi^{-\theta_s - 1} d\varphi &+ M_{o\tilde{d},s} \int_{\varphi_{o,s}^i}^{\infty} q_{o\tilde{d},s}\theta_s(\varphi_{o\tilde{d},s}^*)^{\theta_s} \varphi^{-\theta_s - 1} d\varphi
\end{aligned}$$

Which is equivalent to:

$$
\begin{aligned}
B_{o,s}Q_o = & \sum_{d\neq\tilde{d}} M_{od,s}\frac{E_{d,s}}{P_{d,s}^{1-\sigma_s}}\left(\tau_{od,s}\right)^{-\sigma_s}\left(\varphi^*_{oo,s}\right)^{\sigma_s}\left(\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\sigma_s-\theta_s}{1-\sigma_s}}\Phi_{o,s}^{\sigma_s-\theta_s}-1\right)\\
& -\sum_{d\neq\tilde{d}} M_{od,s}\frac{E_{d,s}}{P_{d,s}^{1-\sigma_s}}\left(\frac{\tau_{od,s}}{\xi_{o,s}}\right)^{-\sigma_s}\left(\varphi^*_{oo,s}\right)^{\sigma_s}\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\sigma_s-\theta_s}{1-\sigma_s}}\Phi_{o,s}^{\sigma_s-\theta_s}\\
& + M_{o\tilde{d},s}\frac{E_{\tilde{d},s}}{P_{\tilde{d},s}^{1-\sigma_s}}\left(\tau_{o\tilde{d},s}\right)^{-\sigma_s}\left(\varphi^*_{oo,s}\right)^{\sigma_s}\left(\frac{f_{o,s}}{f^x_{o\tilde{d},s}}\right)^{\frac{\theta_s}{1-\sigma_s}}\left(\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\sigma_s-\theta_s}{1-\sigma_s}}\Theta_{o,s}^{\theta_s}\Phi_{o,s}^{\sigma_s-\theta_s}-1\right)\\
& - M_{o\tilde{d},s}\frac{E_{\tilde{d},s}}{P_{\tilde{d},s}^{1-\sigma_s}}\left(\frac{\tau_{o\tilde{d},s}}{\xi_{o,s}}\right)^{-\sigma_s}\left(\varphi^*_{oo,s}\right)^{\sigma_s}\left(\frac{f_{o,s}}{f^x_{o\tilde{d},s}}\right)^{\frac{\theta_s}{1-\sigma_s}}\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\sigma_s-\theta_s}{1-\sigma_s}}\Theta_{o,s}^{\theta_s}\Phi_{o,s}^{\sigma_s-\theta_s}
\end{aligned}
$$

Where $B_{o,s} = (\sigma_s - \theta_s)\left(\frac{\sigma_s}{\sigma_s-1}w_{o,s}\right)^{\sigma_s}\frac{1}{\theta_s}$

Without foreign market the expression for aggregate production simplifies to:

$$
\begin{aligned}
\frac{B_{o,s}}{C_{o,s}}Q_o = & \sum_{d\neq\tilde{d}} M_{od,s}\frac{\kappa_{d,s}^{\sigma}I_d}{\kappa_{d,s'}^{\sigma}P_{d,s'}^{1-\sigma}+\kappa_{d,s}^{\sigma}P_{d,s}^{1-\sigma}}\left(\tau_{od,s}\right)^{-\sigma_s}\left(\left(\eta_{o,s}\right)^{\frac{\theta_s+1-\sigma_s}{1-\sigma_s}}\left(\frac{\xi_{o,s}^{1-\sigma_s}}{1-\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\theta_s}{1-\sigma_s}}+1\right)^{\frac{\sigma_s}{\theta_s}}\\
\times & \left(\left(\frac{1-\xi_{o,s}^{1-\sigma_s}}{\eta_{o,s}\xi_{o,s}^{1-\sigma_s}}\right)^{\frac{\sigma_s-\theta_s}{1-\sigma_s}}\left(1-\xi_{o,s}^{1-\sigma_s}\right)-1\right)
\end{aligned}
$$

Where $C_{o,s} = \left(\frac{f_{o,s}}{f^e_{o,s}}\theta_s b_{o,s}^{\theta_s}\left(-\frac{1}{\sigma_s-\theta_s-1}-\frac{1}{\theta_s}\right)\right)^{\frac{\sigma_s}{\theta_s}}$

D Further Results

D.1 Solar production results

Table D.1: Solar production, capacity and firm counts

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
Panel production	2.140***	0.705**	2.513***	3.078***
	(0.471)	(0.341)	(0.525)	(0.702)
Cell production	1.831***	1.298*	2.024***	2.455**
	(0.592)	(0.664)	(0.707)	(1.010)
Cell capacity	1.928***	1.310*	2.066**	2.322*
	(0.672)	(0.709)	(0.842)	(1.197)
Panel firm counts	0.558***	0.146	0.677***	0.806***
	(0.125)	(0.109)	(0.140)	(0.184)
Cell firm counts	0.380**	0.229	0.422**	0.540**
	(0.152)	(0.213)	(0.183)	(0.262)
Observations	3,580	3,580	3,580	3,580

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time: 2004-2013. Each column is one sdid regression. Without controls. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

D.2 Exports results

Table D.2: Exports: number of exporters and non-solar exports

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
Exporters firm count	0.220**	0.046	0.314***	0.400**
	(0.095)	(0.107)	(0.107)	(0.167)
Non solar export value	1.388	−0.736	3.094***	3.560**
	(0.924)	(0.979)	(1.026)	(1.641)
Non solar export volume	1.286*	−0.478	2.423**	2.864*
	(0.766)	(0.654)	(0.948)	(1.579)
Non solar export price	0.786*	−0.308	1.518***	1.824***
	(0.458)	(0.537)	(0.436)	(0.630)

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time: 2004-2016. Each column is one sdid regression. Without controls. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

D.3 Cross-city spillovers

Table D.3: Cross-city spillovers

	(1) All patents	(2) Firm count	(3) Revenue	(4) Panel capacity
Any subsidy in an adjacent city	0.373***	0.099	0.485***	0.385
	(0.096)	(0.055)	(0.177)	(0.263)
Observations	5,049	5,049	4,768	3,210

Notes: * 0.1 ** 0.05 *** 0.01. Dependent variables are reported in columns. Each observation is an admin2 level region and there are 358 admin2 regions in China. This sample here is restricted by dropping the 43 regions that have been treated directly by any subsidy. From the remaining regions, 103 cities' neighbours received any kind of subsidy. Time: 2004-2013 for panel capacity, 2004-2019 for patents, firm count and revenues. Each column is one sdid regression. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

D.4 City-level total solar patents

Table D.4: City-level total solar patents

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
Patent	0.444***	0.114	0.662***	1.029***
	(0.150)	(0.138)	(0.213)	(0.219)
Observations	6,086	6,086	6,086	6,086

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time: 2004-2020. Each column is one sdid regression. Without controls. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

D.5 Learning by doing patents

Table D.5: Learning-by-doing patents

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
Patent	0.365**	0.187	0.604***	0.914***
	(0.149)	(0.186)	(0.235)	(0.377)
Observations	5,728	5,728	5,728	5,728

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region and there are 358 admin2 regions in China. 43 regions are treated by any subsidy. Time: 2004-2020. Each column is one sdid regression. Without controls. The coefficient is the ATT which averages the staggered treatment effect. All regressions without controls

D.6 Productivity Analysis

Table D.6: PRODUCTIVITY OUTCOMES

Panel A	(1)	(2)	(3)	(4)
Period: 2004-2019	*Any subsidy*	*Demand subsidy*	*Production subsidy*	*Innovation subsidy*
Revenue	1.100**	0.190	1.887**	2.670**
	(0.456)	(0.198)	(0.767)	(1.193)
Labor	0.859**	0.249	1.443**	1.832*
	(0.435)	(0.291)	(0.664)	(1.034)
Capital	0.609	−0.130	1.302*	1.858
	(0.408)	(0.198)	(0.767)	(1.193)
Observations	5,728	5,728	5,728	5,728
Panel B	(1)	(2)	(3)	(4)
Period: 2004-2013	*Any subsidy*	*Demand subsidy*	*Production subsidy*	*Innovation subsidy*
Revenue	1.926**	0.285	2.392**	3.058**
	(0.767)	(0.193)	(0.944)	(1.517)
Panel production capacity	2.098***	0.587	2.496***	2.930***
	(0.532)	(0.467)	(0.575)	(0.773)
Labor	1.382**	0.523	1.581*	1.773
	(0.677)	(0.442)	(0.848)	(1.188)
Capital	1.470**	0.310	1.784**	2.307
	(0.711)	(0.282)	(0.905)	(1.426)
Observations	3,580	3,580	3,580	3,580

Notes: *0.1 ** 0.05 *** 0.01. Each observation is a city (admin2 level region) and there are 358 cities in China. 43 cities are treated by a subsidy. The time period of panel A is 2004-2019, and 2004-2013 for panel B. Each column contains one Synthetic Difference In Differences (SDID) estimate of the Average Treatment of the Treated (ATT), which averages the staggered treatment effects across all cohorts (years in which there were solar policies). Column (1) has any solar policy, column (2) the demand (installation) subsidies, column (3) production subsidies and column (4) innovation subsidies. Bootstrapped standard errors below the ATT. All regressions without controls.

D.7 With Controls

Table D.7: Controlling for GDP per capita

	(1) *Any subsidy*	(2) *Demand subsidy*	(3) *Production subsidy*	(4) *Innovation subsidy*
All patent	0.483**	0.226	0.867***	1.001***
	(0.205)	(0.242)	(0.220)	(0.341)
□ Design patents	0.187	0.275	0.240	0.141
	(0.132)	(0.190)	(0.167)	(0.254)
□ Invention/utility model patents	0.527**	0.191	0.960***	1.051***
	(0.213)	(0.241)	(0.232)	(0.361)
• Solar patents	0.523***	0.247	0.802***	0.875***
	(0.191)	(0.230)	(0.204)	(0.339)
• Non-solar patents	0.254	−0.061	0.739***	0.801**
	(0.182)	(0.215)	(0.217)	(0.349)
Firm count	0.210***	0.030	0.380***	0.396***
	(0.081)	(0.031)	(0.125)	(0.138)
Revenue	1.076**	0.170	1.882***	2.557***
	(0.458)	(0.205)	(0.727)	(1.102)
Panel capacity	2.025***	0.531	2.415***	2.848***
	(0.466)	(0.428)	(0.470)	(0.705)
Export value	2.409***	0.577	3.210**	4.041**
	(0.886)	(1.009)	(1.292)	(1.992)
Export volume	2.066**	0.038	2.841**	3.726**
	(0.812)	(0.699)	(1.208)	(1.851)
Export price	0.925**	0.176	1.078**	1.354
	(0.407)	(0.483)	(0.534)	(0.896)
Solar export value	4.515***	1.367*	6.250***	8.967***
	(0.970)	(0.741)	(1.428)	(2.136)
Solar export volume	3.848***	0.905	5.120***	7.231***
	(0.864)	(0.688)	(1.251)	(1.803)
Solar export price	1.485***	0.134	2.001***	3.186***
	(0.422)	(0.379)	(0.665)	(0.833)

Notes: * 0.1 ** 0.05 *** 0.01. Each observation is an admin2 level region. Here we control GDP per capita, and this data is available for 284 cities (update: available for 314 cities now). 43 regions are treated by any subsidy. Time: 2004-2020. Each coefficient is one sdid regression. The coefficient is the ATT which averages the staggered treatment effect.