I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14852

Misappropriation of R&D Subsidies: Estimating Treatment Effects with One-Sided Noncompliance

Philipp Boeing
Bettina Peters

NOVEMBER 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14852

Misappropriation of R&D Subsidies: Estimating Treatment Effects with One-Sided Noncompliance

Philipp Boeing
ZEW, IZA, Peking University, CCER and Tsinghua University

Bettina Peters
ZEW, University of Luxembourg and MaCCI

NOVEMBER 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Misappropriation of R&D Subsidies: Estimating Treatment Effects with One-Sided Noncompliance*

In evaluating the effectiveness of R&D subsidies, the literature has focused on potential crowding out effects, while the possibility of misappropriation of public funds that results from moral hazard behavior has been completely neglected. This study develops a theoretical framework with which to identify misappropriation. Using Chinese firm-level data for the period 2001-2011, we show that misappropriation is a major threat. 42% of grantees misused R&D subsidies for non-research purposes, accounting for 53% of the total amount of R&D subsidies. In a second step, we study the loss of effectiveness of R&D subsidies in stimulating R&D expenditures that is due to misappropriation. We measure the loss in effectiveness by estimating the causal effect of R&D subsidies in the presence of misappropriation using an intention-to-treat (ITT) estimator and comparing it to the ideal situation (without misappropriation) using the complier average causal effect (CACE). We find that China's R&D policy could have been more than twice as effective in boosting R&D without misappropriation. R&D expenditures could have been stimulated beyond the subsidy amount (additionality), but noncompliant behavior has resulted in a moderately strong partial crowding out effect. We find significant treatment heterogeneity by period, subsidy size, industry, and ownership. Notably, the loss in effectiveness has diminished following a policy reform in 2006. Nevertheless, the misappropriation of public funds considerably undermines the impact of R&D policies in China.

JEL Classification: O31, O38, C21, H21

Keywords: R&D subsidies, misappropriation, China, moral hazard, policy evaluation

Corresponding author:
Philipp Boeing
ZEW
L 7, 1
68161 Mannheim
Germany

E-mail: philipp.boeing@zew.de

* We are grateful to Loren Brandt, Maryann Feldman, Dietmar Harhoff, Zhijian Hu, Tomohiko Inui, Jordi Jaumandreu, Xibao Li, Georg Licht, Xielin Liu, Pierre Mohnen, Kazuyuki Motohashi, Kyle Myers, Barry Naughton, Emily Oster, Marcin Piatkowski, Mark Roberts, Mathias Sinning, Pete Suttmeier, John Van Reenen, Rainer Widmann, Martin Wörter, Miaojie Yu and Xiaobo Zhang for their helpful comments and discussions. We thank participants at the DRUID (Copenhagen, 2021), VfS Annual Meeting (Regensburg, 2021), EU-SPRI (Oslo, 2021), EARIE (Barcelona, 2019), IIOC (Boston, 2019), ZEW/MaCCI Conference on the Economics of Innovation and Patents (Mannheim, 2019), Munich Summer Institute (Munich, 2019), Concordi (Sevilla, 2019) and the 40th Anniversary of Science and Technology Reform in China Symposium (Beijing, 2018) as well as participants at seminars at CBS, University of Wuerzburg, Goethe University Frankfurt, Fraunhofer ISI, Peking University, Max Planck Institute Munich, Tokyo University, Gakushuin University, ZEW, Central University of Finance and Economics, Chinese Academy of Science and Technology for Development, Tsinghua University, University of Chinese Academy of Science, and University of Luxembourg. Any errors are those of the authors.

"Only around forty percent of China's research funds are used for research, whereas huge amounts trickle away."[1]

1 Introduction

Most countries offer public funding for research and development (R&D) to spur innovation in firms. The main argument for public R&D subsidies is a suboptimal low level of R&D that is due to a market failure caused by spillovers and financial constraints. A general concern is that public funding is not effective because it might simply crowd out private financing of R&D. Therefore, an increasing literature has evaluated the effectiveness of R&D policies empirically by estimating the treatment effect of R&D subsidies.[2] However, moral hazard behavior of firms is a second threat to the effectiveness of R&D policy, as Takalo et al. (2013) point out. Moral hazard occurs if a firm, after getting an R&D subsidy, decides to misuse public R&D funds for purposes other than those for which they were granted. In general, misappropriation or noncompliance is expected to be more likely under a weak monitoring regime, such as when detecting noncompliance is unreasonably expensive or governments fail to install effective monitoring mechanisms to identify noncompliant behavior. For example, firms that apply for grants from the U.S. Small Business Innovation Research Program indicate they will use the grant for R&D in their applications, but there is no monitoring or enforcement once the firms received the lump sum (Howell 2017). However, sound empirical evidence on whether and to what extent firms misappropriate public R&D funds for non-research purposes and how misappropriation impacts the effectiveness of R&D policy is missing.

This study is the first to address the misappropriation of R&D subsidies, identify it and investigate the consequences of such noncompliant behavior for the effectiveness of R&D policy in stimulating firms' R&D expenditures. The study develops a theoretical framework with which to measure misappropriation based on data that is usually available to researchers. Our empirical application uses Chinese firm-level data for the 2001-2011 period. China is an attractive case for answering our research questions because the Chinese State Council wants the country to increase its innovativeness and become a world leader in science and technology (S&T) by 2050. The 2001-2011 time period is covered by China's 10th and 11th Five-Year S&T Development Plans and, after 2006, was covered by the Mid- to Long-term S&T Development Plan (MLP) which fundamentally realigned China's innovation policy. A major target of the MLP is to increase R&D expenditures by private and state-owned domestic firms. While eligibility criteria differ by R&D program, the overarching emphasis on supporting (high) technology-oriented and innovative firms highlights the MLP's picking-the-winner strategy over an aiding-the-poor strategy. Between 2001 and 2011, the annual amount of funding directed to large- and medium-sized firms tripled from 5 billion RMB to 15 billion RMB, while R&D expenditures increased

[1] This statement is quoted from the China Youth Daily (31st August 2011) and was widely reprinted in domestic and international media outlets.

[2] For a recent survey, see Zuniga-Vicente et al. (2014).

more than sevenfold from 51 billion RMB to 366 billion RMB.[3]

A second reason for our study's focus on China is that misappropriation of R&D subsidies is a major concern in China. The increasing allocation of government funds has been accompanied by deficiencies in funding assignment and monitoring. In many programs, firms proposed to use the grants for R&D in their applications, but in practice there was little monitoring or enforcement after the funds were received, which allows for moral hazard behavior (Cao et al. 2013).[4] This problem has already been identified and addressed in the MLP, as it not only calls for more R&D funding but also for better management of R&D programs, selection and monitoring of grantees, and coordination between programs and agencies to reduce double funding of R&D projects and misallocation and misuse of public funds. In September 2011, public interest was sparked by media reports stating that around 60% of public research funds were misused for non-research purposes. Subsequent investigations by the Ministry of S&T and the Central Commission for Discipline Inspection found that bureaucrats of R&D programs, intermediaries who specialize in subsidy applications, and firms (as final recipients) were involved in misappropriation of R&D subsidies. Confirmation of this anecdotal evidence of substantial misappropriation of R&D subsidies based on a large-scale empirical analysis is the first intriguing finding of this study. We calculate that about 42% of grantees have misappropriated funds, corresponding to 53% of the total amount of R&D subsidies granted. We find three additional stylized facts in our data regarding misappropriation: First, firms either choose (almost) full misappropriation or not to misappropriate any funds which may be rationalized by indivisibilities of R&D projects. Second, we find a substantial decline in misappropriation over time, from 81% (2001) to 18% (2011). This decline emerges especially after 2006, coinciding with the implementation of the MLP. Third, misappropriation is not random but, in line with theory, can be explained by the R&D subsidy level, private internal funds, the rate of return to R&D, sanctions, and the probability of detection.

In addition to developing a theoretical framework with which to measure misappropriation, our second novel contribution is to take firms' noncompliant behavior into account when identifying the causal effect of R&D subsidies on private R&D expenditures. In our data, noncompliance can occur only in firms that receive R&D subsidies (the assigned treatment). Since the data set contains information on all types of R&D subsidies, we can rule out that non-assigned firms do not comply and somehow get a treatment. Therefore, in our application, one-sided noncompliance reflects moral hazard behavior in firms after they receive the grants. R&D policy evaluation studies for China (or any other country) have not accounted for this type of misappropriation of R&D funds.[5] In contrast, Chen

[3]See Figure A3 in the Online Appendix 1.

[4]Online Appendix 1 provides a more detailed description of the institutional background of R&D policy and misappropriation of R&D subsidies in China.

[5]Despite the growing importance of R&D and R&D policy in China, only a few studies have evaluated the causal effect of Chinese R&D subsidies on private R&D expenditures. Empirical evidence so far suggests that the effectiveness of grants increased with the introduction of the MLP in 2006, turning from partial crowding out in the pre-2006 period (Boeing 2016) to additionality in the post-2006 period (Liu et al. 2016; Hu and Deng 2018). However, the finding that R&D policy has become more effective has been established only for high-tech firms and private firms, so the average treatment effect for the overall

et al. (2021) deal with one-sided noncompliance that is the result of adverse selection before the granting decision. They investigate the effects of China's so-called InnoCom program, which awards corporate tax cuts to firms if they increase their R&D intensity above a given threshold. They find that many firms re-label non-R&D expenses as R&D to qualify for the treatment (tax cut) and estimate that almost a quarter of the reported R&D investment is due to re-labeling.

Traditionally, the R&D policy evaluation literature focuses on addressing the selection bias created by the fact that R&D subsidies are not randomly allocated to firms and, even in the counterfactual absence of a treatment, the selected treatment group would usually have higher R&D expenditures than the control group. This difference introduces an upward bias in the estimated effect of the R&D subsidy, and several estimators – such as matching, IV, (conditional) difference-in-differences, and regression discontinuity design – are used to correct for this bias. However, even in an initially ideal setting with a randomized R&D subsidy allocation, noncompliance to funding contract rules usually creates an additional source of selection bias that the standard estimators do not address: The bias arises because supported firms deliberately decide whether to comply or not based on comparing the expected outcome of using the funds for research purposes with the expected outcome from alternative uses. Therefore, one-sided noncompliance requires us to differentiate between the causal impacts of the *assigned* and *actual* treatment. Imbens and Angrist (1994) show that, with *randomized* assignment to treatment, the two effects can be consistently estimated by the intention-to-treat (ITT) effect and the complier average causal effect (CACE). To account for the selection in the R&D subsidy allocation and to mimic an (almost) randomized experiment for grant assignment, we suggest using entropy balancing, a method Hainmueller (2012) proposes, as a first design step in estimating ITT. Self-selection into compliance is then tackled using an IV strategy to estimate the CACE consistently. Identification is based on using the randomized assignment of the ITT as an instrument for the actual treatment (Bloom 1984). From an economic perspective, the ITT shows how effective the R&D policy is in the presence of misappropriation. The CACE, in contrast, shows how effective the policy could have been without misappropriation. While the ITT indicates the *effectiveness*, CACE is a measure for the *efficacy* of the R&D subsidy policy. Both are informative for policymakers. For example, if R&D subsidies fail to induce additional R&D investment, the ITT and CACE can help them understand whether the failure originates from flaws in the design or the implementation of policies.

Our causal analysis reveals four important insights. First, we find a medium-level partial crowding out for ITT, which shows that, taking into account the existing misappropriation behavior, R&D subsidies have increased total R&D expenditures, but by less than the subsidy amount. Second, and most salient, the impact would have been more than twice as large in the absence of misappropriation, as indicated by the CACE, which suggests an increase in total R&D expenditures beyond the subsidy amount (additionality). Taken

population of firms may have been overestimated. In addition, Cheng et al. (2019) investigate output additionality and find that subsidized Chinese firms file more domestic but not international patents and do not yield higher productivity, profits and markups.

together, ITT and CACE show that the design of the R&D policy in China works in principle, but that a better monitoring is advisable to fully exploit the policy's potential. Third, we find significant treatment heterogeneity by period, subsidy size, industry, and type of ownership. In particular, both effectiveness and efficacy significantly improved after the MLP was implemented in 2006, whereas both misappropriation and policy design had rendered R&D subsidies ineffective before that. But still, misappropriation of R&D subsidies considerably undermines the efficacy of Chinese R&D programs. Fourth, the results show output additionality but not behavioral additionality for a number of indicators.

The next section presents the data sources used in the empirical analysis. Section 3 explains our theoretical framework for identifying misappropriation in the data and provides several exercises that validate our measure. Section 4 presents the general identification strategy for estimating the causal effects of R&D subsidies with one-sided noncompliance using ITT and CACE, while section 5 explains our empirical implementation strategy. Our empirical results are presented in section 6 and section 7 provides concluding remarks.

2 Data

Our analysis is based on a data set of all domestic firms listed on the stock exchanges in Shanghai and Shenzhen between 2001 and 2011.[6,7] Because of government stock issuance quotas, the sample consists mainly of large and medium-sized domestic firms from manufacturing industries and the coastal region. Other industries and inland regions are less represented. The balance sheet information is compiled from COMPUSTAT and DATASTREAM and the Chinese databases CSMAR, RESSET, and WIND.

Our two key variables are R&D expenditures and R&D subsidies. Data on R&D expenditures comes from the Chinese database WIND. R&D expenditures consist of the sum of directly expensed outlays (for R&D that is not eligible for capitalization) and the capitalized amount (part of the development costs that are eligible for capitalization, following new accounting standards in 2007). As the coverage of R&D expenditures is incomplete in WIND before 2006, we collected R&D expenditures from the universe of annual reports via the Chinese CNINFO database.

China's Accounting Standards define subsidies as monetary or non-monetary assets obtained from the government, excluding capital investments undertaken by the government as a partial owner of the firm. The total amount of subsidies can be observed in

[6] Only domestic firms are listed on the A-share board of the Shanghai and Shenzhen stock exchanges. According to the China Securities Regulatory Commission, before 2008 foreign ownership could not exceed one-third. Since 2008, an individual foreign investor may not own more than 20% of total shares, and all foreign investors together may not own more than 25% of total shares. Ownership includes direct and indirect investments.

[7] Data for listed firms is commonly used to investigate firms' innovation performance, as in Autor et al. (2020) for the U.S., Aghion et al. (2005) for Europe, and Fang et al. (2018) for China. Other widely used Chinese firm-level data are the Annual Survey of Industrial Enterprises from the National Bureau of Statistics (e.g. Wei et al. 2017), and more recently the Administrative Enterprise Income Tax Records from the Chinese State Administration of Tax (e.g. Chen et al. 2021). However, these two data sets provide information only on total subsidies, not on R&D subsidies, so they are not suitable for our analysis.

firms' income accounts. Before 2007, a separate account for subsidy income in firms' financial statements provides details on the total amount and the type of subsidies (Ministry of Finance 2000), so we can distinguish between R&D and non-R&D subsidies using RESSET data. Instead of disclosing the amount of subsidies by type, since 2007 the financial statements' notes disclose in detail the government subsidies received (Ministry of Finance 2006). Data on total subsidies and on (almost) all individual subsidy transactions is provided in CSMAR. We developed a semi-manual approach to classify all 85,480 subsidy-related accounting transactions available in this database into R&D and non-R&D subsidies. The approach first identifies all R&D subsidies based on a keyword search. However, a given subsidy transaction may simultaneously fulfill both a research purpose and a non-research purpose (like subsidizing investments in general). Given our goal of identifying misuse of R&D subsidies, we employ a conservative definition of an R&D subsidy that automatically corrects for false positives in R&D subsidies by searching for keywords related to non-R&D subsidies in an additional step. Online Appendix 2 explains the approach in more detail.[8] We further subdivide R&D subsidies into strict and broad R&D subsidies, the latter being received for, for example, patents, technology acquisition, technology transformation, and rewards. Thus, we are able to observe all subsidies received by each firm and distinguish accurately between R&D subsidies and non-R&D subsidies. To avoid measurement error, we exclude observations for which the sum of strict, broad, and non-R&D subsidies is smaller or larger than total subsidies or for which total subsidies exceed sales.

In addition to R&D expenditures and R&D subsidies, we observe employment, net fixed assets, sales, age, profitability, and industry affiliation. More details on the measurement of all variables used throughout the study are given in Table A1 in Appendix 1. All variables in monetary values are deflated using China's GDP deflator from the World Bank. As many of China's previously state-owned firms have been privatized since the late 1990s (Hsieh and Song 2015), we account for this ownership transformation by differentiating between four ownership regimes. State ownership is defined as >50% in majority state-owned enterprises (SOEs) and between 50% and >0% in minority SOEs. Privatized firms have a government share of 0%, but were SOEs in prior periods, whereas state ownership in de-novo private firms never exceeded 0%. Finally, to investigate output and behavioral additionality, we match the PATSTAT database, which includes all Chinese invention patents filed since the establishment of China's patent system in 1984 (Boeing et al. (2016) detail the matching routine). We use patent information to calculate a firm's patent stock, changes in the application rate, high-tech IT orientation, university-firm collaboration and employment of foreign scientists.

Our unbalanced panel for the period 2001-2011 consists of 15913 observations with

[8] For an interim comparison with the R&D subsidies that were reported in Fang et al. (2018), we follow their industry selection for 2008-2011, the time period covered in both studies. The share of grantees increases from less than 70% to almost 90% in Fang et al., whereas our measure shows a lower level but similar increase from 21% to 56%. This difference can most likely be explained by our comprehensive semi-manual data classification and the rigorous correction for false positives, which make our subsidy measure more conservative.

non-missing R&D expenditures and R&D subsidies for 2317 firms (Table 1). It covers the manufacturing and service sectors (except for the finance industry). Table A2 in Appendix 2 provides information on the distribution of firms by industry. Table 1 corroborates the extraordinary development of R&D in the period under consideration. The share of R&D performers quadrupled from 14.7% to 63.0%, while their median R&D expenditures increased from 3.0 million RMB to 16.2 million RMB.[9] This change was associated with a rise in the mean and median R&D intensity (R&D expenditures to sales) from around 1.0% to 3.3% and from 0.4% to 2.6%, respectively. The share of grantees that received R&D subsidies also increased sharply, from 6.4% to 43.2%. However, the amount of R&D subsidies per subsidized firm declined over time, as the median R&D subsidy fell from about 1.4 million RMB to 0.8 million RMB. The evolution of the share of subsidized firms over time, on the one hand, and the median and the quartiles of R&D subsidies, on the other, suggest that the expansion of government funding took place along the extensive margin. Finally, R&D subsidies accounted for an average of 10.8% of total subsidies. This shows that government support received through non-R&D subsidies accounts for a multiple of R&D subsidies in China.[10] Overall our sample covers 12.1% of total R&D expenditures and 10.1% of R&D subsidies of large- and medium-sized firms in China in the 2001-2011 period (see Figure A3).

Table 1: R&D expenditures and R&D subsidies

		R&D expenditures[a)]				R&D subsidies[b)]			
Year	Obs.	Firms	P25	Median	P75	Firms	P25	Median	P75
2001	1047	0.147	1.127	2.970	9.855	0.064	0.485	1.422	4.656
2002	1115	0.152	1.488	3.784	10.005	0.076	0.410	1.380	3.991
2003	1168	0.168	1.454	3.696	9.680	0.097	0.222	0.837	3.539
2004	1274	0.181	1.332	3.673	10.161	0.108	0.249	0.655	3.315
2005	1286	0.176	1.560	3.952	12.997	0.107	0.200	0.692	2.441
2006	1417	0.174	1.945	5.184	13.703	0.103	0.176	0.602	2.567
2007	1509	0.229	2.595	7.972	23.281	0.082	0.122	0.649	2.659
2008	1557	0.283	3.357	10.640	25.258	0.150	0.282	1.028	3.090
2009	1567	0.347	5.166	12.829	32.022	0.282	0.326	0.836	2.508
2010	1876	0.568	5.310	13.398	33.002	0.393	0.211	0.744	2.082
2011	2097	0.630	6.927	16.175	37.534	0.432	0.272	0.767	2.441
Total	15913	0.310	3.267	10.392	27.168	0.197	0.250	0.790	2.523

Notes: Monetary values are in million RMB in constant prices of 2005. [a)] Quartiles calculated for R&D performers. [b)] Quartiles calculated for firms that received R&D subsidies.

[9] Using the 2005 RMB-Euro year-end exchange rate, this change corresponds to an increase from 0.311 to 1.693 million Euro.

[10] In contrast to R&D subsidies, we see an increase in the extensive and intensive margin for total subsidies in the 2001-2011 period. The share of firms that received a subsidy of any type increased from 31.7% to 90.0%, while the median subsidy increased from 2.1 to 6.0 million RMB.

3 Misappropriation

We define misappropriation of R&D subsidies as a situation in which a firm does not (fully) spend the assigned subsidy amount on R&D activities. This definition is conservative, as it is not considered misappropriation if the firm uses the subsidy for R&D but spends it on a different R&D project than the one for which it originally received funding. At this stage, we are agnostic about the specific alternative use of misappropriated funds, so the firm (or its managers) may use the money for other firm-related purposes, such as investments in physical capital, or spend it on private consumption. However, from a welfare perspective, the alternative use matters, and section 6.6 provides some empirical evidence on this issue.

3.1 Theoretical Framework

Basic Model. The theoretical framework we set up to explain a firm's incentive to misuse public R&D funds draws on basic theoretical insights using the simple model of a firm's optimal R&D investment by Howe and McFetridge (1976) (see also David et al. 2000; Hottenrott and Peters 2012). We start with a profit-maximizing firm that decides on its optimal level of R&D investment in a situation without any subsidy rd^*. The decision is based on a comparison of the marginal rate of return to R&D mr and the marginal cost of capital mc. Both the marginal rate of return and marginal cost vary with the level of R&D investment rd, but while mr is downward sloping, mc is constant as long as internal finance f is used and is upward sloping if additional, more costly external financing c^{ext} is borrowed. This approach reflects the well-known pecking order for R&D funds in finance, according to which internal means are fully used before a firm draws on external financing. In addition to R&D, the marginal rate of return may depend on the firm's innovative capabilities ic and other firm- and industry-specific variables summarized in the vector x_1, such that $mr = g_1(rd, ic, x_1)$. Similarly, we define $mc = g_2(rd, r^{alt}, f, c^{ext}, x_2)$. The marginal cost of capital reflects the opportunity costs of investing funds in R&D. In addition to the level of R&D, mc depends on the expected returns from alternative non-R&D uses of available funds such as investment in tangible or financial assets r^{alt}, the amount of internal finance f, the costs of external capital c^{ext}, and other firm- and industry-specific variables x_2. A firm invests in R&D if and as long as the marginal rate of return to R&D is larger or equal to the marginal cost of capital. Therefore, the optimal R&D investment without subsidy financing is given by $rd^* = g(ic, r^{alt}, f, c^{ext}, x_1, x_2)$, with rd^* equal to or greater than zero.

Now what happens if a firm receives an R&D subsidy s and can simultaneously decide about (non)compliance? The R&D subsidy increases the amount of internal financial means for R&D from f to f', but now we have two types of internal financing, the public R&D subsidy s and other private internal funds f^{priv}. In contrast to the standard framework, the possibility of noncompliance and the associated risk of detection and sanctioning lead to differences in marginal costs between the two types of internal funds. If instead of spending subsidy s for R&D, the firm decides to spend it on an alternative non-R&D

purpose, it risks being detected and paying sanctioning costs sc with a detection probability $p > 0$. The expected sanctioning costs $E(sc)$ lower the expected net return from an alternative non-R&D use from r^{alt} to $r^{alt'} = r^{alt} - E(sc)$ and, therefore, lower the opportunity costs of investing the subsidy s in R&D activities and, as a result, marginal costs. In contrast, marginal costs remain unchanged for other private internal means f^{priv}. According to the general idea of a pecking order, firms will use those funds that have the lowest opportunity costs first. As a result, we get an extended pecking order in a framework with potential noncompliance of R&D subsidies: The latter are fully used before any other internal funds or, if necessary, external funds are spent on R&D. It is furthermore reasonable to assume that the risk of being detected of misappropriation p and the sanctioning costs sc are highest when the firm spends nothing on R&D and that both shrink as R&D expenditures increase, leading to falling expected sanction costs and hence a rising marginal cost curve in the interval $(0, s)$. This rationale is reflected in the new marginal cost curve mc', shown in Figure 1. The intersection of mc' and mr defines the new optimal R&D investment level $rd^{*'}$.

Important for our first research question is that the extended pecking order allows us to identify misappropriation of R&D subsidies by comparing the optimal R&D investment level $rd^{*'}$ and the R&D subsidy amount s. Conditional on receiving an R&D subsidy s, misappropriation M occurs if the optimal R&D investment level $rd^{*'}$ is lower than the R&D subsidy s:

$$M_{|s>0} = \begin{cases} 0 & \text{if} \quad rd^{*'} \geq s \\ 1 & \text{if} \quad rd^{*'} < s \end{cases} \tag{1}$$

The difference between the optimal R&D investment level $rd^{*'}$ and the R&D subsidy s is a measure for the absolute level of misappropriation m. According to our theoretical framework, a firm's decision to misappropriate R&D funds thus depends on the R&D subsidy level and all arguments that determine the optimal R&D investment with subsidy financing, $rd^{*'}$:

$$M_{|s>0} = h(s, f^{priv}, ic, r^{alt}, sc, p, c^{ext}, x_1, x_2). \tag{2}$$

Our second research question aims at measuring the causal impact of R&D subsidies on R&D expenditures if misappropriation is possible. The theoretical framework also helps to explain the interplay between the change in R&D (full crowding out, partial crowding out or additionality) and misappropriation. In the standard framework without noncompliance, an R&D subsidy s will lead to an increase in the optimal R&D investment level, $rd^{*'}$, if and only if the firm is financially constrained before it receives the subsidy payment (Hottenrott and Peters 2012). This result no longer holds if we allow for misappropriation, as shown in Figure 1. mr_A, mr_B and mr_C depict three alternative marginal rates of return, all of which lead to optimal R&D expenditures, rd_A^*, rd_B^* and rd_C^*, respectively, that are lower than the internal financial means f, indicating no financial constraints without subsidy funding. After a firm receives the R&D subsidy s, three outcomes for initially unconstrained firms are possible: Case $(A, A^{'})$ depicts the situation in which optimal R&D expenditures remain unchanged at zero $(rd_A^{*'} - rd_A^* = 0)$ and subsidy s is fully misappropriated $(m_A = s)$, while case $(C, C^{'})$ is the case in which positive R&D expenditures remain unchanged $(rd_C^{*'} - rd_C^* = 0$; full crowding out) with no misappropriation $(s < rd_C^{*'}$ so that $m_C = 0)$. Novel to this framework is case $(B, B^{'})$, in which the initially unconstrained firm increases its R&D, but by an amount less than the subsidy. Therefore, case $(B, B^{'})$ describes the concurrence of partial crowding out $(0 < rd_B^{*'} - rd_B^* < s)$ and partial misappropriation $(0 < m_B < s)$.

Figure 2 describes possible outcomes for firms that are financially constrained without subsidy funding. As in the standard framework, receiving R&D subsidy s increases the optimal R&D level, but this increase may be accompanied by partial misappropriation, as in case $(D, D^{'})$, or no misappropriation, as in case $(E, E^{'})$. In both cases, the increase in R&D is less than the subsidy level, implying partial crowding out. Additionality or crowding in can take place when subsidies not only increase internal funds but also, for instance, indirectly improve conditions for external financing (Lerner 1999), flattening the slope of the non-horizontal part of the marginal cost curve, as in $mc^{''}$, and the new optimal R&D level $rd_E^{*''}$.

Figures 1 and 2 indicate that compliant behavior may go along with full crowding out $(C^{'})$, partial crowding out $(E^{'})$, or additionality $(E^{''})$, so the causal impact of R&D subsidies has to be determined empirically. Misappropriation should lower the causal impact of subsidies on R&D expenditure as it is associated with full $(A^{'})$ or partial crowding out $(B^{'}$,$D^{'})$.

Extension. Until now, we have assumed that R&D projects are arbitrarily divisible. However, in practice, R&D projects can be indivisible. Firms are often not able to scale down R&D investments at will but need a minimum of financing (González et al. 2005). Figure 3 extends the basic framework by assuming a minimum R&D threshold $rd^{min} > 0$, such that the marginal rate of return is zero for values below the threshold. We focus on the outcome for firms for which the threshold is initially binding, implying that their optimal R&D investment is zero. Figure 3 depicts three firms that have high (H), medium (M) and low (L) innovation capacity and their corresponding marginal rates of return mr_H, mr_M, and mr_L. If they receive a small R&D subsidy $s^{'}$, firms that have high marginal returns to R&D will start investing in R&D, if $f^{'} \geq rd^{min}$, without misappropriating subsidies (optimum $H^{'}$), while zero R&D and full misappropriation is still optimal for firms that have low and medium marginal rates of return $(L^{'}, M^{'})$. With increasing subsidies, more and more firms with medium marginal rates of returns start investing in R&D. Initially, for medium subsidy levels $s^{''}$, they will fully use the subsidy for R&D $(M^{''})$, lowering the likelihood of misappropriation. However, at very high R&D subsidies $s^{'''}$, they lack innovative ideas, so their optimal R&D level falls below the subsidy $(M^{'''})$, suggesting an increase in misappropriation for higher subsidy levels. Accounting for the indivisibilities of

Figure 1: Optimal R&D expenditure and misappropriation for initially unconstrained firms

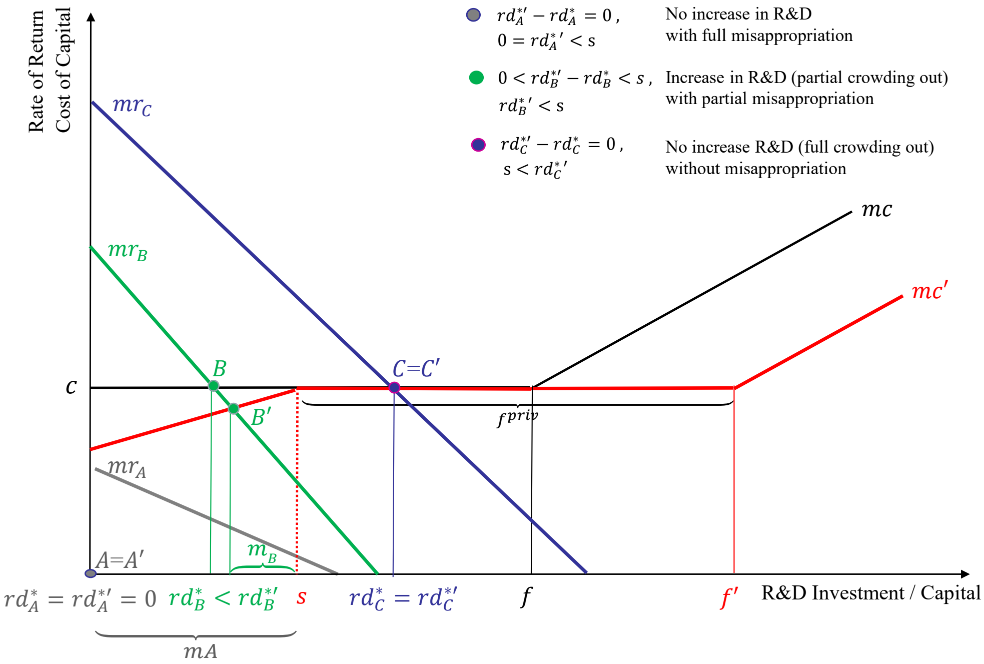
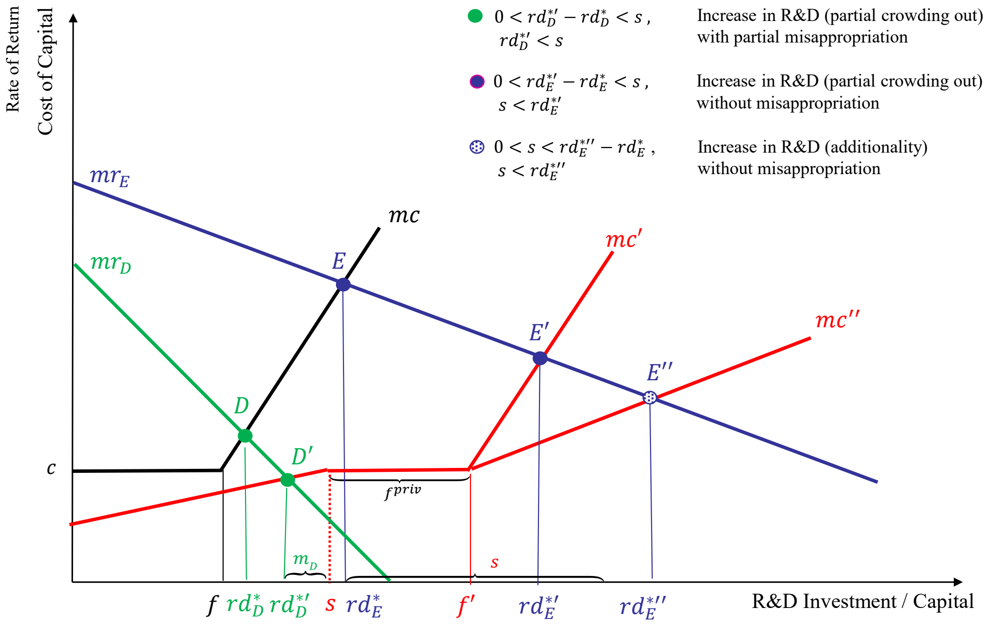


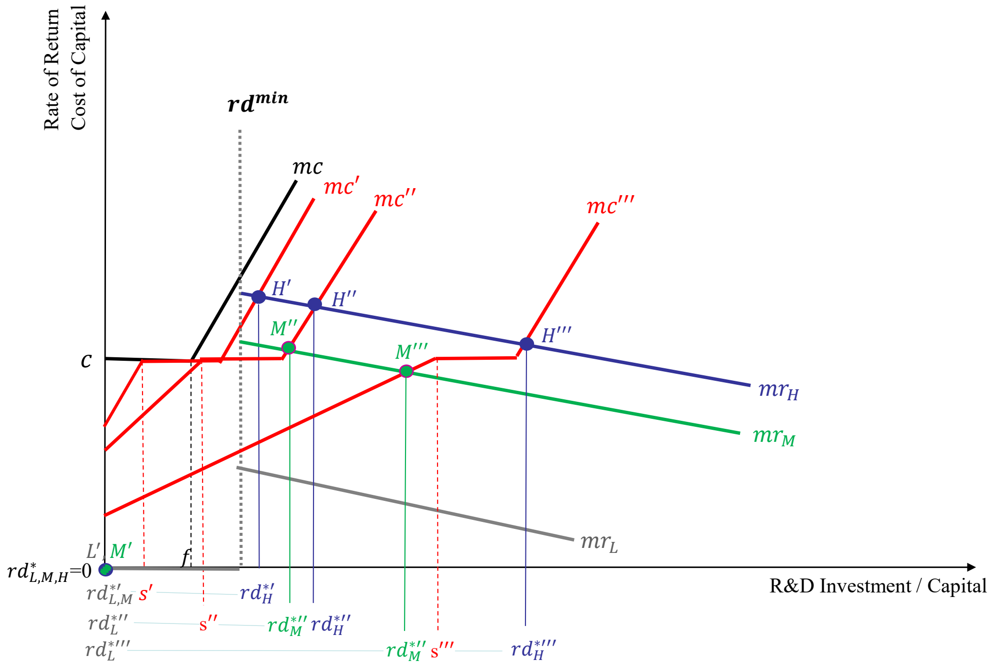
Figure 2: Optimal R&D expenditure and misappropriation for initially constrained firms



R&D projects therefore leads to a U-shaped relationship between R&D subsidies and the likelihood of misappropriation in our framework.

Limitations. Our theoretical framework makes two important simplifying assumptions. First, the amount of the R&D subsidy s is exogenous, and by deciding on the

Figure 3: Initially unconstrained firms and misappropriation with minimum R&D threshold ($rd^{min} > 0$)



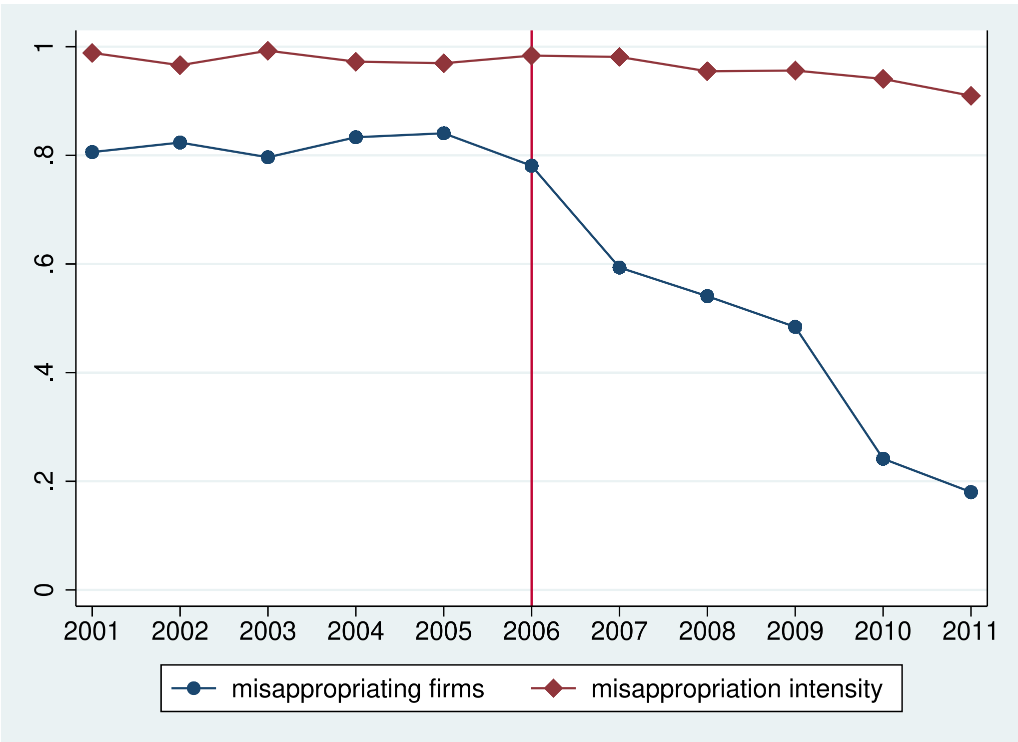
optimal R&D level, firms also decide upon the level of misappropriation. In particular, we do not take into account that the firm might have an a priori intention to misappropriate the subsidy and maximize the amount of the subsidy through fraudulent behavior. Some evidence shows that fraudulent firms are more likely to receive R&D subsidies (Stuart and Wang 2016) and resources obtained through fraudulent means are less likely to be allocated to productive activities like technological innovation (Wang and Li 2014). However, our data does not allow us to identify a priori fraudulent firms.

The second simplifying assumption our theoretical framework makes is that the sanctioning costs sc and the detection probability p are exogenously given. Detecting firms' noncompliance with funding contract rules requires that the government makes monitoring efforts. Monitoring is usually done by the bureaucrats in R&D programs, and such efforts are weakened in a regime with corruptive behavior among bureaucrats. In their theoretical model, Acemoglu and Verdier (2000) show that corrupt bureaucrats are willing to pay subsidies, regardless of the quality of the application, as long as they can keep a proportion of the subsidy as a form of rent extraction. If a firm is matched to a corrupt bureaucrat during the application process, the firm either has to decide to collude or will not receive any funds. After collusion, the corrupt bureaucrat does not monitor compliance to funding contract rules and moral hazard behavior is more likely. In our model, we have implicitly assumed that s is the net subsidy a firm receives after a potential rent extraction λ, hence $s = \tilde{s}(1 - \lambda)$, where $\tilde{s}$ is the gross subsidy.

3.2 Stylized Facts

This section presents stylized facts on the extent of misappropriation. Assuming that the optimal R&D investment level is equal to the observed total R&D expenditure, we calculate misappropriation as the difference between total R&D expenditure and R&D subsidies received, as reported in financial statements. Figure 4 shows the proportion of misappropriating firms over time. Overall, 42% of grantees misused R&D subsidies, which correspond to 53% of the total amount of R&D subsidies. These figures strikingly confirm the anecdotal evidence that misappropriation is a major concern in China. We find two additional intriguing facts regarding misappropriation in our data. First, firms either choose (almost) full misappropriation or choose not to misappropriate any funds. According to our theoretical model, full misappropriation is rationalized by low innovation capabilities and indivisibilities of R&D projects. Figure 4 shows that the average misappropriation intensity (misappropriated R&D subsidies to total R&D subsidies) is largely stable at around 96%, which implies that variation along the intensive margin is of little importance. Second, there is a substantial decline in misappropriation over time, falling from 81% in 2001 to 18% in 2011 along the extensive margin. This decline emerges especially after 2006, which coincides with tougher sanctions for misappropriation and stepped up monitoring efforts in line with the MLP reforms (see Online Appendix 1).

Figure 4: Misappropriation along the intensive and extensive margin



Notes: Misappropriating firms and misappropriation intensity denote the share of firms that have misappropriated R&D subsidies (extensive margin) and the proportion of misappropriated R&D subsidies to total R&D subsidies (intensive margin), respectively. The red line marks the introduction of the MLP, a seminal change in China's innovation policy.

In addition to variation over time, important industry differences are highlighted in

Table A2 in Appendix 2, showing the proportion of R&D performers, grantees and non-compliers at the 2-digit level for manufacturing and the 1-digit level for non-manufacturing industries. The three industries of electronics, machinery & instruments, and pharma & biological products have the highest percentage of R&D performers and are not only highly likely to receive R&D subsidies, but also the least likely to misappropriate them.[11] Around half of the firms in these industries engage in R&D, a quarter to a third receive R&D subsidies, and around a third or fewer misappropriate. The low misappropriation rate is plausible, as these sectors have the highest expected returns from innovation (Peters et al. 2017). This pattern is reversed for the three industries that have the lowest percentage of R&D performers, where a very low likelihood of receiving R&D subsidies corresponds with a high likelihood of misappropriation.

We also explore variation at the provincial level (Figure A1 in Appendix 2). Firms located in developed and industrialized coastal provinces are more likely to receive R&D subsidies but less likely to misappropriate them. In recent years, China's picking-the-winner strategy was somewhat complemented by an aiding-the-poor strategy that directed funds to firms in the less developed western provinces to address inequality in economic development. However, R&D subsidies in the western and northern provinces are associated with higher rates of misappropriation.

3.3 Firm-level Validation

3.3.1 Testing for Potential Measurement Errors

Simply identifying misappropriation based on comparing reported annual R&D expenditures with R&D subsidies could be misleading for two reasons. First, any measurement error in the reported R&D expenditures directly translates to the absolute value of misappropriation. In particular, a firm that recorded fewer R&D expenditures in its financial statements than it actually made would inflate our misappropriation measure. Second, measurement errors may occur because of unknown timing or compositional issues related to the receipt of R&D subsidies.

Regarding the first concern, we have no indication that underreporting of R&D expenditures is a severe problem. Besides the general accounting regulations, China's Securities Regulatory Commission requires listed firms to disclose R&D activities and plans in the Director's Report and New Year's Plan. In addition to the legal requirements, China's R&D tax allowance policy provides incentives to report R&D expenditures. Even when there was misreporting, one would expect firms to report either fewer R&D subsidies, more R&D expenditures, or both. Furthermore, public awareness of misappropriation occurred only after 2011, and it is most likely that firms reported R&D subsidies and R&D expenditures correctly before.[12] If anything, our measure constitutes the lower bound of

[11] This is also observed for *other manufacturing*, which also includes firms with high to medium levels of R&D investments.

[12] Even when it is based on correctly reported R&D subsidies and R&D expenditures, the identification

misappropriation.

To further validate our measure as an indicator of misappropriation that does not pick up non-reported R&D expenditure, we estimate a standard patent production function (Griliches 1990) in which we relate the number of patent applications to the *observed* R&D stock and a set of control variables. Using a Zero-inflated Negative Binomial model to account for overdispersion and a high rate of zero patents, column (1) of Table 2 confirms a highly significant impact of observed R&D on patents. We augment this specification by adding the level of misappropriation in column (2) and a binary indicator of misappropriation in column (3).[13] If these additional variables correctly represent misappropriation, they should have no impact on patents. If they instead measure (at least partially) omitted R&D expenditure that is due to non-reporting, we would expect a positive impact like that for observed R&D. The estimates confirm that neither the level nor the incidence of misappropriation affects the number of patents. Augmenting the specification does not improve the baseline model's goodness of fit and according to the Bayesian information criterion (BIC) the baseline model is preferred. The results remain almost unchanged when we restrict the sample to observations of R&D grantees in columns (4) to (6). The results are also robust to lagged misappropriation and the use of Negative Binominal and Poisson models. Based on the arguments and estimation results, we are fully convinced that unobserved R&D expenditures do not inflate our measure of misappropriation.

To assess the importance of potential measurement errors that are due to unknown timing and compositional issues related to the receipt of R&D subsidies, we additionally calculate four alternative measures, each of which is based on a different assumption. The first measure assumes that the reported R&D subsidy in year t is an advance lump sum payment, and the annual subsidy flows are calculated by allocating the received amount uniformly over an assumed funding period of three years, that is, between t and $t+2$. This measure results in a similar share of 45% of firms that engage in misappropriation. The second measure considers only strict R&D subsidies, so it excludes broad R&D subsidies, and yields 38% of firms that engage in misappropriation. The third measure ignores potential inconsistencies in timing and compares only the sum of R&D subsidies with the sum of R&D expenditures over all years. If the sum of R&D subsidies does not exceed the sum of R&D expenditures, we regard misappropriation in a single year as accounting nuisance and exclude these observations.[14] This measure provides an average lower bound of misappropriating firms at 20%. Until now, we have assumed that the R&D subsidies recorded in the balance sheets equal the net amount received by the firm, that is, less any amount withheld by officials or fees paid to intermediaries. The fourth measure alternatively assumes that recorded R&D subsidies represent gross payments received and firms cannot

of misappropriation is not a trivial accounting exercise. Since it requires replication of our subsidy data classification outlined in Online Appendix 2 in combination with our measurement framework, it is unlikely to be a routine job for auditing agencies.

[13] For this exercise, misappropriation is based on strict R&D subsidies and excludes broad R&D subsidies that are partly determined by the patent count.

[14] While this approach eliminates potential timing inconsistencies that may be incorrectly interpreted as misappropriation (false positives), it may also exclude observations of actual misappropriation (true positives).

Table 2: Effect of observed R&D and misappropriation on patent applications

	All observations			Grantee observations		
	(1)	(2)	(3)	(4)	(5)	(6)
R&D stock $_t$(log)	0.229*** (0.039)	0.230*** (0.039)	0.230*** (0.039)	0.203*** (0.046)	0.201*** (0.048)	0.199*** (0.048)
Misappropriation $_t$(log)	-	0.002 (0.007)	-	-	-0.003 (0.009)	-
Misappropriation $_t$(0/1)	-	-	0.028 (0.109)	-	-	-0.062 (0.128)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Alpha	3.515	3.515	3.515	2.498	2.499	2.500
Wald χ^2	955.1 (36)	954.0 (37)	954.1 (37)	417.0 (36)	416.6 (37)	415.6 (37)
McFadden Pseudo R^2	0.082	0.082	0.082	0.070	0.070	0.070
AIC	51632.0	51633.9	51633.9	13794.2	13796.1	13795.8
BIC	52005.4	52014.8	52014.8	14083.4	14091.1	14090.8
Nonzero observations	5392	5392	5392	1526	1526	1526
Zero observations	7561	7561	7561	877	877	877
Observations	12953	12953	12953	2403	2403	2403

Notes: Results for a Zero-inflated Negative Binominal model. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. Controls include employment, net fixed assets, age, and dummy variables that indicate whether the firm's R&D stock is zero, whether the firm participates in the InnoCom program and whether the firm is located in a province that offers patent subsidies. The zero-inflation model is estimated by logit and includes dummy variables that indicate whether the R&D stock is zero and whether the firm is located in a province that offers patent subsidies as well as year fixed-effects.

keep all R&D subsidies but are forced to return half to corrupt officials and intermediaries. This measure yields a share of 40% of firms that engage in misappropriation. Figure A2 in Appendix 2 impressively shows that all four alternative measures replicate the pattern of our main measure. Furthermore, the average intensive margin remains between 96% and 98% and there is a significant drop after 2006 along the extensive margin.

3.3.2 Decision to Misappropriate R&D Subsidies

Before we use our misappropriation measure to evaluate the effectiveness and efficacy of R&D subsidy policy, we finally check its plausibility using the theoretical framework developed in section 3.1. Conditional on receiving a subsidy, equation (2) describes a firm's decision to misappropriate R&D funds as a function of several variables. We check these theoretical predictions by estimating a two stage probit model with sample selection (Heckprobit). The first stage describes the likelihood of receiving an R&D subsidy, so it accounts for sample selection. Conditional on the firm's receiving an R&D subsidy, the second stage explains the likelihood of misappropriation. Although parameters are identified in theory because of the normal distribution assumption, identification is more robust if exclusion restrictions are imposed (Honoré and Hu 2018). Our identification strategy exploits the panel structure by using the dummy variable for whether a firm received an R&D subsidy in year $t-1$ as an exclusion restriction. The lagged R&D subsidy is likely to affect the current likelihood of receiving funding again, but it should not affect the decision to misappropriate R&D funds once we additionally control for lagged misappropriation behavior. The results shown in Table 3 confirm that lagged R&D subsidies are highly significant in the first stage. Additional estimates show that lagged R&D subsidies are not significant in the second stage after controlling for lagged misappropriation, supporting our identification strategy.

Table 3 reports the second stage results for five specifications. Following equation (2), the table explains misappropriation in column (1) using the current level of R&D subsidies for which we allow for potential non-linearity by specifying a second order polynomial in s. Innovative capabilities (ic) that increase the expected returns to R&D are proxied by two variables: A dummy that indicates whether the firm has prior R&D experience and the log number of patents up to period $t-1$. Private internal funds are measured by firm profitability which is a dummy that equals 1 if firm profits are positive (f^{priv}) in $t-1$, and zero otherwise. We also include firm attributes like the number of employees, net fixed assets, sales, firm age, and ownership observed in $t-1$, as well as year, industry, and province fixed effects. These variables capture variations in the marginal rate of return and marginal cost of capital across firms and, thus, also in the propensity to misappropriate as induced by r^{alt}, c^{ext}, x_1, and x_2.

The results shown in Table 3 largely confirm our theoretical hypotheses. In particular, R&D experience seems to increase the expected rate of return to R&D and lowers the likelihood of misappropriation. Furthermore, more profitable firms have more internal funds, lower cost of capital for internal funds, and, as a result, less incentive to misappropriate. The estimates also show a U-shaped relationship between the subsidy level and misappropriation. We observe a significantly higher likelihood of misappropriation for very low and very high R&D subsidy levels, as predicted by our theoretical framework. The effect for very low subsidy levels can be explained by the indivisibility of R&D projects, whereas the effect for very high subsidy levels is likely to reflect a decreasing rate of return to R&D. Older firms and privatized state-owned enterprises are also more likely to misappropriate. Finally, firms that already misappropriated in $t-1$, but have not been excluded from filing applications in t, are more likely to receive subsidies and misappropriate again, supporting the findings in Stuart and Wang (2016) and Wang and Li (2014).

Columns (2) to (5) additionally account for the effects of variations in sanctions on misappropriation over time and the detection probability across firms by adding monitoring and corruption indicators. Together with tougher sanctions, monitoring efforts were stepped up in the course of the post-2006 MLP reforms. Therefore, in column (2) we include an MLP dummy variable that equals 1 for the period 2007-2011. In addition to this macro-level monitoring indicator, in column (3) we add a firm-level monitoring variable, a dummy variable that takes the value of 1 if the firm has mutual funds investors. Research has shown that, for the U.S. and China, monitoring by institutional investors can be an important mechanism for promoting innovation in firms (Aghion et al. 2013; Rong et al.

2017) and that external monitoring reduces fraud (Chen et al. 2014). The most striking result in column (2) and (3) is the highly significantly negative effect for both variables. The results thus confirm that increased sanctions and monitoring efforts lower the probability of misappropriation.

As argued in section 3.1, corruption weakens monitoring, so it is supposed to increase misappropriation. Therefore, in columns (4) and (5) we also account for variation in corruption across provinces. Based on the argument that rent-seeking in China increases with the size of the government (Naughton 2013, p.80), which Chen et al. (2018) confirm empirically, we hypothesize that the likelihood of misappropriation increases with the average number of bureaucrats per firm, measured by the number of large and medium-sized enterprises (LMEs) in each province. Following the seminal papers by Glaeser and Saks (2006) for the U.S. and Wederman (2004) for China, the number of actual investigations and/or convictions may provide a more precise measure of corruption. We obtain from China's judicial procuratorial system the annual number of criminal cases of abuse of power at the province level, which includes corruption and misappropriation (Wederman 2004). As before, we divide the number of cases by the number of LMEs. We find a positive impact of both corruption measures on the likelihood of misappropriation, with the effect of the number of corruption cases being larger and more significant and dominating the number of bureaucrats when both variables are included.

In a nutshell, this section has identified substantial misappropriation of R&D subsidies, although the trend has been declining since the implementation of the MLP. Furthermore, firms choose either (almost) full or no misappropriation. Finally, misappropriation is not random. In line with the theoretical model, it can be explained by the R&D subsidy level, private internal funds, the rate of return to R&D (R&D experience), sanctions, and the probability of detection, which increases with the monitoring efforts and decreases with corrupt behavior by bureaucrats. In the following, we investigate the consequences of misusing R&D subsidies on the effectiveness of R&D policy in stimulating firms' R&D expenditures.

Table 3: Likelihood of misappropriation

		p			
	Firm attributes	MLP	Mutual fund	Bureaucrats	Corruption
	(1)	(2)	(3)	(4)	(5)
R&D subsidy $_t$ (log)	−0.445***	−0.445***	−0.439***	−0.398***	−0.416***
	(0.135)	(0.135)	(0.136)	(0.135)	(0.133)
R&D subsidy $_t$ (log)^2	0.021***	0.021***	0.021***	0.019***	0.020***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
R&D experience $_{t-1}$ (0/1)	−0.845***	−0.845***	−0.852***	−0.843***	−0.823***
	(0.094)	(0.094)	(0.094)	(0.093)	(0.094)
Patent stock $_{t-1}$ (log)	0.031	0.031	0.031	0.013	0.015
	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)
Profitability $_{t-1}$ (0/1)	−0.184**	−0.184**	−0.140	−0.158*	−0.148
	(0.094)	(0.094)	(0.094)	(0.093)	(0.093)
MLP $_t$ (0/1)		−0.951***	−0.780***	−0.700***	−0.750***
		(0.229)	(0.238)	(0.256)	(0.254)
Mutual fund $_t$ (0/1)			−0.204**	−0.213**	−0.207**
			(0.089)	(0.088)	(0.088)
Bureaucrats/LME $_{p,t}$ (log)				0.116*	−0.046
				(0.060)	(0.079)
Corruption cases/LME $_{p,t}$					0.305***
					(0.102)
Employment $_{t-1}$ (log)	−0.034	−0.034	−0.037	−0.052	−0.051
	(0.043)	(0.043)	(0.043)	(0.043)	(0.042)
Net fixed assets $_{t-1}$ (log)	0.008	0.008	0.015	0.030	0.031
	(0.044)	(0.044)	(0.044)	(0.042)	(0.042)
Sales $_{t-1}$ (log)	0.002	0.002	0.015	0.025	0.026
	(0.047)	(0.047)	(0.048)	(0.045)	(0.046)
Age $_{t-1}$ (log)	0.502***	0.502***	0.502***	0.481***	0.469***
	(0.087)	(0.087)	(0.087)	(0.086)	(0.085)
Minority SOE $_{t-1}$ (0/1)	0.080	0.080	0.089	0.061	0.073
	(0.107)	(0.107)	(0.108)	(0.103)	(0.103)
Privatized $_{t-1}$ (0/1)	0.244**	0.244**	0.245**	0.203**	0.215**
	(0.106)	(0.106)	(0.107)	(0.102)	(0.101)
De-novo private $_{t-1}$ (0/1)	−0.057	−0.057	−0.048	−0.093	−0.074
	(0.104)	(0.104)	(0.105)	(0.099)	(0.099)
Misappropriation $_{t-1}$ (0/1)	0.941***	0.941***	0.936***	0.952***	0.954***
	(0.088)	(0.088)	(0.087)	(0.087)	(0.086)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	No$^{\#}$	No$^{\#}$
Exclusion restr. 1^{st} stage					
R&D subsidy $_{t-1}$ (0/1)	1.304***	1.304***	1.305***	1.318***	1.318***
	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)
ρ	0.658	0.658	0.654	0.673	0.684
	(0.064)	(0.064)	(0.063)	(0.063)	(0.062)
Obs. 1^{st} stage	12953	12953	12953	12953	12953
Obs. 2^{nd} stage	2403	2403	2403	2403	2403

Notes: The subscript p, t indicates variables measured at the province-year level. All other variables are at the firm-year level, except MLP, which is at the year level. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. # Instead of province FE, columns (4) and (5) include the province-year variables log of R&D expenditures/LME and log of GDP per capita as well as their second polynomials in both stages. While the R&D variables are insignificant, the GDP variables are significant. The coefficients in the first stage are 2.478 and -0.153 in column (4) and 4.832 and -0.290 in column (5). The turning point in both columns is close to the mean of 8.183. The results remain robust when we use Heckman and RE Heckman instead of Heckprob or estimate the second stages using only a simple Probit or linear probability model.

4 Identifying the Causal Effects of R&D Subsidies with One-Sided Noncompliance: ITT and CACE

Randomized experiments are considered the gold standard for causal inference, but they are often infeasible in practice. Even if they were feasible, further complications can arise. Most importantly, noncompliance with the *assigned* treatment may occur. In the event of one-sided noncompliance, units that are assigned to a treatment may decide not to comply with the assignment and receive no *actual* treatment.[15,16] In our application, one-sided noncompliance among R&D subsidy recipients is substantial, as shown by the figures on misappropriation of R&D funds presented in section 3. As a result, we have to distinguish between the assigned treatment and the actual treatment. The main problem with noncompliance for causal inference is that the actual treatment is the result of a deliberate choice that very likely takes into account the expectation about the causal effect of the treatment. This post-assignment self-selection process breaks the initial randomization of the assigned treatment and calls the unconfoundedness of the actual treatment into question (Imbens and Rubin 2015).

Consider the randomized assigned treatment Z_i, which takes the value of 1 if firm i is assigned to the treatment group and 0 if it is assigned to the control group.[17] Further, let D_i denote the actual treatment, which takes the value of 1 if firm i actually receives the treatment, and 0 otherwise. In our setting, Z_i indicates whether firm i receives an R&D subsidy and D_i indicates whether the firm uses the R&D subsidy for research purposes. Noncompliance occurs when $Z_i \neq D_i$.

There are two naïve approaches to studying the treatment effect in such a setting. The *as-treated* approach compares the treatment and control group according to their actual treatment status D but ignores that whereas Z is randomly assigned, D is not. The *per protocol* approach simply discards noncompliers ($Z \neq D$) and analyzes the compliers as if they were randomized. However, as compliance is self-selected, the remaining subsample is not representative of the study's population. Thus, both approaches generally fail to provide consistent estimates of the treatment effect (Imbens and Rubin 2015).

In contrast, we differentiate between the intention-to-treat (ITT) and the complier average causal effect ($CACE$) in order to consistently estimate and evaluate R&D subsidies under one-sided noncompliance with funding contract rules. That is, we are interested in two treatment effects: That of the assigned treatment and that of the actual treatment (Imbens and Angrist 1994; Imbens and Rubin 2015). Although the problem of noncompliance

[15] If, in addition, units from the control group are also able to circumvent the non-assignment to receive the treatment, noncompliant behavior may arise among both treated and control units, which is called two-sided noncompliance.

[16] Chen et al. (2021) investigate the effect of the Chinese InnoCom program implemented in 2008 which provides large incentives for R&D investment in the form of a corporate income tax cut. They estimate that about 24% of the increase in reported R&D was due to re-labeling of administrative expenses. In contrast to our application, re-labeling can be considered as one-sided noncompliance among initially non-assigned (non-eligible) control firms in order to actually get the treatment (tax deduction).

[17] To simplify notation, we leave out time subscripts in this section. However, our empirical analysis is based on panel data and includes a time dimension.

in experimental designs is not new and is frequently applied in the biostatistics literature, applications in economic policy evaluation are rather scarce, and they are nonexistent for R&D policy evaluation mainly because it is often difficult to identify noncompliant behavior in the data, even when it exists.[18]

4.1 Intention-to-Treat (ITT) Effect

The ITT effect denotes the causal effect of the assigned treatment Z_i on the outcome variable Y_i. In our application, ITT is the causal effect of a granted R&D subsidy on the growth rate of R&D expenditures. Allowing for heterogenous treatment effects across firms, the individual ITT is the difference in the firm-level outcome variable Y_i from the assigned treatment status Z_i: $ITT_{Yi} = Y_{1i} - Y_{0i}$. Y_{zi} denotes the outcome variable for the assigned treatment status; that is $Y_{zi} = Y(Z_i = z) = Y(z)$ for $z = 0, 1$. However, for each firm, we observe only either Y_{1i} or Y_{0i}. But if the initial assignment to treatment is randomized[19], the average ITT_Y is consistently estimated as the expected difference in the outcome variable Y between the assigned treatment and control groups: $ITT_Y = E(Y_1 - Y_0)$. Consistency holds as long as the stable unit treatment value assumption (SUTVA) holds, which states that one unit's treatment assignment has no causal effect on another unit's outcome.

An additional complication arises in our setting, as it is well known that the R&D application and granting process is generally not random but depends on firm-specific covariates. A key advantage of randomized experiments is that the treated and control groups only randomly differ from one another on all observed and unobserved covariates (Stuart 2010). Matching methods have become a widely used sample design tool to mimic randomization by selecting a control group that is similar to the treatment group in terms of observed covariates. By improving the covariate balance between both groups, the treatment variable comes closer to being independent of the confounding variables. In practice, when the set of covariates is large, finding a well-balanced matched control sample is often time-consuming, and the outcome also depends on the specific matching procedure (Hainmueller and Xu 2013). Entropy balancing, developed by Hainmueller (2012), is an alternative method for achieving covariate balance and mimicking randomization more closely (Athey and Imbens 2017). The idea of entropy balancing is to find a weight for each control observation, so that the set of weights satisfies the desired balance constraints and remains as close as possible in an entropy sense to uniform base weights which retains efficiency in the subsequent analysis (Hainmueller and Xu 2013). Imposing balance constraints means that we simultaneously require that the first, second, and third moments of all covariate distributions in

[18] Notable exceptions are the evaluations of the JTPA training program in Bloom et al. (1997) and the Head Start education program in Kline and Walters (2016). In the R&D context, noncompliance has been neglected in the literature so far. Data that reveals cases in which R&D subsidies are larger than total R&D expenditures has been interpreted as an accounting nuisance that is due to the time allocation of subsidies and per protocol has been used for estimation (González et al. 2005; González and Pazó 2008; Arqué-Castells and Mohnen 2015). For example, using Spanish data, Arqué-Castells and Mohnen (2015) exclude almost 20% of the observations that received public R&D subsidies.

[19] Or more generally if the assigned treatment is unconfounded with the potential outcome.

the weighted control group exactly balance their counterparts in the treatment group. To estimate the ITT consistently and get an (almost) randomized assignment to treatment, we therefore use entropy balancing as an additional first design step. Section 5.3 on the empirical strategy provides further details. Because this approach relies on selection on observables and on unobservable time-invariant confounders, in section 6.3 we examine the sensitivity of our estimates to potential bias stemming from unobserved time-variant controls, using the test recently developed by Oster (2019).

While ITT_Y provides a consistent estimate of the causal effect of the assigned treatment, it ignores the compliance status D_i completely, so it is not informative about the causal impact of the actual treatment. Still, estimating the ITT is important, as it tells us about the *effectiveness* of the treatment when noncompliance (misappropriation) exists.

4.2 Complier Average Causal Effect (CACE)

As explained, the problem in estimating the causal effect of the actual treatment D_i is that compliance is based on self-selection, implying that the actual treatment is confounded with the potential outcome Y_i. In our case, firms with higher expected returns to R&D are more likely to spend R&D subsidies for research purposes. Using ordinary least squares (OLS) in a regression of Y on D without accounting for endogeneity would lead to results that are biased upward. Therefore, we employ an IV strategy using the randomized assigned treatment Z as an instrument (Bloom 1984; Angrist and Imbens 1994; Angrist et al. 1996). The idea is that Z_i predicts D_i, which, in turn, affects outcome Y_i. This identification strategy is valid if three assumptions are met:

First, Z is randomized or more generally unconfounded with potential outcomes of D and Y. This independence assumption allows us to consistently estimate (i) the causal effect of Z on Y (i.e. ITT_Y), which is equivalent to the coefficient of the reduced form equation in an IV setting and (ii) the causal effect of Z on D, which is called ITT_D and is equivalent to the coefficient of the instrument in the first stage of IV. Independence in our application would be violated if policymakers follow a picking-the-winner or aiding-the-poor strategy when allocating R&D subsidies or if they allocate R&D subsidies based on firms' expected compliance behavior. As we already explained for the ITT, we do not expect the R&D subsidy process (and, thus, the instrument) to be initially random but suggest using the entropy balancing method as a first step to get an (almost) randomized assignment to treatment Z_i.

The second assumption that must be met is that the instrument Z_i affects the potential outcome $Y_i(z, d)$ only via the actual treatment D_i. Since the instrument Z_i is excluded from the potential outcome, this is the exclusion restriction. We believe that this restriction holds in our application since it is reasonable to assume that the growth rate of R&D expenditures is affected only by the fact that the firm has actually decided to spend the R&D subsidy grant on R&D projects, but not on the assignment of a grant as such.

The third assumption that must be satisfied is monotonicity, which implies that defying behavior is ruled out.[20] Under one-sided noncompliance, this assumption is fulfilled by design.

Under these assumptions, Bloom (1984) and Angrist and Imbens (1994) show that the IV estimator of Y_i on D_i using randomized Z_i as an instrument is the causal effect of the actual treatment of D_i on outcome Y_i among the subgroup of compliers. This effect is known as the local average treatment effect ($LATE$) or complier average causal effect ($CACE$). $CACE$ can be calculated as the ratio between ITT_Y and ITT_D :

$$\widehat{CACE} = \frac{\widehat{ITT_Y}}{\widehat{ITT_D}} \tag{3}$$

In contrast to ITT, which measures the treatment's effectiveness, CACE measures the efficacy of the treatment in an ideal situation without noncompliance, that is, how effective the R&D subsidy policy could have been without misappropriation. From a policy point of view, knowledge about both effects and their comparison is useful. For instance, if we find that ITT is (close to) zero but CACE is significantly positive, we can conclude that the design of the R&D program works, in principle, by stimulating R&D expenditures, but that policymakers should strive to improving monitoring of grantees. If both ITT and CACE were zero, we would instead argue that the R&D policy is ineffective even in an ideal situation because of the program's design. Overall, the relationship between a policy's effectiveness and its efficacy informs us about the loss in effectiveness that is due to noncompliance.

5 Empirical Strategy

This section explains our empirical strategy for estimating both treatment effects. We start by describing the samples for the ITT, complier, and control groups in sub-section 5.1, followed by some basic descriptive statistics in sub-section 5.2. Sub-sections 5.3 and 5.4 specify the econometric model and explain the estimation strategy used to identify the ITT and $CACE$, respectively.

5.1 Sample of ITT, Compliers, and Controls

The ITT group consists of all firms that received R&D subsidies in year t (and potentially in years $t+n$) but that did not receive R&D support in year $t-1$.[21] To identify the causal impact of the R&D subsidy assignment in a before-after comparison, we exclude firms that were subsidized in years t *and* $t-1$ from the econometric analysis. The ITT group is

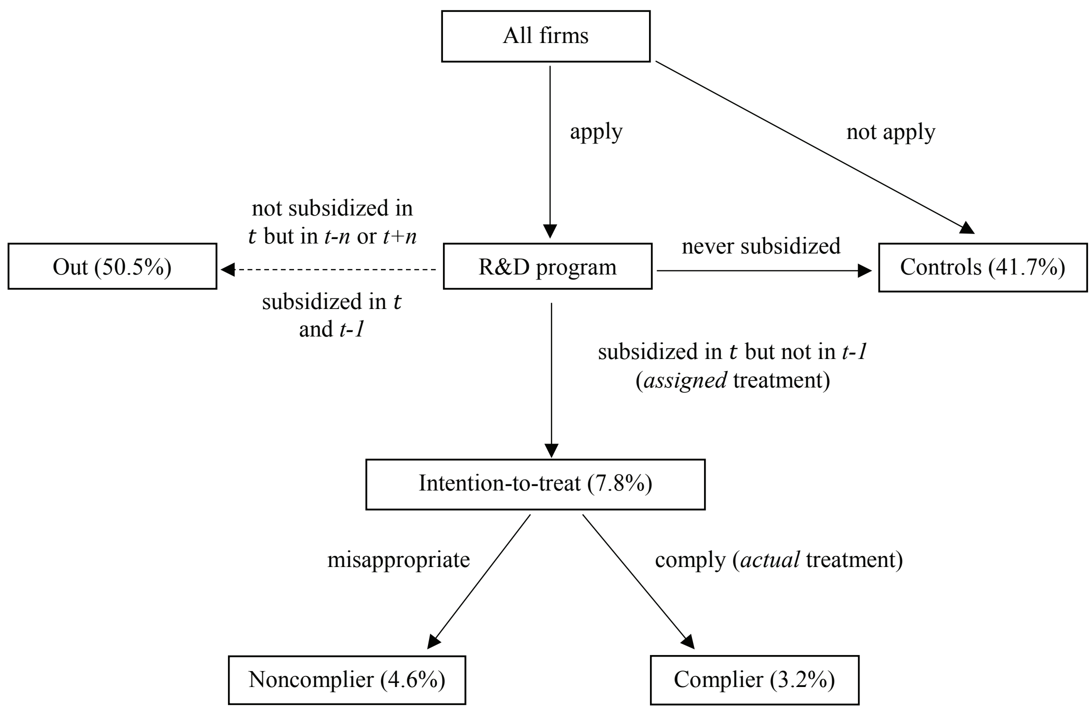
[20] Defiers are firms for which actual and assigned treatment never coincides; that is, they are noncompliant irrespective of whether they are treated or not.

[21] In our unbalanced panel of 15913 observations, 51.4% of subsidized firms report funding in one year, 27.6% and 13.2% of them in two and three consecutive years, respectively, while the remaining 7.8% of subsidized firms receive funding in four to ten consecutive years.

further split into two mutually exclusive groups of firms, those that spend their subsidies on research (compliers) and those that misappropriate them (noncompliers). Figure 5 shows the share of firms in each group. Among the firms in the ITT sample, 41% complied with the assigned treatment. That the percentage of compliers is even slightly lower than the proportion we document in section 3.2 can be explained by the exclusion of firms that received subsidies in years t and t-1, and the fact these firms with successive funding are more likely to spend public funds for research purposes. The group of noncompliant firms comprises full and partial noncompliers, although the majority (90.7%) of them commit full misappropriation.

Firms that never applied for an R&D subsidy program or that applied but were never subsidized build the control group. Since we do not have subsidy application data, we cannot restrict the control group to firms that applied for R&D funding but did not receive it. As explained in more detail in sub-section 5.3, the causal impact of an R&D subsidy is defined as the change in the R&D expenditure between $t-1$ and $t+1$. To eliminate possible long-term or anticipation effects of a program in the control group, we require not only that firms did not receive grants in the three years from $t-1$ to $t+1$ but also that they never received R&D subsidies in other years. Because we observe all R&D subsidies received from any R&D program, we can rule out contamination by direct grants that may lead to substitution bias (Heckman and Smith 1995).

Figure 5: Sample definition of ITT, compliers and control groups



5.2 Descriptive Statistics

Table 4 reports the descriptive statistics for the total sample of all firms and the estimation sample of the assigned treated (ITT), actually treated (compliers), and control observations.[22] Relative to the control group, the ITT group is characterized pre-treatment by higher median employment, net fixed assets, sales and patent stock but lower age. Compared to the overall ITT group, compliers median pre-treatment stock of patents, along with employment and sales, are larger, as is the percentage of profitable firms.

In the year before firms are granted their subsidies, we observe higher average R&D expenditures for compliers (29.7 million RMB) than for ITT firms (13.9 million RMB), which are both significantly larger than the average R&D expenditures of firms in the control group (7.4 million RMB). This result supports the view that both the government's selection of firms and the firms' decision to comply is not random. In the year following a positive grant decision, ITT firms significantly increased their R&D expenditures. In this group, the average log-growth rate of R&D expenditures between t-1 and t+1 is about 2.916 or about 98% per year. Compared to the ITT group, compliers have a higher (3.279) average two-year growth rate, and the control group has a lower one (0.846). The average growth rates of R&D expenditures are high compared to those found in other studies for developed countries, but they are sensible given China's tremendous rise in R&D expenditures at the aggregate level during that period (see Figure A3 and Table 1), and given that many firms started R&D or increased R&D from a very low initial level.[23] In all three groups, the distribution of the R&D growth rate is rather skewed. The median growth rate of R&D is much lower for the ITT and controls groups (0) than it is for compliers (0.705). Interestingly, ITT firms receive a higher average R&D subsidy than compliers do (2.7 compared to 1.8 million RMB), and their R&D subsidy intensity (ratio of R&D subsidy to R&D expenditure among R&D performers) is three times larger than that of compliers (39.0% versus 12.8%). Neglecting selection and misappropriation biases, a simple comparison of the before-and-after average R&D expenditures of the ITT (17.4) and control groups (7.5) shows that R&D subsidies fostered R&D expenditures by 9.9 million RMB (2.1 in the average log change). The corresponding figure for the sub-sample of compliers is 19.5 million RMB (2.4 in the average log change).

In summary, the descriptive statistics suggest that it is important to account for misappropriation and to distinguish between the ITT group and compliers. In addition, it is important to account for differences in the levels of pre-treatment R&D expenditure and other firm characteristics when treatment and control groups are compared.

[22] Compared to section 2, the total sample size is reduced from 15913 to 10433 observations because of taking 2-year lags and dropping observations with missing values for the relevant variables used in the estimation.

[23] High average values are not the result of a few outliers. Winsorizing the data at the 99^{th} percentile yields average growth rates in the ITT, complier and control groups of 2.903, 3.266 and 0.834, respectively.

Table 4: Descriptive statistics

	Estimation sample									Total sample		
	ITT			Complier			Controls					
	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd
Quantitative variables[a]												
R&D expenditures $_{t+1}$(m. RMB)	31.266	1.154	238.149	56.639	16.102	342.825	14.888	0.000	160.105	18.179	0.000	148.996
R&D expenditures $_{t-1}$(m. RMB)	13.866	0.000	153.259	29.679	5.774	236.849	7.380	0.000	108.03	8.567	0.000	109.824
Log-growth in R&D expend $_{t-1}$ to $_{t+1}$	2.916	0.000	7.642	3.279	0.705	7.558	0.846	0.000	5.346	1.415	0.000	6.516
R&D subsidy $_t$ (m. RMB)	2.707	0.556	9.451	1.764	0.482	4.076	0.000	0.000	0.000	0.539	0.000	4.416
R&D subsidy intensity $_t$ (%)[b]	0.390	0.057	1.112	0.128	0.043	0.187	0.000	0.000	0.000	0.563	0.000	19.585
Employment $_{t-1}$	3584	1918	8566	4038	1946	12265	4650	1380	23142	4058	1694	15923
Net fixed assets $_{t-1}$(m. RMB)	964.0	372.8	2487.5	973.0	322.1	2283.4	2886.0	381.2	18274.7	1850.0	392.6	12361.7
Sales $_{t-1}$(m. RMB)	2916.4	1105.6	11198.9	3524.9	1210.2	16509.0	5930.3	819.0	51360.2	4059.4	946.1	33844.2
Age (years)	11.855	11.000	4.303	11.349	11.000	4.322	12.312	12.000	4.250	11.847	12.000	4.330
Patent stock$_{t-1}$	21.188	2.850	88.651	20.629	5.018	43.015	20.058	0.000	151.810	19.223	0.723	121.547
Binary variables												
R&D experience $_{t-2}$ >0	0.471			0.734			0.195			0.311		
Profitability $_{t-1}$	0.816			0.872			0.793			0.820		
Majority SOE $_{t-1}$	0.174			0.155			0.252			0.237		
Minority SOE $_{t-1}$	0.241			0.188			0.264			0.253		
Privatized $_{t-1}$	0.257			0.245			0.229			0.226		
De-novo private $_{t-1}$	0.327			0.412			0.255			0.284		
Number of observations		816			335			4350			10433	

Note: The total sample consists of all observations with nonmissing values for the relevant variables and with R&D expenditure available in years t-1 and t+1. The estimation sample is the subsample of ITT and control observations, leaving out the Out group (Figure 5). [a] R&D expenditure, R&D subsidies, employment, net fixed assets, sales, age and patent stock are not log-transformed. However, for estimation purposes log values are used to account for the skewness of the distribution. [b] R&D subsidy intensity is calculated only for observations with positive R&D expenditures. m. RMB denotes million RMB.

5.3 ITT: Econometric Model and Entropy Balancing

The panel structure of our data allows us to compare R&D expenditures before and after the treatment between treated and control firms. The outcome variable is defined as the growth rate of R&D expenditure between year $t+1$ and $t-1$ (for a similar approach see Einioe 2014; Aerts and Schmidt 2008). We examine changes from the last pre-treatment year to the second treatment year for two interrelated reasons. First, grant decisions are made during the whole year t, so an R&D subsidy is not likely to cover the entire first year and may kick in later, and the timing varies across firms. Second, for multi-annual R&D projects a larger fraction of the costs accrues after the first year. In our data, the mean duration of consecutive support is 1.8 years, and the mean subsidy size in year $t+1$ is 26.1% larger than it is in year t, while the additional increase in year $t+2$ is only 12.4%.

Let $y_{i,t+1}$ denote the log R&D expenditure in year $t+1$. The log-growth rate of R&D expenditure $y_{i,t+1} - y_{i,t-1}$ is assumed to depend on whether the firm received an R&D subsidy in year t (Z_{it}), the firm-specific pre-treatment variables summarized in $X_{i,t-1}$, and industry ϕ_j, year ϕ_t, and industry-year fixed effects ϕ_{jt}, as shown in equation (4):

$$y_{i,t+1} - y_{i,t-1} = \alpha_{0,ITT} + \gamma_{ITT} Z_{it} + X_{i,t-1}\beta_{ITT} + \phi_j + \phi_t + \phi_{jt} + \varepsilon_{it} \tag{4}$$

The vector of pre-treatment characteristics $X_{i,t-1}$ includes the log number of employees and its square term to control for nonlinear firm size effects, log net fixed assets to measure capital, and log age to control for firm-age effects. Furthermore, we account for the availability of internal financial means for paying for R&D projects by including log sales and a dummy variable that is 1 if the firm yields positive profits. Finally, we expect the growth rate of R&D expenditure to depend on a firm's prior innovation activities, which we capture by three variables: Log R&D expenses in year $t-1$ to address the concern that growth rates may vary with pre-treatment levels of R&D investment, as growth in R&D expenditure is likely to be higher for firms that start R&D activities because of R&D-specific set-up costs; a dummy variable R&D experience that takes the value of 1 if the firm has prior R&D experience in year $t-2$, and zero otherwise; and the log of patent stock in year $t-1$ to capture the firm's past innovation success. Unobserved industry time-specific factors like technological opportunities or (expected) demand for innovative technological solutions might also drive a firm's decision to invest in R&D and, hence, its growth in R&D expenditures and its likelihood of receiving R&D subsidies, which would bias OLS estimates. These unobserved industry time-specific factors are controlled for by adding industry-year fixed effects ϕ_{jt}. ε_{it} is an *i.i.d.* error term with mean 0 and variance σ^2_ε. γ_{ITT}, β_{ITT}, the constant $\alpha_{0,ITT}$ and ϕ_j, ϕ_t, and ϕ_{jt} are parameters to be estimated. The main parameter of interest in equation (4) is γ_{ITT}, which measures the average ITT effect of an R&D subsidy on the growth in R&D expenditures (ITT_Y).

As explained in sub-section 4.1, if the initial assignment to treatment Z_{it} were randomized, we could consistently estimate equation (4) with OLS. However, the R&D application and granting process is generally not random but based on firm-specific characteristics, which is confirmed in our data. Table 4 shows that both groups differ significantly in terms of the observed variables. Subsidized firms are significantly larger, younger, and de-novo private, and they have higher sales, more R&D experience, and a larger prior stock of patents. A selection bias arises if a firm's R&D investment decision (partly) depends on the same common variables, which confound selection into treatment Z_{it}. For instance, past innovation success is likely to explain both the likelihood of receiving an R&D subsidy and the growth in R&D expenditures. If all of these common covariates are observable, the dependence between the R&D subsidy and the growth rate in R&D expenditure can be removed by conditioning on these observables.[24]

To achieve covariate balance and to get an (almost) randomized assignment to treatment, we employ entropy balancing as a first design step in the ITT estimation. Table 5 shows the mean, variance and skewness of the covariates $X_{i,t-1}$ in the treatment and control groups before and after balancing. While we see large distributional differences between the two groups for almost all covariates before balancing, the first, second, and third moments of the covariate distributions are virtually identical in the ITT and control groups after entropy balancing.

Entropy balancing has two main advantages over matching that make it particularly attractive in our setting. First, because of the weighting, the method does not discard any observations, and we can use the weights subsequently to estimate the ITT using weighted OLS. Second, and more important, under most circumstances entropy balancing is more bias-reducing in finite samples than matching.[25] Our treatment comes closer to randomization since we obtain a much higher degree of covariate balance. This is achieved because entropy balancing allows us to already impose a large set of balance constraints as part of the procedure to find optimal weights. Getting a quasi-randomized ITT is essential not only at this stage but also for the IV strategy in estimating the CACE. Table 6 reports estimation results for the likelihood of firms' receiving R&D subsidies. The selection into ITT is determined by the rich set of firm-specific covariates summarized in $X_{i,t-1}$, which reflect both the differences in firms' incentives to apply for funding and the eligibility and selection criteria of major R&D programs in China. We also include industry, year, and industry-year fixed effects to control for changes in China's innovation policy and to account for the possibility that (time) patterns of R&D support differ across industries. Columns (1) and (2) report the estimation results before balancing using both a probit model and a linear

[24] Like matching, balancing controls only for selection on observables. But controlling for the observed covariates also implies controlling for the unobserved covariates to the extent that they are correlated with the observed ones. Therefore, concern arises only because of (the portion of) omitted variables that are unrelated to the observed covariates (Stuart 2010; Oster 2019). In sub-section 6.1, we perform a test that Oster (2019) proposes to investigate how selection on unobservables might affect our estimated ITT.

[25] Matching is less bias-reducing unless the distributions of the covariates are ellipsoidally symmetric or are mixtures of proportional ellipsoidally symmetric distributions. For example, ellipsoidal symmetry fails if covariates include binary, categorial, or skewed continuous variables. Even with a good propensity score model, imbalances often remain in finite samples. Using well studied data from the National Supported Work Demonstration Program, Hainmueller and Xu (2013) show that entropy balancing reduces the selection bias in a regression analysis from 41% to 1.8%.

Table 5: Covariate distribution of ITT and control group before and after balancing

Variable	ITT group (N=816)			Control group (N=4350)			t-test on mean difference
	Mean	Var	Skew	Mean	Var	Skew	p-value
				Before balancing			
R&D expend. $_{t-1}$ (log)	6.201	60.850	0.498	1.883	26.630	2.430	p<0.001
R&D experience $_{t-2}$ (0/1)	0.471	0.249	0.118	0.195	0.157	1.538	p<0.001
Employment $_{t-1}$ (log)	7.519	1.345	-0.354	7.125	2.408	-0.162	p<0.001
Net fixed assets $_{t-1}$ (log)	19.780	1.654	0.047	19.730	3.485	-0.234	p=0.406
Sales $_{t-1}$ (log)	20.860	1.602	0.124	20.500	3.009	-0.114	p<0.001
Age $_{t-1}$ (log)	2.295	0.201	-0.845	2.344	0.185	-0.836	p=0.004
Patent stock $_{t-1}$ (log)	1.575	2.388	0.765	0.860	2.001	1.939	p<0.001
Profitability $_{t-1}$ (0/1)	0.816	0.150	-1.633	0.793	0.164	-1.455	p=0.129
Minority SOE $_{t-1}$ (0/1)	0.241	0.183	1.208	0.264	0.194	1.070	p=0.175
Privatized $_{t-1}$ (0/1)	0.257	0.191	1.110	0.229	0.177	1.292	p=0.076
De-novo private $_{t-1}$ (0/1)	0.327	0.220	0.737	0.255	0.190	1.126	p<0.001
				After balancing			
R&D expend. $_{t-1}$ (log)	6.201	60.850	0.498	6.195	60.760	0.500	p=0.986
R&D experience $_{t-2}$ (0/1)	0.471	0.249	0.118	0.470	0.249	0.120	p=0.985
Employment $_{t-1}$ (log)	7.519	1.345	-0.354	7.519	1.345	-0.355	p=0.993
Fixed assets $_{t-1}$ (log)	19.780	1.654	0.047	19.780	1.654	0.047	p=0.995
Sales $_{t-1}$ (log)	20.860	1.602	0.124	20.860	1.601	0.124	p=0.993
Age $_{t-1}$ (log)	2.295	0.201	-0.845	2.295	0.201	-0.844	p=0.993
Patent stock $_{t-1}$ (log)	1.575	2.388	0.765	1.574	2.385	0.766	p=0.986
Profitability $_{t-1}$ (0/1)	0.816	0.150	-1.633	0.816	0.150	-1.633	p=0.997
Minority SOE $_{t-1}$ (0/1)	0.241	0.183	1.208	0.242	0.183	1.208	p=0.997
Privatized $_{t-1}$ (0/1)	0.257	0.191	1.110	0.257	0.191	1.111	p=0.993
De-novo private $_{t-1}$ (0/1)	0.327	0.220	0.737	0.327	0.220	0.738	p=0.992

Note: Entropy balancing is based on the Stata program ebalance from Hainmueller and Xu (2013).

probability model (LPM) that is robust to violations of normality. Even after controlling for industry-year fixed effects, we find significant effects of prior R&D expenditures, R&D experience, employment, sales, firm age, patents and profitability on the likelihood that a firm will receive R&D subsidies. Our specification explains 21% of the variation in the ITT selection. Columns (3) and (4) show the results of re-estimating the likelihood of receiving R&D subsidies after balancing, that is, using the weights for the control group based on entropy balancing. The results impressively show that, after covariate balancing, all variables become insignificant, and the explanatory power is reduced to almost zero.[26] This outcome reflects a quasi-randomized selection into ITT.

5.4 CACE: Econometric Model and IV Estimation

To provide evidence on the efficacy of R&D subsidies on the growth of R&D expenditure in an ideal situation without misappropriation, we estimate the CACE using equation (5):

[26] Using nearest neighbor matching, the coefficients of the covariates are larger than with entropy balancing though also not significant. However, the explanatory power in terms of pseudo R2 is still 3.8%.

Table 6: Non-random and randomized assignment of R&D subsidies

	Non-random assignment (before balancing)		Randomized assignment (using entropy balancing)	
	LPM (1)	Probit (2)	LPM (3)	Probit (4)
R&D expenditure $_{t-1}$ (log)	0.005***	0.003***	0.000	0.000
	(0.002)	(0.001)	(0.002)	(0.002)
R&D experience $_{t-2}$ (0/1)	0.053**	0.045**	0.000	0.000
	(0.021)	(0.017)	(0.039)	(0.039)
Employment $_{t-1}$ (log)	0.106***	0.163***	0.000	0.000
	(0.021)	(0.036)	(0.099)	(0.098)
Employment $^2_{t-1}$ (log)	−0.008***	−0.012***	0.000	0.000
	(0.002)	(0.003)	(0.007)	(0.007)
Fixed assets $_{t-1}$ (log)	−0.004	−0.007	0.000	0.000
	(0.006)	(0.008)	(0.022)	(0.021)
Sales $_{t-1}$ (log)	0.015**	0.022***	0.000	0.000
	(0.006)	(0.008)	(0.022)	(0.021)
Age $_{t-1}$ (log)	−0.063***	−0.055***	0.000	0.000
	(0.021)	(0.019)	(0.046)	(0.045)
Patent stock $_{t-1}$ (log)	0.011*	0.011*	0.000	0.000
	(0.007)	(0.006)	(0.014)	(0.013)
Profitability $_{t-1}$ (0/1)	0.030*	0.026	0.000	0.000
	(0.015)	(0.016)	(0.039)	(0.038)
Minority SOE $_{t-1}$ (0/1)	0.028	0.030	0.000	0.000
	(0.019)	(0.022)	(0.054)	(0.053)
Privatized $_{t-1}$ (0/1)	0.006	0.009	0.000	0.000
	(0.017)	(0.018)	(0.044)	(0.044)
De-novo private $_{t-1}$ (0/1)	0.030	0.026	0.000	0.001
	(0.021)	(0.022)	(0.052)	(0.051)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes
Observations	5166	4573	5166	4573
(Pseudo) R^2	0.212	0.206	0.001	0.000

Notes: Average marginal effects on the likelihood of a firm's receiving a R&D subsidy. Industry-year FE perfectly predict the outcome in probit estimations for 593 observations. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1.

$$y_{i,t+1} - y_{i,t-1} = \alpha_{0,CACE} + \gamma_{CACE} D_{it} + X_{i,t-1}\beta_{CACE} + \phi_j + \phi_t + \phi_{jt} + \nu_{it} \tag{5}$$

In contrast to equation (4), equation (5) uses actual treatment D_{it} on the right-hand side, which is 1 if firm i in year t has spent the subsidy on research projects, and zero otherwise. The potential endogeneity of the actual treatment D_{it} resulting from self-selection can be addressed by using the assigned treatment as an instrument if Z_{it} is randomized. Therefore, we estimate equation (5) using IV in combination with entropy weights.

6 Empirical Results

This section first reports our benchmark results on the ITT and CACE effects of R&D subsidies on the growth rate of R&D expenditures. Sub-section 6.2 investigates heterogeneity in treatment effects across time, subsidy size (payments), industry, and ownership. Various robustness tests are presented in sub-section 6.3, followed by results for alternative outcomes and long-term effects in sub-sections 6.4 and 6.5, respectively. Finally, we investigate alternative uses of misappropriated funds in sub-section 6.6.

6.1 Treatment Effects on R&D Expenditures

Table 7 reports our main results. We estimate ITT and CACE by employing two specifications for each treatment effect. In column (1) we estimate equation (4) without specifying pre-treatment R&D expenditure $y_{i,t-1}$ as an explanatory variable. This model controls for any time-invariant unobserved heterogeneity (fixed effects) in the log level of R&D expenditure. If regression to the mean behavior is present in the data, implying that firms that have a high pre-treatment R&D expenditure tend to have lower R&D growth rates and vice versa, and if firm-specific characteristics in $X_{i,t-1}$ are positively correlated with pre-treatment R&D level $y_{i,t-1}$, the regression results in column (1) are downward biased (Allison 1990). Hence, specification (2) adds pre-treatment R&D expenditure.[27] The results in column (2) show that the pre-treatment level is highly significant and a comparison of the estimated coefficients with the first model confirms a downward bias for almost all coefficients. Therefore, model (2) is our preferred specification. The control variables behave as expected. The R&D growth rate is higher for firms that have R&D experience, but it declines with the pre-treatment level of R&D expenditures. Furthermore, higher pre-treatment sales, patent stock, profitability and younger firm age are associated with higher growth in R&D expenditure. For firm size, we find an inverse U-shaped effect on R&D growth with an estimated turning point of about 2003 employees.

The main parameter of interest γ_{ITT} is a measure of the effectiveness of the R&D subsidy policy when misappropriation of funds occurs. $\gamma_{ITT} \leq 0$ indicates full crowding out which implies that, on average, R&D subsidies do not raise total R&D expenditures. If $\gamma_{ITT} > 0$, public R&D funds increase total R&D investment, which encompasses both, a situation in which total R&D expenditure increases by less (partial crowding out) and by more (i.e. additionality) than the subsidy amount S. Since Z is a binary treatment dummy, γ_{ITT} does not allow us to infer directly how much of R&D expenditure is stimulated by a subsidy of 1 RMB and whether the additional R&D is more or less than the subsidy amount S. However, Einioe (2014) shows that under one additional assumption, γ_{ITT} becomes informative about the degree of partial crowding out and additionality. Let us assume we want to test the null hypothesis of at least 50% crowding out, $h \geq 0.5$, against the alternative that the subsidy leads to less than 50% crowding out. Under the null

[27] This model is equivalent to a dynamic model in which $y_{i,t+1}$ is regressed on $\alpha_0, y_{i,t-1}, X_{i,t-1}, Z_{it}$ and $\phi_j, \phi_t, \phi_{jt}$.

hypothesis the following condition must hold: $Y_{i,t+1} \leq Y_{i,t-1} + (1-h)S$, where h denotes the crowding out rate and $Y_{i,t+1}$ and $Y_{i,t-1}$ denote the post- and pre-treatment levels of total R&D expenditure. Now, assuming that the amount of subsidy S equals a share s of the post-treatment level of total R&D expenditure $Y_{i,t+1}$, and using the maximum subsidy rate of 50% that is usually paid by the government as estimate for s, i.e. $s = 0.5$, we get the following condition: $\log(Y_{i,t+1}) - \log(Y_{i,t-1}) = y_{i,t+1} - y_{i,t-1} \leq \log(1/(1-(1-h)s)) = 0.288$. Thus, under the null hypothesis the log growth of R&D expenditure due to the subsidy must be below this threshold and we can test whether $\gamma_{ITT} \leq \log(1/(1-(1-h)s)) = 0.288$. Similarly, we get a threshold of $\gamma_{ITT} \leq 0.47$ for the null hypothesis of more than 25% crowding out and $\gamma_{ITT} \leq 0.693$ indicates no additionality whereas $\gamma_{ITT} > 0.693$ confirms additionality (no crowding out). The key assumption of a 50% subsidy rate provides a conservative estimate of the threshold value. A lower subsidy rate is associated with a lower threshold on γ_{ITT}.[28] In columns (1) and (2), γ_{ITT} is 0.877, indicating a medium-level partial crowding out, as we reject the null hypothesis of more than 25% crowding out (at the 10% level) but also find no support for additionality.

Next, we estimate the CACE based on equation (5) with weighted 2SLS and employ the randomized Z as an instrument for the endogenous D. The highly significant point estimate in the first stage and the large Kleibergen-Paap F-statistic support the relevance of the IV. The coefficients of control variables are similar. In column (3) and (4), γ_{CACE} is 2.137, so it rejects both full and partial crowding out at the 5% level. That is, we find significant evidence for additionality among compliers. Comparison of γ_{ITT} and γ_{CACE}, which estimate effectiveness and efficacy, respectively, shows that the effect of China's R&D policy could have been more than twice as large if misappropriations had not occurred.

In Table 8, we compare ITT and CACE with three other estimators of the treatment effect. In column (1) we estimate the upward biased average treatment effect on the treated (ATT) based on the non-randomized assigned treatment Z and control for the pre-treatment level of R&D expenditures, firm-level controls, and industry, year, and industry-year fixed effects. Compared to our benchmark ITT estimate (see also column (2)), the coefficient of 1.338 is overestimated and erroneously confirms additionality at the 5% level. Next, we compare the effects of the actual treatment D and in column (3) estimate the as-treated effect, which compares outcomes of compliant assignees with a control group consisting of non-assignees and noncompliant assignees, ignoring that the compliance decision is endogenous. This comparison should reveal an upward bias because compliers have a higher expected outcome than the control group does. Then, in column (4) the per-protocol effect is estimated by excluding the group of observed noncompliers. We find an effect that is similar in size to the as-treated effect. However, both estimates are overestimated by about 50% compared to our benchmark CACE estimate.

[28]The effective subsidy rate might be lower either because the government covers only a lower proportion of R&D costs or because not all R&D projects receive public funding. The assumption of a maximum subsidy rate of 50% seems to be sensible also in the Chinese context; see "National Science and Technology Plan and Special Funds Subsidy Management Regulations" and related explanations.

Table 7: ITT and CACE of R&D subsidies on growth of R&D expenditures

	ITT		CACE	
	(1)	(2)	(3)	(4)
Z	0.877***	0.877***		
	(0.304)	(0.283)		
D			2.137***	2.137***
			(0.727)	(0.674)
Pre-treatment R&D level $_{t-1}$ (log)		−0.664***		−0.694***
		(0.026)		(0.026)
R&D experience $_{t-2}$ (0/1)	−2.685***	2.045***	−3.028***	1.915***
	(0.396)	(0.406)	(0.417)	(0.398)
Employment $_{t-1}$ (log)	1.023	2.600***	0.831	2.479***
	(0.855)	(0.808)	(0.831)	(0.771)
Employment $^2_{t-1}$ (log)	−0.073	−0.171***	−0.060	−0.163***
	(0.062)	(0.060)	(0.061)	(0.057)
Net fixed assets $_{t-1}$ (log)	0.028	−0.092	0.061	−0.064
	(0.214)	(0.191)	(0.211)	(0.186)
Sales $_{t-1}$ (log)	0.232	0.560***	0.209	0.552***
	(0.214)	(0.195)	(0.209)	(0.190)
Age $_{t-1}$ (log)	0.587	−1.355***	0.770**	−1.259***
	(0.368)	(0.326)	(0.369)	(0.316)
Patent stock $_{t-1}$ (log)	0.286**	0.402***	0.287**	0.408***
	(0.131)	(0.117)	(0.129)	(0.114)
Profitability $_{t-1}$ (0/1)	0.187	0.826**	0.102	0.770**
	(0.475)	(0.405)	(0.468)	(0.391)
Minority SOE $_{t-1}$ (0/1)	−0.671	0.094	−0.682	0.117
	(0.508)	(0.458)	(0.500)	(0.443)
Privatized $_{t-1}$ (0/1)	−0.484	−0.366	−0.441	−0.317
	(0.535)	(0.460)	(0.527)	(0.444)
De-novo private $_{t-1}$ (0/1)	0.104	0.509	0.046	0.469
	(0.499)	(0.454)	(0.493)	(0.443)
Crowding-out test (p-value)				
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.026	0.019	0.005	0.003
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.090	0.075	0.011	0.007
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.272	0.257	0.024	0.016
IV 1^{st} stage (Z)			0.411***	0.411***
			(0.017)	(0.017)
KP F-statistic			599.1	608.6
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes
Observations	5166	5166	5166	5166

Notes: Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. The estimate for IV 1st stage (Z) is also the estimate for ITT_D. KP denotes the Kleibergen-Paap Wald F-test on weak instruments (Kleibergen and Paap 2006).

6.2 Heterogeneous Treatment Effects

This section investigates the heterogeneity in treatment effects across time, subsidy size (payments), industry, and ownership. In Table 9, we split the sample into the pre-MLP period (2001-2006) and the MLP period (2007-2011) to determine whether the MLP has improved the design of R&D policy. Particularly outstanding is the result that both γ_{ITT} in

Table 8: Comparison of treatment effects

	Biased ATT (1)	ITT (2)	As-treated (3)	Per-protocol (4)	CACE (5)
Non-randomized Z	1.338*** (0.287)				
Randomized Z		0.877*** (0.283)			
Non-instrumented D			3.291*** (0.374)	3.097*** (0.375)	
Instrumented D					2.137*** (0.674)
Crowding-out test (p-value)					
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.000	0.019	0.000	0.000	0.003
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.001	0.075	0.000	0.000	0.007
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.012	0.257	0.000	0.000	0.016
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes
Observations	5166	5166	5166	4685	5166

Notes: Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

column (1) and γ_{CACE} in column (2) fail to reject total crowding out in the pre-MLP period, which shows that R&D policy was ineffective in stimulating private R&D expenditure in this period and that the policy's ineffectiveness is not due only to noncompliance but also to poor policy design. In contrast, in the MLP period, γ_{ITT} in column (3) shows only mild partial crowding out, so we reject the null hypothesis of more than 25% crowding out at the 5% level but cannot confirm additionality. γ_{CACE} in the MLP period supports significant additionality at the 5% level. Our overall results reported in Table 7 are therefore primarily driven by the MLP period. Overall, the gap between the estimated effectiveness and efficacy narrows over time, but the effect of China's R&D policy during the MLP period could still be more than twice as large in the absence of misappropriation. The key take-away is that the MLP has improved the design of R&D policy significantly.

To test whether the subsidy size affects the effectiveness of R&D policy, we split the ITT group at the median subsidy amount of 0.555 million RMB into treated firms with small and large public R&D funds, respectively. Strikingly, we find stronger growth in R&D expenditure among firms with below-median R&D subsidies than we do for the benchmark model (Table 7). These results reveal additionality for both ITT and CACE. In the group of firms with small subsidies, the R&D policy would, on average, stimulate private R&D in an ideal situation (no misappropriation), but it has done so already in the given situation with misappropriation, albeit at a lower level. In contrast, for high R&D subsidies, γ_{ITT} and γ_{CACE} in columns (7) and (8) are insignificant, suggesting that R&D subsidies that are (too) high fully crowd out private R&D expenditures, even among compliers. We find similar results when we split the ITT group into firms with single and multiple funding payments.[29] Our findings for both subsidy size and payments suggest that not only misappropriation but also overfunding and coordination failure, i.e. misallocation, on the side of R&D programs contribute to low policy effectiveness.

Table 9: Treatment effect heterogeneity by time and subsidy size

	2001-2006		2007-2011		Small R&D subsidy		Large R&D subsidy	
	ITT (1)	CACE (2)	ITT (3)	CACE (4)	ITT (5)	CACE (6)	ITT (7)	CACE (8)
Z	0.341		1.056***		1.307***		0.449	
	(0.406)		(0.350)		(0.336)		(0.351)	
D		2.178		2.147***		2.907***		1.208
		(2.505)		(0.693)		(0.720)		(0.915)
Crowding-out test (p-value)								
$H_0 : \gamma \leq 0.287$	0.448	0.225	0.014	0.004	0.001	0.000	0.322	0.157
$H_0 : \gamma \leq 0.470$	0.625	0.248	0.047	0.008	0.006	0.000	0.476	0.210
$H_0 : \gamma \leq 0.693$	0.807	0.277	0.150	0.018	0.034	0.001	0.757	0.287
IV 1^{st} st (Z)		0.157***		0.492***		0.450***		0.372***
		(0.022)		(0.019)		(0.021)		(0.021)
KP F-statistic		48.7		668.3		471.4		314.8
Pre-tre. R&D	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind.-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2083	2083	3083	3083	4757	4757	4759	4759

Notes: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

In Table 10 we study whether R&D policy effectiveness and efficacy differ between high-tech and low-tech industries. Not surprisingly, 54.0% of high-tech firms performed R&D in the pre-treatment year, compared to only 24.3% of low-tech firms, and among R&D performers, high-tech firms had significantly higher pre-treatment levels of R&D. Firms in high-tech industries, which make up for 10.6% of the sample, are not only more likely to receive public funding than low-tech firms are (32.6% vs 13.8%), they also show a higher compliance rate (52.2% vs. 38.0%). The higher support rate in high-tech industries reflects China's picking-the-winner R&D policy, and the higher compliance rate suggests higher returns from R&D investment. However, our findings also suggest that R&D subsidies do not increase R&D spending in high-tech industries. Both γ_{ITT} and γ_{CACE} in columns (1) and (2) cannot reject full crowding out. Conversely, in low-tech industries R&D grants incentivize firms to invest more in R&D, and we find additionality for compliers. The crowding out effect in high-tech industries may be explained by the inclusion of processing firms that assemble high-tech products but often are not actual R&D performers (Brandt

[29] Between 2007 and 2011, we also observe the *number* of annual R&D subsidy payments. Of all grantees, 57.3% receive single and 42.7% receive multiple payments (up to 35) per year, with an average of 2.2 payments and an average payment value of 336,343 RMB. For single-payment firms both γ_{ITT} and γ_{CACE} confirm additionality at the 10% and 1% significance level, whereas for multi-payment firms γ_{ITT} fails to reject total crowding out and γ_{CACE} only rejects 50% crowding out at the 10% significance level.

and Rawski 2019, p. 27).[30] Subsidizing R&D-performing firms that have no financial constraints may be another reason. Because the choice to engage in R&D is likely to be more grant-dependent in low-tech industries than in high-tech industries, the inducement effect is stronger for those firms (see also González and Pazó (2008)).

Table 10: Treatment effect heterogeneity by industry and ownership

	High-tech industries		Low-tech industries		Private ownership		State ownership	
	ITT (1)	CACE (2)	ITT (3)	CACE (4)	ITT (5)	CACE (6)	ITT (7)	CACE (8)
Z	0.803		0.874***		1.031***		-0.077	
	(0.813)		(0.313)		(0.357)		(0.449)	
D		1.537		2.303***		2.280***		-0.231
		(1.411)		(0.803)		(0.758)		(1.298)
Crowding-out test (p-value)								
$H_0 : \gamma \leq 0.287$	0.256	0.188	0.030	0.006	0.019	0.004	0.791	0.655
$H_0 : \gamma \leq 0.470$	0.336	0.225	0.098	0.011	0.059	0.009	0.888	0.705
$H_0 : \gamma \leq 0.693$	0.444	0.275	0.282	0.023	0.173	0.018	0.957	0.762
IV 1^{st} stage (Z)		0.522***		0.379***		0.452***		0.332***
		(0.040)		(0.018)		(0.020)		(0.025)
KP F-statistic		171.8		422.3		510.4		171.9
Pre-tre. R&D	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ind.-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	546	546	4620	4620	2873	2873	2293	2293

Notes: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

Since ownership transformation is an essential feature of the Chinese economy in our period, columns (5) to (8) study whether the R&D policy is equally effective in private enterprises (POE: including privatized and de-novo private) and state-owned enterprises (SOE: including majority state-owned and minority state-owned). In the pre-treatment period, the share of R&D performers is significantly higher among POEs than it is among SOEs (30.7% vs. 23.8%), and among R&D performers POEs also have significantly higher pre-treatment R&D levels. POEs, which make up 56.2% of the sample, are also more likely to receive public funding than SOEs are (18.7% vs 12.1%), and they have a higher compliance rate (45.0% vs. 33.5%). For POEs, the treatment effects are slightly larger than in our benchmark model but our overall finding that γ_{ITT} shows medium-level partial crowding out, whereas γ_{CACE} supports significant additionality, still holds. This finding is in clear contrast to the results for SOEs. Neither γ_{ITT} in column (7) nor γ_{CACE} in column (8) reject full crowding out, indicating that R&D policy has been completely ineffective

[30]To identify the involvement in processing activities, we observe firms' processing trade activities and firm locations that are in close proximity to processing zones. The share of processing-related firms in high-tech industries is almost twice that in low-tech industries. In high-tech industries, processing-related firms have a significantly lower R&D intensity than other high-tech firms. Taken together, these findings make plausible a lower policy impact.

in stimulating R&D expenditure in SOEs, even among compliers. Given our theoretical model, this finding also suggests that SOEs are not financially constrained because R&D subsidies do not increase the (optimal) level of R&D.

6.3 Robustness Tests

In this section, we scrutinize the robustness of the parameter estimates derived in our benchmark model (Table 7, columns (2) and (4)). Table 11 reports a comprehensive set of robustness tests for γ_{ITT} and γ_{CACE}, addressing a number of potential concerns.

Measurement issues. In columns (1) to (4), we deal with unknown timing or compositional issues related to the receipt of R&D subsidies using the four alternative definitions of misappropriation laid out in section 3.2. We found as stylized fact that most firms fully misuse R&D funds. Our benchmark model does not distinguish between full or partial misappropriation. For robustness, we run models that exclude observations with partial misuse in column (5). In column (6), we strictly adhere to the assumption that the largest possible subsidy is 50% of the total post-treatment R&D expenditures and exclude compliers with an R&D subsidy intensity >50% in year t. In column (7), we exclude R&D experience from the specification because the likelihood of observing prior R&D increases with the number of years a firm remains in the panel. In summary, the results are very robust to these measurement issues: γ_{ITT} never supports additionality (except in column (3)) but shows mild to medium partial crowding out, whereas γ_{CACE} always confirms additionality. Finally, we use R&D intensity (ratio of R&D expenditures to sales) instead of R&D growth as the outcome variable in column (8). The results indicate that R&D subsidies significantly increased R&D intensity. Although the size of the effect is not directly comparable to the other estimates, we find a similar loss in effectiveness, that is, the increase in R&D intensity could also have been more than twice as large as without misappropriation.

Table 11: Robustness tests

	R&D subsidies over 3 years	Strict R&D subsidies	Accounting noise excluded	50% R&D subsidies	Partial misuse excluded	$\leq$ 50% R&D support	Without R&D experience	R&D intensity	Economic conditions	Political uncertainty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Z	0.806***	1.017***	1.306***	0.877***	0.806***	0.817***	0.937***	0.540***	0.854***	0.882***
	(0.307)	(0.295)	(0.324)	(0.283)	(0.287)	(0.284)	(0.281)	(0.158)	(0.283)	(0.282)
Crowding-out test (p-value)										
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.046	0.007	0.001	0.019	0.035	0.031	0.010	-	0.023	0.018
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.137	0.032	0.005	0.075	0.121	0.111	0.048	-	0.088	0.072
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.357	0.136	0.030	0.257	0.347	0.331	0.192	-	0.285	0.251
D	1.839***	2.322***	2.559***	2.017***	1.874***	2.073***	2.283***	1.324***	2.069***	2.152***
	(0.683)	(0.652)	(0.624)	(0.636)	(0.651)	(0.703)	(0.670)	(0.379)	(0.670)	(0.673)
Crowding-out test (p-value)										
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.011	0.001	0.000	0.003	0.007	0.006	0.001	-	0.004	0.003
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.022	0.002	0.000	0.008	0.016	0.011	0.003	-	0.009	0.006
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.047	0.006	0.001	0.019	0.035	0.025	0.009	-	0.020	0.015
IV 1^{st} stage (Z)	0.438***	0.438***	0.510***	0.435***	0.430***	0.394***	0.411***	0.408***	0.413***	0.410***
	(0.017)	(0.018)	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
KP F-statistic	651.1	574.3	751.7	688.6	648.2	560.7	613.6	596.2	611.8	607.4
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ITT firms	655	548	641	816	779	794	816	814	812	816
Complier firms	287	240	327	355	335	313	335	333	335	335
Control firms	4328	5698	3718	4350	4350	4350	4350	4328	4326	4350
Observations	4983	6246	4359	5166	5129	5144	5166	5142	5138	5166

Table 11 continued: Robustness tests

	Corruption	Distance to regulator	Est. before 2007	HNTE participants excluded	Non-R&D subsidies	Winsorized	Random R&D subsidies	Pseudo outcome $y_{t-1} - y_{t-3}$	Pseudo outcome $y_{t-1} - y_{t-3}$ pre-tre. R&D level	Compliers excluded & adjusted
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Z	0.888***	0.884***	0.880***	1.016***	0.549*	0.885***	-0.243	0.006	0.001	-0.422
	(0.284)	(0.296)	(0.284)	(0.318)	(0.284)	(0.285)	(0.184)	(0.344)	(0.364)	(0.356)
Crowding-out test (p-value)										
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.017	0.022	0.019	0.003	0.178	0.018	-	-	-	-
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.071	0.082	0.075	0.015	0.391	0.073	-	-	-	-
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.246	0.260	0.256	0.071	0.694	0.250	-	-	-	-
D	2.156***	2.170***	2.142***	3.447***	1.292**	2.155***	-0.590	0.016	0.004	-
	(0.673)	(0.713)	(0.676)	(0.923)	(0.653)	(0.679)	(0.438)	(0.847)	(0.895)	
Crowding-out test (p-value)										
$H_0 : \gamma \leq 0.287$ ($h \geq 50\%$)	0.003	0.004	0.003	0.000	0.062	0.003	-	-	-	-
$H_0 : \gamma \leq 0.470$ ($h \geq 25\%$)	0.006	0.009	0.007	0.001	0.104	0.007	-	-	-	-
$H_0 : \gamma \leq 0.693$ ($h \geq 0\%$)	0.015	0.019	0.016	0.001	0.179	0.016	-	-	-	-
IV 1^{st} stage (Z)	0.412***	0.407***	0.411***	0.337***	0.425***	0.411***	0.411***	0.397***	0.397***	-
	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.019)	(0.020)	
KP F-statistic	613.6	562.2	602.6	357.6	597.5	611.3	568.9	421.5	414.8	-
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes t-3	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ITT firms	812	766	811	627	816	816	816	599	599	481
Complier firms	335	312	333	211	335	335	335	238	238	0
Control firms	4326	4052	4331	4180	4350	4347	4350	3110	3110	4350
Observations	5138	4818	5142	4807	5166	5163	5166	3709	3709	4831

Notes: We conduct balancing for each subsample to maintain a randomized Z. The sizes of the ITT, complier and control groups change as a result of the alternative definitions. Column (9) and column (11) are estimated without Tibet before 2009, because no LMEs are recorded for it during those years. In column (12) distance is missing for 348 observations, but there is no indication that the missing data is not random. Column (17) does not balance the random treatments. In column (18) and column (19) the outcome is $y_{t-1} - y_{t-3}$ because $y_t - y_{t-2}$ is effected by a treatment in t. The effects of Z and D are likewise insignificant for comparable estimates that use lagged outcome $y_{t-2} - y_{t-4}$ or $y_{t-3} - y_{t-5}$. In column (20), compliers are removed after balancing. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

Omitted variable bias. Entropy balancing corrects for selection bias only on observables. If other unobserved variables affect both the treatment and the outcome variables, our estimates would be biased. We address this concern in two ways. First, we add potential time-variant confounders at the provincial level. Our benchmark model does not control for a firm's location, which might affect both the likelihood of receiving an R&D subsidy and growth in R&D expenditures. Column (9) of Table 11 shows the results of controlling for general economic conditions at the province level by including the log of GDP per capita, the share of loss-making firms, and the log of R&D expenditures per LME. We also control for resource curse by including the log of ensured reserves of coal, China's primary source of energy supply. In coal-rich provinces, government and business priorities may be less innovation-oriented, leading to less public support for R&D and lower growth in R&D expenditure because of higher opportunity costs. In column (10), we specifically control for the influence of local political uncertainty, as proxied by the annual turnover of city-level mayors and party secretaries per province. On the one hand, Feng and Johansson (2017) show that a change in local political leaders is associated with a significant decrease in R&D expenditures. On the other hand, a high turnover of political leaders lowers firm-state connectivity, and Fang et al. (2018) find that firms with weaker political connections receive less R&D subsidies. Column (11) accounts for province-level corruption as a potential confounder. Corruption may not only affect assignment and misappropriation of R&D subsidies but also confound R&D investment because of unreliable intellectual property protection and hence a higher uncertainty to appropriate the returns to innovation. In addition to using province-level location information, in column (12) we also directly control for the firm's distance from the location of the relevant regulator (in logs), which is Beijing for central state-owned firms and the provincial capital for all other firms (distances between 0.5 km and 2283 km). Finally, in column (13) we rule out that the firm's choice of location is endogenous to subsidy policies introduced after the MLP and condition on firms established before 2007. Accounting for firm location in these very different ways leaves γ_{ITT} and γ_{CACE} almost unchanged and corroborates the findings of the benchmark model.

Second, using the test proposed by Oster (2019), we check the overall robustness of our estimated treatment effect that is due to the full set of unobserved control variables. The basic idea is that the selection bias inferred from the incomplete set of observed controls is informative about the bias that is due to unobserved controls. The bias-adjusted treatment effect γ^* can be consistently estimated using equation (6):

$$\gamma^* \approx \tilde{\gamma} - \tilde{\delta}[\dot{\gamma} - \tilde{\gamma}] \frac{R^2_{max} - \tilde{R}^2}{\tilde{R}^2 - \dot{R}^2}. \tag{6}$$

$\dot{\gamma}$ and $\dot{R}^2$ stem from a baseline regression with no additional control variables (uncontrolled model), whereas $\tilde{\gamma}$ and $\tilde{R}^2$ are the corresponding counterparts from a regression with observed controls. Since entropy weights are based on the observed controls as well, we follow Gambaro et al. (2019) in neglecting entropy weights for the baseline regression. Columns (1) and (2) of Table 12 report the corresponding estimation results. R^2_{max} is the (unobserved) R^2 of a hypothetical regression on the full set of observed and unobserved controls. Based on randomized data, Oster (2019) derives a cutoff value for $R^2_{max} = \min(1.3 * \tilde{R}^2, 1)$. $\tilde{\delta}$ is a measure of the degree of proportionality, and $\tilde{\delta} > 1$ ($\tilde{\delta} < 1$) means that selection on unobserved control variables is larger (smaller) than selection on observed control variables. Following Oster (2019) and assuming that $R^2_{max} = \min(1.3 * \tilde{R}^2, 1)$ and $\tilde{\delta} = 1$, we get $\gamma^*_{ITT} = 0.504$, the lower bound of the true treatment effect γ_{ITT}, whereas $\tilde{\gamma}_{ITT} = 0.877$ is the upper bound. Since the lower bound γ^*_{ITT} is within the estimated confidence interval of $\tilde{\gamma}$, we can conclude that our estimated ITT is robust to selection on unobservables. Furthermore, γ^*_{ITT} is larger than 0 and larger than 0.287 and 0.470 but not than 0.693. This finding is consistent with our prior result that we reject crowding out of more than 25% but not mild crowding out of less than 25%. Alternatively, we can derive a bound on δ if we set $\gamma^* = 0$ in equation (6). Column (4) shows that the influence of omitted variables must be more than 2.35 times more important than the observed control factors to explain away the positive estimated ITT. For CACE, this threshold is even larger, suggesting a very high relevance of our observed covariates. Most importantly, the ratio of the bias-adjusted ITT and CACE, and thus the lower bound on the loss of effectiveness that we obtain when we account for selection on observed and unobserved control variables, is of a similar magnitude (0.24) to the loss of effectiveness in our benchmark model (0.41) when we account only for selection on observed variables.

Table 12: Test on omitted variable bias of ITT and CACE

	Uncontrolled model	Controlled model	Bounds of γ	Proportionality δ
γ_{ITT}	2.070***	0.877***	$[0.504, 0.877]$	2.351
	(0.270)	(0.283)		
95% CI		$[0.3226, 1.4322]$		
R^2_{adj}	0.0166	0.4024		
γ_{CACE}	2.251***	2.137***	$[2.102, 2.137]$	61.102
	(0.417)	(0.674)		
95% CI		$[0.8171, 3.4573]$		
R^2_{adj}	0.0089	0.4216		

Substitution bias. Another important source of distortion could be a substitution bias that arises in evaluations of single R&D subsidy programs if members of the control group participate in similar programs (Heckman and Smith 1995). Given that we do not evaluate a single program but include subsidy payments from all R&D programs in our analysis, the bias that results from subsidy-based public support for R&D should be (close to) zero. However, tax-based public support for R&D and non-R&D subsidies might affect our results, although these alternative support policies are not restricted to the control group but are equally accessible by the treated firms. Column (14) of Table 11 addresses tax-based R&D support by excluding participants of the HNTE (InnoCom) program during their participation and one year prior.[31] We expect that HNTE participants may spend ex

[31] Firms that are eligible and funded under the tax-based InnoCom program received a status called High New Technology Enterprise (HNTE). We obtained data on all HNTE participants from the program's official webpage and matched this information to our firms.

ante more on R&D to reach the eligibility criterion of 3% R&D intensity, thus also showing higher R&D growth rates. However, accredited participants already have a relatively high R&D intensity, so their R&D is likely to grow at a slower pace ex post. Thus, HNTE participation among the treated group would, ceteris paribus, lower the estimated treatment effect of R&D subsidies, whereas HNTE participation among the control group would increase it. Increases in γ_{ITT} and γ_{CACE} after the exclusion of HNTE participants suggest that the second effect dominates. However, our main findings of mild crowding out in ITT and additionality in CACE are still confirmed. In contrast, controlling for non-R&D subsidies in column (15) decreases the effect of R&D subsidies (toward the lower bound), suggesting that non-R&D subsidies may also redirect internal funds to R&D investments.[32] After this rigorous test, which is seldom done in the R&D evaluation literature, γ_{ITT} indicates strong partial crowding out (but no full crowding out), while γ_{CACE} shows medium-level partial crowding out since we reject the null hypothesis of more than 50% crowding out at the 10% level.

Outliers and placebo tests. Finally, we perform several standard robustness tests. In column (16), we check the influence of outliers by winsorizing continuous variables at the 1^{st} and 99^{th} percentiles but see no substantial differences. Next, we perform three placebo analyses. First, in column (17), we randomly match Z and the related D indicator (i.e., D follows the randomization of Z) to firms. In column (18), we use exactly the same specification as in our benchmark model but replace the original outcome $y_{t+1} - y_{t-1}$ with the lagged pseudo-outcome just before receiving Z, $y_{t-1} - y_{t-3}$, which is known not to be affected by the treatment (Athey and Imbens 2017). The third placebo test in column (19) differentiates from the second by controlling for the pre-treatment R&D level, in $t-3$, of the pseudo outcome. All three placebo tests confirm that the estimated treatment effects in the benchmark model stem from R&D subsidies and do not result from spurious correlations. As a final robustness test, in column (20) we exclude all compliers. As expected, we find no significant effect of R&D subsidies on the growth of R&D expenditures for noncompliant firms, which suggests that the exclusion restriction we imposed to use randomized Z as a valid instrument is not violated. In conclusion, these tests convincingly support the robustness of the parameters γ_{ITT} and γ_{CACE}.

6.4 Alternative Outcomes

In recent years, the evaluation literature has also increasingly focused on the output and behavioral effects of R&D subsidies. Such effects matter particularly when the supply of R&D inputs is inelastic and a policy-induced demand shock rather increases the R&D costs, e.g. through higher wages of scientists, but does not increase the quantity of R&D activity (Goolsbee 1998). In China, educational reforms led to a steady increase in the number of university graduates during our study period that provided a steady supply

[32] In the estimation sample 55.8% of the observations receive non-R&D subsidies, and the correlation between receiving R&D and non-R&D subsidies is 0.3. Non-R&D subsidies are often provided to improve firm competitiveness by reducing costs, and they show a positive correlation with R&D growth.

of human resources for corporate R&D. However, even an increase in the number of researchers employed may not lead to output growth if research productivity declines (Bloom et al. 2020; Boeing and Huenermund 2020). We therefore extend the analysis from input additionality to include output and behavioral additionality, as shown in Table 13. For each outcome variable, we employ equivalent specifications, as in equations (4) and (5). Columns (1) to (3) show positive direct effects of both ITT and CACE on employment, net fixed assets, and sales, confirming output additionality. Similar to the results for R&D growth, the ratio of CACE to ITT of about 2 to 2.5 suggests a substantial loss in effectiveness that is due to misappropriation, also in terms of output indicators. Strikingly, column (4) shows that there is no effect on labor productivity, suggesting that on average R&D policy fails to increase the output per worker through corporate innovation. Such an increase is a long-term policy goal in China because its labor force has been shrinking since 2012. Insufficient productivity gains may be related to the mission-oriented focus of many R&D programs in China. The literature documents that while mission-oriented support may have a positive impact on sales growth and employment (Steinmueller 2010), it is less likely to impact productivity (Griliches 1995).[33] Column (5) to (8) extend the analysis by investigating more specifically the behavioral impact of R&D subsidies on innovation output. R&D subsidies not only lead to growth in R&D, they are also accompanied by an increase in patenting.[34] However, columns (6) to (8) also reveal that, compared to the control group, neither grantees nor compliers file more high-tech IT patents or more joint applications with universities, or file them with more foreign (non-ethnic Chinese) inventors residing in China. These findings suggest that R&D policy failed to induce an observable shift toward the development of high-tech IT technologies, university-industry collaboration, or employment of foreign scientists in China.

[33] The argument is that mission-oriented research leads more often to product innovation than it does to process innovation. Product innovation may lead to output additionality in terms of higher sales but not to an increase in labor productivity if sales growth is paralleled by employment growth. Another view is that the increasing support of "strategic research for the nation," which often is not provided by the market, comes at the cost of inferior economic outcomes.

[34] Our specification allows for a contemporaneous and one-year lagged effect of R&D subsidies on patent applications, as R&D expenditures usually affect patenting with a short lag (Griliches 1990). We focus on invention patent applications at China's patent office and do not consider the quality of patents as standard measures suffer from policy distortion in China.

Table 13: Alternative outcomes

	Employment	Net fixed assets	Sales	Labor productivity	Patent applications	High-tech IT patent applications	University-industry collaboration	Foreign inventors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z	0.048*	0.125***	0.092***	0.041	0.093**	-0.020	0.000	0.003
	(0.026)	(0.030)	(0.026)	(0.026)	(0.041)	(0.023)	(0.002)	(0.003)
D	0.116*	0.306***	0.224***	0.100	0.226**	-0.049	0.001	0.008
	(0.063)	(0.073)	(0.064)	(0.064)	(0.099)	(0.056)	(0.004)	(0.007)
IV 1^{st} stage (Z)	0.411***	0.409***	0.409***	0.409***	0.411***	0.410***	0.411***	0.411***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
KP F-statistic	600.2	593.8	593.9	593.8	599.1	598.0	599.9	598.5
Pre-treatment outcome level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ITT firms	816	814	814	814	816	816	816	816
Complier firms	335	333	333	333	335	335	335	335
Control firms	4350	4334	4328	4328	4350	4350	4350	4350
Observations	5166	5148	5142	5142	5166	5166	5166	5166

Notes: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

6.5 Long-term Treatment Effects

While the benchmark model focuses on the short-run impact of R&D subsidies, Table 14 enlarges the impact period by two years from $t-1$ to $t+3$ to capture potential long-term effects. Column (1) shows that R&D subsidies induce private R&D spending among grantees also in the long term. The effect is greater than it is in the short term as now both γ_{ITT} and γ_{CACE} confirm additionality at the 5% and 1% levels, respectively. At the same time, the results show that the loss in relative efficiency is larger in the long run, as CACE is about three times larger than ITT. This loss is also confirmed for the R&D intensity in column (2). The long-term findings reinforce the argument for improved monitoring. The results also indicate long-term output additionality regarding employment, net fixed assets and sales. For all three outcomes, we find that the long-term impact of R&D subsidies is stronger than the short-term effect, suggesting that grantees persistently hire more (R&D) employees and have sustained increases in their sales and capital growth. Most strikingly, the lack of productivity improvement in the short run is confirmed in the long run, corroborating the notion that China's selective R&D subsidy policy does not induce productivity gains (Cheng et al. 2019), whereas horizontal R&D tax incentives do, at least to some extent (Chen et al. 2021). The insignificant finding implies that R&D projects selected for government funding do contribute the same to productivity growth as R&D projects financed through the market in the comparison group. The significant effect on patenting observed in the short run vanishes in the long run. This is plausible because the majority of China's R&D expenditures are development-oriented, and firms file patents related to the funded project quickly.[35] Taken together, the evidence shows that R&D subsidies help firms to increase inputs and outputs in the short- and long-run, whereas the efficiency in transformation remains unchanged.

[35] In 2011, China spent 4.7% and 11.8% of total R&D expenditure on basic and applied research, respectively, while the remaining 83.5% was spent on development activities. For comparison, in the U.S. the proportion spent on development was 62.8% in 2011 (OECD 2013).

Table 14: Long-term effects t-1 to t+3

	R&D expenditures	R&D intensity	Employment	Net fixed assets	Sales	Labor productivity	Patent applications	High-tech IT patent applications	University-industry collaboration	Foreign inventors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Z	1.421***	0.871***	0.107**	0.146***	0.126***	0.001	0.057	0.000	0.000	-0.003
	(0.350)	(0.195)	(0.042)	(0.053)	(0.048)	(0.044)	(0.052)	(0.024)	(0.002)	(0.003)
D	4.675***	2.876***	0.352**	0.480***	0.415***	0.002	0.186	0.001	0.001	-0.009
	(1.173)	(0.660)	(0.139)	(0.173)	(0.157)	(0.140)	(0.167)	(0.079)	(0.005)	(0.010)
IV 1^{st} stage (Z)	0.304***	0.303***	0.304***	0.305***	0.304***	0.304***	0.304***	0.301***	0.304***	0.304***
	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
KP F-statistic	247.5	243.9	235.0	236.1	234.3	232.9	235.0	232.6	234.8	234.6
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ITT firms	510	506	510	508	506	506	510	510	510	510
Complier firms	155	154	155	155	154	154	155	155	155	155
Control firms	3338	3311	3338	3322	3311	3311	3338	3338	3338	3338
Observations	3848	3817	3848	3830	3817	3817	3848	3848	3848	3848

Notes: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1. For further details see notes of Table 7.

6.6 Effects of Misappropriation

So far, we have left out for what other purposes the misappropriated R&D subsidies have been used. However, from a public welfare perspective, it matters whether the money is invested in other productive uses or spent on private consumption. We cannot perform a full welfare analysis because we do not know exactly on what firms spend their misappropriated R&D subsidies or how much they spend on which kind of use.

To still gain a deeper understanding of the potential welfare effects, we do two exercises. First, we compare noncompliers with non-treated firms (excluding compliers from the sample) and examine whether the occurrence of misappropriation, measured by the binary misappropriation indicator M, which takes the value of 1 if the firm received R&D subsidies ($Z = 1$) but did not use them for R&D ($D = 0$), affects our outcome measures. According to our argument for the exclusion restriction of the CACE, Z affects the potential outcome Y only through the actual treatment D, hence misappropriated grants should have no effect on R&D expenditures. However, if misappropriated grants are instead used for physical investments, we expect to find a direct effect on net fixed assets and potential indirect effects on other outcomes, such as employment, sales, and labor productivity. Table 15 reports our findings, which are based on randomizing the misappropriation indicator M using entropy balancing to address the endogeneity of M. The results shown in column (1) indirectly confirm our exclusion restriction, as M does not significantly affect R&D expenditures, but it does have a significant short-term direct effect on net fixed assets and indirect effects on sales and labor productivity, as shown in columns (4) to (5). These results suggest that firms use at least part of the misappropriated funds for investments that increase sales and also labor productivity. Columns (6) to (9) show corresponding long-term effects. Interestingly, compared to the control group, firms that misappropriate R&D subsidies keep significantly higher investment levels in the longer run, which is likely to be the result of short-term increases in sales. However, sustained sales growth in the long run goes hand in hand with employment growth, which eats up productivity gains.

In our second exercise, we compare noncompliers and compliers regarding their impact of misused and used funds, respectively, on physical investment. We exploit additional information on the level of misappropriation to obtain an estimate of the elasticity of misappropriated subsidies on physical investments. For the subsample of noncompliers, we regress the log level of investment in net fixed assets on the log level of misappropriated R&D subsidies, firm-specific pre-treatment variables (revenue, profits, firm age, ownership) as well as industry, year and industry-year fixed effects. Using a generalized Heckman model to correct for double selection with respect to receiving subsidies and to misappropriating grants (Tunali 1986), we find an elasticity of 0.044, that a 10% increase in misappropriated funds is associated with a 0.44% increase in physical investment. The estimate of the corresponding elasticity of used funds on physical investment for compliers is very close at about 0.039.[36]

[36] Using bootstrapped standard errors, both effects are significant at the 10% and 5% level, respectively.

Table 15: Short-term and long-term misappropriation effects

	R&D expenditures	Patent applications	Employment	Net fixed assets	Sales	Labor productivity	Employment t-1 to t+3	Net fixed assets t-1 to t+3	Sales t-1 to t+3	Labor productivity t-1 to t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
M	-0.215	0.054	0.043	0.137***	0.138***	0.092***	0.114**	0.180***	0.162***	0.038
	(0.340)	(0.051)	(0.033)	(0.033)	(0.034)	(0.036)	(0.049)	(0.055)	(0.057)	(0.053)
Pre-treatment outcome level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ITT firms	481	343	481	481	481	481	355	353	352	352
Complier firms	-	-	-	-	-	-	-	-	-	-
Control firms	4350	4350	4350	4334	4328	4328	3338	3322	3311	3311
Observations	4831	4693	4831	4815	4809	4809	3693	3675	3663	3663

Notes: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1.

To sum up, our exercises show that we can reject the hypothesis that firms use misappropriated R&D subsidies entirely for private consumption, as at least some of the money is invested and increases output and employment in the long run.

7 Conclusion

This study is the first to address the misappropriation of R&D subsidies that results from firms' moral hazard behavior after receiving the subsidy. We extend the theoretical framework of Howe and McFetridge (1976) to study the impact of R&D subsidies on the level of optimal R&D investment when we allow for the possibility of misappropriation. The existence of an extended pecking order of which R&D funds to use first allows us to identify misappropriation. In terms of the effectiveness of R&D policy in stimulating R&D spending, the model shows that compliance can lead to either full crowding out, partial crowding out, or additionality, while noncompliance leads only to full or partial crowding out. We empirically investigate the phenomenon of misappropriation and evaluate its impact on the effectiveness of R&D subsidy policy using Chinese firm-level data for the period 2001-2011.

Our findings show that misappropriation is a major concern in China. Between 2001 and 2011, about 42% of grantees used R&D funds for purposes other than research, representing 53% of the total amount of R&D subsidies. Three stylized facts stand out. First, firms choose either (almost) full misappropriation or no misappropriation at all, which may be rationalized by the indivisibility of R&D projects. Second, misappropriation declines substantially over time from 81% (2001) to 18% (2011), but still remains a major threat. This decline notably coincides with China's seminal change in industrial and innovation policy through the introduction of the Mid- to Long-term Science & Technology Development Plan (MLP) after 2006 (Naughton 2021). Third, in line with our theoretical framework, we find that firms' misappropriation behavior is determined by subsidy size, private internal funds, the rate of return to R&D, and the expected probability of detection and sanctioning costs. In addition, minimum cost thresholds for carrying out R&D projects lead to a U-shaped relationship between R&D subsidies and the likelihood of misappropriation. While insufficiently small R&D subsidies on average do not help a firm to reach the minimum cost threshold and start the R&D project, excessively large subsidies on average exceed a firm's actual R&D funding needs and also lead to more misappropriation.

In considering the effectiveness of R&D policy, noncompliance requires distinguishing between the causal effect of the *assigned* and *actual* treatment, which can be consistently estimated by the intention-to-treat (ITT) effect and the complier average casual effect (CACE), respectively. The ITT shows how effective the R&D policy is in the presence of misappropriation, while the CACE is a measure for the efficacy of the R&D subsidy policy and it indicates how effective the policy could have been in an ideal situation without misappropriation. If R&D subsidies fail to stimulate additional R&D investments, comparing ITT and CACE can help to understand whether the failure originates from flaws in the policy design or implementation (in the sense of insufficient monitoring), or both. Consistent with our theoretical framework, we find additionality for compliers, that is an increase in R&D expenditures beyond the subsidy amount. However, noncompliance pushes down the policy effect toward moderately strong partial crowding out. Our results emphasize that more than half of China's potential R&D subsidy policy impact in stimulating R&D spending is lost due to misappropriation.

Our findings also suggest substantial treatment effect heterogeneity. Before 2007, we find full crowding out for ITT and CACE, suggesting that both policy design and misappropriation render China's R&D subsidy policy ineffective during this period. However, both design and implementation have significantly improved afterwards. Therefore, our overall results are primarily driven by the post-MLP period. Moreover, we find that irrespective of firms' moral hazard behavior, small (below median) R&D subsidies and single R&D subsidy payments stimulate private R&D on average, while large R&D subsidies and multiple R&D subsidy payments lead to full and strong partial crowding out, respectively, even among compliers. These findings suggest that not only misappropriation, but also overfunding and coordination failure, i.e. misallocation (Wei et al. 2017), on the side of R&D programs contributed to a low policy effectiveness during this period.

Considering heterogeneity across industries, we find a full crowding out effect in high-tech industries even among compliers. According to our theoretical framework, this suggest that high-tech firms in China were not financially constrained in their R&D activities during our study period. This is likely due to a rather low innovation capacity and a high prevalence of processing firms that only assemble high-tech products but hardly perform R&D activities. In contrast, we find stronger inducement effects of R&D subsidies in low- and medium-tech industries, suggesting more grant-dependent R&D choices. A similar dichotomy exists with respect to ownership. Primarily soft budget constraints in SOEs (Poncet et al. 2010) lead to full crowding out effects, while, in contrast, R&D subsidies to private firms, because of their generally more limited financial resources, induce more R&D spending and even additionality for compliers.

Beyond input additionality, we find output additionality for employment, sales, net fixed assets, and patenting; but no evidence of productivity gains and behavioral additionality. For almost all indicators considered, positive effects increase in the long term, but so does the loss in effectiveness due to misappropriation. Our result that misappropriated funds are at least partially invested in physical capital is interpreted as the second best outcome from a welfare perspective.

As emphasized by our findings, accounting for the misappropriation of R&D subsidies allows for a more nuanced evaluation of policy effectiveness. In settings where noncompliance has been ignored, the prior literature actually reports the smaller ITT effects of assigned treatment but not the causal impact of the actual treatment. In contrast, studies that have excluded as accounting nuisance all those cases in which R&D subsidies are larger than total R&D expenditures estimate the upward biased per protocol effect. Therefore, the more rigorous econometric strategy presented in this study may also help to better understand inconclusive findings in the R&D evaluation literature and thereby derive a more consistent overview of prior findings.

One limitation our research design has in common with almost all R&D subsidy evaluation studies is that, unlike in a randomized control experiment, the non-randomly assigned treatment requires ex-post quasi-randomization based on observational data. While using entropy balancing to mimicking randomization more closely, this approach still relies on selection on observables. To assess the potential bias stemming from unobserved confounders, we perform the test of Oster (2019). We find compelling evidence that our results are robust to selection on unobservables, that is, it is highly unlikely that unobservable confounders would turn the estimated effects to zero. Most importantly, the lower bound on the loss of effectiveness that we obtain when we would account for selection on observed and unobserved control variables, is of the same order of magnitude.

Follow-up studies can build on our approach to evaluate R&D policies in other countries, especially when weak monitoring cannot be ruled out. It will be interesting to learn whether the general improvements in policy design and implementation found for China can also be confirmed for other countries. China's R&D policy offers substantial support to stimulate R&D investments in firms and to boost productivity, growth and welfare. Our results show that this policy support has led to an increase in R&D spending and output growth, but they also point to remaining inefficiencies without which the policy would have been more effective. Moreover, the policy has not succeeded in increasing productivity growth, which is all the more worrisome given the evidence of insufficient productivity evolution in recent years (Brandt et al. 2020). China's more recent R&D policy should therefore aim not only to increase R&D inputs, but also to stimulate higher productivity growth.

References

Acemoglu, D. and T. Verdier (2000): "The Choice between Market Failures and Corruption," *American Economic Review*, 90, 194–211.

Aerts, K. and T. Schmidt (2008): "Two for the Price of One? Additionality Effects of R&D Subsidies: A Comparison between Flanders and Germany," *Research Policy*, 37, 806–822.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005): "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 120, 701–728.

Aghion, P., J. Van Reenen, and L. Zingales (2013): "Innovation and Institutional Ownership," *American Economic Review*, 103, 277–304.

Allison, P. D. (1990): "Change Scores as Dependent Variables in Regression Analysis," *Sociological Methodology*, 20, 93–114.

Angrist, J., G. W. Imbens, and D. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455.

Angrist, J. D. and G. W. Imbens (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Arqué-Castells, P. and P. Mohnen (2015): "Sunk Costs, Extensive R&D Subsidies and Permanent Inducement Effects," *Journal of Industrial Economics*, 63, 458–494.

Athey, S. and G. W. Imbens (2017): "The State of Applied Econometrics: Causality and Policy Evaluation," *Journal of Economic Perspectives*, 31, 3–32.

Autor, D., D. Dorn, G. Hanson, G. Pisano, and P. Shu (2020): "Foreign Competition and Domestic Innovation: Evidence from US Patents," *American Economic Review: Insights*, 2, 357–374.

Bloom, H., L. Orr, S. Bell, G. Cave, F. Doolittle, W. Lin, and J. Bos (1997): "The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study," *Journal of Human Resources*, 32, 549–576.

Bloom, H. S. (1984): "Accounting for No-shows in Experimental Evaluation Designs," *Evaluation Review*, 8, 225–246.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020): "Are Ideas Getting Harder to Find?" *American Economic Review*, 110, 1104–1144.

Boeing, P. (2016): "The Allocation and Effectiveness of China's R&D Subsidies - Evidence from Listed Firms," *Research Policy*, 45, 1774–1789.

Boeing, P. and P. Huenermund (2020): "A Global Decline in Research Productivity? Evidence from China and Germany," *Economics Letters*, 197, 109646.

Boeing, P., E. Mueller, and P. Sandner (2016): "China's R&D Explosion - Analyzing Productivity Effects across Ownership Types and over Time," *Research Policy*, 45, 159–176.

Brandt, L., J. Litwack, E. Mileva, L. Wang, Y. Zhang, and L. Zhao (2020): "China's Productivity Slowdown and Future Growth Potential," World Bank Policy Research Working Paper 9298, Washington, DC.

Brandt, L. and T. Rawski (2019): "Policy, Regulation, and Innovation in China's Electricity and Telecom Industries," in *Policy, Regulation and Innovation in China's Electricity and Telecom Industries*, ed. by L. Brand and T. Rawski, Cambridge University Press, 1–51.

Cao, C., N. Li, X. Li, and L. Liu (2013): "Reforming China's S&T System," *Science*, 341, 460–462.

Chen, H., F. Schneider, and Q. Sun (2018): "Size, Determinants, and Consequences of Corruption in China's Provinces: The MIMIC Approach," CESifo Working Paper No. 7175, Munich.

Chen, J., D. Cumming, W. Hou, and E. Lee (2014): "Does the External Monitoring Effect of Financial Analysts Deter Corporate Fraud in China?" *Journal of Business Ethics*, 134, 727–742.

Chen, Z., Z. Liu, J. C. S. Serrato, and D. Y. Xu (2021): "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 111, 2065–2100.

Cheng, H., H. Fan, T. Hoshi, and D. Hu (2019): "Do Innovation Subsidies Make Chinese Firms More Innovative? Evidence from the China Employer Employee Survey," NBER Working Paper w25432, Cambridge, MA.

David, P., B. Hall, and A. A. Toole (2000): "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence," *Research Policy*, 29, 497–529.

Einioe, E. (2014): "R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding Based on the ERDF Population Density Rule," *Review of Economics and Statistics*, 96, 710–728.

Fang, L. H., J. Lerner, C. Wu, and Q. Zhang (2018): "Corruption, Government Subsidies, and Innovation: Evidence from China," NBER Working Paper w25098, Cambridge, MA.

Feng, X. and A. Johansson (2017): "Political Uncertainty and Innovation in China," Stockholm School of Economics Asia Working Paper No. 44.

Gambaro, L., J. Marcus, and F. Peter (2019): "School Entry, Afternoon Care and Mothers' Labour Supply," *Empirical Economics*, 57, 769–803.

Glaeser, E. L. and R. E. Saks (2006): "Corruption in America," *Journal of Public Economics*, 90, 1053–1072.

González, X., J. Jaumandreu, and C. Pazó (2005): "Barriers to Innovation and Subsidy Effectiveness," *RAND Journal of Economics*, 36, 930–950.

González, X. and C. Pazó (2008): "Do Public Subsidies Stimulate Private R&D Spending?" *Research Policy*, 37, 371–389.

Goolsbee, A. (1998): "Does Government R&D Policy Mainly Benefit Scientists and Engineers?" *American Economic Review*, 88, 298–302.

Griliches, Z. (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28, 1661–1707.

——— (1995): "R&D and Productivity: Econometric Results and Econometric Measurement Issues," in *Handbook of the Economics of Innovation and Technological Change*, ed. by P. Stoneman, Oxford, UK and Cambridge, MA: Wiley-Blackwell, 52–89.

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

Hainmueller, J. and Y. Xu (2013): "Ebalance: A Stata Package for Entropy Balancing," *Journal of Statistical Software*, 54, 1–18.

Heckman, J. J. and J. A. Smith (1995): "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, 9, 85–110.

Honoré, B. E. and L. Hu (2018): "Selection without Exclusion," mimeo.

Hottenrott, H. and B. Peters (2012): "Innovative Capability and Financing Constraints for Innovation: More Money, More Innovation?" *Review of Economics and Statistics*, 94, 1126–1142.

Howe, J. D. and D. G. McFetridge (1976): "The Determinants of R & D Expenditures," *The Canadian Journal of Economics*, 9, 57–71.

Howell, S. T. (2017): "Financing Innovation: Evidence from R&D Grants," *American Economic Review*, 107, 1136–1164.

Hsieh, C.-T. and Z. Song (2015): "Grasp the Large, Let Go of the Small: the Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, 2015, 295–366.

Hu, A. G. Z. and Y. Deng (2018): "Does Government R&D Stimulate or Crowd Out Firm R&D Spending? Evidence from Chinese Manufacturing Industries," *Economics of Transition*, 27, 497–518.

Imbens, G. W. and J. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Imbens, G. W. and D. B. Rubin (2015): *Causal Interference for Statistics, Social, and Biomedical Science: An Introduction*, NY: Cambridge University Press.

Kleibergen, F. and R. Paap (2006): "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97–126.

Kline, P. and C. R. Walters (2016): "Evaluating Public Programs With Close Substitutes: The Case of Head Start," *Quarterly Journal of Economics*, 131, 1795–1848.

Lerner, J. (1999): "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program," *Journal of Business*, 72, 285–318.

Liu, X., X. Li, and H. Li (2016): "R&D Subsidies and Business R&D: Evidence from High-tech Manufacturing Firms in Jiangsu," *China Economic Review*, 41, 1–22.

Ministry of Finance (2000): "Accounting Regulations for Enterprises: Subsidy Income," .

——— (2006): "Accounting Standards for Business Enterprises: No. 16 - Government Subsidies," .

Naughton, B. (2013): *Wu Jianling: Voice of Reform in China*, MIT Press.

——— (2021): *The Rise of China's Industrial Policy 1978 to 2020*, Universidad Nacional Autónoma de México.

OECD (2013): *OECD Science, Technology and Industry Scoreboard*, Paris.

Oster, E. (2019): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 37, 187–204.

Peters, B., M. J. Roberts, V. A. Vuong, and H. Fryges (2017): "Estimating Dynamic R&D Choice: An Analysis of Costs and Long-run Benefits," *RAND Journal of Economics*, 48, 409–437.

Poncet, S., W. Stein, and H. Vandenbussche (2010): "Financial Constraints in China: Firm-level Evidence," *China Economic Review*, 21, 411–422.

Rong, Z., X. Wu, and P. Boeing (2017): "The Effect of Institutional Ownership on Firm Innovation: Evidence from Chinese Listed Firms," *Research Policy*, 46, 1533–1551.

Steinmueller, W. E. (2010): "Economics of Technology Policy," in *Handbook of the Economics of Innovation*, ed. by B. Hall and N. Rosenberg, North Holland, Amsterdam: Elsevier, 1181–1218.

Stuart, E. (2010): "Matching Methods for Causal Inference: A Review and a Look Forward," *Statistical Science*, 25, 1–21.

Stuart, T. and Y. Wang (2016): "Who Cooks the Books in China, and Does it Pay? Evidence from Private, High-Technology Firms," *Strategic Management Journal*, 37, 2658–2676.

Takalo, T., T. Tanayama, and O. Toivanen (2013): "Estimating the Benefits of Targeted R&D Subsidies," *Review of Economics and Statistics*, 95, 255–272.

Tunali, I. (1986): "A General Structure for Models of Double-election and an Application to a Joint-Migration/Earnings Process with Remigration," *Labor Economics*, 8, 235–282.

Wang, Y. and J. Li (2014): "Fraud and Innovation: Is There a Cheater's Discount?" mimeo.

Wederman, A. (2004): "The Intensification of Corruption in China," *The China Quarterly*, 180, 895–921.

Wei, S. J., Z. Xie, and X. Zhang (2017): "From "Made in China" to "Innovated in China": Necessity, Prospect, and Challenges," *Journal of Economic Perspectives*, 31, 49–70.

Zuniga-Vicente, J. A., C. Alonso-Borrego, F. J. Forcadell, and J. I. Galan-Zazo (2014): "Assessing the Effect of Public Subsidies on Firm R&D Investment: A Survey," *Journal of Economic Surveys*, 28, 36–67.

Appendix 1: Variable Description

Table A1: Variable description

Variable	Measurement[a,b]
Firm level	
R&D expenditures	Expenditures on research and development (R&D) (log)
Total subsidies	Monetary or non-monetary assets obtained from the government, including tax refund, but excluding capital investments undertaken by the government as a partial owner of the firm (log)
R&D subsidies	Direct subsidies for strict and broad R&D identified using the semi-manual classification approach (see Online Appendix 2) (log)
Strict R&D subsidies	Direct subsidies for R&D (log)
Broad R&D subsidies	Direct subsidies for patents, technology acquisition, technology transformation, and rewards (log)
Non-R&D subsidies	Total subsidies minus R&D subsidies (log)
Employment	Number of employees (log)
Net fixed assets	Fixed asset costs minus accumulated depreciation of fixed assets minus impairment of fixed assets (log)
Sales	Total operating revenue (log)
Age	Number of years since establishment (log)
Profitability	1 if operating profits are positive
Patent applications	Number of invention patent applications at China's patent office (SIPO) (log)
Patent stock	Patent stock in year t is measured as the patent stock in year $t-1$ depreciated by 15% plus the invention patent applications in t (log)
High-tech IT patent applications	Number of invention patent applications that have at least one high-tech IT IPC class, according to the high-tech international patent classification (IPC) by EUROSTAT (log)
University-industry collaborations	Number of invention patent applications that list a university as a co-applicant (log)
Foreign inventors	Number of invention patent applications where at least one inventor has a Chinese address and a non-Chinese family name[c] (log)
State ownership	Categorial variable based on the percentage of shares owned by the state[d](x). *Majority state-owned enterprises*: $x \in [100, 50)$; *minority state-owned firms*: $x \in [50, 0)$; privatized firms: $x = 0$ in year t and $x \in [100, 50)$ for any prior year; *de-novo private firms*: $x \in 0$ in all years
Mutual fund	1 if the firm has domestic mutual fund investors
Distance of firm to relevant regulator	Distance in kilometers between firm's headquarter and the provincial or national capital (measured at middle points of respective 4-digit postcode areas)
High New Technology Enterprise (HNTE) participant	1 if firm is eligible and funded under the tax-based InnoCom program
Province level	
Bureaucrats per LME	Number of civil servants to number of large and medium-sized enterprises (LMEs)[e]
Corruption cases per LME	Number of investigated corruption cases against civil servants to number of LMEs
R&D expenditures per LME	Average R&D expenditure of LMEs

Table A1 continued: Variable description

Variable	Measurement[a,b]
Share of loss-making firms	Number of loss-making industrial enterprises above designated size[f] to total number of industrial enterprises above designated size
Provincial GDP per capita	Gross regional product by province to resident population by province (year-end)
Turnover of city-level mayors and party secretaries	Sum of annual changes of city-level mayors and party secretaries
Ensured reserves of coal	Ensured reserves of coal (100 million tons)
Industry level	
High-tech firm	Firm operates in a high-tech industry[g]

[a] All variables relate to year t if not stated otherwise and all variables in monetary values have been deflated using China's GDP deflator from the World Bank.

[b] In order to deal with values of 0 when taking a log transformation, we added 1 to all values of the following variables and respective sub-categories: R&D expenditures, total subsidies, patent applications, distance of firm to relevant regulator, ensured reserves of coal.

[c] *Foreign inventors.* An inventor is classified as foreign if the family name is different from typical mainland China family names. Thus, our classification only recognizes non-ethnic Chinese foreigners working in China but disregards ethnic Chinese returnees. Due to the high complexity in inventor name disambiguation to reliably differentiate between Chinese returnees and non-returnees, and little prior work in this area, we refrain from this exercise and our measure should be interpreted as a lower bound. As inventor names in PATSTAT are recorded not in Chinese characters but Pinyin, the official romanization system for Standard Chinese in mainland China, we use the following list of 287 typical family names for the identification of Chinese family names and consider all other names as foreign: Ai, An, Ang, Ao, Ba, Bai, Ban, Bao, Bei, Bi, Bian, Bie, Bin, Bing, Bo, Bu, Cai, Cang, Cao, Cen, Ceng, Cha, Chai, Chan, Chang, Chao, Che, Chen, Cheng, Chi, Chong, Chou, Chu, Chuang, Chun, Ci, Cong, Cui, Da, Dai, Dan, Dang, Dao, De, Deng, Di, Diao, Ding, Dong, Dou, Du, Duan, Dun, Duo, E, Er, Fa, Fan, Fang, Fei, Fen, Feng, Fu, Gai, Gan, Gao, Ge, Gen, Geng, Gong, Gou, Gu, Guan, Guang, Guanghua, Guangpu, Gui, Guo, Ha, Hai, Han, Hang, Hao, He, Hei, Heng, Hong, Hou, Hu, Hua, Huai, Huan, Huang, Huangfu, Hui, Huo, Jang, Ji, Jia, Jian, Jiang, Jiao, Jie, Jin, Jing, Jiu, Ju, Jun, Kai, Kan, Kang, Ke, Kong, Kou, Kuai, Kuang, Kuo, Lai, Lan, Lang, Lao, Le, Lei, Leng, Li, Lian, Liang, Liao, Lin, Ling, Liu, Long, Lou, Lu, Luan, Lui, Lun, Luo, Lv, Ma, Mai, Man, Mang, Mao, Me, Mei, Men, Meng, Mi, Miao, Min, Ming, Miu, Mo, Mou, Mu, Na, Nai, Nan, Ni, Nian, Nie, Ning, Niu, Nong, Ou, Ouyang, Pan, Pang, Pei, Peng, Pi, Pian, Piao, Ping, Pu, Qi, Qian, Qiang, Qiao, Qin, Qing, Qiu, Qu, Quan, Que, Ran, Rao, Ren, Rong, Ru, Ruan, Rui, Sai, Sang, Sha, Shan, Shang, Shao, She, Shen, Sheng, Shi, Shou, Shu, Shuai, Shui, Si, Sima, Song, Su, Sui, Sun, Sun', Suo, Tai, Tan, Tang, Tao, Teng, Ti, Tian, Tiao, Tie, Tong, Tu, Tuan, Wan, Wang, Wei, Wen, Weng, Wo, Wong, Wu, Xai, Xi, Xia, Xian, Xiang, Xiao, Xie, Xin, Xing, Xiong, Xiu, Xu, Xuan, Xue, Xun, Yan, Yang, Yao, Ye, Yi, Yin, Ying, Yong, You, Yu, Yuan, Yue, Yun, Zai, Zan, Zang, Zen, Zeng, Zha, Zhai, Zhan, Zhang, Zhao, Zhe, Zhen, Zheng, Zhi, Zhong, Zhou, Zhu, Zhuang, Zhuo, Zi, Zong, Zou, Zu, Zuo.

[d] *State ownership.* Domestic shares are known as A-shares. Not all A-shares are publicly tradable, but at least 25% must be tradable when a firm is listed. Nontradable A-shares comprise state shares, legal person shares, and employee shares. State shares are held by the central government, local governments, and solely SOEs. Legal person shares are held by other domestic institutions including SOEs that are not solely state-owned. Ownership structures in China are highly concentrated and the largest shareholder effectively controls the firm (Rong et al. 2017).

[e] *Industrial large and medium-sized enterprises.* Refers to statistically relevant industries (mining, manufacturing, and the production and supply of electricity, gas and water). Industrial LMEs are defined as firms with at least 300 employees, 30 million RMB revenue and 40 million RMB assets. In September 2011, these thresholds were adjusted to at least 300 employees and 20 million RMB revenue. Note that in regressions we only use related variables until 2010, the year before the adjustment.

[f] *Industrial enterprises above designated size.* Refers to any state-owned and non-state-owned industrial enterprises with annual main business revenue of 5 million RMB or more. In 2006 China's National Bureau of Statistics re-defined the term to refer to any industrial enterprises with annual main business revenue of 5 million RMB or more, excluding state-owned enterprises whose main business revenue fell beneath this threshold starting from 2007. Starting from January 2011, the threshold for categorization as an enterprise above designated size in China was increased, with the requirement for annual main business revenue of 20 million RMB. Note that in regressions we only use related variables until 2010, the year before the adjustment.

[g] *High-tech industry.* We follow the high-tech definition of China's National Bureau of Statistics. The industry codes of the Chinese Securities Regulation Commission (CSRC), version 2001, are in parentheses: Electronic Devices and Components Manufacturing (C51); Other Electronic Equipment Manufacturing (C57); Medical Equipment Manufacturing (C7340); Aviation and Space Craft Manufacturing (C7530); Electric Equipment and Machinery (C76); Instruments, Meters, Cultural and Clerical Machinery (C78); Medicine Manufacturing (C81); Communication and Correlative Equipment Manufacturing (G81); Computer and Correlative Equipment Manufacturing (G83).

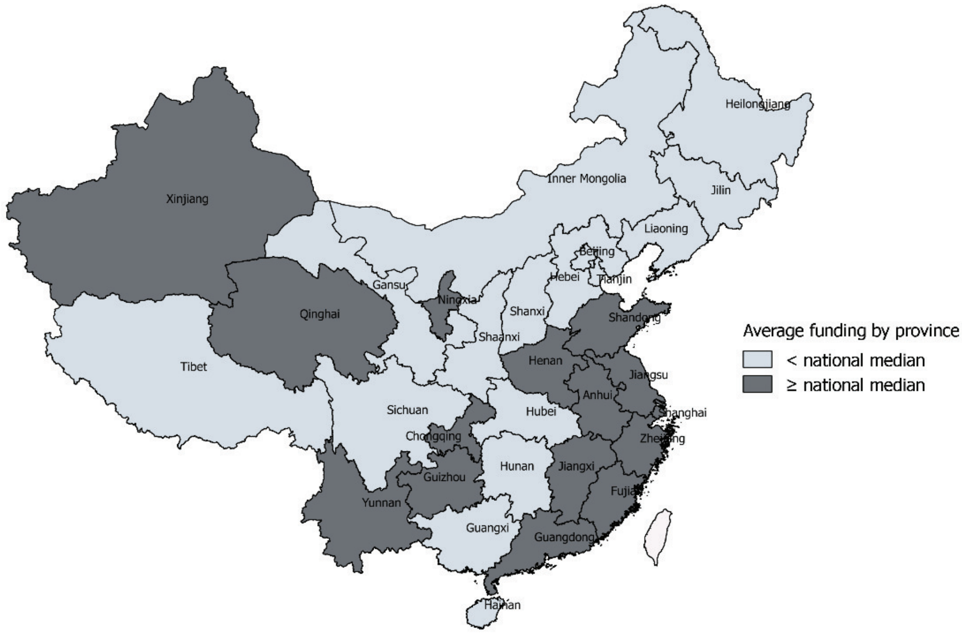
Appendix 2: Misappropriation by Industry, Provinces and Time

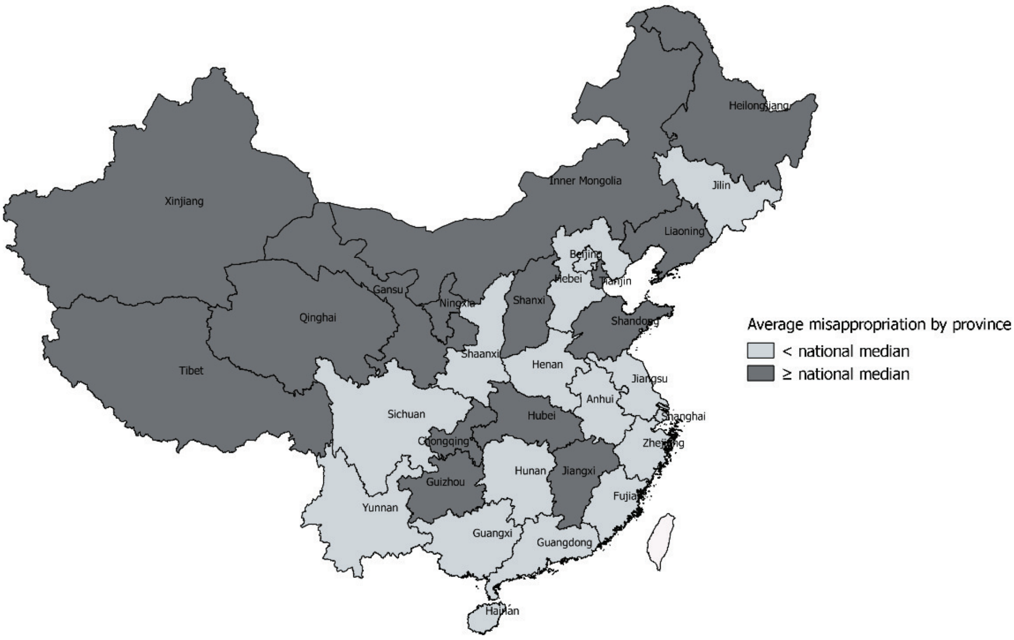
Table A2: R&D expenditures, R&D subsidies and misappropriation by industry

Industry	R&D performers	Grantees	Noncompliers	Obs.
Agriculture	0.189	0.166	0.725	0.019
Mining	0.271	0.115	0.462	0.021
Manufacturing: food & beverages	0.253	0.178	0.672	0.045
Manufacturing: textiles & apparel	0.265	0.244	0.534	0.038
Manufacturing: wood & furniture	0.264	0.236	0.471	0.005
Manufacturing: paper & printing	0.300	0.185	0.379	0.020
Manufacturing: petro-chemistry & plastics	0.369	0.223	0.393	0.113
Manufacturing: electronics	0.539	0.324	0.308	0.039
Manufacturing: metal & non-metals	0.334	0.180	0.388	0.092
Manufacturing: machinery & instruments	0.537	0.300	0.285	0.168
Manufacturing: pharma & biological products	0.496	0.265	0.335	0.064
Manufacturing: other	0.495	0.333	0.182	0.006
Utilities	0.086	0.046	0.833	0.041
Construction	0.315	0.176	0.542	0.017
Transport, storage, and postal services	0.046	0.049	0.781	0.041
Information technology	0.472	0.311	0.360	0.058
Wholesale and retail trades	0.073	0.093	0.755	0.071
Real estate	0.033	0.050	0.808	0.066
Social services	0.113	0.086	0.500	0.028
Communication and culture	0.096	0.106	0.909	0.007
Conglomerates	0.122	0.173	0.754	0.043
Total	0.310	0.197	0.417	1.000

Notes: The table displays the share of R&D performers and grantees relative to all firms and of non-compliers relative to grantees. The 2-digit level for manufacturing industries and the 1-digit level for non-manufacturing industries are displayed according to the CSRC 2001 industry classification.

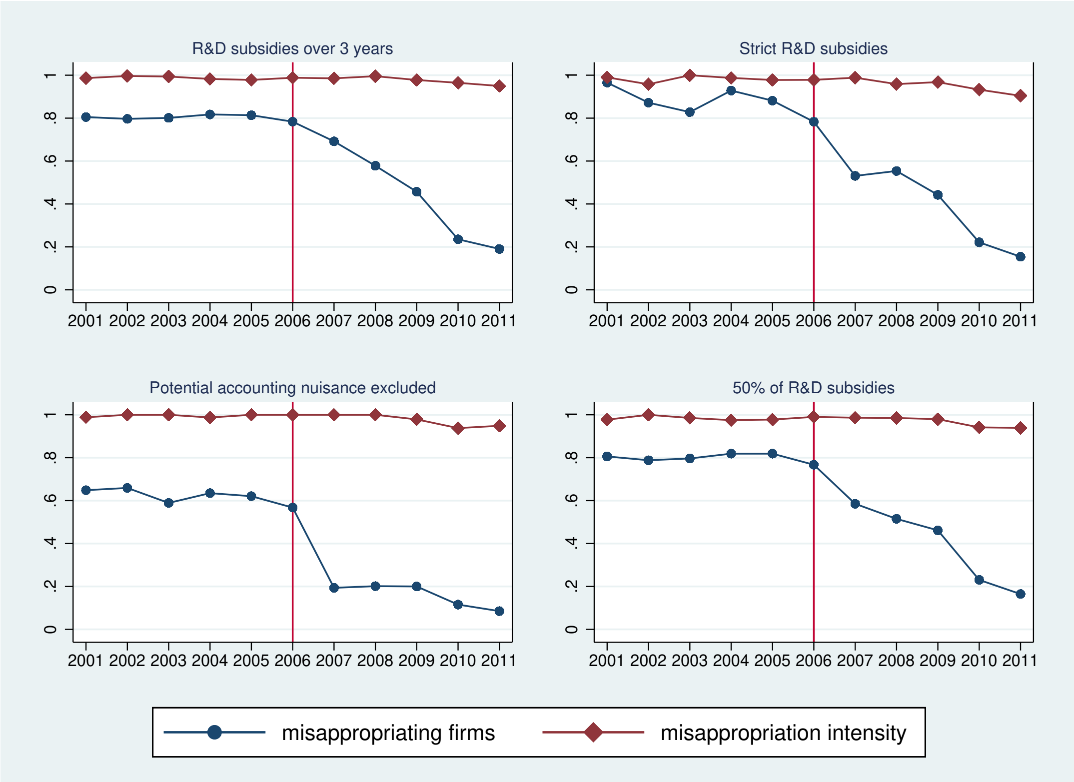
Figure A1: Funding and misappropriation rates by province





Notes: In contrast to the large acreage of Western provinces, only 3% of observations are located in Xinjiang, Qinghai and Tibet.

Figure A2: Development of misappropriation over time using alternative definitions



Notes: Misappropriating firms and misappropriation intensity denote the share of firms that have misappropriated R&D subsidies (extensive margin) and the proportion of misappropriated R&D subsidies to total R&D subsidies (intensive margin), respectively

Online Appendix 1: Institutional Background on R&D Policy and Misappropriation in China

OA1.1 R&D Policy

The State Council aims to transform China into a world leader in science and technology (S&T) before 2050 and invests heavily in innovation policy. The period 2001-2011, underlying our study, is covered by the 10^{th} and 11^{th} Five-Year S&T Development Plans (2001-5 and 2006-11) and the seminal Mid- to Long-term S&T Development Plan (MLP) (2006-20). The MLP's agenda proposes a more integrated innovation policy than prior plans and consists of 99 support policies. In contrast to Five-Year Plans, the MLP also lists more detailed development goals and provides relatively clear guidelines for implementation.[37] The MLP has been consistently followed up since its implementation in 2006 and has recently again been referenced during the Chinese People's Political Consultative Conference in 2019, which emphasizes its overarching role for China's innovation policy. The introduction of the MLP also marked a general change in China's overall industrial policy. Before 2006, China had "very little of it, and what it had was rarely even implemented, much less in an effective way" (Naughton 2021, p.47). However, the "Chinese approach to industrial policy made a 180 degree turn after 2006" and "targeted subsidies quickly became a permanent part of the policy mix" (Naughton 2021, p.66). Over time, administrative structures to prioritize and administer grantees were set up.

A first-order target of the MLP is to increase R&D expenditures of domestic firms. However, it seeks not only to allocate more funds, but also to improve the management of R&D programs. This second goal relates to the selection and monitoring of grantees and coordination between various programs and agencies in order to reduce redundancies, misallocation, and misappropriation of public funds. Finally, the MLP involves a shift toward more mission-oriented funding, and firms preferentially receive funding for R&D projects that align with the government's explicit innovation agenda (Cao et al. 2013). In addition to the amendment of existing and introduction of new R&D programs, other related regulations and policies, like the accounting regulations for R&D expenditures, were implemented after 2006 to incentivize R&D.[38]

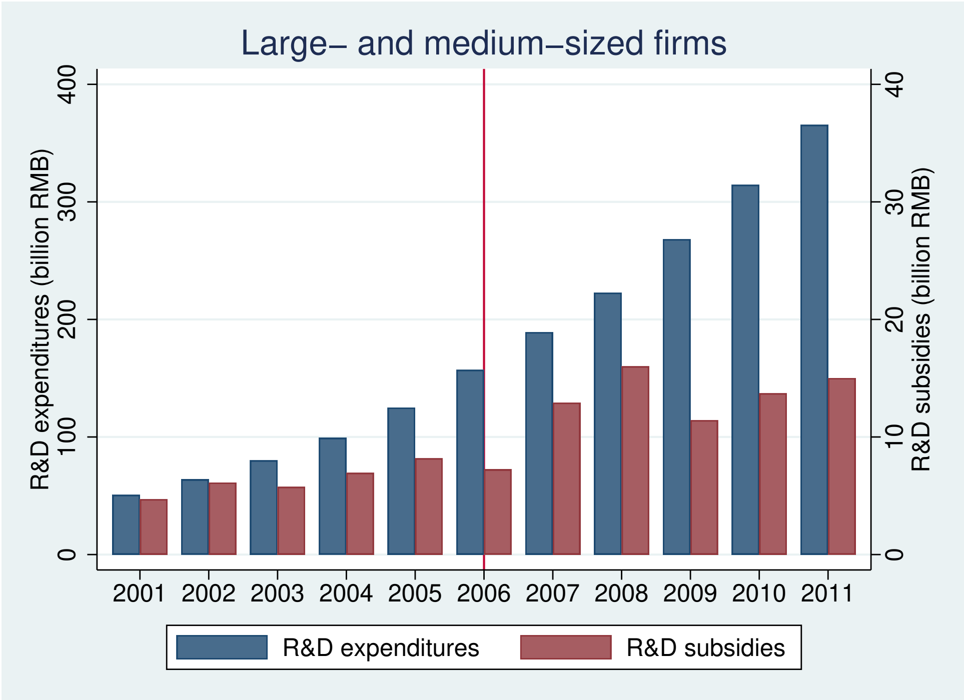
These changes in innovation policy have been accompanied by a tremendous increase in R&D subsidies. Between 2001 and 2011, the annual amount of funding directed to large- and medium-sized firms tripled from 5 billion RMB to 15 billion RMB, while R&D expenditures increased more than sevenfold from 51 billion RMB to 366 billion RMB (see Figure A3). In 2013, China's R&D intensity (ratio of gross expenditures on R&D to GDP)

[37] Each support policy is associated with a lead person in one of the ministries involved. For example, Policy No. 62 "To Develop a Finance Supporting Policy for Encouraging the Innovation of Enterprises" was supervised by Zhang Shaochun from the Ministry of Finance (in cooperation with the National Development and Reform Commission and the Ministry of S&T) and was to be implemented in December 2006.

[38] The accounting regulations for R&D expenditures were amended by the Ministry of Finance: "Accounting Standards for Business Enterprise, No. 6 – Intangible Assets"; see in particular Articles 7 to 9.

exceeded that of the EU28, but has not yet reached the ratio of the U.S. (2019: 2.23% vs. 3.07%; OECD 2021).

Figure A3: China's business R&D expenditures and R&D subsidies



Source: China's National Bureau of Statistics. R&D expenditures and R&D subsidies are in real prices using China's GDP deflator from the World Bank.

The firms in our sample receive *direct R&D subsidies* from programs administered by national ministries, sub-national agencies (e.g. departments and bureaus at the province, prefecture, and county level), and non-classified agencies. Major national R&D programs include the National High-Tech R&D Program (the 863 Program), the National Key Technologies Program, and the State Basic R&D Program (the 973 Program) (see Online Appendix 2 for further details).[39] In principle, all private and state-owned firms may concurrently apply for funding. While eligibility criteria differ by program, the support of (high) technology-oriented and innovative firms is generally emphasized and highlights a picking-the-winner strategy instead of an aiding-the-poor strategy. Based on the raw data for 2007 to 2011, at least 5.38% of individual transactions come from national sources and at least 59.82% come from sub-national sources; while 34.79% are not identified. The average national transaction is more than two times larger than the average sub-national

[39]The *863 Program* aims at increasing firms' innovative capacity. It is in place since 1986, and it has been amended in 2006 and 2011. Independent legal entities registered for more than one year with high capacity for scientific research may apply for R&D projects with a maximum duration between 1.5 and 3 years. There is no upper limit for the grant which is payed as a lump sum in the first year. The *National Key Technologies Program* focuses on solutions for technological problems in social life. It started in 1983 and has been amended in 2006. The maximum duration of funded R&D projects is 3 to 5 years, with a mid-term evaluation for projects exceeding 3 years. Grants generally cover 40% or more of the project's total cost. The *973 Program* supports basic research, is in place since 1997 and has been amended in 2006, 2008, and 2010. Projects can be funded for up to 5 years, and the maximum grant size is 100 to 300 million RMB. Firms receive the payment for the first 2 years together, while the subsequent payment structure is project specific.

transaction: 2.3 vs. 0.9 (nominal) million RMB. Grantees receive between 1 and 35 annual payments, with a median of 2 payments and a mean of 2.8 payments. Among national and sub-national transactions, the average amount received for technology transformation is the highest, followed by technology acquisition, R&D, and patents.

In addition to changes in direct R&D subsidy policies, amendments in the *R&D tax deduction* scheme occurred. Between 1996 and 2002, the super-deduction of 50% of R&D expenditures from taxable income was limited to state-owned and collective industrial enterprises. In 2003, all domestic industrial enterprises with sufficient accounting, auditing, and taxation standards became eligible. As part of the MLP, eligibility was expanded to all enterprises in 2006, although the timing of the subsequent implementation varied across provinces (Sun et al. 2018). In 2008, the State Administration of Taxation provided a unified and simplified framework for the implementation of this R&D tax incentive in China.[40]

Also in 2008, a major corporate tax reform eliminated the dual-track system based on domestic/foreign ownership. A common corporate tax rate of 25% was introduced – replacing a base rate of 33% for all domestic enterprises and a preferential tax rate for foreign-owned enterprises of between 15% and 24%. The InnoCom program awards (private and state-owned) high-tech firms a preferential corporate tax rate of 15% if they have an R&D intensity above a given threshold. The 2008 reform changed the threshold from "a common R&D intensity of 5%, to a size-dependent threshold with a lower hurdle for medium and large firms, 4% and 3%, respectively, and a larger hurdle of 6% for small firms" (Chen et al. 2021, p. 2070).

After 2010, the Strategic Emerging Industry (SEI) initiative became an important component of China's industrial/innovation policy. The SEI strategy targets new industries that present an opportunity for leapfrog latecomer development. Since this policy gained relevance only toward the very end of our study period, we omit further details and refer to Naughton (2021) for a full-fledged discussion.

Despite China's ambitious and generous innovation policy, the internal assessment of the economy's innovation capacity was more modest. In 2014, General Secretary Xi Jinping summarized remaining challenges as follows: "The foundation of China's science and technology innovation is still not solid. The ability of independent innovation, especially the original creativity, is not strong. The situation that core technologies in key areas are under the control of other countries has not fundamentally changed." (People's Daily 2014).

OA1.2 Misallocation and Misappropriation of R&D Subsidies

The large expansion of R&D subsidies was accompanied by several deficiencies. First, regarding the allocation of funds "there is no uniform, national quality control standard,

[40]The framework is laid out in the "Administrative Measures for the Pre-tax Deduction of Enterprise Research and Development Expenses."

nor is there much exchange of information about projects funded across different agencies" as pointed out by Cao et al. (2013, p.460). In such a system, it is more likely that the allocation of R&D subsidies actually does not comply with the program-specific selection rules. The decision to grant subsidies is typically in the hands of individual government officials, rather than peer reviewers and expert panels, which creates the opportunity of accepting bribes and extracting rents from firms (Fang et al. 2018). Critics point out that relations with government officials are more important than research quality to obtain major grants (Shi and Rao 2010). Furthermore, in such a scattered funding landscape without sufficient coordination and information exchange, it becomes more likely that firms seek duplicate funding for the same R&D project from different sources. Second, in addition to misallocation, misappropriation of R&D subsidies occured. Firms propose to use the grants for R&D in their application, but in practice there is little monitoring or enforcement once they receive the funds (Cao et al. 2013). All in all, the steady increase in government budgeting in combination with the lack of coordination and transparency in allocation and subsequent monitoring has led to excess, overlap, and rent-seeking in funding (Cao et al. 2013, Sun and Cao 2014).

In September 2011, public interest was sparked by media reports stating that around 60% of public research funds were misused for non-research purposes.[41] The correctness of the figure was quickly challenged by the Research Propaganda Department of the Chinese Science and Technology Association (September 2011). However, according to subsequent investigations by the Ministry of S&T and the Central Commission for Discipline Inspection, government officials responsible for the administration of national and sub-national R&D programs, intermediaries specialized in subsidy applications, and firms as final recipients were involved in the misuse of funds. In October 2013, S&T Minister Wan Gang still described the state of research funding in China as a "malignant problem" (People's Daily 2013), and in March 2014, the Central Commission for Discipline Inspection announced that it was planning a new round of inspections, including sending a special inspection team to the Ministry of S&T (The Economist 2014).

Inspection groups and accounting agencies detected fraud in more than a third of investigated cases (Central Commission for Discipline Inspection 2015). In one case, fifty officials from the S&T Bureau of Guangdong Province were investigated for taking bribes from firms in exchange for R&D subsidies (The Economist 2014). In Foshan, a city in Guangdong, officials and intermediaries kept 30% of the subsidies handled (The Economist 2014). Intermediaries specialized in public funding and political relationship-building cooperated with misappropriating firms and kept 20% to 50% of the subsidies as consulting fees (Xinhua 2014). Reportedly, misappropriating firms sought to maximize public grants by overstating actual project costs, and then used the R&D subsidies almost entirely for non-research purposes. In 2016, the Ministry of S&T again commented on the original allegations and pointed out that in recent years the use of funds has been generally in line with international practice (People's Daily 2016).

[41] This statement is quoted from the China Youth Daily (31st August 2011) and was widely reprinted in domestic and international media outlets.

OA1.3 Sanctions

During the period 2001 to 2011, sanctions for misappropriation were continuously specified in China's major R&D programs, for example, in the National High-Tech R&D Program (the 863 Program), the National Key Technologies Program, and the State Basic R&D Program (the 973 Program). The initial regulations, which were in force since 2001 or earlier, stated: "For the act of falsifying, intercepting, misappropriating, and squeezing the funds of the project, etc., administrative and economic penalties shall be imposed on the responsible person of the project and the subject (sub-topic). Based on the circumstances, the relevant departments can take measures such as reporting criticism, stopping funding, terminating the project, or disqualifying the project."

Consistent with the MLP's goal of reducing the misappropriation of public funds, a more comprehensive and precise set of sanctions was introduced in September/October 2006 and remained unchanged through the end of our study period. The new sanctions include the immediate stop of funding and termination of the R&D project. In case of confirmed misappropriation, firms will be suspended from applications for national scientific research projects within the next three years, and a public announcement will be made. If a crime is committed, the case will be passed on and handled by the relevant judicial organs, hence, investigations are not only carried out by bureaucrats of the related R&D program, but also include criminal investigations by prosecuting authorities. Altogether, this shows that sanctions became considerably more prohibitive after 2006.

OA1.4 Monitoring

While sanctions are intended to prevent misappropriation, the expected sanctioning costs also depend on the likelihood of detection. Monitoring and evaluations are common practices to detect misappropriation, and the expected sanctioning costs increase with these efforts. Along with tougher sanctions, monitoring efforts have also increased in line with post-2006 MLP reforms. While this has contributed to a decline in misappropriation, the reduction was still not sufficient. In 2014, the Director of Guangzhou's S&T Bureau still stated that in the case of "corruption in the research system, the problem is certainly not the allocation of too many funds, but the misappropriation of funds." (Xinhua 2014). That same year, the State Council and the Ministry of S&T once more advocated improvements in the fund management and evaluation of research programs[42]. They formulated a set of actions that should be taken to, among other things, "(i) clearly define the missions of national R&D programs, (ii) separate the areas of funding, research, and performance evaluation for the sake of checks and balances and accountability, (iii) apply different standards to the evaluation of different types of R&D activities, and (iv) make the reward systems more open and transparent." (Cao et al. 2013).

[42]State Council, 2014, Guofa [2014] No. 11, Opinions on the reform and strengthening of the Central Government's scientific research programs and fund management; Ministry of Science and Technology, Ministry of Finance and National Development and Reform Commission, 2016, Guokefazheng [2016], No. 382, Notice on technological evaluations (for trial implementation).

Additional References

Central Commission for Discipline Inspection (2015): "Inspection Tour Points to Necessary Reform to Prevent Research Funding to Disappear in "Black Holes" (in Chinese)," .

China Youth Daily (2011): "Research Funding Gave Birth to How Many Rich Man (in Chinese)," (August 31, 2011).

Chinese Science and Technology Association, Research Propaganda Department (2011): "Declaration (in Chinese)," (September 8, 2011).

Ministry of Science and Technology, Ministry of Finance and National Development and Reform Commission (2016): "Guokefazheng [2016], No. 382, Notice on Technological Evaluations (for Trial Implementation) (in Chinese)," .

National High-Tech R&D Program (2001): "Special Fund Management Measures," .

——— (2006): "Special Fund Management Measures," .

People's Daily (2013): "Scientific Research Corruption: The Ministry of Communications Misappropriated 186 Million Yuan to Pay Wage Subsidies (in Chinese)," (October 14, 2013).

——— (2014): "Xi Jinping: Speech at the 17th Academic Conference of the Chinese Academy of Sciences and the 12th Academic Conference of the Chinese Academy of Engineering (in Chinese)," (June 10, 2014).

——— (2016): "Where Did the Trillions of Research Funding Go? (in Chinese)," (February 15, 2016).

Shi, Y., and Y. Rao (2010): "China's Research Culture," *Science*, 329, 1128.

State Council (2014): "Guofa [2014] No. 11, Opinions on the Reform and Strengthening of the Central Government's Scientific Research Programs and Fund Management (in Chinese)," .

Sun, Y., and C. Cao (2014): "Demystifying Central Government R&D Spending in China," *Science*, 345(6200), 1006–1008.

Sun, Z., Z. Lei, and Z. Yin (2018): "Innovation Policy in China: Nationally Promulgated but Locally Implemented," *Applied Economics Letters*, 25(21), 1481–1486.

The Economist (2014): "R&D in China, Research and Embezzlement," (March 20, 2014).

Xinhua (2014): "Guangdong Science and Technology Corruption Case Involves Over 50 People, Intermediary Responsible for Pulling Corruption (in Chinese)," (April 31, 2014).

Online Appendix 2: Subsidy Data Classification

We use a semi-manual approach to classify all grant payments into the three categories strict, broad, and non-R&D subsidies. Relevant keywords are obtained by manually screening the raw data and by identifying the category based on various information, e.g. the aim of funding "research and development" or the source of funds "National High Technology Research and Development Program".

OA2.1 Strict R&D Subsidies

First, we identify strict R&D subsidies based on the following keywords.

OA2.1.1 Expenses for R&D, innovation, and science and technology

创新innovation, 新型new design, 新产品new products, 科学science, 科技technology, 科研research, 研发research and development, 研究research, 研制development, 技术technology, 技改technical change, 技术优化technical optimization, 成果转化transformation and conversion of scientific and technological achievements, 科技保险science and technology insurance

OA2.1.2 Expenses for R&D-related training, education, and collaboration

课题research project (often related to universities), 产学研industry-university research collaboration, 实验室laboratory, 院士academician (of Chinese Academy of Sciences or Engineering), 博士后postdoctoral, 引智talent recruitment, 引进智力introduction of intelligence, 智力引进intelligence introduction, 人才推进talent promotion, 英才talent

OA2.1.3 R&D support programs and policies

863 National High Technology Research and Development Program, 973 National Basic Research Program, 131 Leading Researcher/Scientist/Engineer/Technologist Program, 火炬Torch Program, 星火Spark Program, 孵化(abbreviation of 科技孵化器) Science and Technology Incubator Program, 支撑(abbreviation of 国家科技支撑计划) National Key Technology R&D Program, 朝阳产业Sunrise Industry Program, 小巨人Little Giant of Technology Enterprises Program, 科技型中小企业创新基金Technology-based Small and Medium-sized Enterprise Innovation Fund Program

OA2.2 Broad R&D Subsidies

Second, we identify broad R&D subsidies which include grants for patents, technology acquisition, technology transfer, and rewards, based on the respective keywords.

OA2.2.1 Patents

专利patent, 发明invention, 专利申请patent application, 授权patent grant, 官费application fees, PCT, 软件著作权software copyright, 著作权copyright, 知识产权intellectual property

OA2.2.2 Acquisition of foreign technology and experts

国外智力/外国智力/国外专家/外国专家/外智foreign talents/experts, 国外技术/外国技术foreign technology, 国外设备/外国设备/进口设备foreign/imported equipment, 引进国外/引进国际/引进外国/购买外国先进/购买外国先进/技术进口/进口先进技术advanced technology introduction/purchase from abroad

OA2.2.3 Technological transformation

技术改造/技改/技术改/挖潜/改造technology transformation and improvement

OA2.2.4 Rewards for R&D and patents

奖励/表彰/奖reward, 考核examination, 优势企业dominating enterprise, 示范企业(patent) model enterprises, 企业认定recognition of (high-tech) enterprise

OA2.3 Correction of False Positives

Third, we automatically correct for false positives in strict and broad R&D subsidies by searching for keywords related to non-R&D subsidies.

OA2.3.1 Non-R&D subsidies

贴息/贷款soft/free loan, 税收优惠/税优惠/税收返还/税返还/纳税/增值税/退税tax reduction, 出口exports, 管理创新innovation in management, 企业培育development of enterprise, 节能energy conservation, 水利water conservation, 用电/供电electricity supply, 标准化standardization, 商标/名牌registered trademark, 房租/房补housing subsidies, 参展/展位exhibition, 房地产/土地land use, 固定资产fixed assets, 上市奖励/上市补助/上市资助/补偿public listing reward/subsidies, 市场拓展market expansion, 保增长economic growth maintenance, 贡献contribution (to tax income/economy), 扩产production expansion, 质量quality, 金融危机financial crisis, 灾后/救灾disaster relief, 排污pollution emission , 物流logistics and transportation, 就业employment, 社保social insurance, 整治industry regulation, 发展金enterprise development fund, 城市建设city development, 文化产业cultural industry

OA2.4 Manual Check

Fourth, we perform a manual check of every subsidy amount that was classified as strict or broad R&D subsidy by our keyword-based matching algorithm. It follows an assignment of any misclassified item into the correct group, i.e. strict R&D, broad R&D, and non-R&D.

OA2.5 Error Rate

As a final test we randomly draw 1000 observation and again check the accuracy of our semi-manual classification. We identify 25 errors and yield an acceptable error rate of 2.5%.