Effectiveness and Efficacy of R&D Subsidies: Estimating Treatment Effects with One-sided Noncompliance

Philipp Boeing[1,2] and Bettina Peters[1,3,4]

Preliminary draft

January 2019

In evaluating the effectiveness of R&D subsidies, the literature so far has completely neglected the possibility of misappropriation of public funds. This paper contributes to the literature by evaluating the causal effect of R&D subsidies on R&D expenditures when monitoring is weak and misappropriation takes place due to moral hazard behavior. Our analysis is based on Chinese firm-level data for the period 2001-2011. Misappropriation is a major concern as we calculate that 42% of grantees misused R&D subsidies, corresponding to 53% of the total amount of R&D subsidies. In a setting with one-sided noncompliance to funding contract rules, we differentiate between the intention-to-treat (ITT) effect and the complier average causal effect (CACE). The ITT shows how effective the R&D policy was in practice when misappropriation exists. The CACE, in contrast, depicts how effective the policy could have been without misappropriation and thus is a measure for the efficacy of the R&D subsidy policy. Combining entropy balancing and IV methods to estimate both ITT and CACE, the ITT results show mild partial crowding out of R&D expenditures. Most strikingly, however, the CACE turns out to be more than twice as large as the ITT and confirms additionality of R&D subsidies. Thus, misappropriation of R&D subsidies considerably undermines the efficacy of Chinese R&D programs.

[1] Centre for European Economic Research (ZEW), Mannheim, Germany.
[2] Peking University, China Center for Economic Research (CCER), Beijing, China.
[3] University of Luxembourg, Luxembourg.
[4] MaCCI Mannheim Centre for Competition and Innovation, Germany.

Acknowledgement: We are grateful to Loren Brandt, Georg Licht, Xielin Liu, Kazuyuki Motohashi, Berry Naughton, Pete Suttmeier and Xiaobo Zhang for their helpful comments and discussions. We thank participants at the Creativity and Innovation & Entrepreneurship Symposium (Maastricht, 2018) as well as at seminars at Peking University (2018), Tsinghua University (2018), University of Chinese Academy of Science (2018), University of Luxembourg (2018), Chinese Academy of Science and Technology for Development (2019), and Max Planck Institute Munich (2019).

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

1 Introduction

Most countries worldwide 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 due to market failure rooted in spillovers and financial constraints. However, public funding might also simply crowd out private financing of R&D. An increasing literature has empirically evaluated the effectiveness of R&D policies by estimating the treatment effect of R&D subsidies (for surveys see David et al. 2001 and Zuniga-Vicente et al. 2014). But the literature so far has completely neglected the possibility of noncompliance among supported firms when evaluating R&D programs. Takalo et al. (2013) point out that with monetary treatments moral hazard temptations are certainly possible in practice, and this may lead to the misappropriation of R&D subsidies. For example, in the US Small Business Innovation Research Program firms propose to 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). In general, governments may find it hard to preclude misappropriation due to missing, inefficient or unreasonably expensive monitoring mechanisms.

In identifying the causal effect of R&D subsidies, selection bias is a primary concern since R&D subsidies are hardly 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 introduces an upward bias of the estimated effect of the R&D subsidy. Matching, IV or (conditional) Diff-in-Diff estimators have been used to address this selection bias. However, even in a setting with a randomized R&D allocation, noncompliance to funding contract rules usually creates an additional source of selection bias which is not addressed by standard estimators. This bias occurs because firms deliberately decide whether to comply or not based on the expected outcome of using the funds for research purposes compared to alternative uses.

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

This study is the first to address and identify misappropriation of R&D subsidies and to investigate the consequences of such noncompliant behavior on the effectiveness of R&D policy in stimulating firms' R&D expenditures. We use Chinese firm-level data for the period 2001-2011. China is an attractive case to study this question for at least two reasons. First, the Chinese State Council wants the country to become innovative by 2020 and a world leader in science and technology by 2050. To increase firms' R&D expenditures policymakers have substantially enlarged their R&D subsidy programs at national and sub-national levels. The time period 2001-2011 is covered by the 10th and 11th Five-Year Science and Technology Development Plans and, after 2006, by the Mid- to Long-term Science and Technology Development Plan (MLP) which fundamentally changed China's innovation policy. A major target of the MLP is to increase R&D expenditures of domestic firms, and private as well as state-owned firms may apply for public R&D funding. While eligibility criteria differ by R&D program, the emphasis on (high) technology-oriented and innovative firms highlights a picking-the-winner instead of aiding-the-poor strategy. Between 2001 and 2011 the annual amount of increasingly mission-oriented funding directed to large- and medium-sized firms quintupled from 4 to 21 billion RMB, amounting to a period-sum of 123 billion RMB, while R&D expenditures increased more than tenfold from 44 billion RMB to 503 billion RMB (see Figure A1 in Appendix 1).[2]

The second and even more persuasive reason is that misappropriation of R&D subsidies is a major concern in China. The increasing allocation of government funds was accompanied by deficiencies in funding assignment and monitoring. In many programs firms proposed to use the grant for R&D in its application, but in practice there was little monitoring or enforcement once funds are received, which allows for moral hazard behavior (Cao et al. 2013). This problem was already identified and tackled in the MLP as it not only calls for more funds 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 the misallocation of public funds. Although the MLP led to improvements, these were still not sufficient. In September 2011 public interest was sparked by media reports stating that around 60% of public research funds were misused for non-research purposes. According to subsequent investigations by the Ministry of Science and Technology and the Communist Party's Central Commission for Discipline Inspection, administrators of R&D programs, intermediaries

[2] A more detailed description of the Chinese R&D policy and misappropriation of R&D subsidies can be found in Appendix 1.

specialized in subsidy applications, and firms as final recipients were involved in misappropriation of R&D subsidies. The confirmation of 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 misappropriated funds, corresponding to 53% of the total amount of R&D subsidies. Regarding misappropriation we find three additional stylized facts in our data: First, firms either chose (almost) full misappropriation or not to misappropriate any funds which may be rationalized by indivisibilities of R&D projects. Second, there is a substantial decline in misappropriation over time from 78% (2001) to 18% (2011) along the extensive margin of misappropriation. This decline emerges especially after 2006 which coincides with the implementation of new innovation policies. Third, misappropriation is not random but depends on policy induced changes in returns to R&D at the macro level, industry characteristics, as well as prior R&D investments and profitability at the firm level.

Despite the growing importance of R&D and R&D policy in China, there are only a few studies that have evaluated the causal effect of R&D subsidies on private R&D expenditures in China. Instead of a single R&D program all studies observe aggregated grants received from one or more programs. Boeing (2016) estimates the average treatment effect on the treated (ATT) by combining propensity score matching (PSM) with a difference-in-differences (DiD) estimator. Observing domestic listed firms between 2001 and 2006, he finds a partial crowding-out effect. On average, one public RMB crowds out half a private RMB. Liu et al. (2016) estimate the ATT by applying PSM to cross-sectional survey data of high-tech manufacturing firms in Jiangsu province. For the year 2012 they find that grantees increase private R&D expenditures by 14.3%. Hu and Deng (2018) focus on survey data for large- and medium-sized private-owned manufacturing firms which are observed between 2007 and 2011. They estimate the ATT by combining PSM and DiD estimators and find that treated firms almost double private R&D expenditures compared to the pre-treatment year. Because all three studies use different data sources, it is difficult to make direct comparisons.[3] Nonetheless, several implications stand out. First, it seems that the effectiveness of grants has increased with the introduction of the MLP after 2006. Second, even for the pre-2006 period, Boeing (2016) rejects crowding-out for high-tech firms and privatized firms which is supportive of the findings by Liu et al. (2016) and Hu and Deng (2018). Third, in the post-2006 period only high-tech and

[3] Liu et al. (2016) use data from the Development and Reform Commission of Jiangsu Province and Hu and Deng (2018) from the Science and Technology Division of China's National Bureau of Statistics. Both data sources are generally not accessible for other researchers which is a limitation in terms of comparison and replication.

private firms are evaluated which may result in more positive treatment effects compared to the average treatment effect for the population of firms.

Neither the R&D policy evaluation studies for China nor any other country has accounted for potential misappropriation of R&D funds. In our data, noncompliance can only occur among firms which are assigned to a treatment (R&D subsidy). 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. In a setting with one-sided noncompliance to funding contract rules, we have to differentiate between the causal impact of the assigned and actual treatment. Angrist and Imbens (1994) have shown that with randomized assignment to treatment, the effects can be consistently estimated by the intention-to-treat (ITT) effect and the complier average causal effect (CACE). In order to account for the selection in the R&D subsidy allocation and to randomize the assignment of the ITT, we suggest combining ITT and CACE with entropy balancing as a first estimation step, a method recently proposed by Hainmueller (2012).

The ITT shows how effective the R&D policy was in practice when misappropriation exists. The CACE, in contrast, depicts how effective the policy could have been without misappropriation and is thus is a measure for the efficacy of the R&D subsidy policy. Both are informative for policymakers. If R&D subsidies e.g. fail to induce additional R&D investment, it is important for them to better understand whether the failure originates from flaws in the design or the implementation of policies. Our ITT results show mild partial crowding out, i.e. total R&D expenditure have increased but by less than the R&D subsidy. Most strikingly, however, the CACE turns out to be more than twice as large as the ITT and suggests additionality which implies an increase of total R&D expenditures beyond the subsidy amount. Furthermore, both effectiveness and efficacy have significantly improved after the MLP implementation in 2006. Until 2006 both misappropriation and policy design had rendered R&D subsidies ineffective. But still, misappropriation of R&D subsidies considerably undermines the efficacy of Chinese R&D programs.

In the next section, we present the data sources used in the empirical analysis. In section 3 we explain our novel approach how to identify misappropriation in the data, and we provide several exercises to validate our measure of misappropriation. Section 4 presents the general identification strategy how to estimate causal effects of R&D subsidies with one-sided noncompliance using ITT and CACE whereas 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 consisting of all domestic firms listed at the stock exchanges in Shanghai and Shenzhen between 2001 and 2011.[4] Due to government stock issuance quotas the sample mainly includes domestic large and medium-sized firms from manufacturing industries and the coastal region, whereas other industries and inland regions are represented to a lesser extent.[5] The balance sheet information is compiled from COMPUSTAT and DATASTREAM and the Chinese databases CSMAR, RESSET, and WIND.[6] As the coverage of R&D expenditures in these databases is rather incomplete before 2007, complementary information on reported R&D expenditures has been collected from the universe of annual reports accessible 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. Before 2007, financial statements included a single account for subsidy income as well as mandatory notes on the different types of subsidies received.[7] Since 2007, subsidy income has been included in the non-operating income and the available information on different types of subsidies is considerably reduced.[8] To have the same detailed information for the period 2007-2011 as in the earlier years, we developed a semi-manual approach to classify all 85480 subsidy-related accounting transactions available in the CSMAR database into R&D and non-R&D subsidies (see Appendix 3 for more details). We further subdivide R&D subsidies into strict and broad R&D subsidies, the latter being received for patents, technology acquisition, technology transfer, and rewards. Thus, we very accurately distinguish between R&D and non-R&D subsidies and in total observe all subsidies received

[4] Data for listed firms is commonly used to investigate the innovation performance of firms, e.g. see Autor et al. (2017) for the US; Aghion et al. (2005) for Europe; and Fang et al. (2018) for China. Another commonly used data source in the Chinese context is the Annual Survey of Industrial Enterprises from the National Bureau of Statistics, e.g. Wei et al. (2017), and a more rarely used data source is the Administrative Enterprise Income Tax Records from the Chinese State Administration of Tax, e.g. Chen et al. (2018). However, the latter two data sets only provide information on total subsidies but not R&D subsidies and are thus not appropriate for our analysis.

[5] Only domestic firms are listed on the A-share board of the stock exchanges of Shanghai and Shenzhen. According to the definition of the China Securities Regulatory Commission (2002; 2006), a firm is considered domestic if the percentage of total shares held by foreign parties does not exceed 20%.

[6] The time period is China's fiscal year, which starts on January 1st and ends December 31st (see China Accounting Law 1985, Chapter 2, Article 11).

[7] China Securities Regulatory Commission (2000): "Regulation No. 2 disclosure guideline for the content and format of the annual report for the public offering of companies".

[8] China Accounting Standard Committee (2006): "Accounting standards No. 16–government subsidies".

by each firm. To avoid measurement error, we exclude observations when the sum of strict, broad, and non-R&D subsidies is larger than total subsidies. We also exclude observations for which total subsidies exceed sales.

In addition to R&D and subsidy information, we also observe employment, fixed assets, sales, age, profitability, ownership type[9], and industry affiliation. All variables in monetary values have been deflated using China's GDP deflator from the World Bank. Finally, we merge information in patent documents to calculate a firm's patent stock[10], changes in the application rate, high-tech IT orientation, university-firm collaboration and employment of foreign scientists. We observe all Chinese invention patents filed since the establishment of China's patent system in 1984 in the PATSTAT database and match these information to our panel data (Boeing et al. 2016 details the matching routine).

Our unbalanced panel for the period 2001-2011 includes 15911 observations with non-missing R&D expenditures and R&D subsidies for 2317 firms (Table 1). It covers the manufacturing and service sector (except for the finance industry). Table A2 in Appendix 2 provides information on the firm distribution by industry. Table 1 corroborates the extraordinary development of R&D in the period under consideration. The share of R&D performers quadruples from 14.7% to 63.0% while their median R&D expenditures increased from 3.0 million RMB to 16.2 million RMB. This corresponded to a rise of the mean and median R&D intensity (R&D expenditures to sales) from around 1.0% to 3.3% and 0.4% to 2.6%, respectively. The share of grantees receiving broad 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. The median R&D subsidy fell from about 1.4 to 0.8 million RMB. The development 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 hand suggests that the expansion of government funding rather took place along the extensive margin. The period-sum of R&D subsidies amounts to 12.4

[9] Since the late 1990s many of China's previously state-owned firms have been privatized (Hsieh and Song 2015). To accommodate for this ownership transformation, we differentiate between four ownership regimes. State ownership is >50% in *state-owned enterprises* (SOEs) and between 50% and >0% in *minority state-owned firms*. *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%.

[10] The 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. (Hall et al. 2005). We use patent applications instead of granted patents as they are a more timely measure of inventive activity, whereas granted patents only become observable several years later after successful examination. While in the US all applications are automatically examined, the lag between application and grant date can be much longer in Europe, Japan and China due to the deferral of the request for examination.

billion nominal RMB which corresponds to an annual income of 294706 Chinese scientists.[11] This furthermore implies that our sample covers 12.2% of total R&D expenditures and 10.1% of total R&D subsidies of large- and medium-sized firms in China (see Figure A1).

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

Year	Obs.	R&D expenditures a)				R&D subsidies b)			
		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	1268	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	1508	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	1566	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	15911	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 subsidized firms.

3 Misappropriation

3.1 Theory

We define misappropriation of R&D subsidies as a situation in which a firm does not (fully) spend the assigned subsidy amount for R&D activities. At this stage, we are agnostic about the specific alternative use. That is, the firm (or their managers) may either use the money to finance other firm-specific productive purposes like physical investments or spend it on private consumption. From a welfare perspective, the alternative use matters and in section 6.5 we provide some empirical evidence on this issue.

We set up a theoretical framework to explain a firm's incentive to misuse public R&D funds that 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, Hall 2002 and Hottenrott and Peters 2014). Let's assume a profit-maximizing firm that decides upon its optimal level of R&D investment without any subsidies. Its decision is based on a comparison of the marginal rate of return to R&D (MRR) and marginal cost of capital (MCC). Both, MRR

[11] Calculated as R&D subsidies divided by the annual average income of Chinese scientists. The time series "Annual average income of urban units' employees in the sector of scientific research and technical service" is accessible at China's National Bureau of Statistics.

and MCC vary with the level of R&D investment ($R\&D$) but while MRR is downward sloping, MCC is constant as long as internal finance is used and upward sloping if additional more costly external finance is borrowed. This 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 finance. In addition to $R\&D$, MRR may depend on the firm's innovative capabilities (IC) and other firm- and industry-specific variables summarized in the vector X'_1. That is, we have $MRR = f_1(R\&D, IC, X_1')$. Similarly, we define $MCC = f_2(R\&D, R^{non-R\&D}, IF, c^{ext}, X_2')$. The MCC reflects the opportunity costs of investing funds in R&D vs. non-R&D projects, and thus depends on, among others, the expected returns to other uses of available funds such as investment in tangible or financial assets ($R^{non-R\&D}$) the amount of internal finance (IF), 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. The optimal R&D investment without subsidy financing is thus given by: $R\&D^* = f(IC, R^{non-R\&D}, IF, c^{ext}, X_1', X_2')$. Note that $R\&D^*$may be equal or greater zero.

Now what happens if a firm gets an R&D subsidy S and can simultaneously decide about compliance? The R&D subsidy increases the amount of financial means for R&D (IF), but we can now think of having two types of internal finance, the public R&D subsidy (S) and other private internal funds (IF^{priv}). We extend the idea of a pecking order for R&D funds and make the mild hierarchical assumption that R&D subsidies are fully used before any other internal and if necessary external funds are spent on R&D. It is rational to assume that if the firm decides to invest in R&D it will use R&D subsidies for financing first because if it instead uses other internal funds for R&D projects first and R&D subsidies for non-R&D projects, it risks being detected and paying sanctioning costs (SC) with a detection probability $p > 0$. Thus, other internal funds have higher opportunity costs and therefore higher MCC (depending also on p and SC) than R&D subsidies. Only if the probability p that misuse of R&D subsidies is detected and sanctioned goes to zero, the opportunity costs of R&D subsidies and other internal funds are about the same and a firm would be indifferent between both types of funding which one to use first. In this situation, the potential for moral hazard is at a maximum. An increase in the amount of financial means for R&D, IF, shifts the MCC to the right. Whether this also boosts the new optimal R&D investment level with subsidy financing, $R\&D^{**}$, depends on whether the firm was financially constrained before the subsidy payment or not. Hottenrott and Peters (2014) showed that optimal R&D investment only increases if firms were initially financially constrained implying insufficient internal financial means.

More importantly for our research question, the assumption of the extended pecking order allows us to identify misappropriation of R&D subsidies simply by comparing the new optimal R&D investment level $R\&D^{**}$ 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 $R\&D^{**}$ is lower than the R&D subsidy S :

$$M|_{S>0} = \begin{cases} 0 & if \;\; R\&D^{**} \geq S \\ 1 & if \; R\&D^{**} < S \end{cases} \tag{1}$$

The difference between the optimal R&D investment level $R\&D^{**}$ and the R&D subsidy amount S is a measure for the absolute value of misappropriation. According to theory, a firm's decision to misappropriate R&D funds thus depends on the R&D subsidy level and all arguments determining the optimal R&D investment with subsidy financing, $R\&D^{**}$:

$$M|_{S>0} = g\left(IC, R^{non-R\&D}, S, IF^{priv}, SC, p, c^{ext}, X_1', X_2'\right) \tag{2}$$

Note, in our theoretical model the R&D subsidy amount is exogenous and by deciding upon the optimal R&D level, firms also decide upon the level of misappropriation. In particular, we do not take into account that a priori the firm might have an intention to misappropriate and that it evinces fraudulent behavior to maximize the expected subsidy amount. Our data does not allow us to identify a priori fraudulent firms. For the Innofund program, Stuart and Wang (2016) found that fraudulent firms who overstate financial figures are more likely to get funding. For the same program, Wang and Li (2014) showed that resources obtained through fraudulent means are less likely to be allocated to productive activities such as technological innovation.

Second, the detection probability is also exogenously given. Detecting firms' non-compliant behavior to funding contract rules requires monitoring effort by the government. Government monitoring is usually done by bureaucrats. Monitoring efforts are weakened in a regime with corruptive behavior among bureaucrats. Acemoglu and Verider (2000) explicitly develop a theoretical model with firms with good or bad technology (here application), corrupt and honest bureaucrats, and a welfare-maximizing government. Bureaucrats randomly evaluate firms and should only subsidize good applications. They show that corrupt bureaucrats in general are willing to pay subsidies regardless of the quality of the application as long 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 evaluation process, both "collude". After collusion, the corrupt bureaucrat does not monitor compliance to funding contract rules and the firm is exposed to the highest moral hazard for misappropriation. Note that in addition to the government, external investors might also have an incentive to monitor firms in order to prevent them from using the funding for less productive purposes.

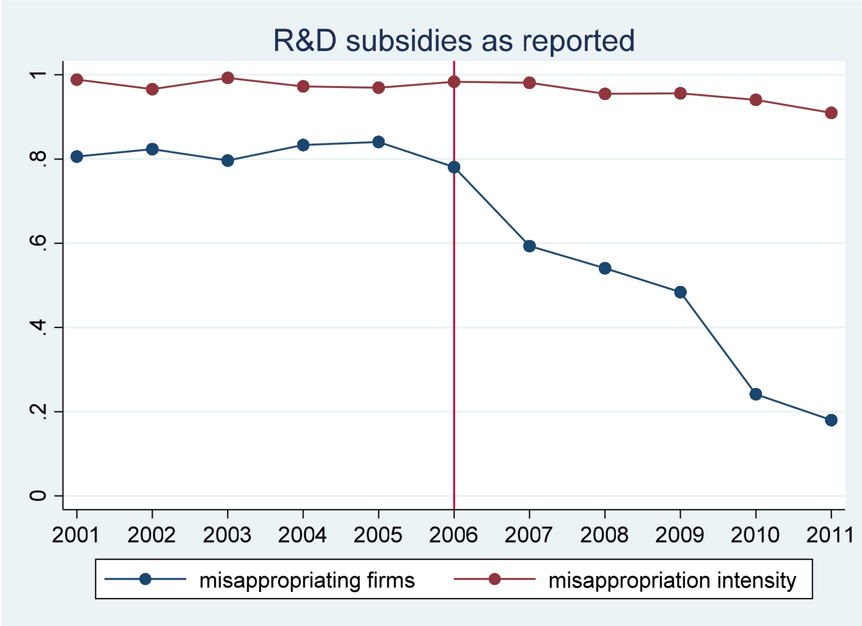
Finally, the theoretical model assumes that R&D projects are arbitrarily divisible. In practice, however, indivisibilities of R&D projects exists. Firms are often not able to arbitrarily scale down R&D investments but need a minimum of financing (Gonzales et al. 2005). The larger the deviation between the actual subsidy amount received and the one applied for, e.g. because bureaucrats kept their "fair" share, the more likely the actual subsidy falls below the minimum threshold. As a result, the firm is not able to conduct the research project and misuse the funding. That is, indivisibilities of R&D projects might lead to unintended misappropriation.

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 total R&D expenditure, we calculate misappropriation as difference between total R&D expenditure and received R&D subsidies[12] as reported in financial statements. Overall, 42.0% (1313 out of 3127 subsidized firm-year observations) of grantees misused R&D subsidies, corresponding to 52.5% of the total amount of R&D subsidies. These figures strikingly confirm anecdotal evidence that misappropriation is a major concern in China. Regarding misappropriation, we find two additional intriguing facts in our data. First, firms either chose (almost) full misappropriation or not to misappropriate any funds which may be rationalized by indivisibilities of R&D projects. Figure 1 shows that the average misappropriation intensity (misappropriated R&D subsidies to total R&D subsidies) is very stable around 95.9% which implies that variation along the intensive margin is of lesser relevance. Second, there is a substantial decline in misappropriation over time from 80.6% (2001) to 18.0% (2011) along the extensive margin of misappropriation. This decline emerges especially after 2006, as indicated by the red reference line, which coincides with the implementation of the MLP.

[12] We assume that the amount of R&D subsidies recorded in the balance sheets equals the net amount received by the firm, i.e. the amount that is actually available to the firm. However, if the recorded R&D subsidies equal the gross amount received by firms, including the amount actually withheld by officials or fees paid to intermediaries, we would inflate our measure along the intensive margin. As most misappropriating firms are non-R&D performers, the extensive margin of misappropriation hardly changes if we interpret gross as net.

Figure 1: Misappropriation along the intensive and extensive margin



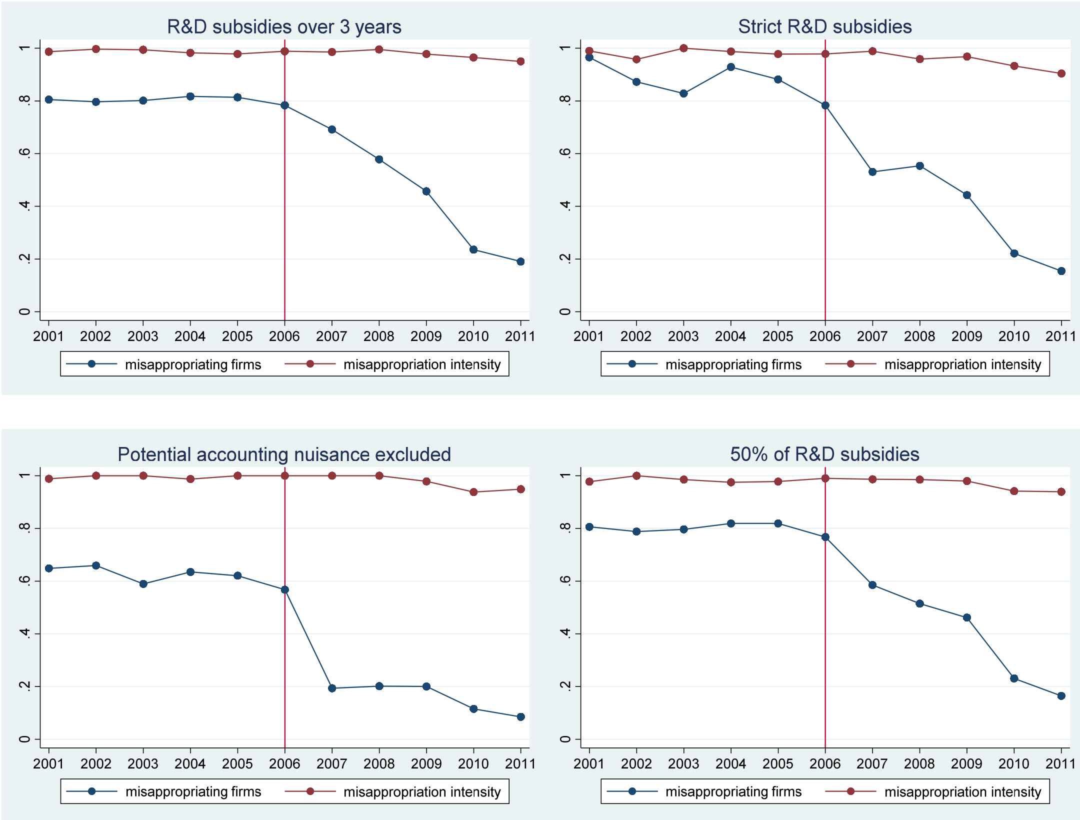
Notes: Misappropriating firms and misappropriation intensity denote the share of firms having misappriated R&D subsidies (extensive margin) and the share of misappropriated R&D subsidies to total R&D subsidies (intensive margin), respectively.

Simply identifying misappropriation based on comparing annual reported R&D expenditure with R&D subsidies could be misleading because of unknown timing or compositional issues related to the receipt of R&D subsidies. Hence, we additionally calculate four alternative measures based on different assumptions. The first alternative measure assumes that firms receive a lump-sum in t and the annual subsidy flows are calculated by allocating the received amount uniformly over an assumed funding period of three years, i.e. between t and $t + 2$, resulting in 45.4% (1895/4174) misappropriating firms. The second measure only considers strict R&D subsidies excluding grants received for patents, technology acquisition, technology transfer, and yields 37.9% (849/2239) misappropriating firms. The third measure ignores any timing inconsistencies but only compares 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.[13] This measure provides the lower bound of misappropriating firms at 19.6% (442/2256). The fourth measure assumes that firms cannot keep all R&D subsidies accounted for in their financial statements but are obliged to return 50% to corrupted officials and intermediaries (see Appendix 2 for related information). This yields a share of 40.4%

[13] While this approach eliminates potential timing inconsistencies that we may incorrectly interpret as misappropriation (false positives), it may also exclude observations with actual misappropriation (true positives).

(1264/3127) misappropriating firms. Figure 2 shows that all four alternative measures replicate the pattern of our main measure: the intensive margin is between 90.4% and 99.0% and there is a significant drop after 2006 along the extensive margin.

Figure 2: Alternative measures of misappropriation



In addition to relevant variation over time, there are important industry differences as highlighted in Table A1. It shows the proportion of R&D performers, subsidized and misappropriating firms at the 2-digit level for manufacturing and the 1-digit level for non-manufacturing industries. The three industries *electronics*, *machinery & instruments*, and *pharma & biological products* have the highest percentage of R&D performers and not only exhibit a high likelihood of receiving R&D subsidies but also the lowest likelihood of misappropriation.[14] Around half of the firms in these industries perform R&D, a quarter to a third receive R&D subsidies, and around a third or less misappropriate. The low misappropriation rate is plausible as these sectors have the highest expected returns to innovation (Peters et al. 2017). This pattern is precisely reversed for the three industries with

[14] This is also the case for the industry *other manufacturing* which also includes firms that produce defense-oriented goods.

the lowest percentage of R&D performers, where a very low likelihood of receiving R&D subsidies corresponds with a very high likelihood of misappropriation.

This pattern is largely resembled at the provincial level (Figure A2). Firms located in developed and industrialized coastal provinces have a higher likelihood of receiving R&D subsidies but a lower likelihood of misappropriation. In recent years, China's picking-the-winner strategy was somewhat complemented by an aiding-the-poor strategy that directed funds to firms in less developed Western provinces. However, misappropriation is more likely in Western and Northern provinces.

Public awareness for misappropriation only occurred after 2011 and it is most likely that firms reported R&D subsidies and R&D expenditures correctly before.[15] However, even if there was misreporting, one would expect firms to report either less R&D subsidies, more R&D expenditures, or both, and thus our measure constitutes the lower bound of misappropriation. Finally, China's recent anti-corruption campaign initiated in November 2012 is unlikely to determine firm behavior in our study period.

3.3 Validation

Before we use our misappropriation measure to evaluate the effectiveness and efficacy of R&D subsidy policy, we further check its plausibility by examining the variation in misappropriation across subsidized firms. Conditional on receiving subsidies, equation (2) describes a firm's decision to misappropriate R&D funds as a function of different variables. We thus estimate a two stage probit model with sample selection (Heckprobit). The first stage describes the likelihood of receiving an R&D subsidy and thus accounts for sample selection. Conditional on receiving an R&D subsidy the second stage explains the likelihood of misappropriation. Table A2 reports the second stage results for four different specifications. In all specifications we include general firm attributes like the number of employees, firm fixed assets, sales, age and, ownership as well as year, industry and province FE. They capture variations in the marginal rate of return and marginal cost of capital across firms and thus also in the propensity to misappropriate induced by $R^{non-R\&D}, c^{ext}$, X_1' and X_2'. Misappropriation is furthermore explained by 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

[15] Even based on correctly reported R&D subsidies and R&D expenditures, the identification of misappropriation is not a trivial accounting exercise. It essentially requires the replication of our subsidy data classification outlined in Appendix 3 in combination with our measurement framework which is unlikely to be a routine job for auditing firms.

expected returns to R&D are proxied by two variables: a dummy indicating 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 (IF^{priv}).

Before commenting on our findings, we explain our identification strategy since having a valid exclusion restriction is key for uncovering the true parameter values. Due to the panel structure of our data, we use the dummy variable whether a firm got an R&D subsidy in year t-1 as exclusion restriction. Lagged R&D subsidy is likely to affect the current likelihood of getting funding. However, it should not affect the current decision to misappropriate R&D funds once we have additionally controlled for lagged misappropriation behavior. Results in Table A2 show that lagged R&D subsidies are indeed highly significant in the first stage. Additional estimates further show that lagged R&D subsidies are not significant in the second stage after controlling for lagged misappropriation, supporting our identification strategy.

Results in Table A2 largely confirm our hypotheses set forth in section 3.1. In particular, prior 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 costs of capital for internal funds and as a result lower incentives to misappropriate. The estimates also show a u-curve relation between the subsidy level and misappropriation. We observe a significantly higher likelihood of misappropriation for very low and very high R&D subsidy levels. The first effect can be explained by indivisibilities of R&D projects whereas the second effect likely reflects decreasing rate of returns to R&D. Finally note that older firms, firms with more fixed assets and privatized state-owned enterprises are also more likely to misappropriate.

Columns (2) and (3) additionally account for variations in the detection probabilities across firms by adding monitoring and corruption indicators, respectively. Monitoring effort varies over time. Improving the monitoring was one key element of the MLP and thus we include a MLP dummy variable that equals one for the period 2007-2011. In addition to this macro-level monitoring indicator, we include a firm-level monitoring variable. The latter is a dummy variable that is one if the firm has mutual funds investors. It has been shown that these institutional investors are more likely to monitor the firm in order to prevent misuse of R&D grants for unproductive purposes. Results in column (2) show a significantly negative effect for both variables. The results thus confirm that higher monitoring effort lowers the probability of misappropriation.

Columns (3) additionally accounts for variation in corruption across provinces. We include the number of cases of investigations against bureaucrats divided by number of large and medium-sized firms (LMEs) to measure corruption as argued in the seminal paper by Glaeser and Saks (2006). The results confirm a significantly positive effect of the number of investigated bureaucrats per LMS on misappropriation. As the opportunity for rent-seeking by bureaucrats increases with the number of bureaucrats and results in less money left over for R&D subsidies, we hypothesize that the amount of R&D subsidies received by firms decrease with the average number of bureaucrats per firm in each province (Naughton, p.80, 2013). HWe additionally control for any size effect by adding the total number of bureaucrats per LME. The results show a significantly positive effect of the number of investigated bureaucrats per LMS on misappropriation. /// Bureaucrats per LME: Argument follows the rationale provided by Jinglian Wu, one of China's most influential economists (Naughton, p.80, 2013). Our results however, do not confirm this hypothesis.

Finally, instead of using the sum of R&D subsidy levels across all programs, we alternatively use the mean subsidy level per R&D project. Column (4) shows that we find a similar u-curve relation between the average subsidy level per R&D project and the likelihood of misappropriation.

In a nutshell, our misappropriation indicator behaves plausible at all levels. We find that misappropriation is not random but depends on policy induced changes in returns to R&D at the macro level, industry characteristics, as well as prior R&D investments and profitability at the firm level.

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

Randomized experiments are seen as the gold standard to causal inference in many fields. In practice, however, they are often not feasible. But even if they were, further complications can arise. Most importantly, noncompliance to the *assigned* treatment may exist. On the one hand, units who are assigned to a treatment may decide not to comply with the assignment and actually get no treatment. A situation in which only the assigned treatment units can decide to comply or not is called one-sided noncompliance. If in addition, units from the control group are also somehow able to circumvent the non-assignment and actually get the treatment, noncompliant behavior arises among treated and control units and is thus called two-sided noncompliance. In our application *one-sided* noncompliance among R&D subsidy grantees is a major concern as we will show in more detail in section 4. Around 42 % of firms deliberately choose not to comply with the assigned treatment and spend R&D subsidies for other non-research purposes. We will call this one-sided noncompliant behavior in the empirical application also as misappropriation of R&D subsidies. As a result of the noncompliance, we have to distinguish between the *assigned* treatment and *actual* treatment. The main problem of noncompliance for causal inference is that the *actual* treatment is the result of deliberate choices by units that most likely take into account expectations 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 1 if firm i is assigned to the treatment group ($Z_i = 1$) and 0 if it is assigned to the control group ($Z_i = 0$).[16] Further, let D_i denote the *actual* treatment which takes 1 if firm i actually receives the treatment and 0 otherwise. In our setting Z_i indicates whether firm i gets an R&D subsidy grant and D_i whether the R&D subsidy has been used for research purposes. Noncompliance occurs when $Z_i \neq D_i$.

There are two naïve approaches to study the treatment effect which are sometimes conducted in such settings. 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 non-compliers ($Z \neq D$) and analyzes the compliers as if they were randomized. However, as compliance is self-selected, the remaining subsample is not representative of the study population. Thus both approaches in general fail to provide consistent estimates of the treatment effect (Imbens and Rubin 2015).

On the contrary, we will differentiate between the intention-to-treat (ITT) group effect and the complier average causal effect (CACE) in order to consistently estimate and evaluate R&D subsidies under one-sided noncompliance to funding contract rules. That is, we are actually interested in two treatment effects, the one of the assigned treatment and the one of the actual treatment. In the following, we explain the rationale, estimation and interpretation of both treatment effects. We follow the approach by Angrist and Imbens (1994) and Imbens and Rubin (2015).[17] Although the problem of noncompliance in experimental designs is not new and applied in the biostatistics literature, there are only few applications in economic policy

[16] To simplify notation, we leave out any time subscripts in this section. Our empirical analysis, however, is based on panel data and includes a time dimension as well.
[17] See also Angrist, Imbens and Rubin (1996), Imbens and Wooldridge (2009) and Angrist and Pischke (2009).

evaluation studies and none in the R&D policy literature we are aware of because it is often hard to identify noncompliant behavior in the data even if 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 it is the causal effect of an assigned or granted R&D subsidy on the growth rate of R&D expenditure. Allowing for heterogenous treatment effects across firms, the individual ITT is the difference in the unit-level outcome variable Y_i by assigned treatment status Z_i: $ITT_{Y,i} = Y_{1i} - Y_{0i}$. Y_{zi} denotes the outcome variable for assigned treatment status, i.e. $Y_{zi} = Y(Z_i = z) = Y(z)$ for $z = 0{,}1$. For each firm, however, we either only observe 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 group: $ITT_Y = E(Y_1 - Y_0)$. Except for the stable unit treatment value assumption (SUTVA) saying that there is no causal effect of one unit's treatment assignment on another unit's outcome, there is no need for additional assumptions to get consistency of ITT in case of random assignments. As it is well known, however, an R&D application and granting process is generally not based on a random experiment.[20] To consistently estimate the ITT, we thus suggest using entropy balancing as an additional first design step in the ITT estimation to get an (almost) randomized assignment to treatment (see section 5 on the empirical strategy for more details).

While ITT_y provides a consistent estimate of the causal effect of the assigned treatment, it ignores the compliance status D_i completely and is thus not informative about the causal impact of the actual treatment. Still, it is interesting to calculate the ITT as it tells us about the *effectiveness* of the treatment when noncompliance exists. That is, the ITT shows how effective the R&D policy is in practice when misappropriation of funds occurs.

[18] Notable exceptions are the evaluations of the JTPA training program in Bloom et al. (1997) or the Head Start education program in Kline and Walters (2016). In the R&D context, noncompliance has completely been neglected in the literature. If data reveals cases in which R&D subsidies are larger than total R&D expenditures, this has been interpreted as accounting nuisance due to the time allocation of subsidies and in fact per protocol has been used for estimation, see Gonzalez et al. (2005), Gonzales and Pazo (2008) and Arques-Castells and Mohnen (2015) which all used Spanish data. Arques-Castells and Mohnen (2015) excluded 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.

[20] The innovation voucher program in the Netherlands in 2004 is one of the few exemptions (see Cornet et al. 2006)

4.2 *Complier Average Causal Effect (CACE)*

We are also interested in the *causal effect of the actual treatment* D_i. As explained above the problem is that compliance is based on self-selection implying that the actual treatment is confounded with the potential outcome Y_i. In our case, it is likely that firms with higher expected returns to R&D are more likely to spend R&D subsidies for research purposes. Using OLS in a regression of Y on D without accounting for endogeneity would lead to upward biased results. We thus employ IV using the randomized assigned treatment Z as instrument (see Bloom 1984 and Angrist et al. 1996). In this setting it turns out that the IV estimator is the local average treatment effect (LATE) which only uses information for a subgroup of firms, the compliers.

Two key ideas are important to understand the identification strategy in this setting (Bloom et al. 1984 and Angrist and Imbens 1994). First, not only Y but also D is a *potential* outcome of assigned treatment status Z. Let D_{1i} be i's *potential* actual treatment status if $Z_i = 1$, i.e. $D_{1i} = D(Z_i = 1)$ and similarly $D_{0i} = D(Z_i = 0)$. If two-sided noncompliance exists D_{1i} and D_{0i} can both take values $\in \{0,1\}$, whereas in our case of one-sided noncompliance D_{0i} can only be 0. The *observed* (realized) actual treatment is D_i can be expressed as $D_i = D_{0i} + (D_{1i} - D_{0i})Z_i$. Furthermore, let $Y_i(z,d)$ denote the potential outcome of Y for firm i. Since both Z and D are binary, we generally have four potential outcomes: $\{Y_i(1{,}1), Y_i(1{,}0), Y_i(0{,}0), Y_i(0{,}1)\}$ which reduces to three in case of one-sided noncompliance leaving out $Y_i(0{,}1)$.

Second, if two-sided noncompliance exists, we can generally *partition the population* of firms into 4 subgroups based on their *compliance behavior* in different assignment states: compliers for whom actual and assigned treatment would always coincide (D_{1i} =1 and D_{0i} =0), always-takers who would actually get treated independent of the assignment (D_{1i} =1 and D_{0i} =1), never-takers who would never get treated irrespective of their assignment (D_{1i} =0 and D_{0i} =0) and defiers for whom actual and assigned treatment would never coincide (D_{1i} =0 and D_{0i} =1). Always-takers, never-takers and defiers do not comply in at least one situation and can thus also be summarized as noncompliers. For each firm i, however, we only observe one assignment status, i.e. either $Z_i = 1$ or $Z_i = 0$. But in our application non-compliant behavior in case of non-assignment is ruled out, so that there are no always-takers and defiers and we are left with only compliers and never-takers (noncompliers). How does this partition of the population relate to our *observed* assigned and actual treatment and control groups? Under one-sided noncompliance three possible states of (Z_i, D_i) are observed: $\{(1{,}1), (1{,}0), (0{,}0)\}$.

Among the assigned treatment group, firms with observed $(1,1)$ must be compliers since in case of no assigned treatment they could only comply, too. Similarly, $(1,0)$ must be never-takers. In the assigned control group, however, firms with $(0,0)$ can be either compliers or never-takers as their compliance behavior in case of assigned treatment is unobserved. Hence, we cannot identify the compliance type of each firm from the observed data alone[21] and thus also not the size of the subgroups in the population. We can only distinguish compliers and noncompliers in the subgroup of firms assigned to treatment $(Z_i = 1)$. Under one-sided noncompliance noncompliers are not informative about the effect of the actual treatment as they never get actually treated. Compliers are the only subgroup from which we can obtain information about the effect of the actual treatment and which are thus used for identifying the causal effect.

For firm i the causal effect of the actual treatment D_i given its realized assignment Z_i is $Y_i(Z_i, 1) - Y_i(Z_i, 0)$. The problem for estimating the average causal effect is that D_i is self-selected. We solve the endogeneity issue by using Z_i as an instrumental variable. 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: $[D_{1i}, D_{0i}, Y_i(z,d)] \perp Z_i$. This *independence* assumption allows us to consistently estimate (i) the causal effect of Z on Y, i.e. the 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 which is equivalent to the coefficient of the first stage in IV. Since conditional and unconditional expectations are the same under independence, the following holds for the reduced form effect $ITT_Y = E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0] = E[Y_i(1, D_{1i})|Z_i = 1] - E[Y_i(0, D_{0i})|Z_i = 0] = E[Y_i(1, D_{1i}) - Y_i(0, D_{0i})]$ and the first stage effect $ITT_D = E[D_i|Z_i = 1] - E[D_i|Z_i = 0] = E[D_{1i}|Z_i = 1] - E[D_{0i}|Z_i = 0] = E[D_{1i} - D_{0i}]$. 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 when they allocate R&D subsidies based on firm's expected compliance behavior. We already explained for the ITT that we do not expect the R&D granting process and thus the instrument to be initially random but we suggest using the

[21] There is a conceptual difference between compliers who would comply under both alternative assignments and the firms for which we actually observe that they comply with their realized assignment status.

entropy balancing method as a first step to get an (almost) randomized assignment to treatment Z_i (Hainmueller 2012).

Second, the instrument Z_i affects the potential outcome $Y_i(z, d)$ only via the actual treatment D_i or to put it differently, the potential outcome $Y_i(z, d)$ does only depend on the actual treatment irrespective of the assigned treatment: $Y_i(0, d) = Y_i(1, d)$. As a result we can define $Y_{1i} = Y_i(0,1) = Y_i(1,1)$ and $Y_{0i} = Y_i(0,0) = Y_i(1,0)$ which in turn allows us to express the observed (realized) outcome as $Y_i = Y_{0i} + (Y_{1i} - Y_{0i})D_i$. Since the instrument Z_i is excluded from the potential outcome, this is also called the *exclusion restriction*. We believe that the exclusion restriction also holds in our application since it is reasonable to assume that the growth rate of R&D expenditure is only affected by the fact whether 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. Using these two assumptions, the IV estimator can be written as:

$$\tau_{IV} = \frac{ITT_Y}{ITT_D} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]}$$
$$= \frac{E[Y_{0i} + (Y_{1i} - Y_{0i})D_i|Z_i = 1] - E[Y_{0i} + (Y_{1i} - Y_{0i})D_i|Z_i = 0]}{E[D_{0i} + (D_{1i} - D_{0i})Z_i|Z_i = 1] - E[D_{0i} + (D_{1i} - D_{0i})Z_i|Z_i = 0]}$$
$$= \frac{E[Y_{0i} + (Y_{1i} - Y_{0i})D_{1i}] - E[Y_{0i} + (Y_{1i} - Y_{0i})D_{0i}]}{E[D_{1i}] - E[D_{0i}]}$$
$$= \frac{E[(Y_{1i} - Y_{0i})(D_{1i} - D_{0i})]}{E[D_{1i} - D_{0i}]}$$

Note that $D_{1i} - D_{0i}$ can take on values $\in \{1,0,-1\}$: 1 for compliers, 0 for always-takers or never-takers and -1 for defiers. Hence we can also write the expectation in the numerator as weighted sum of the conditional expectations by subgroups times the share of the subgroup. The third key assumption that has to hold is monotonicity in a sense that $D_{1i} \geq D_{0i}$ for all firms i. This implies that defying behavior is ruled out. In a one-sided noncompliance case this assumption is fulfilled by design. If we additionally take into account that always-takers also do not exist under one-sided noncompliance, Bloom (1984) and Angrist and Imbens (1994) have shown that the IV estimator of Y_i on D_i using randomized Z_i as instrument is the causal effect of the actual treatment of D_i on outcome Y_i among the subgroup of compliers, known as local average treatment effect ($LATE$) or complier average causal effect ($CACE$).

$$\tau_{IV} = \frac{E[(Y_{1i} - Y_{0i})(D_{1i} - D_{0i})]}{E[D_{1i} - D_{0i}]} = \frac{E[Y_{1i} - Y_{0i}|D_{1i} - D_{0i}]P[D_{1i} - D_{0i}]}{1 \cdot P[D_{1i} > D_{0i}] + 0 \cdot P[D_{1i} = D_{0i}]}$$

$$= \frac{E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}]P[D_{1i} > D_{0i}] + E[Y_{1i} - Y_{0i}|D_{1i} = D_{0i}]P[D_{1i} = D_{0i}]}{P[D_{1i} > D_{0i}]}$$

$$= \frac{\tau_{CACE} \cdot P[Complier] + \tau_{NACE} \cdot P[Noncomplier]}{P[Complier]}$$

$$= \frac{\tau_{CACE} \cdot P[Complier] + 0 \cdot P[Noncomplier]}{P[Complier]} = \tau_{CACE}$$

Note that $CACE$ and $NACE$ are local effects defined on the subgroup of compliers and noncompliers, respectively. As already explained, the proportion of compliers in the population is unknown. But since the instrument is randomly assigned, the proportion of compliers among the $Z_i = 1$ group is representative for all compliers. Since $D_{1i} > D_{0i}$ only holds true for observed actual treatment $D_i = 1$, we can also write: $\tau_{CACE} = E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}] = E[Y_{1i} - Y_{0i}|D_i = 1]$. That is, $CACE$ is equivalent to the average treatment effect on the actually treated in the case of one-sided noncompliance since we do not have any always-takers. In the empirical analysis we estimate $CACE$ as IV estimator, i.e. as the ratio of ITT_Y from the reduced form and ITT_D from the first stage regression of 2SLS.

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

In contrast to ITT which measures the effectiveness, $CACE$ measures the *efficacy* of the treatment in an ideal situation without noncompliance. That is, $CACE$ depicts 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 relevant. For instance, if we find that ITT is (close to) zero but $CACE$ is significantly positive, than we can conclude that the design of the R&D program in principle works on stimulating R&D expenditures but that policymakers should strive for improving the monitoring of R&D subsidies. 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 design. Overall, the relation of the effectiveness to the efficacy informs us about the loss in effectiveness due to noncompliance.

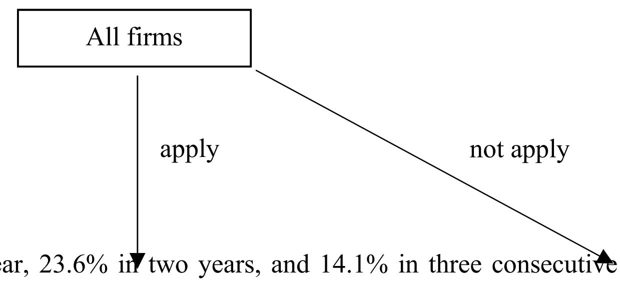
5 Empirical Strategy

5.1 *Sample of ITT, Compliers and Controls*

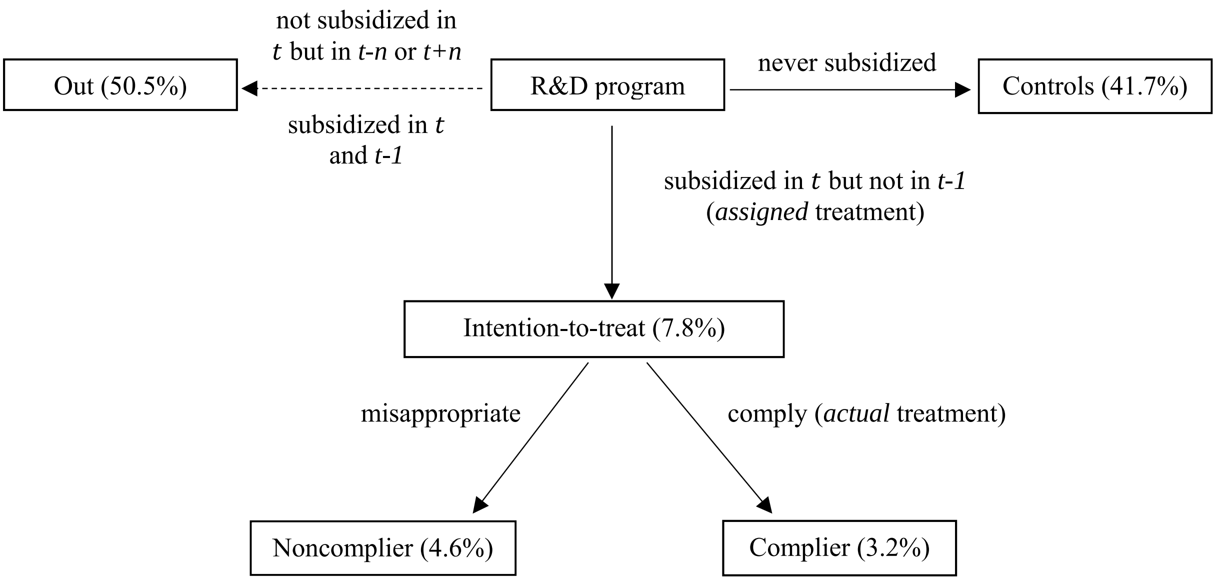
Due to noncompliance we have to distinguish between assigned and actual treatment. The assigned treatment group, also called intention-to-treat (ITT) group, consists of all firms that received R&D subsidies in year t (and potentially in years $t+n$) but did not receive any R&D support in year $t-1$.[22] To properly identify the causal impact of the R&D subsidy assignment in a before-after comparison, we exclude firms that were subsidized in year t and $t-1$ from the econometric analysis. The ITT group is further separated into two mutually exclusive groups of firms that actually spend the subsidy for research (compliers) and those that misappropriate the funds (noncompliers). Figure 3 shows the share of firms in each group. Among the ITT firms, 41% complied to the assigned treatment which is even slightly lower than what we document in section 4 and can be explained by the fact that multiple subsidy receivers are excluded that are more likely to spend public funds for research purposes. Note the group of noncompliant firms comprises full and partial non-compliers though the vast majority (90.7%) conducts full misappropriation.

Firms that never applied for an R&D subsidy program or applied but were never subsidized build the control group. Unfortunately, we do not have subsidy application data and hence cannot restrict the control group to firms that applied for R&D funding but did not get it. As explained in more detail in subsection 5.3, the causal impact of an R&D subsidy is defined as the change in the R&D expenditure between $t-1$ and $t+1$. In order to eliminate possible long-term or anticipation effects of a program among the control group, however, we not only require that firms did not receive any grants in the consecutive three years from $t-1$ to $t+1$ but also require that control firms 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 3: Sample definition of intention-to-treat, compliers and control groups



[22] 51.2% of firms report funding in one year, 23.6% in two years, and 14.1% in three consecutive years. The remaining 11.1% of firms received funding in up to eight consecutive years.



5.2 *Descriptive Statistics*

Table 2 reports descriptive statistics for the total sample of all firms and the estimation sample consisting of the assigned treated (ITT), actual treated (compliers) and control observations.[23] Pre-treatment differences between both treatment and control groups are rather obvious. Relative to the control group, the ITT group had on average lower pre-treatment employment, fixed assets, and sales but a similar age, patent stock and profitability, whereas the compliers' average pre-treatment employment, fixed assets, sales, and profitability were higher compared to the ITT group.

In contrast, in the year before being granted the subsidy we observe the highest R&D expenditure for compliers (29.7 million RMB), followed by ITT firms (13.9 million RMB) and the control group (7.4 million RMB). This result supports the view that both the selection of firms by the government and the firm's 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, so roughly 1.458 or 145.8% per year. Compared to the average two-year growth rate of the ITT group, compliers have a higher (3.279) and the control group a lower one (1.415). The average growth rates of R&D expenditures are very high compared to other studies for

[23] Compared to section 4, the total sample has reduced from 15911 to 10430 observations due to taking 2-year lags and dropping observations with missing values for the relevant variables used in the estimation.

industrialized countries. They are, however, sensible given the tremendous rise in R&D expenditure at the aggregate level in China during that period (see Figure A1 and Table 1) and given the fact that many firms started R&D or increased R&D from a very low initial level.[24] 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 with 0 for ITT and controls and 0.706 for compliers. Comparing the R&D subsidy amount and R&D subsidy intensity, it turns out that the ITT firms got on average higher R&D subsidies than compliers (2.7 compared to 1.8 million RMB) and that their R&D subsidy intensity (among R&D performers) is three times larger (39.0% compared 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 group (7.5) shows that R&D subsidies have fostered R&D expenditures by 9.9 million RMB (2.1 in the average log change). The corresponding figure for the subsample of compliers are 19.5 million RMB and 2.4 in the average log change.

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

[24] High average values are not the result of a few outliers. Winsorizing the data at the 99th percentile yields average growth rates of 2.903, 3.266 and 0.835, respectively.

Table 2: Descriptive statistics

	Estimation sample									Total sample		
	ITT			Complier			Controls					
	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd
Quantitative variables												
R&D expenditures $_{t+1}$ (Mio RMB)	31.266	1.154	238.149	56.639	16.102	342.825	14.898	0.000	160.160	18.185	0.000	149.017
R&D expenditures $_{t-1}$ (Mio RMB)	13.866	0.000	153.259	29.679	5.774	236.849	7.385	0.000	108.067	8.570	0.000	109.840
Log-growth in R&D expenditures $_{t-1 \text{ to } t+1}$	2.916	0.000	7.642	3.279	0.705	7.558	0.847	0.000	5.348	1.415	0.000	6.517
R&D subsidy $_t$ (Mio 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 $_t$ (%) [a)]	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	4653	1381	23150	4059	1695	15925
Fixed assets $_{t-1}$ (Mio RMB)	3443.2	1593.8	10290.8	3698.9	1471.4	14166.4	7236.6	1641.9	51377.7	5020.5	1663.6	29076.3
Sales $_{t-1}$ (Mio RMB)	2916.4	1105.6	11198.9	3524.9	1210.2	16509.0	5934.4	819.9	49064.7	4060.6	946.3	33849.0
Age	11.855	11.000	4.303	11.349	11.000	4.322	12.311	12.000	4.251	11.846	12.000	4.330
Patent stock $_{t-1}$	21.188	2.850	88.651	20.629	5.018	43.015	20.052	0.000	151.862	19.220	0.723	121.564
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		
Minority state-owned $_{t-1}$	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			4347			10430	

Note: Total sample consists of all observations with nonmissing values for the relevant variables and with R&D expenditure available in both years $t-1$ and $t+1$. The estimation sample is the subsample of ITT and control observations leaving out the Out group (see Figure 3). [a)] R&D subsidy intensity is calculated for a total sample of 2795 observations with positive R&D expenditures, of which 371 are in the ITT group, 334 in the complier group, 612 in the control group, and 1812 in the out group.

5.3 *ITT: Econometric Model and Entropy Balancing*

As set forth in section **Fehler! Verweisquelle konnte nicht gefunden werden.**, we are interested in estimating the causal effect of both the assigned and actual treatment of an R&D subsidy, Z and D, respectively. In this subsection, we specify the econometric model and explain the estimation strategy used to identify the intention-to-treat (ITT) effect in more detail. Subsection 5.4 continues with the complier average causal effect ($CACE$).

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 $\mathrm{t}-1$ (for a similar approach see Einiö 2014 or 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 that it is likely that an R&D subsidy does not cover the entire first year and may kick in very lately. Furthermore, the timing varies across firms. Second, for multi-annual R&D projects a larger fraction of the cost accrues in the second year. In our data, the mean successive support duration is 1.8 years and the mean R&D subsidy in $t+1$ is 34.5% larger than in t while the additional increase in $t+2$ only accounts for 5.9%.

Let $y_{i,\,t+1}$ denote 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}, firm-specific pre-treatment variables summarized in $X_{i,t-1}$ and industry-year fixed effects ϕ_{jt} in the following way:

$$y_{i,\,t+1}-y_{i,\,t-1}=\alpha_{0,ITT}+\phi_{jt}\alpha_{ITT}+X_{i,t-1}\beta_{ITT}+\gamma_{ITT}Z_{it}+\varepsilon_{it} \qquad (1)$$

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 fixed assets as a measure of capital and log age to control for firm age effects. Furthermore, we account for the availability of internal financial means for R&D projects by including log sales and a dummy variable that is 1 for a firm with 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: First, we include log R&D expenses in year $t-1$ to address the potential concern that growth rates may vary with pre-treatment levels of R&D investment. Specifically, growth in R&D expenditure is likely to be higher for firms starting R&D activities because of R&D-specific set-up cost. Thus, we add a dummy variable R&D experience$_{t-2}$ which equals 1 if firms have prior R&D

experience in year $t-2$ and 0 else. Third, the log patent stock in year $t-1$ captures firms' past innovation success.[25] Furthermore, unobserved industry-time specific factors like technological opportunities or (expected) demand for innovative technological solutions might drive a firm's decision to invest in R&D and hence its growth rate of R&D expenditure as well as its likelihood to get R&D subsidies which would render OLS estimates biased. We control for these unobserved industry-time specific factors by adding a vector of industry-year fixed effects ϕ_{jt} where j indicates the industry. ε_{it} is an i.i.d. error term with mean 0 and variance σ_{ε}^{2}. $\alpha_{0,ITT}$ denotes the constant and α_{ITT}, β_{ITT} and γ_{ITT} are parameters to be estimated. The main parameter of interest in equation (1) is γ_{ITT} which measures the average intention-to-treat effect of an R&D subsidy on the growth of R&D expenditures (ITT_Y).

As explained in section 2, if the ITT, i.e. the initial assignment to treatment Z_{it}, was randomized, we could consistently estimate equation (1) with OLS. However, the R&D application and granting process is generally not random but is based on firm-specific characteristics. The fact that selection into ITT is also not random in our data is confirmed in Table 3 showing that both groups differ significantly in observed variables. Subsidized firms are significantly larger, younger and de-novo private and they have higher sales, more frequently R&D experience and a larger prior patent stock. A selection bias arises if a firm's R&D investment decision (partly) depends on the same common variables confounding selection into treatment Z_{it}. Past innovation success, for instance, is likely to explain both the likelihood of getting an R&D subsidy and the growth in R&D expenditure. 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.

A key advantage of randomized experiments for estimating ITT is that the treated and control groups only randomly differ from one another on all observed and unobserved covariates (Stuart 2009). Matching methods have become a widely used sample design tool to mimic randomization as best as possible by selecting a control group that is similar to the treatment group on all observed covariates. By improving the covariate balance between both groups, the treatment variable becomes closer to being independent of the confounding variables. In practice, finding a well-balanced matched control sample is often time-consuming

[25] Since a non-negligible proportion of firms does not have any patents, we use a transformed patent stock by adding 1 patent to each firm's patent stock.

in case of a large set of covariates and the outcome also depends on the specific matching procedure (Hainmueller and Xu 2013).

Entropy balancing, recently developed by Hainmueller (2012), is an alternative method to achieve covariate balance. We propose using entropy balancing as a first design step in the ITT estimation to get an (almost) randomized assignment to treatment. 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 for 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 the weighted control group exactly balance their counterparts in the treatment group. Table 3 shows the mean, variance and skewness of the covariates $X_{i,t-1}$ in the treatment and control group before and after balancing. While we see large differences between both groups for almost all covariates and all moments before balancing, the first, second and third moments of the covariate distributions are (almost) identical in the ITT and control group after entropy balancing.

Table 3: Covariate distribution of ITT and control group before and after balancing

Variable	ITT group (N=816)			Control group (N=4347)			t-test on mean difference
	Mean	Variance	Skewness	Mean	Variance	Skewness	p-value
				Before balancing			
R&D experience $_{t-2}$ (0/1)	0.471	0.249	0.118	0.195	0.157	1.537	p<0.001
Employment $_{t-1}$ (log)	7.519	1.345	-0.354	7.126	2.409	-0.163	p<0.001
Fixed assets $_{t-1}$ (log)	21.300	1.013	0.520	21.340	1.733	0.605	p=0.307
Sales $_{t-1}$ (log)	20.860	1.602	0.124	20.500	2.989	-0.095	p<0.001
Age $_{t-1}$ (log)	2.295	0.201	-0.845	2.343	0.185	-0.835	p=0.004
Patent stock $_{t-1}$ (log)	1.575	2.388	0.765	0.858	1.998	1.944	p<0.001
Profitability $_{t-1}$ (0/1)	0.816	0.150	-1.633	0.793	0.164	-1.450	p=0.138
Minority state-owned $_{t-1}$ (0/1)	0.241	0.183	1.208	0.264	0.194	1.073	p=0.185
Privatized $_{t-1}$ (0/1)	0.257	0.191	1.110	0.229	0.177	1.291	p=0.078
De-novo private $_{t-1}$ (0/1)	0.327	0.220	0.737	0.255	0.190	1.125	p<0.001
				After balancing			
R&D experience $_{t-2}$ (0/1)	0.471	0.249	0.118	0.471	0.249	0.120	p=0.985
Employment $_{t-1}$ (log)	7.519	1.345	-0.354	7.519	1.345	-0.355	p=0.992
Fixed assets $_{t-1}$ (log)	21.300	1.013	0.520	21.300	1.012	0.520	p=0.996
Sales $_{t-1}$ (log)	20.860	1.602	0.124	20.860	1.601	0.123	p=0.991
Age $_{t-1}$ (log)	2.295	0.201	-0.845	2.295	0.201	-0.844	p=0.992
Patent stock $_{t-1}$ (log)	1.575	2.388	0.765	1.573	2.385	0.766	p=0.985
Profitability $_{t-1}$ (0/1)	0.816	0.150	-1.633	0.816	0.150	-1.633	p=0.997
Minority state-owned $_{t-1}$ (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 by Hainmueller and Xu (2013).

Entropy balancing has two main advantages compared to matching which makes it particularly attractive in our setting. First, due to the weighting the method does not discard any observations, and we can subsequently use the weights in order to estimate the treatment effect in equation (1) using weighted OLS. Table 4 reports the distribution of weights in the control group.

Table 4: Distribution of entropy-based weights within control group

mean	sd	min	p5	p10	p25	p50	p75	p90	p95	max	N
0.188	0.328	0.000	0.002	0.002	0.020	0.074	0.200	0.486	0.799	4.355	4347

Second and more importantly, under most circumstances entropy balancing is more bias-reducing in finite samples than matching.[26] Thus, our treatment gets closer to randomization since we obtain a much higher degree of covariate balance. This is achieved because entropy balancing allows us to impose a large set of balance constraints already as part of the procedure to find optimal weights. Getting an (almost) randomized ITT is not only important at this stage but is critical also for the IV strategy in estimating the $CACE$. Table 5 reports estimation results for the likelihood of getting 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 differences in firm's incentives to apply for funding and the eligibility and selection criteria of major R&D programs in China. We additionally include industry-year fixed effects in order to control for changes in China's innovation policy and to account for the fact that time patterns of R&D support may differ across industries.[27] Columns (1) and (2) report the estimation results before balancing using both a probit model and a linear probability model (LPM) that is robust against violations of normality. Even after controlling for industry-year fixed effects, we find significant effects of R&D experience, age, profitability, and ownership attributes on the likelihood of getting R&D subsidies. Our specification explains 22.1% of the variation in the ITT selection. Columns (3) and (4) re-estimates the likelihood of getting R&D subsidies after balancing, i.e. using the weights found for the control group using entropy balancing. The results impressively show

[26] Matching is less bias reducing unless the distributions of the covariates are ellipsoidally symmetric or mixtures of proportional ellipsoidally symmetric distributions. Ellipsoidal symmetry fails e.g. 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 43% to 1.8% (see also LaLonde 1986 and Dehejia and Wahba 1999).

[27] Like matching, balancing only controls for selection on observables. But note that controlling for the observed covariates also implies controlling for the unobserved covariates to the extent that they are correlated with the observed ones. Thus, the only unobserved covariates of potential concern are those unrelated to the observed covariates (Stuart 2009).

that after covariate balancing all variables became insignificant and the explanatory power is reduced to almost zero.[28] This outcome reflects a quasi-randomized selection into ITT.

Table 5: Non-random and randomized assignment of R&D Subsidies

	Non-random assignment (before balancing)		Randomized assignment using entropy balancing	
	LPM (1)	Probit (2)	LPM (4)	Probit (5)
R&D experience $_{t-2}$ (0/1)	0.084***	0.066***	-0.000	-0.000
	(0.022)	(0.017)	(0.038)	(0.038)
Employment $_{t-1}$ (log)	0.108***	0.168***	0.000	0.000
	(0.022)	(0.368)	(0.100)	(0.098)
Employment^2 $_{t-1}$ (log)	-0.008***	-0.013***	-0.000	-0.000
	(0.002)	(0.003)	(0.007)	(0.007)
Fixed assets $_{t-1}$ (log)	-0.001	-0.000	-0.000	-0.000
	(0.011)	(0.014)	(0.035)	(0.034)
Sales $_{t-1}$ (log)	0.015*	0.020*	0.000	0.000
	(0.008)	(0.011)	(0.027)	(0.026)
Age $_{t-1}$ (log)	-0.071***	-0.063***	-0.000	-0.000
	(0.021)	(0.019)	(0.045)	(0.044)
Patent stock $_{t-1}$ (log)	0.013*	0.012**	-0.000	-0.000
	(0.007)	(0.006)	(0.013)	(0.013)
Profitability $_{t-1}$ (0/1)	0.034**	0.030*	0.000	0.001
	(0.016)	(0.016)	(0.038)	(0.038)
Minority state-owned $_{t-1}$ (0/1)	0.032*	0.034	0.000	0.000
	(0.019)	(0.022)	(0.053)	(0.052)
Privatized $_{t-1}$ (0/1)	0.007	0.010	0.000	0.000
	(0.017)	(0.019)	(0.045)	(0.044)
De-novo private $_{t-1}$ (0/1)	0.031	0.027	0.001	0.001
	(0.021)	(0.022)	(0.052)	(0.051)
Industry-year FE	Yes	Yes	Yes	Yes
Observations	5163	4570	5163	4570
(Pseudo) R2	0.207	0.203	0.001	0.000

Note: Average marginal effects on the likelihood of getting an R&D subsidy using a linear probability model (LPM) in columns (1) and (3) and probit model in columns (2) and (4). Standard errors are clustered at the firm level. Industry-year FE perfectly predict outcome in probit estimation for 593 observations. ***, ** and * indicate significance at the 1%, 5% and 10% level.

5.4 CACE: Econometric Model and IV Estimation

In order to provide evidence on the efficacy of R&D subsidies on the growth of R&D expenditure in an ideal situation without misappropriation of funds, we estimate the complier average causal effect utilizing equation (2):

$$y_{i,\,t+1} - y_{i,\,t-1} = \alpha_{0,CACE} + \phi_{jt}\alpha_{CACE} + X_{i,t-1}\beta_{CACE} + \gamma_{CACE}D_{it} + \nu_{it}. \quad (2)$$

In contrast to equation (1), we use actual treatment D_{it} on the right-hand side which is 1 if firm i in year t has complied to the contract funding rules and has spent the subsidy on research

[28] 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%.

projects and 0 else. As explained in section 2, compliance is likely to be endogenous because of self-selection which, however, can be addressed by using the assigned treatment Z_{it} as instrument given Z_{it} is randomized. The last subsection has shown that initially the allocation of R&D subsidies, is not random, however, using entropy balancing as a first design step we get an almost randomized ITT variable Z_{it}. Thus, we estimate equation (2) using IV in combination with entropy weights.

6 Results

In this section, we first report ITT and $CACE$ effects of R&D subsidies on the growth rate of R&D expenditures in subsection 6.1. Hereafter, we show robustness tests in subsection 6.2 and results for alternative outcomes and long-term effects in subsection 6.3.

6.1 Treatment Effects on R&D Expenditures

Table 6 reports our main results. We first estimate the ITT by running OLS on equation (1) where the R&D subsidy treatment Z is randomized based on entropy balancing in a first step. We use two different specifications. In column (1) we estimate equation (1) without using pre-treatment R&D expenditure y_{it-1} as explanatory variable. This model controls for any time-invariant unobserved heterogeneity (fixed effects) at the firm, industry, and province level in the log level of R&D expenditure. If regression to the mean behavior is present in the data, implying that firms with a high pre-treatment R&D expenditure tend to have lower R&D growth rates and vice versa, and firm-specific characteristics in X_{it-1} are positively correlated with pre-treatment R&D level y_{it-1}, regression results in column (1) are downward biased (Allison 1990). Hence, specification (2) adds pre-treatment R&D expenditure.[29] Results in column (2) show that the pre-treatment level is highly significant and a comparison of the estimated coefficients between column (1) and (2) confirms a downward bias for almost all coefficients. Thus, model (2) is our preferred specification. The control variables behave as expected. The R&D growth rate is higher for firms that have prior R&D experience but it declines in the pre-treatment level of R&D expenditures. Furthermore, higher pre-treatment sales, patent stock and profitability is associated with a higher growth in R&D expenditure. For firm size, we find an inverse u-shape effect on R&D growth with an estimated inflection point of 2220 employees.

The main parameter of interest, however, is the coefficient γ_{ITT} of the assigned R&D subsidy treatment Z. It is a measure for the effectiveness of the R&D subsidy policy in China

[29] This model is equivalent to a dynamic model in which $y_{it+2}$is regressed on $\alpha_0, y_{it-1}, X_{i-1}, Z_{it}$ and ϕ_{jt}.

given 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 $\gamma_{ITT} > 0$, the R&D support scheme increases total R&D investment. This encompasses both, a situation in which total R&D expenditure increases by less (*partial crowding out*) and by more (*additionality*) than the subsidy amount S. Since Z is a binary treatment dummy, γ_{ITT} does not allow us to directly infer how many RMB 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. Einiö (2014) shows, however, 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 25% crowding out, $h \geq 0.25$, against the alternative that the subsidy leads to less than 25% crowding out. Under the null hypothesis the following condition must hold: $Y_{i,t+1} \leq Y_{i,t-1} + (1-h) \cdot S$, where h denotes the crowding out rate and $Y_{i,t+1}$ and $Y_{i,t-1}$ denote the post- and pre-treatment level 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% which is usually paid by the government as estimate for s, i.e. $s = 0.5$, we get the following condition: $\log Y_{i,t+1} - \log Y_{i,t-1} = y_{i,t+1} - y_{i,t-1} \leq \log\left(\frac{1}{1-(1-h)\cdot s}\right) = 0.47$. 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 $\gamma_{ITT} \leq \log\left(\frac{1}{1-(1-h)\cdot s}\right)$=0.47. Similarly, we get a threshold of $\gamma_{ITT} \leq 0.288$ for 50% crowding out and $\gamma_{ITT} \leq 0.693$ for no crowding out (additionality). 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}.[30]

In column (1), γ_{ITT} is 0.521 and we can reject the null hypothesis of full crowding-out only at the 10% level. Once we additionally control for heterogeneity in pre-treatment levels of R&D expenditures in column (2), γ_{ITT} increases to 0.881. We can reject full crowding out at the 1% level and more than 50% crowding out at the 5% level.

Next we estimate the CACE based on equation (2) with 2SLS and employ the randomized Z as an instrument for the endogenous D. The highly significant point estimate in the first stage and the Kleibergen-Paap F-statistic support the relevance of the IV. In column (3) γ_{CACE} is 1.269 and rejects 50% crowding out at the 10% level. Once we control for pre-

[30] 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 get public funding.

treatment levels of R&D expenditures in column (4), γ_{CACE} increases to 2.187 and rejects crowding-out and neutral effects at the 5% level – hence additionality is confirmed for compliers. Because of our setting with one-sided noncompliance we obtain an (almost) identical γ_{CACE} when we divide γ_{ITT} by the difference in compliance rates between ITT group and control group $(0.881/(0.41/1) = 2.144)$. The coefficients of control variables are similar, except that employees have no inflection point. The comparison of γ_{ITT} and γ_{CACE}, which estimate effectiveness and efficacy, shows that the effect of China's R&D policy could be more than twice as large if misappropriation had not occurred.

Table 6: ITT and CACE of R&D subsidies on growth of R&D expenditures

	ITT		CACE	
	(1)	(2)	(3)	(4)
Z	0.521*	0.881***		
	(0.305)	(0.282)		
D			1.269*	2.187***
			(0.728)	(0.686)
Pre-treatment R&D level $_{t-1}$ (log)		-0.663***		-0.694***
		(0.026)		(0.027)
R&D experience $_{t-2} > 0$ (0/1)	-2.480***	2.115***	-2.684***	1.978***
	(0.396)	(0.412)	(0.417)	(0.405)
Employment $_{t-1}$ (log)	-0.807	2.250***	-0.717	2.162***
	(0.901)	(0.830)	(0.878)	(0.792)
Employment^2 $_{t-1}$ (log)	-0.060	-0.146**	-0.054	-0.140
	(0.065)	(0.061)	(0.063)	(0.059)
Fixed assets $_{t-1}$ (log)	0.055	-0.194	0.106	-0.117
	(0.335)	(0.305)	(0.330)	(0.300)
Sales $_{t-1}$ (log)	0.189	0.614***	0.159*	0.582**
	(0.266)	(0.238)	(0.260)	(0.233)
Age $_{t-1}$ (log)	0.661*	-1.337***	0.768**	-1.245***
	(0.373)	(0.326)	(0.374)	(0.316)
Patent stock $_{t-1}$ (log)	0.260**	0.396***	0.260**	0.402***
	(0.131)	(0.117)	(0.128)	(0.114)
Profitability $_{t-1}$ (0/1)	0.068	0.839**	0.012	0.778**
	(0.474)	(0.397)	(0.467)	(0.383)
Minority state-owned $_{t-1}$ (0/1)	-0.663	0.033	-0.673	0.048
	(0.503)	(0.459)	(0.495)	(0.445)
Privatized $_{t-1}$ (0/1)	-0.474	-0.384	-0.450	-0.338
	(0.531)	(0.458)	(0.522)	(0.443)
De-novo private $_{t-1}$ (0/1)	0.094	0.446	0.061	0.406
	(0.494)	(0.451)	(0.486)	(0.441)
Crowding-out test (p-value)				
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.221	0.018	0.088	0.003
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.433	0.073	0.163	0.006
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.286	0.253	0.214	0.015
IV 1st stage (Z)			0.421***	0.403***
			(0.017)	(0.017)
KP F-statistic			593.8	569.7
Firm FE	Yes	No	Yes	No
Industry-year FE	Yes	Yes	Yes	Yes
Observations	5163	5163	5163	5163

Note: ***, ** and * indicate significance at the 1%, 5% and 10% level. Standard errors are clustered at the firm level. The estimate for IV 1st stage (Z) is also the estimate for ITT_D. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

In Table 7, we compare ITT and CACE with three other treatment effects estimators. In column (1) we estimate an upward biased 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 fixed effects. Hence, the coefficient of 1.337 is overestimated and erroneously confirms additionality at the 5% level. In comparison, γ_{ITT} in column (2) is much lower and rejects 25% crowding-out only at the 10% level. Next we compare effects of the *actual* treatment D. In column (3) we estimate the *as-treated* effect, which compares outcomes of compliers (D=1, Z=1) with a control group that includes noncompliers (D=0, Z=1), and is upward biased because noncompliers have a lower expected outcome than compliers. In column (4) the *per-protocol* effect is estimated by excluding observable noncompliers (D=0, Z=1) but not unobservable noncompliers in the control group (Z=0). The *per-protocol* is smaller than the *as-treated* bias, because excluded noncompliers (D=0, Z=1) have a lower expected outcome than the control group (Z=0). Both coefficients are significantly larger than γ_{CACE} in column (5) and erroneously confirm additionality at the 1% level.

Table 7: Comparison of treatment effects

	Biased ATT (1)	ITT (2)	As-treated (3)	Per-protocol (4)	CACE (5)
Non-randomized Z	1.337*** (0.287)				
Randomized Z		0.881*** (0.282)			
Non-instrumented D			3.294*** (0.374)	3.117*** (0.375)	
Instrumented D					2.187*** (0.686)
Crowding-out test (p-value)					
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.000	0.018	0.000	0.000	0.003
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.001	0.073	0.000	0.000	0.006
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.012	0.253	0.000	0.000	0.015
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes
Observations	5163	5163	5163	4682	5163

Note: Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

6.2 *Heterogeneous Treatment Effects*

In this section we investigate the heterogeneity of treatment effects to answer two questions: (1) has the MLP improved the R&D policy design and (2) do high-tech and private firms exhibit higher returns to R&D subsidies? In Table 8 we split the sample into the pre-MLP period 2001-6 and the MLP period 2007-11. In the pre-MLP period both γ_{ITT} in column (1) and γ_{CACE} in column (2) fail to reject total crowding-out, which shows that policy ineffectiveness is not only due to noncompliance but also due to poor policy design. In the MLP period γ_{ITT} in column (3) rejects 25% crowding-out at the 5% level and γ_{CACE} in column (4) confirms additionality at the 5% level. While the gap between the (imprecisely) estimated effectiveness and efficacy narrows over time, the effect of China's R&D policy during the MLP period could still be more than twice as large without misappropriation. The key take-away from this comparison over time is that the MLP has improved the design of R&D policy significantly.

Table 8: ITT and CACE in 2001-6 and 2007-11

	2001-2006		2007-2011	
	ITT	CACE	ITT	CACE
	(1)	(2)	(3)	(4)
Z	0.294		1.096***	
	(0.403		(0.350)	
D		1.895		2.270***
		(2.511)		(0.707)
Crowding-out test (p-value)				
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)			0.010	0.003
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)			0.037	0.005
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)			0.125	0.013
IV 1^{st} stage (Z)		0.155***		0.483***
		(0.022)		(0.019)
KP F-statistic		48.4		619.6
Pre-treatment R&D level	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes
Observations	2082	2082	3081	3081

Note: We conduct balancing for each subsample to maintain a randomized Z. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level.

In Table 9 we focus on high-tech firms, which are more R&D-oriented than non high-tech firms, and private firms, which face more capital constraints than state-owned firms. Among the firms in the estimation sample 11.3% operate in high-tech industries and 32.5% are grantees – of which 51.5% are compliers. Among firms in non high-tech industries 13.9% are grantees – of which 38.3% are compliers. The higher support rate of high-tech firms is reflective of China's picking-the-winner R&D policy and the higher ratio of compliers rate suggests a higher profit-maximizing level of R&D among supported high-tech firms in comparison to supported non high-tech firms. Both γ_{ITT} and γ_{CACE} in columns (1) and (2) are larger than for non high-tech firms in columns (3) and (4), but less significant, which may be due to the low number of observations.

To account for ownership regimes we partition the sample into private (including privatized and de-novo private) and state-owned (including state-owned and minority state-owned) firms. Among the 55.6% private firms 18.7% are grantees – of which 45.0% are compliers. State-owned firms have a comparatively lower grant rate of 12.1% and a lower compliance rate of 34.1%. For private firms, γ_{ITT} in column (5) rejects crowding-out at the 5% level and γ_{CACE} in column (6) confirms additionality at the 1% level. For state-owned firms, interestingly, neither γ_{ITT} in column (7) nor γ_{CACE} in column (8) can reject full crowding-out. The comparison across firm types emphasizes that a more selective funding of high-tech and private firms may result in higher returns to R&D grants. Further, it also shows that the estimated treatment effects for high-tech firms in Liu et al. (2016) and private firms in Hu and Deng (2018) most likely exceed the average treatment effect for the population of firms.

Table 9: ITT and CACE of R&D subsidies on growth of R&D expenditures

	High-tech firms		Non high-tech firms		Private firms		State-owned firms	
	ITT	CACE	ITT	CACE	ITT	CACE	ITT	CACE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z	1.274*		0.839***		0.975***		-0.227	
	(0.696)		(0.306)		(0.306)		(0.439)	
D		2.528*		2.230***		2.570***		-0.691
		(1.302)		(0.809)		(0.780)		(1.291)
Crowding-out test (p-value)								
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.079	0.043	0.038	0.008	0.009	0.002		
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.124	0.057	0.118	0.014	0.032	0.004		
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.202	0.079	0.320	0.028	0.108	0.010		
IV 1^{st} stage (Z)		0.504***		0.376***		0.442***		0.328***
		(0.041)		(0.019)		(0.020)		(0.025)
KP F-statistic		153.5		410.5		473.1		170.5
Pre-treatment R&D level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	526	526	4637	4637	2872	2872	2291	2291

Note: We conduct balancing for each (sub) sample to maintain a randomized Z. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level.

6.3 Robustness Tests

In Table 10 we report robustness tests for γ_{ITT} and γ_{CACE}. First, we address measurement issues. In column (1) we assume that grantees do not receive an annual grant payment in every year of the R&D project but receive a lump sum in t and expense it evenly over the years t to $t+2$.[31] In column (2) we use strict instead of broad R&D subsidies to rule out a possible inflation of the misappropriation measure[32] and in column (3) we exclude all observations where we may mistake accounting nuisance for misappropriation.[33] This setting only retains noncomplier observations when accounting noise, as an alternative explanation for misappropriation, can be rejected. For column (4) we assume that firms only have 50% of R&D subsidies at their disposal, while 50% have to be returned to officials and intermediaries. In column (5) we exclude 37 observations with partial instead of full misuse and in column (6) we strictly adhere to the assumption that the largest possible subsidy is 50% of the total post-treatment R&D and exclude 22 compliers with an R&D subsidy intensity >50% in year t. Finally, in column (7) the R&D experience is excluded, because the likelihood to observe prior R&D increases with the number of years a firm remains in the panel. In summary, the results are robust: γ_{ITT} either rejects 25% or 50% crowding-out and γ_{CACE} always confirms additionality.

Second, we focus on potential time-variant confounders of R&D subsidies and R&D expenditures at the provincial level. In column (8) we include the log of GDP per capita, the share of loss-making firms, as well as the log of R&D expenditures divided by acreage to control for general economic conditions. In column (9) we specifically control for the influence of political uncertainty by accounting for the annual turnover of city-level majors and party secretaries per province, which varies between 0 and 23 changes per province-year. In column (10) we control for corruption and resource curse as potential confounders. The log of investment in construction and the log of civil servants to large- and medium-sized firms[34] control for the potential of corruption and the log ratio of convicted civil servants to civil servants is assumed to measure successful anti-corruption efforts. Resource curse is proxied by

[31] In the total sample, Z=1 observation increase from 1710 to 2543. In the estimation sample, due to the sample definition specified in Figure 3, the composition of observations with Z=1 and Z=0 changes from 816 and 4347 to 655 and 4325, respectively. The decrease in γ_{ITT} and γ_{CACE} suggest that the alternative 3-year timing structure explains less of the variation in the outcome.

[32] In the total sample, Z=1 observation decrease from 1710 to 1172. In the estimation sample, this translate into a composition of 548 observations with Z=1 and 5543 observations with Z=0, respectively.

[33] 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. If this also leads to the exclusion of *actual* noncompliers from the ITT group, which we cannot rule out, this will introduce upward bias (similar to the *per-protocol* setting).

[34] According to Wu Jinglian, a Chinese economists, the likelihood for rent extraction of officials from firms increases in the officials-to-firm ratio.

the log of explored net coal resources in a given year. The results confirm the robustness of our specification, as γ_{ITT} and γ_{CACE} remain almost unchanged.

Third, we control for other policies to account for substitution bias. In column (11) we account for the MLP macro-shock and include an indicator that turns from zero to one after 2006. As policies under the MLP were generally accompanied by a higher degree of decentralization, in column (12) we interact the MLP indicator with the firm's log distance to the location of the relevant regulators, which is Beijing for central state-owned firms and the provincial capital for all other firms, resulting distances between 0.5 km and 2283 km. In column (13) we rule out that the firm's location choice is endogenous to support policies introduced after the MLP and condition on firms established before 2007. The results in columns (11) to (13) remain robust. In column (14) we exclude (expected) participants of the HNTE program [35] to account for two opposing effects: expected participation may be accompanied by a higher growth rate of R&D to reach eligibility, whereas accredited participants have reached a 3% R&D-intensity and will stabilize this intensity (in accordance with sales growth), which may correspond with a lower R&D-growth rate. Increases in γ_{ITT} and γ_{CACE} after the exclusion of (expected) HNTE participants suggest that the second effect dominates. Finally, in column (15) controlling for non-R&D subsidies decreases the effect of R&D subsidies – suggesting some mild substitution bias. [36] Nonetheless, even after this rigorous test, which usually is not done in the R&D evaluation literature, γ_{ITT} still rejects full crowding out at the 5% level and γ_{CACE} rejects 25% crowding-out at the 10% level.

Fourth, we perform several standard tests. In column (16) we winsorize continuous variables at the 1 and 99 percentiles but fail to note substantial differences. In column (17) we exclude compliers (Z=1, D=1) and can reject the hypothesis that noncompliance (Z=1, D=0) has a significant effect on the outcome, which suggests that the exclusion restriction is not violated. In column (18) we randomly match Z and the related D indicator to firms and confirm that the treatment effect actually comes from R&D subsidies but not from spurious correlation. The unconfoundedness of the IV, e.g. whether pre-treatment trends – trends before actually receiving Z – were different between the ITT or compliers groups and the control group, is

[35] We obtain data on all HNTE participants from the program's official webpage and match these information to our firms.

[36] In the total sample 62.4% and in the estimation sample 55.8% of observation receive non-R&D subsidies. The correlation between R&D subsidies and non-R&D subsidies in the estimation sample is 0.3. Non-R&D subsidies, which are often provided to firms in situations of financial distress or, more generally, to improve firm competitiveness by cost reduction, show a positive correlation with R&D-growth, which suggests that such subsidies may also redirect internal funds towards R&D investments.

tested in column (19) and the hypothesis of common trends is not rejected. Column (20) shows results for a specification where the outcome is the change in R&D-growth from year $t-2$ to year $t+1$. This specification quantifies the effect of R&D subsidies net of potential pre-treatment trends and rejects 25% crowding-out at the 5% level for γ_{ITT} and confirms additionality at the 1% level for γ_{CACE}. In conclusion, these test support the robustness of our results.

Table 10: Robustness tests I

	R&D subsidies over 3 years (1)	Strict R&D subsidies (2)	Accounting noise excluded (3)	50% net R&D subsidies (4)	Partial misuse excluded (5)	≤50% R&D support (6)	Without R&D stock (7)	Economic development (8)	Political uncertainty (9)	Corruption and resource curse (10)
Z	0.791***	1.053***	1.334***	0.881***	0.817***	0.819***	0.942***	0.880***	0.885***	0.686***
	(0.307)	(0.296)	(0.324)	(0.282)	(0.286)	(0.283)	(0.284)	(0.282)	(0.282)	(0.260)
Crowding-out test (p-value)										
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.051	0.005	0.001	0.018	0.032	0.030	0.011	0.018	0.017	0.020
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.148	0.025	0.004	0.073	0.113	0.109	0.049	0.074	0.071	0.081
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.365	0.112	0.024	0.253	0.332	0.328	0.191	0.254	0.248	0.271
D	1.788***	2.226***	2.665***	1.972***	1.922***	2.112***	2.403***	2.184***	2.197***	2.133***
	(0.677)	(0.603)	(0.627)	(0.619)	(0.657)	(0.713)	(0.711)	(0.686)	(0.685)	(0.678)
Crowding-out test (p-value)										
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.013	0.001	0.000	0.003	0.006	0.005	0.001	0.003	0.003	0.003
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.026	0.002	0.000	0.008	0.014	0.011	0.003	0.006	0.006	0.007
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.053	0.006	0.001	0.019	0.031	0.023	0.008	0.015	0.014	0.017
IV 1st stage (Z)	0.442***	0.473***	0.501***	0.447***	0.425***	0.388***	0.392***	0.403***	0.403***	0.406***
	(0.018)	(0.018)	(0.019)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)
KP F-statistic	636.5	674.7	691.2	706.7	620.6	529.7	540.9	575.5	568.4	583.2
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-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4980	6091	4356	5163	5126	5141	5163	5163	5163	5135

Note: We conduct balancing for each (sub) sample to maintain a randomized Z. Column (10) is estimated without Tibet because there are no LMEs recorded for the respective years. The log of civil servants to large- and medium-sized firms and the log ratio of convicted civil servants to civil servants are winsorized at the 1st and 99th percentile. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

Table 10: Robustness tests II

	MLP shock (11)	Distance to regulator (12)	Established before 2007 (13)	(Expected) HNTE participation excluded (14)	Non-R&D subsidies (15)	Winsorized (16)	Non-compliers (17)	Random R&D subsidies (18)	Pre-treatment trend outcome $y_{t-1} - y_{t-2}$ (19)	Treatment effect net of pre-treatment outcome $y_{t+1} - y_{t-2}$ (20)
Z	0.881***	0.869***	0.817***	1.167***	0.562**	0.888***	-0.402	0.007	0.228	1.110***
	(0.282)	(0.297)	(0.286)	(0.319)	(0.282)	(0.280)	(0.357)	(0.180)	(0.235)	(0.351)
Crowding-out test (p-value)										
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.018	0.025	0.017	0.003	0.165	0.016				0.010
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.073	0.090	0.070	0.015	0.372	0.068				0.034
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.253	0.277	0.246	0.069	0.322	0.243				0.118
D	2.187***	2.177***	2.205***	3.511***	1.348**	2.205***		0.018	0.568	2.755***
	(0.686)	(0.729)	(0.690)	(0.937)	(0.661)	(0.681)		(0.428)	(0.572)	(0.852)
Crowding-out test (p-value)										
H0: $\gamma \leq 0.287$ ($h \geq 50\%$)	0.003	0.005	0.003	0.000	0.054	0.002				0.002
H0: $\gamma \leq 0.470$ ($h \geq 25\%$)	0.006	0.010	0.006	0.001	0.092	0.005				0.004
H0: $\gamma \leq 0.693$ ($h \geq 0\%$)	0.015	0.021	0.014	0.001	0.161	0.013				0.008
IV 1^{st} stage (Z)	0.403***	0.399***	0.403***	0.332***	0.417***	0.403***		0.412***	0.403***	0.403***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.017)		(0.017)	(0.017)	(0.017)
KP F-statistic	569.7	522.0	564.7	348.0	556.8	568.6		587.9	569.7	569.7
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-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5163	4815	5139	4804	5163	5163	4828	5163	5163	5163

Note: We conduct balancing for each (sub) sample to maintain a randomized Z. In column (12) we were unable to obtain the distance measure for 84 firms, however, these are missing at random. In column (17) compliers are removed after balancing. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

6.4 *Alternative Outcomes and Long-term Effects*

In Table 11 we estimate the ITT and CACE effects of R&D subsidies on alternative outcomes to extent the analysis from input additionality towards output and behavior additionality.[37] Related to input additionality, γ_{ITT} and γ_{CACE} in column (1) and (2) not only confirm that grantees have significantly higher R&D growth but also that this growth is over proportional to the (induced) growth in sales. Columns (3) to (5) show positive direct effects on employment, fixed assets, and sales, confirming output additionality. Interestingly, column (6) shows that there is no effect on labor productivity which may be related to the mission-oriented focus of many R&D programs. The literature documents that while mission-oriented support may have a positive impact on growth and employment (Steinmueller 2010) it is less likely to have an impact on productivity (Griliches 1995).[38] Column (7) shows that the growth in R&D is accompanied by an increase in patenting.[39] However, in columns (8) to (10) the patent measures also reveal that, in comparison to the control group, neither grantees nor compliers file more high-tech IT patent, file more joint applications with universities or foreign scientists residing in China. This suggests that China's R&D policy has yet to induce an observable shift towards the development of strategic core technologies, university-industry collaboration, or contributions of foreign scientists employed in Chinese firms.

In Table 12, we extent the impact period by two years from $t-1$ to $t+3$ to capture potential long-term effects. In columns (1) γ_{ITT} confirms additionality at the 5% level and γ_{CACE} at the 1% level. While this shows that R&D subsidies induce private R&D spending among grantees in the long-term, the effect is more than three times higher in the absence of misappropriation (this is also confirmed in column (2) for R&D intensity) – which provides a strong argument for improved monitoring. Further, γ_{CACE} doubles in the long-term but γ_{ITT} increases less than twofold. While the long-term effects for employment and sales exceed short-term effects, the long-term effects for all other outcomes are insignificant. This suggests that grantees may persistently hire more (R&D) employees and have a sustained increase in their sales, whereas there is no persistent difference in capital-growth. The missing long-term effect

[37] We omit a regressor in the estimated equation if the variable is (part) of the outcome.

[38] This is because mission-oriented research more often leads to product instead of process innovation. Product innovation may lead to output additionality due to more sales but not to an increase in labor productivity if sales-growth is paralleled by employment-growth. Process innovation, on the other hand, is directly reflected in productivity-growth if the cost reduction is not fully passed on to lower prices but also to higher mark-ups.

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

on patenting is plausible as the majority of China's R&D expenditures are development-oriented[40]: firms may quickly file patents related to the funded project and focus on commercialization hereafter.

[40] In 2011, China's expenditure for basic and applied research was 4.7% and 11.8% of GERD, respectively, while the remaining 83.5% were accounted for by development expenditures (Sun and Cao 2014).

Table 11: Alternative outcomes

	Natural growth rate R&D expenditures	R&D intensity	Employment	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	0.132***	0.541***	0.053**	0.038*	0.098***	0.042	0.086**	-0.023	0.001	0.001
	(0.043)	(0.157)	(0.026)	(0.023)	(0.026)	(0.027)	(0.042)	(0.023)	(0.002)	(0.001)
D	0.327***	1.347***	0.130**	0.092*	0.241***	0.103	0.210**	-0.055	0.001	0.002
	(0.105)	(0.384)	(0.063)	(0.054)	(0.064)	(0.064)	(0.100)	(0.056)	(0.004)	(0.002)
IV 1st stage (Z)	0.403***	0.401***	0.411***	0.409***	0.409***	0.409***	0.411***	0.410***	0.411***	0.401***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
KP F-statistic	569.7	554.9	594.8	588.7	588.8	588.1	592.2	590.2	592.7	578.3
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-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5163	5163	5163	5145	5139	5139	5163	5163	5163	5163

Note: We conduct balancing for each (sub) sample to maintain a randomized Z. In column (1) the natural growth rate of R&D expenditures is calculated as ((1+R&D expenditures $_{t+1}$) – (1+R&D expenditures $_{t-1}$)) / (((1+R&D expenditures $_{t-1}$) + (1+R&D expenditures $_{t+1}$)) / 2). Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

Table 12: Long-term effects $t-1$ to $t+3$

	R&D expenditures	R&D intensity	Employment	Fixed assets	Sales	Labor productivity	Patent applications	High-tech patent applications	University-industry collaboration	Foreign inventors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Z	1.392***	0.839***	0.117***	0.025	0.135***	0.001	0.052	-0.004	0.001	-0.000
	(0.346)	(0.193)	(0.044)	(0.042)	(0.049)	(0.045)	(0.052)	(0.025)	(0.002)	(0.000)
D	4.648***	2.798***	0.389***	0.081	0.444***	0.003	0.171	-0.014	0.002	-0.001
	(1.182)	(0.663)	(0.147)	(0.136)	(0.162)	(0.144)	(0.170)	(0.082)	(0.005)	(0.001)
IV 1st stage (Z)	0.299***	0.300***	0.300***	0.305***	0.304***	0.304***	0.304***	0.301***	0.304***	0.305***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
KP F-statistic	236.3	231.9	230.0	234.2	232.4	230.9	232.9	230.6	232.6	232.7
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-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3846	3815	3846	3828	3815	3815	3846	3846	3846	3846

Note: Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. KP denotes the Kleibergen-Papke Wald F-test on weak instruments (Kleibergen and Papke 2006).

6.5 *Short-term and Long-term Misappropriation Effects*

So far we have been agnostic about the specific alternative use of misappropriated R&D subsidies. However, from a welfare perspective it matters whether the money is invested or used for private consumption by managers. To analyze this issue we estimate the effect of misappropriated R&D subsidies on several outcome measures. For this exercise we first exclude (observed) compliers from the sample and only maintain firms for which the following assigned and actual treatment combinations (Z_i, D_i) are observed: $\{(1,0), (0,0)\}$. Hereafter we randomize Z_i based on entropy balancing. According to the exclusion restriction Z_i affects the potential outcome $Y_i(z, d)$ only via the actual treatment D_i. Hence, misappropriation of R&D subsidies, which we call M_i, should have no effect on the potential outcome when $Y_i(z, d)$ is measured in terms of R&D expenditures. However, if M_i is used for physical investment we would expect to find an effect on fixed assets and other outcomes such as employment, sales and labor productivity.

In column (1) of Table 13 we first investigate the exclusion restriction. As we fail to find a significant effect of M_i on R&D expenditures the exclusion restriction is not rejected. M_i has a significant short-term effect on fixed assets in column (3), sales in column (4) and labor productivity in column (5). In addition, we estimate long-term effects on employment in column (6) and labor productivity in column (8). Based on these results it seems that in the short term firms use misappropriated R&D subsidies for investments, which increases sales and hence labor productivity. In the long run sustained sales growth is accompanied by growth in employment. However, misappropriating firms fail to maintain a higher growth rate of total assets or productivity in the long run. Most importantly, we can reject the hypothesis that misappropriated R&D subsidies are entirely used for consumption. At least some of the money is invested and increases employment and output in the long term.

Table 13: Short-term and long-term misappropriation effects

	R&D expenditures	Employment	Fixed assets	Sales	Labor productivity	Employment $t-1$ to $t+3$	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)
M	-0.334	0.045	0.052**	0.135***	0.087**	0.115**	0.038	0.150***	0.024
	(0.332)	(0.032)	(0.025)	(0.033)	(0.034)	(0.047)	(0.044)	(0.055)	(0.051)
Pre-treatment outcome level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4828	4828	4828	4806	4806	3691	3673	3661	3661

Note: We conduct balancing for each (sub) sample to maintain a randomized Z. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level.

7 Conclusion

This paper is the first to evaluate an R&D subsidy policy in a regime in which monitoring is weak and firm considerably misappropriate funds. Our analysis shows that misappropriation is a major concern in China as we calculate that 42% of grantees misused R&D subsidies, corresponding to 53% of the total amount of R&D subsidies. In a setting with one-sided noncompliance to funding contract rules, we differentiate between the intention-to-treat (ITT) effect and the complier average causal effect (CACE) in order to evaluate R&D policy in China. The ITT shows how effective the R&D policy is in practice when misappropriation exists. The CACE, in contrast, depicts how effective the policy could have been without misappropriation and thus is a measure for the efficacy of the R&D subsidy policy. Combining entropy balancing and IV methods to estimate both ITT and CACE, the ITT results show mild partial crowding out of R&D expenditures. Most strikingly, however, the CACE turns out to be more than twice as large as the ITT and confirms additionality of R&D subsidies. Thus, misappropriation of R&D subsidies considerably undermines the efficacy of Chinese R&D programs.

Literature

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P. 2005. Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics* 120, 701-728.

Angrist, J.D., Imbens, G.W., Runbin, D.B. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91, 444-455.

Angrist, J.D., Pischke, J.S. 2009. *Mostly harmless econometrics*. Princeton University Press, Princeton.

Arque-Castells, P. Mohnen, P. 2015. Sunk cost, extensive R&D subsidies and permanent inducement effects. *Journal of Industrial Economics* 63, 458-494.

Autor, D., Dorn, D., Hanson, G., Pisano, G., Shu, P. 2017. *Foreign Competition and Domestic Innovation: Evidence from U.S. Patents*, NBER Working Paper w22879, Cambridge, MA.

Bloom, H.S. 1984. Accounting for no-shows in experimental evaluation design. *Evaluation Review* 8, 225-246.

Bloom, H.S., Orr, L.L., Bell, S.H., Cave, G., Doolittle, F., Lin, W., Bos, J.M. 1997. The benefits and cost of JTPA Title II-A Programs: Key findings from national job training partnership act study. *The Journal of Human Resources* 32, 549-576.

Boeing, P. 2016. The allocation and effectiveness of China's R&D subsidies – Evidence from listed firms. *Research Policy* 45, 1774-1789.

Boeing, P., Mueller, E., Sandner, P. 2016. China's R&D explosion – Analyzing productivity effects across ownership types and over time. *Research Policy* 45, 159-176.

Chen, Z., Liu, Z., Serrato, J., Xu, D. 2018. *Notching R&D investments with corporate income tax cuts in China*, NBER Working Paper w24749, Cambridge, MA.

David, P., Hall, B., Toole, A. 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, *The Review of Economics and Statistics* 96, 710–728.

Fang, L., Lerner, J., Wu, C., Qi, Z. 2018. *Corruption, government subsidies, and innovation: Evidence from China*. Harvard Business School Working Paper 19-03, Cambridge, MA.

Gonzalez, X., Jaumandreu, J., Pazo, C. 2005. Barriers to innovation and subsidy effectiveness. *RAND Journal of Economics* 36, 930-950.

Gonzalez, X., Pazo, C. 2008. Do public subsidies stimulate private R&D spending? *Research Policy* 37, 371-389.

Griliches, Z., 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 28, 1661–1707.

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., Xu, Y. 2013. ebalance: A Stata package for entropy balancing. *Journal of Statistical Software* 54, 1-18.

Hall, B.H. 2002. The Financing of Research and Development. *Oxford Review of Economic Policy* 18, 35–51.

Hall, B.H., Jaffe, A., Trajtenberg, M. 2005. Market value and patent citations. *RAND Journal of Economics* 36, 16-38.

Heckman, J.J., Smith, J. 2005. Assessing the case for social experiments. *Journal of Economic Perspectives* 9, 85-110.

Howell, S. 2017. Financing Innovation: Evidence from R&D grants. *American Economic Review* 107, 1136-1164.

Hsieh, C., Song, Z. 2015. *Grasp the large, let go the small: The transformation of the state sector in China*. NBER Working Paper w210006, Cambridge, MA.

Hu, A.G.Z., Deng, Y. 2018. Does government R&D stimulate of crowd out firm R&D spending? Evidence from Chinese manufacturing industries. *Economics of Transition* 26. 1-26.

Imbens, G.W., Angrist, J.D. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, 467-475.

Imbens, G.W., Rubin, D.B. 2015. *Causal interference for statistics, social, and biomedical science: An introduction*. Cambridge University Press, NY.

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47, 5–86.

Kline, P. and Walters, C. (2016), Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics* 131, 1795-1848.

Liu, X., Li, X., Li, H. 2016. R&D subsidies and business R&D: Evidence from high-tech manufacturing firms in Jiangsu. *China Economic Review* 41, 1-22.

Peters, B., Roberts, M.J., Vuong, V.A., Fryges, H., 2017. Estimating dynamic R&D choice: An analysis of costs and long-run benefits. *RAND Journal of Economics* 48, 409-437.

Sun, Y., Cao, C. 2014. Demystifying central government R&D spending in China. *Science* 345, 1006-1008.

Takalo, T., Tanayama, T., Tiovanen, O., 2013. Estimating the benefits of targeted S&D subsidies, *Review of Economics and Statistics* 95, 255-272.

Wei, S., Xie, Z., Zhang, X. 2017. From “Made in China” to “Innovated in China”: Necessity, prospect, and challenges. *Journal of Economic Perspectives* 31, 49-70.

Zuniga-Vicente, J.A., Alonso-Borrego, C., Forcadell, F.J., Galan-Zazo, J.I. 2014. Assessing the effect of public subsidies on firm R&D investment: A survey. *Journal of Economic Surveys*. 28, 36–67.

Appendix 1: R&D Policy and Misappropriation

R&D Policy

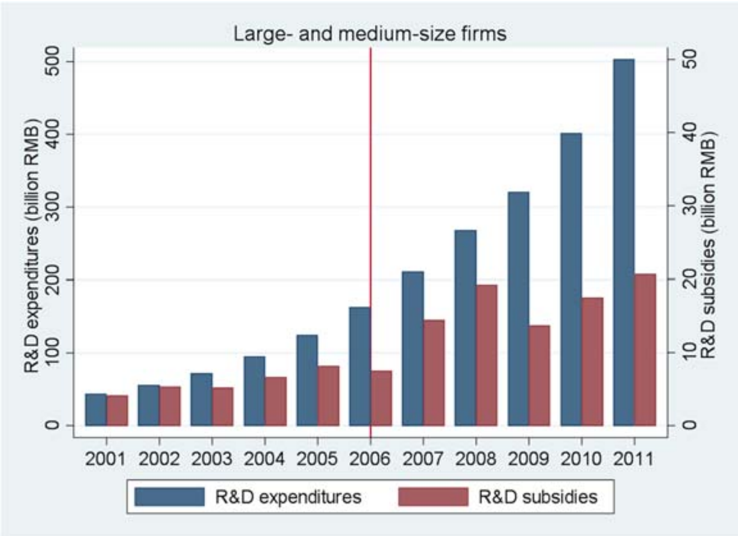
The State Council wants China to become an innovative country by 2020 and a world leader in science and technology by 2050. Against this target, China's ratio of gross expenditures for R&D to GDP has already overtaken the ratio of the European Union and in gross R&D expenditures China is projected to overtake the USA around 2020. In order to stimulate business R&D expenditure the Chinese government invests heavily in innovation policy. The time period 2001 to 2011, underlying this analysis, is covered by the 10th and 11th Five-Year Science and Technology Development Plans (2001-5 and 2006-11) and the Mid- to Long-term Science and Technology Development Plan (MLP) (2006-20). The MLP's seminal agenda proposes a more integrated innovation policy than before. In contrast to Five-Year Plans, the MLP lists more detailed development goals and provides a relatively clear guideline for implementation.[41] A first-order target of the MLP is to increase R&D expenditures of domestic firms.

The MLP not only calls for more funds but also a better management of R&D programs, related to the selection and monitoring of grantees and coordination between various programs and agencies, to reduce redundancies and misallocation of public funds. In addition to R&D programs, other regulations and policies were implemented after 2006 to incentivize R&D.[42] Between 2001 and 2011 the annual amount of R&D subsidies directed to large- and medium-sized firms in China quintupled from 4 to 21 billion RMB (Figure A1), amounting to 123 billion RMB, while R&D expenditures increased more than tenfold from 44 billion RMB to 503 billion RMB. It is also worth noting that under the MLP the proportion of mission-oriented funding has increased (Cao et al. 2013).

[41] Each of the 99 support policies 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 Science and Technology) and was to be implemented in December 2006.

[42] E.g. the accounting regulations for R&D expenditures were amended by the Ministry of Finance: "Enterprise Accounting Standards No. 6 - Intangible Assets" – see in particular Articles 7 to 9. An example for a tax-based program is the High New Technology Enterprise (HNTE) program which is administered by the Ministry of Science and Technology and was launched in 2008.

Figure A1: China's business R&D expenditures and R&D subsidies



Source: China's National Bureau of Statistics. RMB are in nominal prices.

The firms we observe received direct R&D subsidies from programs administered by national ministries, sub-national agencies (e.g. departments and bureaus at the province, municipal, city, district, and county level), and non-classified agencies (which also provided funds for defense-related projects). Major national R&D programs include the State High-Tech R&D Program (the 863 Program), the National Key Technologies Program, and the State Basic R&D Program (the 973 Program).[43] 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 instead of aiding-the-poor strategy. Support is offered for science and technology oriented research (both basic and applied), the development of new products, the transformation

[43] The 863 Program has the aim to increase China's innovative capacity, is in place since 1986, and 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 has the aim to solve technological problems in social life, is in place since 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, and 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. The maximum duration of funded projects is 5 years, the maximum grant size is 100 to 300 million RMB, and firms receive the payment for the first 2 years together, while the subsequent payment structure is project specific. [In summary, it seems that for projects with a short duration, e.g. 1 to 3 years, a lump sum is paid in year 1. For longer projects, one payment occurs in the first year and subsequent payments later on.]

and improvement of existing technology, university-industry collaborations, and the attraction and training of human capital.

During the time period we study, the locus of China's R&D was relocated from public research institutes towards firms and this shift coincided with the privatization of state-owned firms until the mid-2000s. At that time, public funding generally supported the transition of R&D from public research institutes towards the corporate sector. After 2006, the MLP initiated a more mission-oriented approach and firms preferentially received funding for R&D projects that concurred with the government's more explicit innovation agenda.

Misappropriation

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).[44] In the case of national R&D programs, for instance, funds are transferred from the central government via the Ministry of Finance (primary funder) to various agencies, e.g. the Ministry of Science and Technology, (secondary funders), to various R&D programs and then to the final recipient who performs the R&D project. It is not obvious whether R&D programs are actually abiding by their stated selection rules. Critics point out that relations with government officials are more important than research quality to obtain major grants (Shi and Rao 2010). In many programs the firm proposes to use the grant for R&D in its application, but in practice there is little monitoring or enforcement once the firm received the funds. Effectiveness analysis by government agencies, e.g. the National Center for Science and Technology Evaluation, are in their infancy and confound causal interpretations.[45]

In September 2011 public interest was sparked by media reports stating that around 60% of public research funds were misused for non-research purposes.[46] According to subsequent

[44] Cao et al. (2013) point out that "there is no uniform, national quality control standard, nor is there much exchange of information about projects funded across different agencies. Despite having respective research priorities, these national R&D programs likely overlap, under the administration of different ministries and organizations (even within the same ministry) that may set the same goals […]."

[45] Because the National Center for Science and Technology Evaluation is a subordinate agency of the Ministry of Science and Technology, its political capacity to conduct independent and objective evaluations is questionable.

[46] This statement is quoted from the China Youth Daily (31st August 2011) and was widely reprinted in domestic and international media outlets. The correctness of the figure was subsequently challenged by the Research Propaganda Department of the Chinese Science and Technology Association (September 2011). Nonetheless, in October 2013 Minister Wan Gang still described the state of research funding in China as a "malignant problem" (Renmin Net 2013) and in March 2014 the Communist Party's Central Commission for Discipline Inspection announced that it planned a new round of inspections, including sending a special inspection team to the Ministry of Science and Technology (The Economist 2014). In 2016 the Ministry of

investigations by the Ministry of Science and Technology and the Communist Party's 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.[47] Inspection groups and accounting agencies detected fraud in more than a third of investigated cases (China Communist Party's Central Commission for Discipline Inspection 2015). Reportedly, misappropriating firms sought to maximize public grants by overstating actual project costs and then used R&D subsidies almost entirely for non-research purposes.

According to the newly appointed Director of Guangzhou's Science and Technology Bureau, as for "corruption in the research system, the problem certainly is not the allocation of too much funds but the misappropriation of funds." (Xinhua 2014). Concerned about the role of misappropriation for low returns on the government's R&D investment, the State Council formulated a set of actions to be taken, including to "(i) clearly define missions of national R&D programs, (ii) to separate entities of funding, research, and performance evaluation for the sake of checks and balances and accountability, (iii) to apply different standards to the evaluation of different types of R&D activities, and (iv) to make the reward systems more open and transparent." (Cao et al. 2013). This shows that while the MLP's earlier call for better fund management in R&D programs may have led to improvements, it was still not sufficient. More recent policies by the State Council and the Ministry of Science and Technology explicitly address the fund management and evaluation of research programs.[48]

Additional References

Cao, C., Li, N., Li, X. Liu, L. 2013. Reforming China's S&T system. *Science* 341, 460−462.

China Association for Science and Technology, Research Propaganda Department, 2011. *Statement*. (September 8th 2011).

Science and Technology 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).

[47] In one example, 50 officials from the Science and Technology Bureau of Guangdong Province were investigated for taking bribes from firms in exchange for R&D subsidies (The Economist 2014). The decision to grant subsidies is typically in the hands of individual government officials, rather than peer reviewers and expert panels, and this creates opportunities to accept bribes and extract rents from firms (Fang et al. 2018). 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).

[48] 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).

China Youth Daily, 2011, *Research funding gave birth to how many rich man* (in Chinese), (August 31st 2011).

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

Communist Party's Central Commission for Discipline Inspection, 2015. *Inspection tour points to necessary reform to prevent research funding to disappear in "black holes"* (in Chinese).

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).

People's Daily, 2013. *Scientific research corruption: The Ministry of Communications misappropriated 186 million yuan to pay wage subsidies* (in Chinese), (October 14th 2013).

People's Daily, 2016. *Where did the trillions of research funding go?* (in Chinese), (February 15th 2016).

Shi, Y., Rao, Y. 2010. China's Research Culture. *Science* 329, 1128.

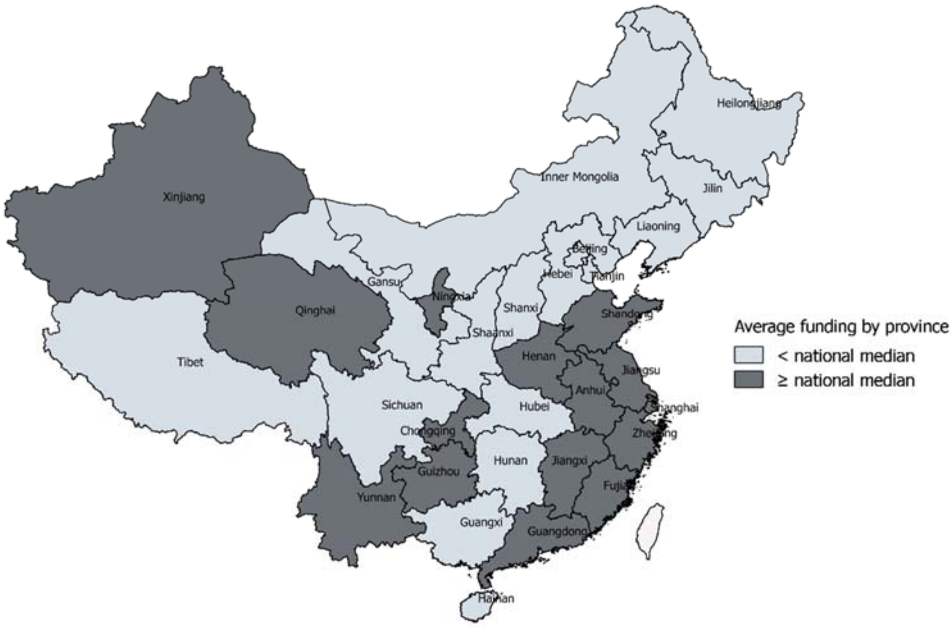
State Council, 2014, Guofa [2014] No. 11, Opinions on the reform and strengthening of the Central Government's scientific research programs and fund management (in Chinese).

The Economist, 2014. *R&D in China, research and embezzlement*, (March 20th 2014).

Xinhua, 2014. *Guangdong science and technology corruption case involves over 50 people, intermediary responsible for pulling corruption* (in Chinese), (April 31st 2014).

Appendix 2: Misappropriation: Stylized Facts and Validation

Figure A2: Funding and misappropriation rates by province



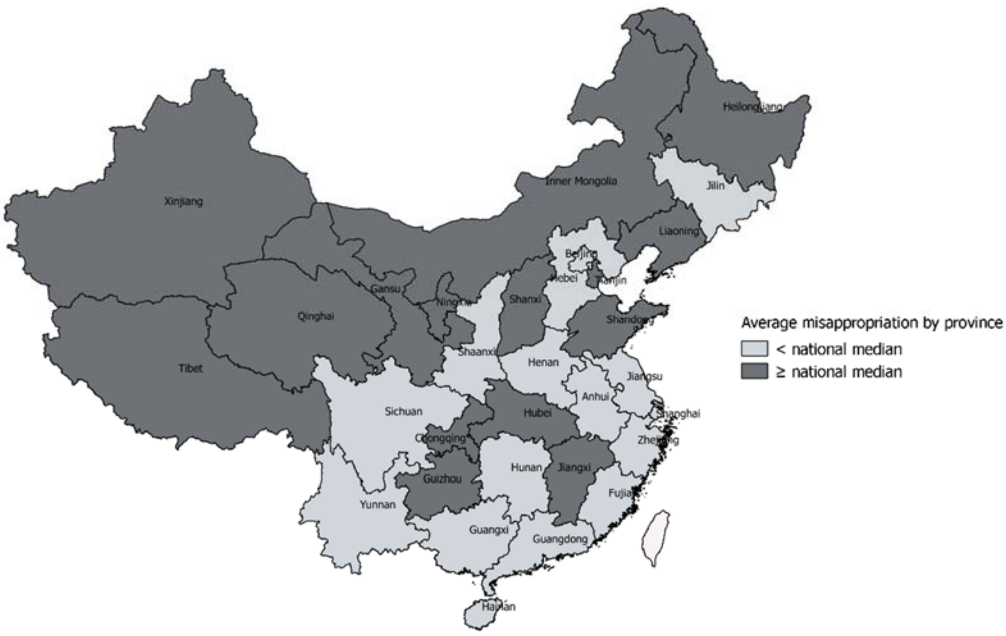


Table A1: R&D expenditures, R&D subsidies and misappropriation by industry

Industry	R&D performers	R&D subsidies	Misappropriation	Obs.
	%	%	%	%
Agriculture	0.189	0.166	0.725	0.019
Mining	0.271	0.115	0.462	0.021
Manufacturing: food & beverages	0.254	0.179	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	0.038
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: R&D performers and recipients of R&D subsidies relative to all firms. Misappropriating firms relative to recipients of R&D subsidies. 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.

Table A2: Likelihood of misappropriation, 2nd stage of Heckprobit model

	Firm attributes	Monitoring	Bureaucratic corruption	R&D project mean subsidy
	(1)	(2)	(3)	(4)
MLP $_{t}$ (0/1)		-0.739***	-0.721***	
		(0.238)	(0.253)	
Mutual fund $_{t}$ (0/1)		-0.218**	-0.225**	
		(0.090)	(0.089)	
Bureaucrats per LME $_{p,t}$ (log)			-0.041	
			(0.085)	
Investigated bureaucrats per LME $_{p,t}$			0.296***	
			(0.101)	
GDP per capita $_{p,t}$ (log)			-5.059***	
			(1.491)	
GDP per capita $_{p,t}$ (log)^2			0.301***	
			(0.094)	
R&D project mean subsidy $_{t}$ (log)				-0.391**
				(0.156)
R&D project mean subsidy $_{t}$ (log)^2				0.020***
				(0.006)
R&D subsidy $_{t}$ (log)	-0.435***	-0.425***	-0.399***	
	(0.136)	(0.136)	(0.134)	
R&D subsidy $_{t}$ (log)^2	0.020***	0.020***	0.019***	
	(0.005)	(0.005)	(0.005)	
Misappropriation $_{t-1}$ (0/1)	0.943***	0.938***	0.954***	0.891***
	(0.087)	(0.087)	(0.086)	(0.095)
R&D experience $_{t-1}$ (0/1)	-0.840***	-0.847***	-0.819***	-0.782***
	(0.094)	(0.093)	(0.093)	(0.105)
Employment $_{t-1}$ (log)	-0.037	-0.041	-0.050	-0.047
	(0.041)	(0.042)	(0.040)	(0.047)
Fixed assets $_{t-1}$ (log)	0.059	0.088	0.110	0.110
	(0.071)	(0.072)	(0.069)	(0.069)
Sales $_{t-1}$ (log)	-0.033	-0.034	-0.028	-0.024
	(0.062)	(0.062)	(0.060)	(0.069)
Age $_{t-1}$ (log)	0.500***	0.501***	0.468***	0.610***
	(0.086)	(0.086)	(0.083)	(0.107)
Patent stock $_{t-1}$ (log)	0.030	0.029	0.012	0.027
	(0.023)	(0.023)	(0.022)	(0.025)
Profitability $_{t-1}$ (0/1)	-0.188**	-0.142	-0.150	-0.132
	(0.094)	(0.094)	(0.093)	(0.104)
Minority state-owned $_{t-1}$ (0/1)	0.081	0.091	0.070	-0.061
	(0.106)	(0.107)	(0.102)	(0.129)
Privatized $_{t-1}$ (0/1)	0.274**	0.248**	0.224**	0.161
	(0.106)	(0.107)	(0.102)	(0.123)
De-novo private $_{t-1}$ (0/1)	-0.059	-0.050	-0.081	-0.151
	(0.103)	(0.105)	(0.098)	(0.120)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	No	Yes
Exclusion restriction 1st stage				
R&D subsidy $_{t-1}$ (0/1)	1.302***	1.303***	1.315***	1.194***
	(0.058)	(0.058)	(0.058)	(0.063)
Rho	0.661	0.656	0.686	0.676
	(0.063)	(0.063)	(0.062)	(0.071)
Observations 1st stage	12951	12951	12951	7175
Observations 2nd stage	2403	2403	2403	1850

Notes: $_{p,t}$ indicates variables measured at the province-year level. All other variables measured at the firm-year level, except MLP, which is measured at the year level. Column (3) includes the log of R&D expenditures per LME at the province-year level and its second polynomial in the first and second stage. Column (4) is restricted to the years 2007 to 2011 and accounts for individual transactions of R&D subsidies per firm-year. Standard errors are clustered at the firm level. ***, ** and * indicate significance at the 1%, 5% and 10% level. The results remain robust when we estimate first and second stages by Heckman and only the second stages by a simple Probit or linear probability model.

Appendix 3: 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 identify 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".

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

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

   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 technical optimization, 优化 technology transformation and improvement, 成果转化 transformation and conversion of scientific and technological achievements, 科技保险 science and technology insurance

   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

   1.3 R&D support policies and programs

   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 Technology-based Small and Medium-sized Enterprise Program

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

   2.1. Patents

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

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

2.3. Technological transformation

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

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

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

3.2. 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

4. 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.

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