Cash on the Table? Imperfect Take-up of Tax Incentives and Firm Investment Behavior

Wei Cui, Jeffrey Hicks, Jing Xing

Cash on the Table? Imperfect Take-up of Tax Incentives and Firm Investment Behavior

Abstract

We investigate whether investment incentives work in less developed countries by exploiting the introduction of accelerated depreciation (AD) for fixed asset investment in China as a natural experiment. In contrast to the large positive impact of similar tax incentives in the U.S. and U.K. found in recent studies, we document that AD was ineffective in stimulating Chinese firms' investment. Further, using confidential corporate tax returns from a large province, we find that firms fail to claim the AD benefit on over 80% of eligible investments. Firms' take-up is significantly influenced by their taxable positions and tax sophistication. Moreover, resources of local tax authorities help improve awareness of the policy. Our study contributes to the understanding of conditions under which tax incentives for investment can be effective.

JEL-Codes: H200, H300.

Keywords: tax incentives, investment, take-up, tax administration.

Wei Cui
Peter A Allard School of Law
University of British Columbia
Vancouver / BC / Canada
cui@allard.ubc.ca

Jeffrey Hicks
Vancouver School of Economics
University of British Columbia
Vancouver / BC / Canada
jeffhicks05@gmail.com

Jing Xing*
Antai College of Economics and Management
Shanghai Jiao Tong University / China
jing.xing@sjtu.edu.cn

*corresponding author

Tax incentives for investment are widely used around the world today. They have enjoyed a particularly long history in developed countries. The U.S., for example, introduced accelerated depreciation (AD) in 1954 and investment tax credits in 1962 (Auerbach, 1982). Yet studies have only recently offered convincing evidence for positive investment responses to such policies (House and Shapiro, 2008; Maffini et al., 2019; Ohrn, 2018; Zwick and Mahon, 2017).[1] Meanwhile, evidence persists that substantial frictions may dampen policy effectiveness. In the U.S., 40-60% of firms did not claim bonus depreciation on eligible investments (Kitchen and Knittel, 2016). Sources of friction recently examined include losses (Edgerton, 2010), accounting rules that counteract tax rules (Edgerton, 2012; Graham et al., 2017), and compliance costs (Kitchen and Knittel, 2016). Such frictions continue to fuel critiques of AD and similar policies (Bazel and Mintz, 2019).

This paper analyzes AD's recent introduction in China. China has seen more capital investment than any other country in the last decade (Chen et al., 2019). Until recently, Chinese tax policy had generally relied on lower tax rates to encourage investment. Facing declining investments and budgetary concerns, the government introduced AD in 2014 as a potentially better-targeted tool. The AD rules were more generous for a subset of industries, delivering benefits comparable to that of U.S. bonus depreciation or U.K. first-year capital allowances. We apply a difference-in-differences (DiD) approach to a data set of confidential corporate tax returns from a large province to study the causal effect of AD on investment. We find that Chinese firms showed limited investment response, in contrast to large responses in the U.S. and the U.K.

Relying on unique features of the tax return data, we uncover one fact that may underlie AD's failure to stimulate investment: firms failed to claim AD on over 80% of eligible investment. We evaluate two complementary explanations for why firms forgo the tax benefits.

---

[1] Hassett and Hubbard (2002) noted that even in the early 1990s, "almost no economist believed that the investment demand elasticity was much different from zero".

One is that widespread losses reduce the value of AD. While we find significant disincentive effects of tax losses on claiming AD, even after accounting for this, the take-up rate of AD remains dismal. The second explanation is that, due to poor policy publicity and a lack of prior exposure, many firms are unaware or fail to grasp the policy's benefits. Consistent with this idea, larger firms and those with more tax expertise were more likely to claim AD.

This raises the question of whether better information transmission can improve take-up. In China, taxpayers rely heavily on tax administrators for knowledge about tax policy (Cui, 2015). Each firm is assigned to a local tax bureau responsible for securing compliance and educating taxpayers. Consistent with the hypothesis that resource-constrained tax bureaus will devote less time to taxpayer education, we find that administrators' workload negatively predicts take-up. Firms further away from their tax bureau are also less likely to claim AD, suggesting that tax administrators are better able to convey policies to nearby firms. A series of robustness checks, placebo tests, and heterogeneity results reinforce our interpretation of the effects as the consequence of information transmission from tax administrators to firms.

Finally, we test whether better-informed firms increased investment because of AD, or simply claimed advantageous deductions *ex post*. We find that the largest 5% of firms both are more likely to claim AD than the rest of the sample and displayed a significant investment response. However, in the rest of the sample, proxies for awareness do not predict greater responsiveness. The ineffectiveness of China's AD policy for the large segment of small and medium-sized firms may thus be over-determined: other factors may also have held back investment. While taxpayer understanding of an incentive is not a sufficient condition for it to influence behavior, it is a necessary condition, and we show that even the satisfaction of this basic condition cannot be assumed.

Our study relates to several strands of literature. First, we offer new evidence for the phenomenon of imperfect take-up of tax benefits and explore its causes. Imperfect take-up of AD benefits received frequent comment in the U.S. until the 1980s (Wales, 1966; Auerbach,

1982). It was also persistent in other corporate tax systems (Kanniainen and Sodersten, 1994) where its causes were not well-understood (Aarbu and Mackie-Mason, 2003; Forsling, 1998; Gronberg, 2015). In implementing similar incentives in emerging economies, it is unsurprising that similar puzzles of non-responsiveness should resurface. Second, we contribute to the literature on the role of tax administration in policy implementation (Dabla-Norris et al., 2020; Goodspeed et al., 2013). In contrast to the prior literature's emphasis on the effect of tax administration on tax evasion, we find that tax administration resources can also increase firm engagement with new policies. Our study thus adds to the discussion of information transmission and expertise in tax policy implementation (Abeler and Jäger, 2015; Chetty et al., 2013; Graham et al., 2017). Finally, we contribute to studies of investment incentives in developing countries (Chen et al., 2019).

Concurrent work by Fan and Liu (2020) also examines AD's introduction in China. They find modestly larger investment effects than our study, but their estimates are still smaller than recent estimates from developed countries. Our study differs from theirs in several significant ways. First and most important, they use a nationwide taxpayer survey and do not observe claimed AD deductions. In comparison, we draw on actual corporate tax returns from one large province, and are able to document low take-up and analyze its drivers. Second, Fan and Liu (2020) use data ending in 2015. As a result, they study only firms eligible for the 2014 AD policy and estimate treatment effects using a single post-treatment year. Our tax return data ends in 2016, which allows us to study both the 2014 and 2015 cohorts of firms eligible for AD—the latter encompasses substantially more firms. Third, their sample consists of mostly large firms, while our sample covers a wide range of the firm size distribution. Fourth, they explore taxpayer non-compliance as an explanation of the muted investment response. We view this as complementary to our analysis, and provide evidence that the effect of information frictions is independent of non-compliance.

The paper proceeds as follows. Section 1 provides background on China's adoption of AD

and compares the policy to investment incentives studied in recent scholarship. Section 2 describes our data. Section 3 analyzes firms' investment responses and Section 4 documents firms' take-up of AD. Section 5 investigates what factors influence take-up of AD and whether firms aware of the policy increased investment. Section 6 concludes by discussing the policy implications of our findings.

1 Accelerated Depreciation: Policy Background

AD is a familiar tax policy tool in developed countries. Section 179 expensing and bonus depreciation studied in the recent U.S. literature (Kitchen and Knittel, 2016; Zwick and Mahon, 2017) are only the latest episodes in a long history of similar incentives.[2] In contrast, AD is a new policy instrument in Chinese taxation. The basic depreciation rules in the 2008 Enterprise Income Tax Law (EITL) are extremely simple and provide only for straight-line depreciation and five fixed asset classes, with asset lives ranging from 3 to 20 years.[3] Before the 2014 policy change, AD was permitted only to correct serious errors in the classification of assets for economic depreciation.[4] Moreover, claims were subject to scrutiny by tax administrators and supposed to be verified by field audits. In fact, before 2014 there was no entry for claiming AD separately from regular depreciation on schedule A105080 of the corporate income tax return, where firms report depreciation and amortization deductions.

The Ministry of Finance announced the new AD policy on September 24, 2014, and the State Administration of Taxation issued more detailed rules in October and November

---

[2]The U.S. first introduced AD in 1954 by allowing taxpayers to use the double declining balance (DDB) and the sum-of-the-year's-digits (SYD) methods, and Section 179 expensing (the most accelerated form of depreciation) became available to small businesses in 1958 (Guenther, 2018). As uniform depreciation schedules were introduced in the 1960s and 1970s, options for choosing shorter useful lives were also offered (Auerbach, 1982). This culminated in the Accelerated Cost Recovery System in 1981. Many other industrialized countries have likewise long used AD as a policy tool (Forsling, 1998).

[3]Unlike the tax depreciation rules, Chinese accounting rules allow DDB and SYD depreciation.

[4]In 2012, firms in software development and integrated circuits industries were permitted to access AD, but these firms were relatively few.

2014. Table 1 illustrates the matrix of policies. Effective from January 1, 2014, all firms regardless of industry could immediately expense newly purchased fixed assets with unit value under 5,000 CNY, and newly purchased instruments and machinery with unit value under 1 million CNY used exclusively for R&D.[5] For purchases with unit values greater than 1 million CNY and used exclusively for R&D, firms could also claim AD. A subset of industries were eligible for even more generous AD. Starting from January 1, 2014, firms in six industries could elect to depreciate any newly purchased fixed assets over 60% of the normal asset life (or, alternatively, use the DDB or SYD method), regardless of the size and purpose of the investment.[6] Moreover, Small and Micro Profit Enterprises (SMPEs)[7] in the six industries could immediately expense investments on instruments and machinery partially used for R&D and with unit values under 1 million CNY. In September 2015, these incentives were extended to four additional industries for asset purchases made on or after January 1st, 2015. All AD policies announced in 2014 and 2015 were introduced as permanent measures and they remain in force as of 2021.[8]

As shown in Table 1, the preferential AD policies targeted primarily certain manufacturing industries. As these AD incentives were not limited in types of fixed assets, they were more generous than those available to other industries.

Figure A1 in the Online Appendix plots the search intensity index for the phrase "accelerated depreciation for fixed assets investment" (in Chinese) from the search engine Baidu during the period 2014/01-2016/12. There is a clear jump in search intensity in the week of the 2014 policy announcement, and very little search activity before, indicating that the

---

[5]1 CNY = 0.16 USD during this period.

[6]Under China's AD rules, taxpayers may choose one method from shortened asset life, DDB, and SYD. We discuss the 40% reduction in asset life as the main tax benefit, although the SYD method may yield faster depreciation for long-lived assets.

[7]SMPEs are firms with (1) total assets under 30 million or 10 million CNY depending on industry, (2) total employees under 100 or 80 depending on industry, and (3) taxable income of less than 300,000 CNY.

[8]In 2019, the Chinese government extended the preferential treatment given to the 2014 and 2015 industries to all manufacturing industries, also on a permanent basis.

2014 policy was likely unexpected. We plot the search intensity index for the phrase "tax reporting" (in Chinese) during the same period for comparison. The two indices co-move, suggesting that search for AD information coincides with return filing.

China's AD policy announced in 2014 and 2015 resembles earlier U.S. policies, such as the shortening of statutory asset lives under the Accelerated Cost Recovery System (ACRS) in 1981[9] and the Modified ACRS (MACRS) in place since 1986. The benefit of AD policy in China is close to that available under MACRS, which permits DDB depreciation.[10] Panel A of Table 2 separately calculates the present value (PV) of depreciation deductions under the regular schedule, under AD using 60% asset life, and under DDB (the rough equivalent of MACRS). Denoting $Z_{t,k}$ as the depreciation deduction in year t, the deductions' PV is:

$$Z_k = Z_{1,k} + \sum_{j=2}^{L_k} \frac{Z_{t,k}}{(1+r)^j} \tag{1}$$

where r is the risk-adjusted discount rate. We use a 7% rate which is used by Zwick and Mahon (2017) and Maffini et al. (2019) and likely a lower bound for Chinese firms. Table 2's middle three columns show that for an asset with regular asset life of 5 years ($L_k = 5$), the difference in PV between the regular and accelerated regimes is \$5.9 on a \$100 investment, which generates \$1.46 in tax savings at the standard 25% tax rate.

The second-to-last column of Table 2 (Panel A) shows the PV of depreciation under the DDB method. For assets of 5 years or longer, the PV of deductions is greater under shortened useful life than under DDB. This allows the calculation in the last column of an implied bonus depreciation factor relative to DDB depreciation.[11] The most favorable

---

[9]Table 1 of Gravelle (1982) shows that the simple average (across asset classes) of the extent by which ACRS shortened asset lives was 43.67%, very close to the 40% reduction under China's AD policy.

[10]However, (i) China does not have a half-year rule and depreciation begins in the month when asset is placed in service; (ii) MACRS allows switch-over from DDB to straight-line depreciation when that is faster, but China does not.

[11]U.S. bonus depreciation allows firms to deduct a percentage of the asset value immediately and the remaining portion according to MACRS (i.e. DDB depreciation).

Chinese AD rules are thus equivalent to allowing first-year bonus depreciation at 21% or 22% (for 5- and 10-year assets), smaller than the 30% bonus factor in U.S. legislation in 2002.

To quantify AD-induced tax savings for our firm sample, we assume that each firm allocates new investment dollars in proportion to its current asset holdings. Denoting $V_{i,k}$ as firm i's holding of type k assets, the average tax value of depreciation deductions is:

$$\tau Z = \frac{1}{N} \sum_i \tau_i \sum_k \frac{V_{i,k}}{\sum_k V_{i,k}} Z_k \tag{2}$$

where τ_i is firm i's tax rate. While the standard tax rate in China is 25%, SMPEs faced 20% and 10% rates during this period, and High and New Technology Enterprises (HNTEs) faced a 15% rate. Using each firm's observed tax rate on their 2013 tax returns, Panel B of Table 2 reports that AD increased Z by $9.5 on a $100 investment, which is greater than the effect of bonus depreciation studied in Zwick and Mahon (2017) and that of the U.K. first-year capital allowances studied in Maffini et al. (2019). However, tax savings (τZ) are attenuated by China's lower statutory rates: on average $1.7 on a $100 investment. The user cost of capital (UCC), $(1 - \tau Z)/(1 - \tau)$, declined by .022, or 2.1 percent following AD, which is on the lower bound of the benefits of bonus depreciation but higher than the U.K. first-year capital allowances. These figures quantify the effect of the preferential AD benefits that were only available to firms in the targeted industries.

Two major tax policy changes during the period we study may also have affected firms' investment decisions. The first is the lowering of the effective corporate tax rate for certain SMPEs. However, in our analysis sample, the prevalence of SMPEs is the same in the treatment and control groups, and therefore would not confound the effect of AD. The second is the expansion of the value added tax (VAT) to the service sector, which may increase the investment incentives of firms in service industries. In our research design, firms in both the treated and control groups are from sectors already subject to the VAT. They

were thus not directly affected by the VAT expansion, and any indirect effect of VAT reform may reasonably be assumed to be common across treated and control groups.

2   Data and Sample Descriptions

Our analyses use a novel administrative data set from a large and prosperous province. The data is extracted from the comprehensive database used by the provincial tax agency for all of its activities, including taxpayer risk assessment and inspections. The data covers the period 2010-2016 and includes de-identified information for firms of all sizes and sectors. It contains a large number of variables from the annual corporate income tax return, income statements and balance sheets, as well as a taxpayer registry. The data is relied on by tax administrators and backed by genuine legal obligations borne by taxpayers.

Specifically, our data set includes entries on Schedule A105080 of the corporate income tax return where taxpayers report asset-specific AD. For each of five different asset classes, there are four fields that report, respectively, the sum and the individual values of three forms of AD: (1) immediate expensing; (2) AD introduced in 2014/5; and (3) the minor forms of AD in place prior to 2014/5. These fields were not available on tax returns before 2014. Schedule A105080 also reports a firm's fixed asset stock by type, which allows us to calculate asset compositions. Moreover, information from the main Schedule A100000 provides observations of firms' current tax loss positions, whether they have loss carry forwards, and their statutory tax rates. The data also includes financial statement variables, including the stock of fixed assets net of accounting depreciation.

Another novel feature of our data is that it provides a 9-digit geographic code (the neighborhood level) for each taxpayer, and identifies the tax bureau directly in charge of each firm.[12] We manually collect the physical address of each tax bureau. Using the 9-digit

---

[12]During the period we study, the general configuration of tax bureaus remained relatively stable.

area code to proxy for each firm's location and combining it with the location of the tax bureau, we calculate the geographic distance between each firm and its tax bureau. We also manually search the websites of each tax bureau to obtain information on staff size.[13] Pooling all this information together, we can analyze how features of tax administration influence the effectiveness of AD.

We obtain from the tax returns 4,547 firms in the 2014 targeted industries, 17,721 firms in the 2015 targeted industries, and 8,419 firms in non-targeted manufacturing industries, for which we observe non-missing necessary financial information.

Table 3 provides summary statistics of key variables in our analyses for each group as of 2013. Firms in the 2015 targeted industries are similar to those in the non-targeted industries in most dimensions. Firms in the 2014 targeted industries appear more different: as they are more "high tech" than traditional manufacturing industries (Table 1), they are younger, smaller, faster growing, more likely to be an HNTE, hold more cash and are less likely to borrow[14]. For all three groups, a large proportion of firms are loss-making and carry a substantial amount of loss carry-forward.

3   Did AD Stimulate Investment?

3.1   Empirical Strategy

We start by examining whether AD's introduction stimulated investment. Since the AD provisions were more generous for the targeted than for the non-targeted industries, policy variation arises as long as investment is not mainly driven by small purchases or purchases of equipment used exclusively for R&D.[15] Denote $y_{i,t}$ as a measure of firm i's investment in

---

[13]The websites do not disclose historical information on staff size by year. But assuming staff size to be highly persistent over time, the 2018 information will accurately proxy bureau resources in 2014-2016.

[14]As proxied by whether they claim deductions for interest payments.

[15]Few firms claimed R&D super deductions in the tax returns, which indicates infrequent R&D investment.

year t, D_i as an indicator for being in a targeted industry, and $Post_t$ as an indicator for years after the policy implementation. Firms in 2014 targeted industries are not in the control group for the 2015 treatment, nor vice versa.[16] The baseline DiD specification is:

$$y_{i,t} = \alpha_t + \alpha_i + \beta \times Post_t \times D_i + \epsilon_{i,t,g} \tag{3}$$

The coefficient of interest, β, captures the difference in $y_{i,t}$ between targeted and non-targeted firms under the AD regime, relative to the difference before AD's implementation. We control for firm and year fixed effects (α_i, α_t).

Our outcome $y_{i,t}$ is based on the stock of fixed assets net of accounting depreciation. To relate changes in asset stock to investment expenditure, we assume that assets accumulate according to the following model:

$$K_t = I_t + (1 - \gamma)(K_{t-1} - S_t) \tag{4}$$

K_t denotes fixed assets net of accounting depreciation at the end of year t, I_t denotes new asset purchases made in t, γ is the depreciation rate on existing assets, and S_t denotes asset dispositions in t. Our first outcome variable is $Ln(K_t) - Ln(K_{t-1}) \approx \frac{K_t - K_{t-1}}{K_{t-1}} = \frac{I_t}{K_{t-1}} - \frac{(1-\gamma)S_t}{K_{t-1}} - \gamma$, which differs from I_t/K_{t-1} in two ways. First, the change in asset stock incorporates accounting depreciation γ. However, firm fixed effects will absorb γ if it is constant over time for each firm.[17] Second, our measure incorporates S_t (which we do not observe). If older assets already subject to substantial depreciation are more likely to be disposed of, S_t should be small. Under these assumptions, the estimate of β, using $Ln(K_t) - Ln(K_{t-1})$ as the outcome variable, approximates the effect on I_t/K_{t-1}. For robustness, we present results using $Ln(K_t)$ as the outcome, which measures asset accumulation rather than

---

[16] For the 2014 targeted industries, the post period is 2014 to 2016. For the 2015 targeted industries, the post period is 2015 to 2016.

[17] Tax depreciation is distinct from accounting depreciation, so the AD policy will not directly affect γ.

asset growth. This also gives us one more year before treatment that helps with testing for parallel pre-treatment trends, especially for the 2014 reform.

Table 3 reveals differences in observed characteristics between targeted and non-targeted firms, particularly the 2014 group.[18] We therefore construct a control group of firms more comparable with targeted firms using a two-step matching procedure. First, we restrict the control group to non-targeted manufacturing firms. Second, using Coarsened Exact Matching (CEM) (Iacus et al., 2011), we match targeted and non-targeted firms on their i) age, ii) profit margin ($\frac{\text{Total Profit}}{\text{Sales}}$), iii) total assets, and iv) revenue growth, all as of 2013.[19] By matching on pre-AD characteristics, we lessen the likelihood that treatment and control firms experience different shocks during the policy. Table A1 in the Online Appendix reports summary statistics for the matched sets of firms both before and after matching. Before matching, the targeted and non-targeted groups are statistically different in all the matching variables in 2013. After matching, firms become more similar on both matched and unmatched characteristics.

3.2 Baseline Results

Figure 1 reports the evolution of $Ln(K_t) - Ln(K_{t-1})$ and $Ln(K_t)$ for the matched treated and control groups. We estimate a dynamic version of equation (3) as $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 2013} \beta_s \times \mathbb{1}\{t = s\} \times D_i + \epsilon_{i,t,g}$. We then plot α_t for the control group and $\alpha_t + \beta_t$ for the treatment group and the associated 95% confidence intervals.[20] For both treatment cohorts, asset growth declined during 2013-2016 consistent with China's declining GDP growth during

---

[18]Firm fixed effects α_i will account for level shifts in $y_{i,t}$ caused by these differences. but not account for differences in how firms respond to AD policy. The ideal experiment assigns different AD generosity to firms with the same investment responsiveness.

[19]Matching on age selects control firms at similar points in their life-cycle; on profit margin, firms of similar productivity; total asset stock, firms of similar size; and revenue growth, firms similar in growth trajectories.

[20]The coefficient β_t represents the difference between treated and control in year t relative to the baseline difference in 2013. Adding α_t allows us to illustrates level trends as opposed to just relative changes.

this period.[21] Pre-treatment trends are also parallel, as best illustrated by Panels C and D.

Table 4 presents estimates of equation (3), with all coefficients scaled by 100. All regressions employ DFL re-weighting (DiNardo et al., 1996) based on total sales, a common approach for controlling for changes in the composition of firms over time (Yagan, 2015; Zwick and Mahon, 2017).[22] Standard errors are clustered at the three-digit industry level to account for within-industry correlations and serial correlation over time within a firm. Columns (1) and (2) show that AD did not lead to higher asset growth for the treated group: the point estimates in both are negative and insignificant. In columns (3) and (4), point estimates for the effect of AD on $Ln(K_t)$ are positive but statistically insignificant.

The AD provisions available only to the targeted industries reduced the UCC by 2.1 percent. We derive an upper bound on the elasticity of the investment rate (I_t/K_{t-1}) as $\eta = \frac{\frac{\hat{\beta}_u}{I/K_{Pre}}}{.021}$, where $\hat{\beta}_u$ is the upper bound on the 90% confidence interval of β and $\frac{I}{K}_{Pre}$ is the average investment rate before the policy changes in the targeted industries. In section 3.1, we showed that $\frac{I_t}{K_{t-1}} = \frac{K_t - K_{t-1}}{K_{t-1}} + \frac{(1-\gamma)S_t}{K_{t-1}} + \gamma$. We observe $\frac{K_t - K_{t-1}}{K_{t-1}}$ in our data: .069 and .049 for the 2014 and 2015 groups, respectively, in the pre-AD period (Table 4). We do not observe $\frac{(1-\gamma)S_t}{K_{t-1}} + \gamma$, and so rely on an estimate from Qiu and Wan (2019) of .095 for the same years.[23] This allows us to estimate $\frac{I}{K}_{Pre}$: .164 and .144 for the 2014 and 2015 cohorts, respectively. The resulting 90% upper bounds on η are .1 and 6.25 for the 2014 and 2015 policies, respectively. These are below the elasticity estimates from Maffini et al. (2019) (8.3-9.9) and Ohrn (2018) (6.5).[24] Alternatively, Online Appendix A.3 uses a two-stage least

---

[21]China's GDP growth declined from 7.86% in 2012 to 6.85% in 2016.

[22]Online Appendix A.2 details the construction of the weights. The results are highly similar without re-weighting.

[23]Qiu and Wan (2019) use aggregate data from the National Bureau of Statistics (NBS) to calculate depreciation rates. We average their annual rates for the period 2010 to 2013. The 9.5% reported by Qiu and Wan (2019) is similar to the 9% assumed by Liu et al. (2019).

[24]We can also compute the elasticity of capital stock with respect to the UCC changes. Based on the point estimates in columns 3-4, this elasticity is 0.57 for the 2014 reform, and 0.32 for the 2015 reform—considerably lower than the theoretical elasticity of 1 (Hall and Jorgenson, 1967), and also lower than what recent literature finds (Bond and Xing, 2015).

squares approach to estimate the elasticities, yielding very similar results.

We conduct a series of robustness checks in the Online Appendix A.4. The first four columns in Table A3 report results using alternative outcome variables: columns (1) and (2) use the change in the capital stock, $K_t - K_{t-1}$, normalized by sales in the base period, while columns (3) and (4) use $Ln(K_t) - Ln(K_{t-1})$ with K measured at original costs from the tax return rather than net-of-depreciation from the balance sheet. Columns (5) and (6) control for linear time trends at the three digit-industry level in our baseline model. We exclude firms that claimed R&D deductions in columns (7) and (8), and remove observations where the firm was eligible for SMPE status in columns (9) and (10). In Figure A2, we show dynamic DiD estimates from alternative specification and sample choices, including: (1) removing DFL re-weighting, (2) removing both matching and DFL re-weighting, and (3) estimating based on a balanced sample. Finally, Online Appendix A.5 shows dynamic DiD estimates when the 2014 and 2015 treated industries are combined into a single treatment group. In all these additional checks, we continue to find little impact of AD on firm asset growth.

3.3 Heterogeneity Across Firms and Asset Types

To explore firm heterogeneity in responses, we first execute a series of split sample regressions. The sample splits are based on firm characteristics measured as of 2013. We then estimate equation (3) for each sub-sample separately, using $Ln(K_t) - Ln(K_{t-1})$ as the outcome.[25] We do this for eight firm characteristics, for both the 2014 and 2015 targeted industries, resulting in 32 estimates of β. Panel A of Figure 2 plots each estimate, the associated 95% confidence intervals, and the p-values from the test of the treatment effects being the same across the sub-samples.

The first four heterogeneity cuts are based on whether (i) the firm is in tax losses, (ii)

---

[25]Figure A4 in the Online Appendix reports results using $Ln(K_t)$ as the outcome variable—the results are similar.

the average life of the firm's asset portfolio is above or below the sample median, (iii) the firm's cash-to-revenue ratio is above or below the sample median, and (iv) the firm claimed interest deductions on their tax return (the last two variables proxy for cash and financing constraints). Among the 2014 treatment group, the only statistically relevant finding is that firms that claimed interest expenses responded more than those did not. Among the 2015 group, responsiveness is not different across sub-samples.

Since AD is more beneficial for long-lived assets, it should stimulate investment in long-lived assets more. The null result in Panel A of Figure 2 regarding asset life is thus surprising. To further explore heterogeneity by asset life, we construct $Ln(K_{i,k,t}) - Ln(K_{i,k,t-1})$ using fixed assets measured at historical cost for each of the five asset classes k and estimate equation 3 separately for each k.[26] Panel B of Figure 2 plots the estimated βs and confidence intervals. We do not detect significant heterogeneity across asset types.

A second set of heterogeneity cuts in Panel A of Figure 2 examines proxies for firm tax sophistication (firm size and HNTE status) and access to tax administrators. We discuss the interpretation of these characteristics further in Section 5. There is little consistent heterogeneity based on these sample cuts, with the exception of firm size: for both treatment groups, the estimated treatment effect appears to be bigger for larger firms.

To further explore size heterogeneity, we use pre-treatment total assets to split firms into quartiles, and estimate the treatment effect for each quartile. Figure 3 reports the results. Estimated treatment effects only weakly increase with firm size.[27] However, when we narrow the sample further to the largest 5% firms, we obtain a strongly positive treatment effect for the 2015 treated group and a positive but imprecise estimate for the 2014 group (less than 130 firms belong to the top 5% sample for the 2014 group). Appendix A.7 discusses these results in more detail, including cautions that need to be exercised in interpreting them as

---

[26]We do not observe the net-of-depreciation value at the asset class level.

[27]Figure A5 shows very similar results when splitting by revenue rather than total assets.

causal evidence of investment response. We note that in both Zwick and Mahon (2017) and Maffini et al. (2019), smaller firms show greater responsiveness to AD.

3.4 External Validity

To ensure that the province we study is not an outlier, we present trends using a nationally representative sample of firms from the Orbis database. Orbis collects balance sheet information from Chinese firms, including the value of fixed assets net of accounting depreciation. Panels A and B of Figure A8 plot the time trends of $Ln(K_t) - Ln(K_{t-1})$ for the targeted industries and non-targeted manufacturing firms using Orbis data. First, the nationwide data show a downward trend in $Ln(K_t) - Ln(K_{t-1})$ leading up to the treatment, mimicking the provincial trends in Figure 1 observed in our data. Second, neither figure exhibits a detectable increase in investment among firms in the targeted industries.

In concurrent work, Fan and Liu (2020) find modest treatment effects for the 2014 AD policy. Appendix A.9 discusses in detail how their analysis differs from ours. There are four potentially significant sources of difference. First, they use a national survey of firms which comprises mostly of large and medium-size firms: the average revenue of their sample is approximately the average revenue of the top quartile of our sample. Even within their sample of already-large firms, they find that larger firms responded more. Second, their data ends in 2015, leaving only one post-treatment year. Figure A9 shows that, in our data, treatment effects are larger if we consider only 2015, rather than 2015-2016. Third, their survey data directly reports capital expenditure. In Appendix A.9, we develop an approximate correspondence between the treatment effects on their outcome, $Ln(I_t)$, and ours using $Ln(K_t)$. In that framework, their estimates are within our confidence intervals. Fourth, there are differences in empirical specifications. Most importantly, Fan and Liu (2020) include service industries in the control group (we include only manufacturing firms), which produces slightly larger treatment effects in some specifications in our data.

15

4 Take-up of Accelerated Depreciation

Results from the previous section suggest that AD did not significantly stimulate invest-ment among targeted firms. Using tax returns, we extend the investigation to the take-up of AD. We start by documenting trends in claims of AD. Denote $AD_{i,k,t}$ as the amount of AD deduction claimed by firm i for asset class k in year t. Panel A of Figure 4 plots the fraction of firms with positive AD amounts ($\sum_k AD_{i,k,t} > 0$). At the peak, fewer than 20% of firms reported positive AD amounts. Claims spiked in 2015, consistent with the policy announcement occurring late in 2014. Figure B1 in the Online Appendix shows that over 80% of firms with positive AD amounts in 2015 were first-time claimers. Targeted firms have higher claim rates than control firms in all years, consistent with the broader scope of eligible assets. Table B1 further shows claim rates in each targeted industry. The claim rate was highest in manufacturers of i) instruments, (ii) computer, communications and other electronic equipment, and (iii) special equipment–all belonging to the 2014 targeted group.

Panel B of Figure 4 shows the 1st to 3rd quartile values of AD deductions. The main tax advantage offered to non-targeted firms was immediate expensing on purchases with unit value of 5,000 CNY or less. Consequently, the median amount claimed by these firms is below 10,000 CNY in all years. The values of AD claimed among the 2014 and 2015 targeted firms were generally higher, especially in 2016.

Panels C and D of Figure 4 consider the *take-up* of AD, defined as the likelihood of claiming AD conditional on making eligible purchases. We do this at the asset class level. We define *claiming* more narrowly here as a firm reporting a positive year-over-year increase in the AD amount for a given asset class k: $C_{i,k,t} = 1$ if $AD_{i,k,t} > AD_{i,k,t-1}$.[28] We also

---

[28]We choose this approach, rather than $AD_{i,k,t} > 0$, because AD reported in year t may be attributable to asset purchases from previous periods. Both methods are imperfect. In practice, all AD claims in 2014 and the majority in 2015 are first-time claims, therefore the definitions are equivalent in those years. Even in 2016, around 50% of AD claimers did not have positive AD deductions in the previous year. In line with this, we obtain similar results when claiming is defined based on $AD_{i,k,t} > 0$.

define *investing* narrowly as a year-over-year increase in the stock of asset type k measured at original cost: $I_{i,k,t} = 1$ if $K_{i,k,t} > K_{i,k,t-1}$.[29] Take-up is claiming conditional on investing: $C_{i,k,t} = 1 \mid I_{i,k,t} = 1$. Panel C plots the average take-up rate in each year averaged over the five asset classes. Firms in the targeted industries could, in theory, have 100% take-up. However, the take-up rate ranges from 2.5% to 17.5%. Panel D shows that the low take-up rate is observed for all asset classes. Buildings and structures have the lowest take-up rates, despite being the longest-lived assets.

Imperfect take-up is present for both large-scale and small-scale investment. Figure B2 in the Online Appendix plots the distribution of $K_{i,k,t} - K_{i,k,t-1}$ separately for $C_{i,k,t} = 1$ and $C_{i,k,t} = 0$. The distributions are similar but with a modest rightward shift for the $C_{i,k,t} = 1$ distribution. In Table B2, we calculate the tax savings foregone on unclaimed investment. Total tax savings foregone by the 2014 and 2015 treatment groups, respectively, is 176 million CNY (30 million USD) and 543 million CNY (90 million USD). The average forgone savings per firm, for the 2014 and 2015 groups respectively, is 75,000 CNY and 57,000 CNY (12,000 and 9,120 USD). Table B2 also indicates that the total tax revenue cost to the government due to actual AD claims is 26 million and 48 million CNY respectively.

Imperfect take-up of AD is well-documented in developed countries, though explanations for it are offered only in passing.[30] It is especially worth noting that China's AD rules are simple by comparison to advanced economies. Chinese tax law still contains neither rules for

---

[29]Assets at the type k level are only observed at historical cost in our data. The resulting measure of investment is a conservative indicator for new purchases since a firm that simultaneously purchases a new asset and disposes of an old one may report a year-over-year decrease in the asset stock. Our benchmark estimations in the following section exclude firms that had negative "investment" but claimed AD.

[30]Wales (1966) suggests that taxpayer learning, the presence of losses or insufficient net income may all explain observed imperfect utilization of AD benefits in the US in 1950s. Many small businesses also failed to choose AD under the 1971 US Asset Depreciation Range system: accounting complexity and prior taxpayer non-compliance were suggested as possible causes (Auerbach, 1982). Subsequent U.S. literature tended to emphasize incentives, rather than knowledge, as constraints on AD's effectiveness. For example, Gordon et al. (1987) document that U.S. individual investors adopt straight-line depreciation for 60% of their investments in structures and forgo AD benefits. They suggest that incentives for churning real property may explain the under-utilization of AD.

recapture of excessive deductions nor "anti-churning" rules–all rules that the U.S. had put in place by the time ACRS was adopted. Nor is there any book-tax conformity requirement that would hinder adopting tax AD. It is thus unlikely that unique aspects of Chinese law explain low AD take-up. Studies of imperfect take-up that examine compliance costs and information transmission are more likely to be germane to understanding Chinese firms' response. Kitchen and Knittel (2016) show that firms in U.S. states where tax returns are less harmonized with the Internal Revenue Code display lower take-up of bonus depreciation, suggesting that compliance costs hinder take-up. Zwick (2021), examining a different tax benefit (loss carryback), shows that take-up is affected by the professional advice firms receive. We will examine in the next section an important but often-neglected mechanism of information transmission–tax administrators–as a potential determinant of take-up.

5   What Affects Take-up?

We examine two complementary explanations of low take-up. One is that widespread losses, and the tax law's treatment of losses, render the benefit of AD small or even negative relative to regular depreciation. The other is that, due to poor policy publicity and a lack of prior exposure, firms were unaware or did not understand AD's benefits. We disentangle these narratives by examining how loss positions, firm tax sophistication, and tax administration influence take-up. We focus on targeted industries only. Control industries could only claim AD for a subset of investments, so their limited take-up is less informative.

5.1   Firm-Level Characteristics

We first examine how take-up correlates with firms' taxable income before any AD deduction is applied. Panel A of Figure 5 plots the take-up rate of AD against uncensored taxable

income normalized by total assets.[31] We normalize taxable income by total asset stock to control for the effect of firm size on take-up. Take-up rates are less than 5% for firms in taxable loss, then begin to rise steeply with income after the zero-income benchmark. This suggests a strong negative effect of tax losses on take-up. The upward slope when taxable income is positive can be explained if normalizing by total assets is insufficient to reduce the size effect. A firm with more taxable income is also more likely to claim AD because it is less likely to have taxable losses in the future.

The present value of AD is likely to be smaller for firms with larger stock of unclaimed tax losses as they may take a longer time to become profitable. We construct the stock of unused tax losses for each firm by beginning with observed tax losses in 2010, the first period of our data, and summing the accumulation of losses thereafter. Panel B of Figure 5 plots the lagged stock of taxable losses against the take-up rate of AD, and shows a negative correlation between the stock of unused taxable losses and take-up.[32]

Low take-up may also reflect non-trivial compliance costs of claiming. Such costs can arise when firms must learn about new tax incentives, and when general guidance from tax authorities is limited. Firms with greater tax sophistication may be advantaged in coping with this cost. One crude measure of tax sophistication is the size of a firm–larger firms are more likely to employ dedicated accountants and tax experts. Panel C of Figure 5 shows a strong positive correlation between the firm's total assets and take-up. Alternative measures of firm size such as business revenue yield similar results.[33]

Prior experience with complex tax incentives may also enhance firms' tax sophistication.

---

[31]We define uncensored taxable income as total book profit reported on the tax return plus the net value of book-tax adjustments (excluding AD adjustments) minus exempt income and claimed loss carry-forwards.

[32]In China, tax losses generally can be carried forward for only five years and not back, a much less generous treatment than in other countries. This suppresses the incentives of firms with large unused tax losses to claim AD.

[33]As a by-product of large firms better availing themselves of AD's tax savings than small firms, the policy creates a size-based benefit. Figure B3 in the Online Appendix plots the average ratio of tax savings from AD over firm revenue, across the firm size distribution. Tax savings among the largest firms is almost double that of the smallest firms.

In China, various tax incentives given to firms with HNTE status pre-date AD policy (Chen et al., 2020). To obtain HNTE status, firms need not only to satisfy requirements based on R&D intensity, but also to comply with extensive procedures for claiming tax benefits. Such firms are likely to have invested in the capacity to claim tax preferences and thereby better-positioned to take advantage of new incentives. Consistent with this conjecture, Figure 5 Panel D shows that HNTEs are twice as likely to claim AD than non-HNTE firms.[34]

Finally, the absolute value of claiming AD grows with the size of investment. In the presence of fixed costs of claiming AD, the probability of claiming should rise with investment size. Consistent with this hypothesis, as shown previously, Figure B2 in the Online Appendix shows a modest positive relationship between investment size and take-up.

We jointly estimate how these covariates $X_{i,t}$ predict take-up. We restrict to firms with $I_{i,k,t} = 1$, estimate a probit model of $C_{i,k,t}$, and present estimates of the average marginal effects: $\frac{\partial P(C_{i,k,t}=1|I_{i,k,t}=1)}{\partial X_{i,t}}$. The goal is to examine the predictors of take-up in a reduced-form manner without fully-specifying the dynamic nature of investment and claiming decisions. As an extension, we estimate a selection model that accounts for some forms of selection bias that could arise due to correlated propensities to both invest and claim AD.

Table 5 presents the results. All columns report average marginal effects scaled by 100.[35] Columns (1) and (2) report the estimated effects of each firm characteristic on the conditional probability of claiming. Column (1) includes two-digit industry, year, and asset class fixed effects. Column (2) adds prefecture fixed effects to control for geographic variation. Confirming the graphical evidence, losses negatively predict take-up. Based on column (2), firms in current year loss positions (before AD deductions) are 2.85 percentage points (p.p), or 45 percent (2.85/6.36), less likely to claim AD on eligible investments. The stock of unused

---

[34]While HNTE firms face a lower statutory rate, they are also frequent users of a 50% super-deduction for R&D related expenses. Such super-deduction would raise the value of one additional dollar of AD to a level similar to other taxpayers.

[35]Standard errors are computed using the delta-method and clustered at the two-digit industry level.

tax losses at the end of the previous year also negatively correlates with take-up probability.

Results regarding proxies for tax sophistication also confirm the graphical evidence. Column (2) indicate that a 100% increase total asset stock is associated with a 0.63 p.p. increase in the probability of claiming on a base claiming rate of 6.36%. HNTE firms are 1.8 p.p (28%) more likely to claim AD relative to non-HNTE firms. A doubling of the size of eligible investment $(K_{i,k,t} - K_{i,k,t-1})$ increases the take-up probability by .09 p.p.[36]

By restricting to firms with $I_{i,k,t} = 1$, selection bias could arise if the idiosyncratic propensity to invest is correlated with the propensity to claim AD. In column (3), we examine whether accounting for such selection matters by implementing the standard two-stage choice model (Heckman, 1979)(discussed in Online Appendix C). If there is selection bias, the estimates presented in column (3) should differ from those in column (2). However, the point estimates are mostly unchanged.

5.2   The Role of Tax Administration

AD was a new policy in China and less straightforward than tax rate cuts. Consequently, the policy may have lacked salience for most firms. In this context, while firms' own tax sophistication affect the take-up of AD, tax administrators could also play a crucial role. In addition to ensuring taxpayer compliance, China's front-line tax administrators are responsible for informing taxpayers of the content of tax law (Cui, 2015). Indeed, most taxpayers rely on officials in nearby tax offices instead of third-party professionals to learn about rules applicable to their businesses. Major tax policy changes are commonly accomplished by campaigns to spread knowledge of the changes among taxpayers. Disseminating information about policies that are politically important are often a part of civil servants' performance

---

[36]Since our measure of investment based on changes in asset stock is conservative, the estimated effect of investment size on take-up is likely attenuated downwards.

metrics. In contrast, general media coverage of tax policies is weak.[37] For all these reasons, tax administrators role in transmitting policy information takes on singular importance.

We consider two aspects in tax administration that potentially bear on policy information transmission: geographical distance between firms and their assigned tax bureaus, and the staffing level of each bureau. First, information transmission may be weakened if firms and tax administrators are distant from each other. Proximity to regional tax offices has been used in research in other countries as a proxy for the degree of information transmission from tax administrators to firms (McKenzie and Seynabou Sakho, 2010). Though Chinese tax agencies began to make extensive use of social media and smart phone apps to publicize policies in recent years, such use was not yet prevalent in the years we study.[38] This leads to our first hypothesis that being located closer to the local tax bureau increases the awareness and therefore use of tax incentives. In our data, each firm is assigned to the jurisdiction of a tax bureau. Utilizing the area codes of the firm and its local tax bureau, we calculate the geographic distance for each firm-bureau pair.[39] Figure B4 in the Online Appendix presents the histogram of measured distances in our sample, which displays rich variation across firms.

Second, information delivery is likely to be constrained by the resources of local tax bureaus. We obtained data on the number of staff for a subset of local tax bureaus, available from their official websites. We divide this number into the number of firms assigned to the bureau in 2013. A higher ratio of firms to bureau staff in a given jurisdiction suggests a greater workload for administrators and weaker capacity to provide guidance to firms. This, in turn, may predict lower take-up.

Panels E and F of Figure 5 plot the correlation between take-up and the two tax bureau

---

[37]The adoption of tax policy in China largely bypasses legislatures and the legal system, and consequently lacks the venues of publicity and outreach that tax policies possess in democratic countries.

[38]Even today technologies often complement, instead of substitute for, in-person taxpayer services.

[39]We observe the firm's bureau assignment as of 2017. Firms generally do not change this assignment. However, if some bureaus changed location between 2014 and 2016, then the 2017 information will contain measurement error.

characteristics, respectively. Consistent with both hypotheses, firms further away from their bureaus have lower claim rates, as do firms assigned to lower-resourced tax bureaus. Table 6 presents estimates of the average marginal effects of these two proxies. All columns control for the firm characteristics in Table 5 and two-digit industry fixed effects. We also control for prefecture fixed effects to account for unobservable geographic factors that affect take-up. Column (3), our baseline specification, shows that a 100% increase in staffing predicts a 1.89 p.p., or a 27 percent (1.89/6.85), increase in take-up. A 100% increase in distance predicts a 0.76 p.p, or 11 percent, decrease in take-up.

A firm's distance to its local tax bureau may correlate with the proximity to other important government institutions. As a placebo test, we examine whether take-up correlates with the firm's distance to the nearest district or county People's Government office (the executive branch's central office). Column (4) shows that the estimated effect of this placebo is indistinguishable from zero. This corroborates the view of local tax bureaus as the relevant source of tax-related information. Additionally, in Column (5), we include tax bureau fixed effects such that the effect of distance is identified solely on within-bureau variation. The estimated effect is largely unchanged.

One concern is that less tax-compliant firms may be less likely to respond to tax incentives (Fan and Liu, 2020) and may also locate further from tax bureaus. In column (6), we include three indicators of firms' non-compliance tendency: the gap between the firm's statutory tax rate (STR) and their measured effective tax rate (ETR), the ratio of fixed assets to total assets, and the ratio of profits to business revenue. Firms with much lower ETR than STR may be less compliant; firms with more fixed assets may be more compliant since it is easier to verify their assets (Gordon and Li, 2009); and more profitable firms may be more compliant (Cai and Liu, 2009). While these are only rough proxies for tax compliance, column (6) shows that none predicts take-up. Nor does including them in the regression materially

23

change the estimated coefficients on the tax administration variables.[40]

Another possible confounding factor is that firms located in industrial parks, or in urban areas, may have more opportunities to learn about AD from other firms. Tax offices may also be set up specifically in such locations to deal with greater demand. In column (6), we include indicators for whether a firm is located in an industrial park or in an urban area. Neither are significantly associated with take-up.

Finally, firms with lower tax sophistication should benefit more from increased access to tax administrators, while firms with greater tax sophistication should be less affected. To test this, column (7) interacts the tax administration variables with an indicator for HNTE status. We expect HNTE firms to be less reliant upon tax administrators for understanding AD. Indeed we find the marginal effects of both bureau staff resources and firm distance to tax bureaus attenuated towards zero for HNTE firms.

5.3   Local Accounting Resources

External tax professionals help firms optimize tax strategies (Zwick, 2021). We use two measures of local accounting resources to proxy firms' access to third-party tax expertise. The first is the number of accountants among the working population in each district.[41] The second is the ratio of accounting firms to the total number of firms in each district. We obtain information on accounting firms and their physical addresses from a certified online platform that provides information on companies' registration and credit record. Table B4 shows the effects of each measure on take-up. The estimated coefficients on accountants per worker and on accounting firm density in the same area are both statistically indistinguishable from 0. Somewhat puzzling, there is a negative and significant association between take-up and the ratio of accountants per firm in our tax data. Overall, these results suggest that external

---

[40]In Table B3 of the Online Appendix, we do not find evidence that firms located further away from tax offices, or those facing a higher ratio of firms to tax bureau staffs, are more likely to avoid tax.

[41]We calculate these based on the 2010 Population Census conducted by the National Statistics Bureau.

accounting resources are less important in China than in the US. This further highlights the role of local tax administration in information transmission.

5.4   Did Knowledge Lead to Greater Investment?

Our analysis suggests that information frictions were a primary driver of low take-up. If such frictions were ameliorated, would investment responsiveness have increased? One way to address this question is to test whether traits that make firms better informed also predict greater investment responsiveness (β from equation (3)). Returning to Panel A of Figure 2, the last eight rows present split-sample estimates of β for four proxies of policy awareness: HNTE status, firm size, distance to a tax bureau, and tax bureau resource level. Neither tax administration measure predicts greater investment responsiveness, nor does HNTE status. There is, however, weak evidence that larger firms responded more positively. As reported in Section 3.3, only the largest 5% of firms in the targeted industries appear to have increased investment significantly, relative to comparable control firms. These firms are also more likely to claim AD (See Figure A7 in the Online Appendix). Thus, AD appears to have been more effective for this small group of very large firms.

We can also examine whether, regardless of size, firms that claimed AD on their tax returns experienced greater asset growth relative to non-claimers. Claimers overcame information frictions at least during tax filing. But firms that invested due to non-tax reasons may also be more likely to claim AD whereas firms without investment, by definition, cannot. This creates a positive selection bias, making the comparison descriptive rather than causal. We estimate the following DiD model:

$$y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 2013} \theta_s \times \mathbb{1}\{t = s\} \times C_i + \epsilon_{i,t,g} \tag{5}$$

where C_i is one if the firm reported positive AD deductions any time after the treatment.

Figure 6 plots θ_s for two samples: the full set of treated firms and the subset of treated firms with positive asset growth ($K_t > K_{t-1}$) at least once during the AD years. A positive θ_s in the post-treatment years could be driven either by selection where firms claim AD on infra-marginal investment, or by AD-induced increases in investment. Restricting to firms with $K_t > K_{t-1}$ partially controls for the selection effect. In Panels A and B where the outcome variable is $Ln(K_t) - Ln(K_{t-1})$, we observe that θ_s is indistinguishable from zero in all years. In Panels C and D where the outcome variable is $Ln(K_t)$, the comparison between claimers and non-claimers in general displays pre-treatment trends. Conditional on positive asset growth, claimers and non-claimers are not different. Thus, the average claiming firm most likely claimed AD on inframarginal investment.

Overall, our analyses indicate that knowledge of the AD incentive alone was insufficient to stimulate investment for most firms. Other factors may have rendered the policy ineffective. China's GDP growth slowed considerably during our sample period and there was high economic uncertainty (Huang and Luk, 2020), which could dampen investment incentives (Guceri and Albinowski, 2021). Larger firms have more resources to cope with such an uncertain environment (Ghosal and Loungani, 2000), but for the vast majority of smaller firms, it could have been difficult for AD incentives to overcome such drags on investment, even if awareness increased. Nonetheless, even within the top 5% of firms with positive asset growth, the claiming rate of AD is mediocre (less than 40%), indicating substantial information friction. Hence, improving the salience of AD incentives even just among large firms may be fruitful.

6 Conclusions

Using confidential corporate tax returns from a large province, we document that the introduction of AD in China failed to meaningfully stimulate investment, in contrast to

recent studies of similar policies in the US and UK. We further document that firms did not claim AD on over 80% of cases of eligible investments. Tax losses, tax sophistication, and access to tax administration all strongly predict firms' likelihood of taking up AD benefits.

Our findings–in terms of both the low take-up for a significant tax benefit, and the impact of tax administrator accessibility on such take-up–are likely not unique to the Chinese context. Imperfect take-up of AD been documented in other countries (see Section 4 above). It is also well-recognized that the accounting and legal professions play an important role in transmitting information about the content of law and policy (OECD, 2019). The success of such professions, however, likely depends on the existence of sizeable populations of firms that are sufficiently large and profitable to outsource compliance tasks. In less developed countries, these professions tend not to flourish. This sharply reduces available channels for policy information transmission. Tax administrators may become the main "tax professionals" disseminating tax knowledge. Where professional resources are even less abundant than in China, tax administrators may play a still more crucial role in educating taxpayers.

One response to constraints on information transmission about tax policy is to make tax law less complex. But there are limits on how "simplified" tax law can become. Restricting tax policy instruments only to tax rate variations clearly has disadvantages. The Chinese AD policy and the tax return schedule on which AD benefits are claimed are already relatively simple. If taxpayers fail to apply even simple rules, it may be important to examine instead how tax law information can be more effectively transmitted. The use of newly available digital technology is certainly one direction for exploration (OECD, 2019). Designing incentives for tax administrators so that they are adequately motivated to educate taxpayers even while trying to raise revenue is another possible direction. Both take on special significance when traditional types of tax professionals may not easily materialize.

References

Aarbu, Karl Ove and Jeffrey K Mackie-Mason, "Explaining Underutilization of Tax Depreciation Deductions: Empirical Evidence from Norway," *International Tax and Public Finance*, 2003, *10* (2003), 229–257.

Abeler, Johannes and Simon Jäger, "Complex Tax Incentives," *American Economic Journal: Economic Policy*, aug 2015, *7* (3), 1–28.

Auerbach, Alan J, "The New Economics of Accelerated Depreciation," *Boston College Law Review*, 1982, *23* (5), 1327–1355.

Bazel, Philip and Jack Mintz, "Policy Forum: Is Accelerated Depreciation Good or Misguided Tax Policy?," *Canadian Tax Journal*, 2019, *61*, 41–55.

Bond, Stephen and Jing Xing, "Corporate taxation and capital accumulation: Evidence from sectoral panel data for 14 OECD countries," *Journal of Public Economics*, 2015, *130*, 15–31.

Cai, Hongbin and Qiao Liu, "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms," *The Economic Journal*, 2009, *119* (537), 764–795.

Chen, Zhao, Jiang Zian, Liu Zhikuo, and Juan Carlos Suarez Serrato, "Tax Policy and Lumpy Investment Behavior: Evidence from China's VAT Reform," *NBER Working Paper 26336*, 2019, *13* (1), 1689–1699. http://www.nber.org/papers/w26336.

_ , **Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Yi Xu**, "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, 2020, pp. 1–65.

Chetty, Raj, John N Friedman, and Emmanuel Saez, "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, dec 2013, *103* (7), 2683–2721.

Cooper, Michael and Matthew Knittel, "Partial Loss Refundability: How Are Corporate Tax Losses Used?," *National Tax Journal*, 2006, *59* (3).

Cui, Wei, "Administrative decentralization and tax compliance: A transactional cost perspective," *University of Toronto Law Journal*, 2015, *65* (3), 186–238.

Dabla-Norris, Era, Florian Misch, Duncan Cleary, and Munawer Khwaja, "The quality of tax administration and firm performance: evidence from developing countries," *International Tax and Public Finance*, 2020, *27*, 514–551.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, sep 1996, *64* (5), 1001.

Edgerton, Jesse, "Investment incentives and corporate tax asymmetries," *Journal of Public Economics*, 2010, *94* (11-12), 936–952.

— , "Investment, Accounting, and the Salience of the Corporate Income Tax," *NBER Working Paper 18472*, 2012. https://www.nber.org/papers/w18472.

Fan, Ziying and Yu Liu, "Tax Compliance and Investment Incentives: Firm Responses to Accelerated Depreciation," *Journal of Economic Behavior and Organization*, 2020, *176*, 1–17.

Forsling, Gunnar, "Utilization of tax allowances," *Finnish Economic Papers*, 1998, *11* (2).

Ghosal, Vivek and Prakash Loungani, "The differential impact of uncertainty on investment in small and large businesses," *Review of Economics and Statistics*, 2000, *82* (2), 338–343.

Goodspeed, Timothy, Jorge Martinez-Vazquez, and Li Zhang, "Do Companies View Bribes as a Tax? Evidence on the Tradeoff between Corporate Taxes and Corruption in the Location of FDI," in Clemens Fuest and Zodrow George R, eds., *Critical Issues in Taxation and Development*, The MIT Press, 2013, pp. 45–64.

Gordon, Roger and Wei Li, "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 2009, *93* (7-8), 855–866.

Gordon, Roger H, James R Hines, and Lawrence Summers, "Notes on the Tax Treatment of Structures," in Lawrence Summers, ed., *The Effects of Taxation on Capital Accumulation*, Chicago: University of Chicago Press, 1987, pp. 223–257.

Graham, John R., Michelle Hanlon, Terry Shevlin, and Nemit Shroff, "Tax rates and corporate decision-making," *Review of Financial Studies*, 2017, *30* (9), 3128–3175.

Gravelle, Jane G, "Effects of the 1981 Depreciation Revisions on the Taxation of Income from Business Capital," *National Tax Journal*, 1982, *35*, 1–20.

Gronberg, Sami, "Tekevätkö yritykset korotettuja poistoja?," Technical Report, VATT Institute for Economic Research, Helsinki 2015.

Guceri, Irem and Maciej Albinowski, "Investment responses to tax policy under uncertainty," *Journal of Financial Economics*, 2021, pp. 1–63.

Guenther, Gary, "The Section 179 and Section 168(k) Expensing Allowances: Current Law and Economic Effects," *Congressional Research Service Report*, 2018. https://fas.org/sgp/crs/misc/RL31852.pdf.

Hall, Robert E and Dale W Jorgenson, "Tax policy and investment behavior," *The American Economic Review*, 1967, *57* (3), 391–414.

Hassett, Kevin A. and Glenn R. Hubbard, "Tax policy and business investment," *Handbook of Public Economics*, 2002, *3*, 1293–1343.

Heckman, James J., "Sample Selection Bias as a Specification Error," *Econometrica*, jan 1979, *47* (1), 153.

House, Christopher L and Matthew D Shapiro, "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *American Economic Review*, 2008, *98* (3), 737–768.

Huang, Yun and Paul Luk, "Measuring economic policy uncertainty in China," *China Economic Review*, 2020, *59* (October 2019), 101367.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Multivariate Matching Methods That Are Monotonic Imbalance Bounding," *Journal of the American Statistical Association*, 2011, *106* (493), 345–361.

Kanniainen, Vesa and Jan Sodersten, "Cost of monitoring and corporate taxation," *Journal of Public Economics*, 1994, *2727* (94).

Kitchen, John and Matthew Knittel, "Business Use of Section 179 Expensing and Bonus Depreciation, 2002-2014," *Office of Tax Analysis Working Paper 110*, 2016, (October). https://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/WP-110.pdf.

Liu, Li, Ben Lockwood, Miguel Almunia, and Eddy H F Tam, "VAT Notches, Voluntary Registration, and Bunching: Theory and UK Evidence," *Review of Economcis and Statistics*, 2019, *Just Accep.*

Maffini, Giorgia, Jing Xing, and Michael P. Devereux, "The Impact of Investment Incentives: Evidence from UK Corporation Tax Returns," *American Economic Journal: Economic Policy*, 2019, *11* (3), 361–389.

McKenzie, David and Yaye Seynabou Sakho, "Does it pay firms to register for taxes? The impact of formality on firm profitability," *Journal of Development Economics*, 2010, *91* (1), 15–24.

OECD, *Tax Administration 2019: Comparative Information on OECD and Other Advanced and Emerging Economies*, Paris: OECD Publishing, 2019.

Ohrn, Eric, "The effect of corporate taxation on investment and financial policy: Evidence from the DPAD," *American Economic Journal: Economic Policy*, 2018, *10* (2), 272–301.

Qiu, Qiqi and Junmin Wan, "Depreciation Rate by Perpetual Inventory Method and Depreciation Expense as Accounting Item," *Fukuoka University Center for Advanced Economic Study (CAES) Working Paper 2019-012*, 2019. http://www.econ.fukuoka-u.ac.jp/researchcenter/workingpapers/WP-2019-012.pdf.

Wales, Terence J, "Estimation of an Accelerated Depreciation Learning Function," *Journal of the American Statistical Association*, 1966, *61* (316), 995–1009.

Yagan, Danny, "Capital tax reform and the real economy: The effects of the 2003 dividend tax cut," *American Economic Review*, 2015, *105* (12), 3531–3563.

Zwick, Eric, "The Costs of Corporate Tax Complexity," *American Economic Journal: Economic Policy*, May 2021, *13* (2), 467–500.

_ **and James Mahon**, "Tax policy and heterogeneous investment behavior," *American Economic Review*, 2017, *107* (1), 217–248.

Table 1: **Policy Overview**

	Unit Value < 5000	5000 < Unit Value < 1 Million	Unit Value > 1 Million
Targeted Industries	Immediate expensing for all fixed asset purchases	(1) Accelerated depreciation for all fixed asset purchases (2) Immediate expensing for <u>equipment</u> that is used <u>exclusively</u> for R&D (3) *Small Firms Only:* Immediate expensing for <u>equipment</u> that is <u>shared</u> for R&D	Accelerated depreciation for all fixed asset purchases
All Firms	Immediate expensing for all fixed asset purchases	Immediate expensing for <u>equipment</u> that is used <u>exclusively</u> for R&D	Accelerated depreciation for <u>equipment</u> that is used <u>exclusively</u> for R&D

2014 Targeted Industries

Biopharmaceutical manufacturing
Manufacturing of professional equipment
Manufacturing of railway, vessel, aerospace and other transportation equipment
Manufacturing of computer, communication, and other electronic equipment
Manufacturing of instruments and apparatus
Information transmission, software, and information technology services

2015 Targeted Industries

Light industry; Textile; Machinery; Automobile

Non-Targeted Manufacturing Industries

Wine, beverage, tea, and tobacco processing and manufacturing
Petroleum processing, coking and nuclear fuel processing
Chemical raw materials and chemical manufacturing
Rubber and plastic products industry
Non-metallic mineral products industry
Ferrous and non-ferrous metal smelting and rolling processing industry
Other Manufacturing
Waste resource comprehensive utilization industry
Metal products, machinery and equipment repair industry

Note: This table illustrates the provisions for AD and immediate expensing implemented in 2014 and 2015 and lists the set of targeted industries and non-targeted manufacturing industries. Section 1 describes the AD provisions for the targeted industries.

Table 2: **Depreciation schedules and Cost of Capital Changes**

Panel A: Regular and Accelerated Asset Depreciation Schedules

Assets	Asset Life Regular	AD	PV Deductions ($r = 7\%$) Regular	AD	Diff.	DDB	Implied BD Factor
Buildings and Structures	20	12	0.566	0.708	0.142	0.643	N.A.
Production Equipment	10	6	0.752	0.850	0.098	0.808	0.220
Furniture, Tools, Appliances	5	3	0.877	0.936	0.059	0.919	0.210
Transportation	4	3	0.906	0.936	0.030	0.945	N.A.
Electronic Equipment	3	2	0.936	0.967	0.031	0.971	N.A.

Panel B: Average Change in Cost of Capital in Tax Data Sample

	China Regular	AD	Change	Recent Studies Change in Zwick (2017)	Change in Maffini (2019)
Z	.741	.836	.095	.048 to .078	.033
τZ	.132	.150	.017	.017 to .027	.011
$\frac{1-\tau Z}{1-\tau}$	1.060	1.038	-.022	-.026 to -.042	-.016

Note: Panel A illustrates the depreciation schedules for each asset class. The first two columns show the asset lives for each asset class under the regular depreciation schedule and the AD schedule introduced in 2014 and 2015. The next three columns show the present value (PV) of deductions for a dollar of investment under the regular depreciation schedules, the AD schedule, and the difference between the two. We use a discount rate of 7%. The sixth column reports the PV of depreciation using the double declining balance (DDB) method and assuming there is no salvage value. This approximates U.S. MACRS but ignores the half-year rule and the switch to straight-line depreciation allowed under U.S. law. Lastly, we calculate the implied bonus depreciation (BD) factor obtained by equating the PV of depreciation under BD + DDB to that under AD. We only calculate this for the 5-year and 10-year assets, since U.S. BD is not available for the 20-year asset, and DDB is faster than AD for 3- and 4-year assets. Panel B quantifies the effect of the regular and AD schedules at the firm level. The first row reports the average PV of depreciation deductions (Z) with and without AD. The second row reports the PV of the tax savings due to these deductions (τZ). The third row presents estimates of the UCC under each depreciation scheme ($\frac{1-\tau Z}{1-\tau}$). We assume that each firm allocates a dollar of new investment in proportion to their pre-AD asset holdings. Denoting $V_{i,k}$ as firm i's holdings of type k assets, define $Z_i = \sum_k \frac{V_{i,k}}{\sum_k V_{i,k}} Z_k$ and $Z = \frac{1}{N} \sum_i Z_i$, where Z_k is the PV of deductions for a dollar of investment in asset class k. We define τZ as $\frac{1}{N} \sum_i \tau_i Z_i$ where τ_i is firm i's tax rate. Similarly, $\frac{1-\tau Z}{1-\tau}$ is the average across firms of $\frac{1-\tau_i Z_i}{1-\tau_i}$. Our calculations of Z under AD are based on the 40% reduction in useful life provisions for purchases greater than 5,000 CNY in the targeted industries. The calculations ignore immediate expensing on purchases of less than 5,000 CNY and the immediate expensing provisions for R&D-related purchases. The last two columns report the analogous figures for AD policies studied in Zwick and Mahon (2017) and Maffini et al. (2019).

Table 3: **Full Sample Descriptive Statistics 2013**

	Non-Targeted Industries		2014 Targeted Industries		2015 Targeted Industries	
	Mean	N	Mean	N	Mean	N
Total Assets	5516	8419	3649	4547	4497	17721
Fixed Assets Net of Depreciation	468	8419	298	4547	417	17721
Fixed Assets Historical Cost	1529	8419	902	4547	1350	17721
Average Useful Life	11.41	7114	8.88	3311	11.56	15620
Business Revenue	4699	8384	2188	4510	3403	17670
Revenue Growth	-.06	7940	.12	4092	-.04	16895
Taxable Income	114	8418	128	4547	80	17721
Percent In Tax Losses	.28	8419	.4	4547	.27	17721
Tax Loss Stock	227	2510	81	1867	166	5109
Cash Holdings / Total Assets	.11	8417	.26	4544	.11	17716
Claimed Interest	.29	8419	.2	4547	.33	17721
Age	15.99	8419	10.52	4547	15.76	17721
Employees	65.4	8419	50.09	4547	74.04	17721
State or Collectively Owned	.12	8419	.04	4547	.08	17721
High and New Tech Enterprise	.08	8419	.14	4547	.06	17721
Small or Micro Enterprise (SME)	.61	8419	.66	4547	.61	17721
Distance to Tax Bureau (km)	12.98	6869	9.16	3222	12.64	14437
Firms per Tax Administrator	206	6740	270	3119	250	14719

Note: This table displays means for select firm characteristics among targeted industries and manufacturing firms in non-targeted industries for the full sample of firms observed in 2013. Total Assets, Fixed Assets, Business Revenue, Taxable Income, and Tax Loss Stock are all in CNY10,000s. The mean of Tax Loss Stock is calculated conditional on having a positive amount. The average asset life duration is calculated among the subset of firms with positive reported values for at least one asset type. All continuous variables are winsorized at the 1st and 99th percentiles within each group of firms.

Table 4: **Did AD Stimulate Growth in Firms' Asset Stock?**

	$Ln(K_{i,t}) - Ln(K_{i,t-1})$		$Ln(K_{i,t})$	
	(1)	(2)	(3)	(4)
	2014	2015	2014	2015
Treat × Post	-3.45	-0.59	1.20	0.68
	(2.10)	(1.50)	(3.45)	(3.04)
	[-6.95,0.04]	[-3.07,1.89]	[-4.54,6.95]	[-4.35,5.71]
N	10013	34304	12890	44252
Treated Firms	1055	5868	1078	6030
Untreated Firms	1583	3208	1638	3308
Dep. Var Mean	6.90	4.85	1447.05	1454.85

Note: This table reports the estimated treatment effects based on Equation (3). In columns (1) and (2) the outcome variable is $Ln(K_t) - Ln(K_{t-1})$. In columns (3) and (4) the outcome is $Ln(K_t)$. Both are winsorized at the 1st and 99th percentiles within each year separately for the treated and control groups. The coefficients are scaled by 100. The 90% confidence interval is shown in square brackets. Standard errors are clustered at the three-digit industry level and shown in parenthesis. There are 44 clusters in the control groups, and 38 and 90 clusters in the 2014 and 2015 treated groups, respectively. The average of the outcome variable for the pre-treatment period (2011 to 2013) among the treated industries is shown in the "Dep. Var. Mean" row and scaled by 100. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: **Predictors of Claiming: Firm-Level Characteristics**

	Pr(Claim | Invest)		
	(1)	(2)	(3)
	Probit Model		Selection Model
In Tax Loss Before AD	-2.79***	-2.85***	-2.73***
	(0.45)	(0.43)	(0.43)
Ln(Tax Loss Stock $_{t-1}$)	-0.09***	-0.08**	-0.08**
	(0.03)	(0.03)	(0.03)
Ln(Total Assets $_{t-1}$)	0.56***	0.63***	0.42***
	(0.11)	(0.11)	(0.14)
High and New Technology Enterprise	2.24***	1.80***	1.83***
	(0.61)	(0.58)	(0.59)
$Ln(K_{t,k} - K_{t-1,k})$	0.12***	0.09**	0.09**
	(0.04)	(0.04)	(0.04)
Two-Digit Industry FE	Yes	Yes	Yes
Prefecture FE	No	Yes	Yes
N	48663	48663	171846
Mean of Outcome Var.	6.36	6.36	6.36

Note: This table reports the estimated average marginal effects of firm-level characteristics on the probability of claiming AD conditional on having purchased eligible investment ($P(C_{i,t,k} = 1|I_{i,t,k} = 1)$) as outlined in Section 5. Columns (1) and (2) present estimates from the probit model and column (3) reports estimates from the selection model, both described in Section 5. Coefficients are scaled by 100. All time-varying explanatory variables are measured in year $t - 1$. Ln(Tax Loss Stock $_{t-1}$) is the natural logarithm of (Tax Loss Stock $_{t-1}$ +1) to account for zeros. All continuous covariates are winsorized at the 1st and 99th percentiles within each year. Standard errors are computed using the delta-method and are clustered at the three-digit industry level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
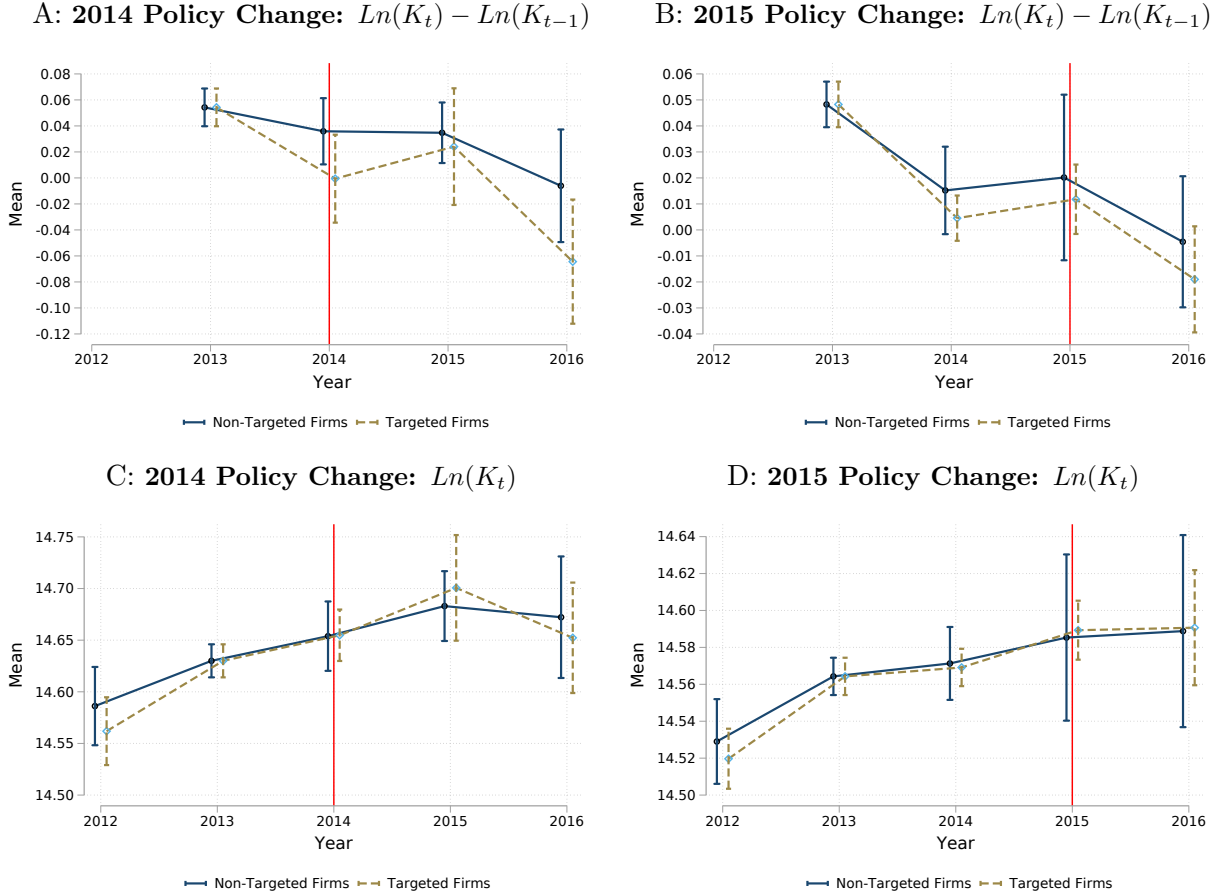
Table 6: **Predictors of Claiming: Tax Administration**

	Claim | Invest						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(# Firms / # Tax Administrators)	-0.95*** (0.23)		-1.89*** (0.35)			-1.91*** (0.44)	-1.95*** (0.36)
Log(Distance to Tax Bureau)		-0.40** (0.20)	-0.76*** (0.26)		-0.48** (0.21)	-0.73** (0.28)	-0.99*** (0.33)
Log(Distance to Government Office)				0.24 (0.18)			
Firm Headquarters in Industrial Park						0.43 (0.64)	
Urban Bureau						-0.24 (0.64)	
STR minus Cash ETR						-0.85 (0.57)	
Fixed Assets / Total Assets						0.85 (0.80)	
Total Profit / Business Revenue						0.57 (0.38)	
Ln(# Firms / # Tax Admin.) × HNTE							0.28 (0.58)
Ln(Distance to Tax Bureau) × HNTE							0.98 (0.69)
Firm Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Two-Digit Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tax Bureau FE	No	No	No	No	Yes	No	No
N	42905	40841	34374	46693	40360	32140	34374
Mean of Outcome Var.	6.59	6.65	6.85	6.31	6.73	6.90	6.85

Note: This table reports the estimated average marginal effects of tax bureau characteristics on the probability of claiming AD conditional on having purchased eligible investment ($P(C_{i,t,k} = 1 | I_{i,t,k} = 1)$) using the probit model described in Section 5. Coefficients are scaled by 100. All continuous covariates are winsorized at the 1st and 99th percentiles within each year for each of the treatment groups. Standard errors are computed using the delta-method and are clustered at the three-digit industry level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1: **Growth of Fixed Assets Net of Depreciation**

A: **2014 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

B: **2015 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

C: **2014 Policy Change:** $Ln(K_t)$

D: **2015 Policy Change:** $Ln(K_t)$

Note: Panels A and B plot trends in $Ln(K_t) - Ln(K_{t-1})$ while Panels C and D plot trends in $Ln(K_t)$. In all four, K is fixed assets net of accounting depreciation as reported on balance sheets. Trends are plotted separately for the matched treated and control firms. The control firms are those in non-targeted manufacturing industries and matched as described in Section 3. For each panel, we estimate a dynamic version of equation (3) as $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 2013} \beta_s \times \mathbb{1}\{t = s\} \times D_i + \epsilon_{i,t,g}$. We then plot α_t for the control group and $\alpha_t + \beta_t$ for the treatment group and the associated 95% confidence intervals. The vertical red lines indicate the timing of the AD policy announcement and implementation.

Figure 2: **Heterogeneous Responses to the Reform?**

A: **By Firm Characteristics**
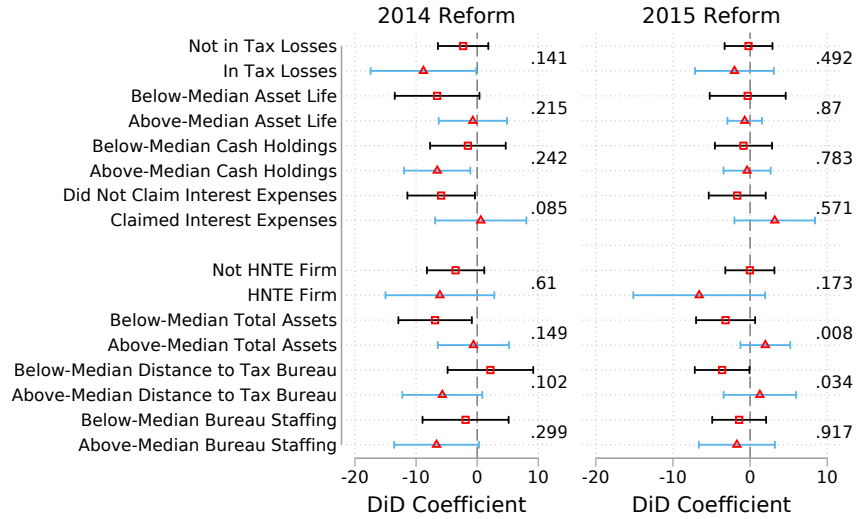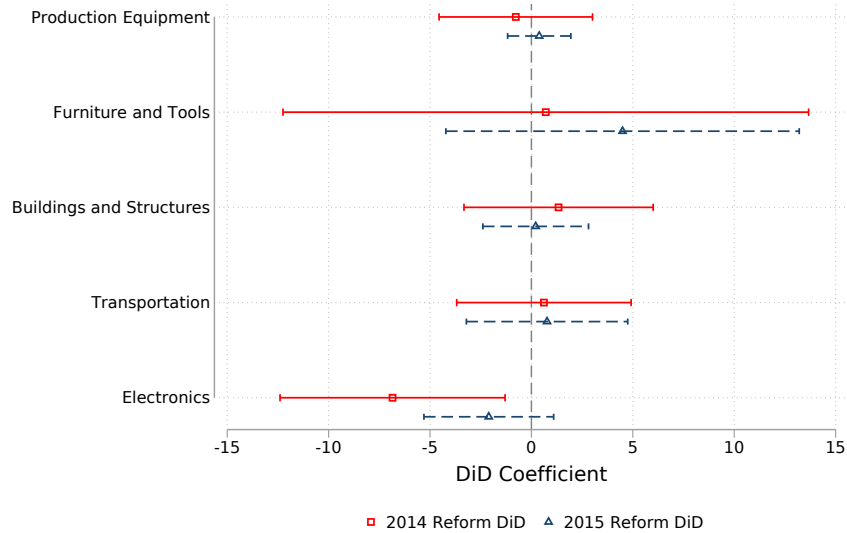
B: **By Asset Type**

□ 2014 Reform DiD   △ 2015 Reform DiD

Note: Panel A plots estimates of β and 95% confidence intervals from equation (3) for subsets of the sample. For each of eight different firm characteristics X, we split the sample into above and below median when X is continuous, or by $X = 0$ and $X = 1$ when X is binary. The first four characteristics proxy for a firm's readiness to invest, and the how beneficial AD is for the firm. The last four characteristics proxy for a firm's informational awareness and sophistication. P-values for the test of the treatment effects being the same across sub-samples are displayed to the right of the confidence intervals. Panel B plots estimates of β for the full sample of firms but estimated separately for each asset class. In both panels, the outcome is $Ln(K_t) - Ln(K_{t-1})$. In panel A, K is fixed asset stock measured net of accounting depreciation from balance sheets. In Panel B, since net-of-depreciation measures are not observed at the asset level, K is measured at original cost from tax depreciation return, and the unit of observation is an asset type-firm-year triplet. Estimates are scaled by 100 to be consistent with the estimates reported in Table 4.
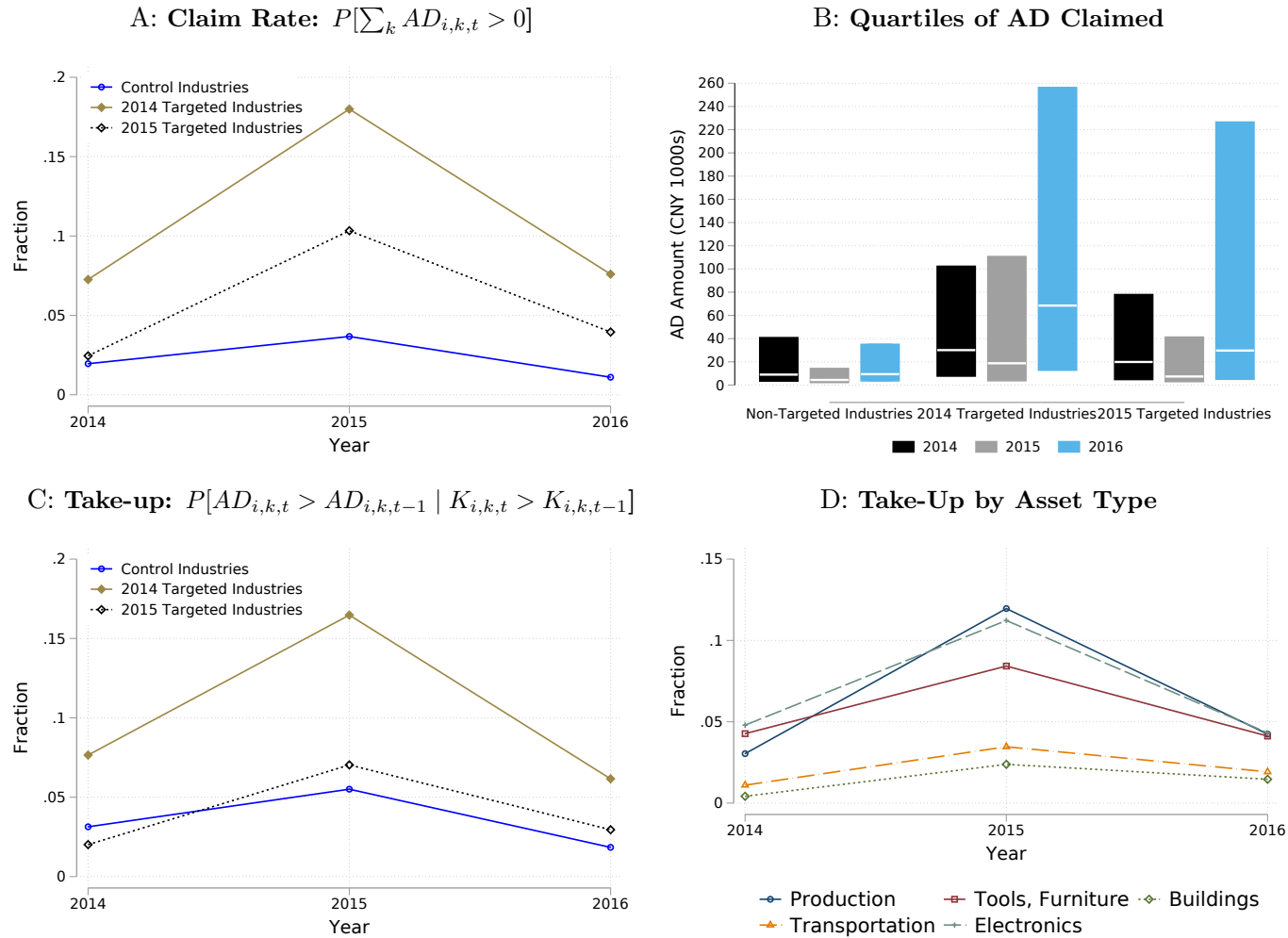
Figure 3: **Responsiveness by Pre-Reform Asset Quartiles**

A: **Outcome:** $Ln(K_t) - Ln(K_{t-1})$ ⟶ B: **Outcome:** $Ln(K_t)$
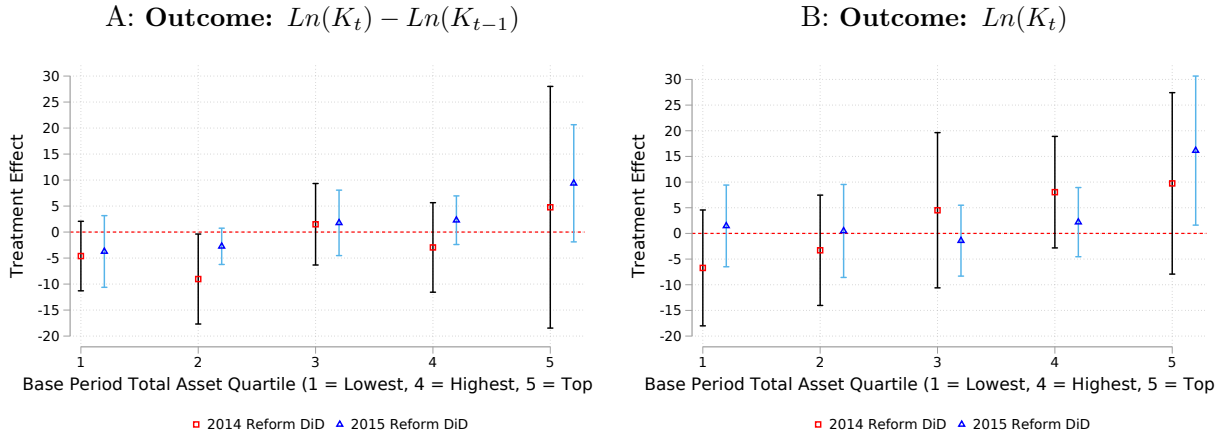
Quartile	Left Endpoints		Number of Firms	
	2014	2015	2014	2015
1	0.00	0.00	666	2318
2	6.13	5.39	669	2287
3	16.18	13.25	655	2244
4	53.98	40.09	653	2227
5	381.68	295.88	128	444

Note: Panels A and B split firms into quartiles based on pre-treatment average total assets (1 = smallest, 4 = largest). Each then plots estimates of β from equation 3 and 95% confidence intervals for each quartile. The fifth bin (5) plots β for the largest 5% of firms. Estimates are scaled by 100 to be consistent with the estimates reported in Table 4. K_t is fixed assets net of accounting depreciation. The table lists the quartile's left-endpoints, in 1,000,000CNY, and the number of firms in each.

Figure 4: **Take-Up of Accelerated Depreciation**

A: **Claim Rate:** $P[\sum_k AD_{i,k,t} > 0]$

B: **Quartiles of AD Claimed**

C: **Take-up:** $P[AD_{i,k,t} > AD_{i,k,t-1} \mid K_{i,k,t} > K_{i,k,t-1}]$

D: **Take-Up by Asset Type**
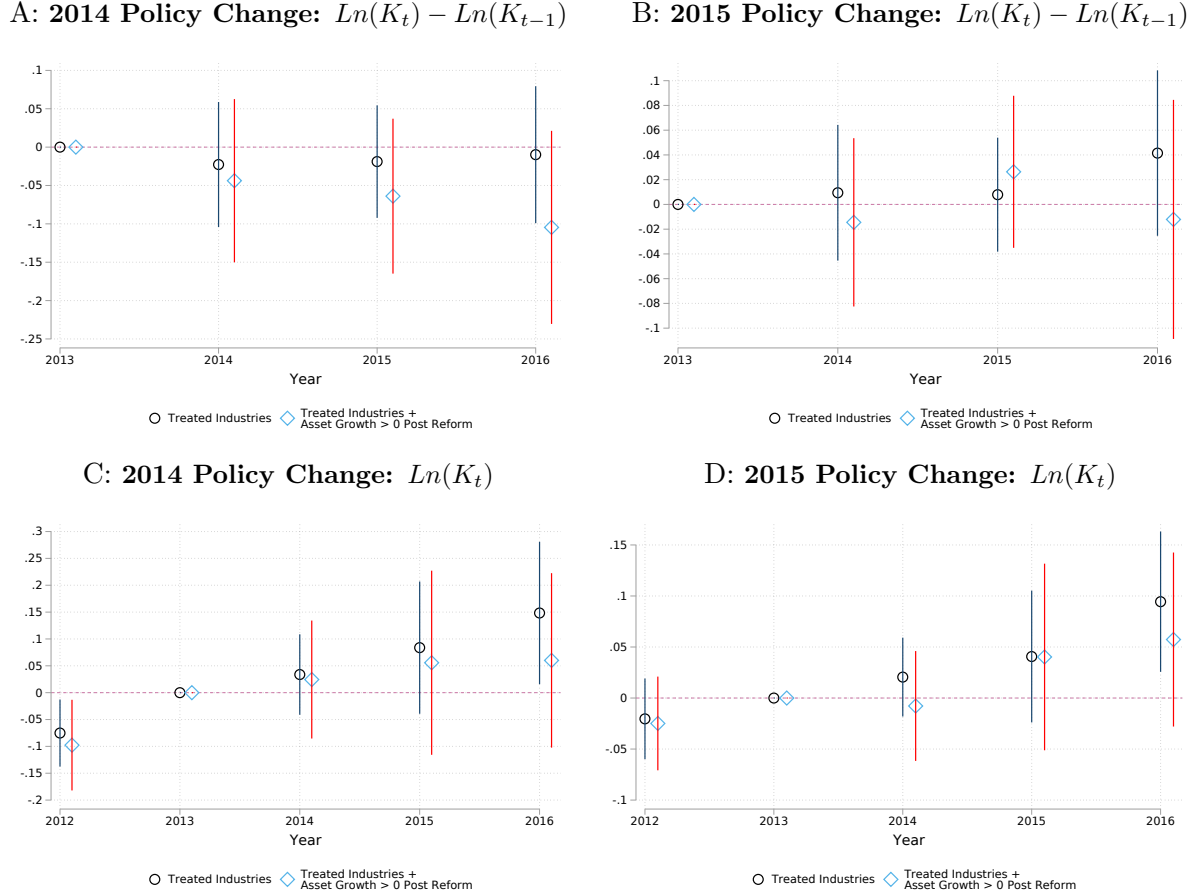
Note: This figure plots AD claiming trends during the period 2014-2016. Firms are divided into the 2014 targeted group, the 2015 targeted group, and the non-targeted group. Panel A plots the percent of firms within each group that claimed AD in at least one asset class. Panel B plots the median and interquatile range of AD amounts claimed. The top and bottom of the bars indicate the 25th and 75th percentiles, respectively, and the white divider on each bar denotes the median. Panel C calculates the take-up rate conditional on having eligible investment. For each asset class, we restrict to firms with a year-over-year increase in the asset stock (at historical cost), then calculate the percent with a year-over-year increase in AD deductions. We then average this take-up rate across the five asset classes. Panel D plots the take-up rate within each asset class.

Figure 5: **Predictors of AD Take-Up**
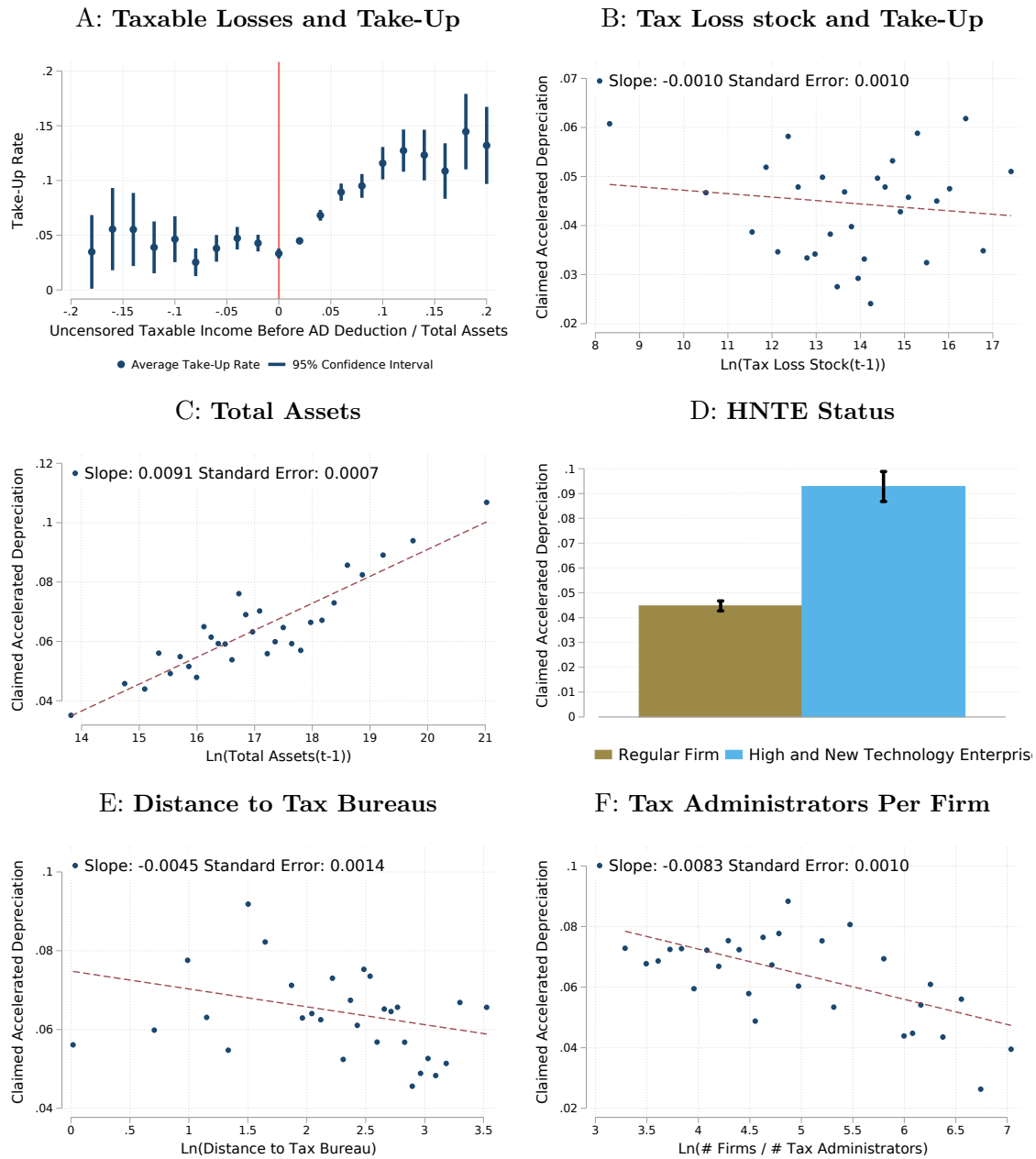
Note: This figure restricts the sample to firms that belong to the 2014 or 2015 targeted industries. Each observation is a firm-year-asset type triplet. Within each triplet, we retain firms with an increase in the asset stock relative to the previous period. Panel A plots the probability of claiming any AD against uncensored taxable income dived by total assets. Panels B through E plot the correlation between take-up of AD and a firm characteristic (both described in the main text) after controlling for two-digit industry, asset type, and year fixed effects.

43

Figure 6: **Growth of Fixed Assets by Claiming Status**

A: **2014 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

B: **2015 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

C: **2014 Policy Change:** $Ln(K_t)$


D: **2015 Policy Change:** $Ln(K_t)$

Note: This figure examines whether firms that claimed AD experienced differential increases in investment after the treatment. Defining C_i as a time-invariant indicator for whether the firm claimed AD in the post-implementation period (2014 - 2016 for the 2015 policy cohort, 2015 to 2016 for the 2015 cohort), we estimate the descriptive regression $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 0} \beta_s \times 1\{t = s\} \times C_i + \epsilon_{i,t,k}$ for the targeted industries. The figures plot estimates of β_s. The second series restricts the sample to firms that had at least one year of positive asset growth in the post-implementation period before estimating β_s. Standard errors are clustered at the three digit industry level.

44

Appendix for Online Publication Only

A   Additional Difference-in-Difference Results

A.1   Search Intensity

Figure A1: **Search Engine Queries for Accelerated Depreciation**
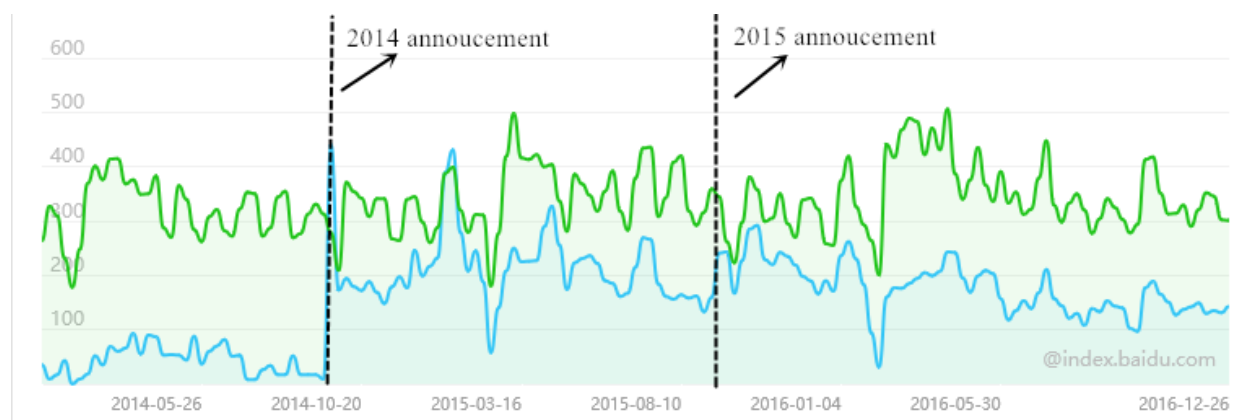
Note: This figure plots the intensity indices of search engine queries from the Chinese website Baidu during the period January, 2014-December, 2016. The shaded blue series is the index for the key words "fixed assets accelerated depreciation" and the green series is the index for the key words "tax reporting". Both key words are in Chinese. The dashed black lines indicate the weeks for the 2014 and 2015 AD policy announcements.

A.2   Matching and Regression Weights

Regression weights for the DiD analysis are constructed in two steps. First, we match on base year observables using Coarsened Exact Matching (CEM) (Iacus et al., 2011) as described in the main text and calculate weights following their recommended procedure. CEM first coarsens the matching variables by sorting values of each variable into mutually exclusive bins. It then creates *strata*, which are simply interactions of the bins. For example, with two matching variables, each grouped into two bins, CEM would create four strata from the product of each binned variable. Targeted and control firms are then exactly matched based on which of the four cells they belong in.[42]

The regression weights are constructed as follows. Denote N_c and N_d the total number of matched control and treated firms. For each strata s, we calculate the number of control

---

[42]Before CEM matching, we restrict the treatment and control samples to have a common support for each of the matching variables and drop firm-year observations with zero business revenue.

firms N_c^s and the number of treated firms N_d^s. Treated firms are given a regression weight h_i of one. Control firms are assigned regression weights proportional to the number of control firms, relative to the number of treated firms, in the strata: $h_i = \frac{N_c}{N_d} \frac{N_d^s}{N_c^s}$. Table A1 shows the descriptive statistics among the matched treatment and control groups.

Second, we use the re-weighting method of DiNardo et al. (1996) to flexibly control for changes in the firm distribution between control and treated industries, as in Yagan (2015) and Zwick and Mahon (2017). The re-weighting procedure proceeds as follows. First, we create ten bins (b) corresponding to the deciles of the revenue distribution for the control firms in the year before the policy implementation (base year). The DFL weights ($w_{i,t,g,b}$) for firm i, with revenue in bin b, in group g (where a group is a treatment status-year pair) are:

$$w_{i,t,g,b} = h_i \times \frac{\sum_{i' \in b \,\cap\, i' \in \underline{g}} h_{i'}}{\sum_{i' \in b \,\cap\, i' \in g} h_{i'}} \times \frac{\sum_{i' \in g} h_{i'}}{\sum_{i' \in \underline{g}} h_{i'}} \tag{6}$$

where $\underline{g}$ is the control group in the base year. These weights capture changes in the distribution of firm size (revenue) over time. One may be concerned that this attenuates the treatment effect towards zero if investment growth caused by the AD policy is correlated with revenue growth. In our setting, results are highly similar without DFL re-weighting – using h_i rather than $w_{i,t,g,b}$.

Table A1: **Matched Difference-in-Difference Sample**

Panel A: 2014 Change	Not Matched					Matched				
	Targeted		Control			Targeted		Control		
	Mean	N	Mean	N	p	Mean	N	Mean	N	p
Total Assets	5304	3334	6179	7114	0	7950	1014	9780	1497	0
Profit Margin	-.27	3146	-.08	6965	0	-.03	1014	-.04	1497	.4
Revenue Growth	.3	2730	.05	6424	0	.07	980	.05	1458	.4
Age	12.13	3334	16.12	7114	0	16.12	1014	15.91	1497	.28
Asset Growth Net of Depreication	.07	1616	.04	3530	.04	.07	1014	.04	1497	.16
Fixed Assets Net of Depreciation	443.14	3334	521.29	7114	.03	1041.72	1014	1309.34	1497	.01
Fixed Assets Historical Cost	1362.8	3334	1767.91	7114	0	1987.16	1014	2856.14	1497	0
Business Revenue	3346	3319	5228	7106	0	5241	1014	7551	1497	0
Taxable Income	178.62	3334	137.46	7113	0	327.46	1014	359.29	1497	.39
Tax Loss Stock	39	3334	51	7114	0	33	1014	57	1497	0
Percent In Tax Losses	.34	3334	.26	7114	0	.2	1014	.21	1497	.58
Average Useful Life	8.89	3334	11.41	7114	0	10.92	1014	11.44	1497	0
State or Collectively Owned	.04	3334	.1	7114	0	.04	1014	.06	1497	0
Small or Micro Enterprise (SME)	.57	3334	.57	7114	.82	.4	1014	.41	1497	.7
Distance to Bureau (km)	9.81	2364	13.13	5760	0	10.74	707	12.11	1115	0
Firms/Bureau Employees	274	2408	216	5762	0	348	825	321	1222	0

Panel B: 2015 Change	Not Matched					Matched				
	Targeted		Control			Targeted		Control		
	Mean	N	Mean	N	p	Mean	N	Mean	N	p
Total Assets	5096	15620	6013	7114	0	5993	5575	6590	3023	0
Profit Margin	-.04	15334	-.05	6965	0	-.03	5575	-.03	3023	.82
Revenue Growth	.05	14288	.03	6424	0	.04	5376	.03	2926	.22
Age	15.8	15620	16.14	7114	0	16.15	5575	16.13	3023	.84
Asset Growth Net of Depreciation	.05	7943	.04	3530	.33	.05	5575	.04	3023	.27
Fixed Assets Net of Depreciation	467.14	15620	495.95	7114	.16	892.17	5575	894.59	3023	.95
Fixed Assets Historical Cost	1591.48	15620	1709.75	7114	.04	1823.92	5575	1896.19	3023	.46
Business Revenue	3893	15601	4928	7106	0	4611	5575	5588	3023	0
Taxable Income	96.31	15620	124.54	7113	0	140.9	5575	159.96	3023	.09
Tax Loss Stock	40	15620	49	7114	0	42	5575	51	3023	0
Percent In Tax Losses	.25	15620	.26	7114	.9	.23	5575	.23	3023	.47
Average Useful Life	11.56	15620	11.41	7114	0	11.41	5575	11.17	3023	0
State or Collectively Owned	.07	15620	.1	7114	0	.05	5575	.06	3023	0
Small or Micro Enterprise (SME)	.58	15620	.57	7114	.11	.52	5575	.51	3023	.46
Distance to Bureau (km)	12.71	12669	13.13	5760	0	13.09	4098	12.61	2260	.02
Firms/Bureau Employees	257	13116	216	5762	0	390	4514	330	2471	0

Note: This table reports descriptive statistics of key variables in 2013 for the unmatched and matched samples using non-targeted manufacturing firms as the control group. Dollar values are reported in 10,000 CNY. Firm-year observations with zero fixed assets are excluded. The matching process is described in Section 3. The first five columns report the mean, sample size, and p-value for the difference of means for the targeted and control firms without matching. The last five columns report the same using the set of matched firms. Profit margin is the ratio of total pre-tax profit over business revenue. The stock of tax losses are the unclaimed tax losses accumulated since 2010. Average useful life is the dollar-weighted average tax-life of firms' asset holdings. All continuous variables are winsorized at the 1st and 99th percentiles within each group.

A.3   Cost of Capital Regression

This section presents an alternative approach to calculating user cost elasticities that uses the policy reform as an instrument in the following specification:

$$y_{it} = \gamma_t + \gamma_i + \eta Ln(\frac{1 - \tau_{i,t}Z_{i,t}}{1 - \tau_{i,t}}) + \epsilon_{it} \tag{7}$$

where $Ln(\frac{1-\tau_{i,t}Z_{i,t}}{1-\tau_{i,t}})$ is the UCC (introduced in section 1) and η is the effect of the UCC on y. We instrument for the UCC using the policy change:

$$Ln(\frac{1 - \tau_{i,t}Z_{i,t}}{1 - \tau_{i,t}}) = \pi_t + \pi_i + \pi D_i \times Post_t + u_{it} \tag{8}$$

where the instrument is $D_i \times Post_t$. Table A2 reports the estimate of η and the 95 percent confidence intervals. Note, this coefficient should be negative if investment increases when the tax component of the user cost of capital decreases. For both the 2014 and the 2015 treatments, the confidence intervals for the point estimates contain zero, which is consistent with the DiD results reported in Table 4 of the paper.

The estimates of η for $Ln(K_t)$ are elasticities of the capital stock with respect to the UCC — a 1% increase in the UCC causing a .59% and .4% decrease in the capital stock, for the 2014 and 2015 policy cohorts respectively.

The estimates of η for $Ln(K_t) - Ln(K_{t-1})$ are semi-elasticities of the UCC with respect to the investment rate (I_t/K_{t-1}) (recall, we argued in Section 3 that the treatment effect on $Ln(K_t) - Ln(K_{t-1})$ closely approximates the treatment effect on I_t/K_{t-1}). To convert the semi-elasticity to an elasticity, η must be divided by the average I_t/K_{t-1} (in the pre-period). As shown in Section 3, $\frac{I_t}{K_{t-1}} = \frac{K_t - K_{t-1}}{K_{t-1}} + \frac{(1-\gamma)S_t}{K_{t-1}} + \gamma$. We observe $\frac{K_t - K_{t-1}}{K_{t-1}}$ in our data: .069 and .0485 for the 2014 and 2015 groups, respectively, from in the pre-AD period (Table 4). We do not observe $\frac{(1-\gamma)S_t}{K_{t-1}} + \gamma$, and so rely on an estimate from Qiu and Wan (2019) of .0953 for the same years to construct estimates of the average pre-period I_t/K_{t-1} of .164 and .144 respectively. We divide these into the 90% bound of η reported in Table A2 (-.41 and -1.14) to get bounds on the elasticity of 1.1 and 3.7 (absolute values, as in the main text).

Table A2: **User Cost Regression**

	$Ln(K_t) - Ln(K_{t-1})$		$Ln(K_t)$	
	2014	2015	2014	2015
Log User Cost	1.86	0.15	-0.59	-0.40
	[-0.41,4.12]	[-1.14,1.43]	[-4.34,3.16]	[-3.02,2.22]
N	9343	31880	12018	41091
Treated Firms	1014	5615	1046	5818
Untreated Firms	1499	3041	1566	3176

Note: This table presents estimates of equation (7). Standard errors are clustered at the three-digit industry level. 90% confidence intervals are shown in square brackets.

## A.4 Specification and Outcome Robustness

Table A3: **Impact of AD on Investment–Robustness Checks**

|	$\frac{K_{t,net}-K_{t-1,net}}{Sales_{2010-2013}}$		Historical Cost		Time $\times$ Ind		Excl. RD Claimers		Excl. SMPE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	2014	2015	2014	2015	2014	2015	2014	2015	2014	2015
Treat $\times$ Post	-0.57	0.27	-0.88	0.05	-1.48	1.07	-3.13	0.03	0.23	2.48
	(0.58)	(0.35)	(1.86)	(0.80)	(3.39)	(2.80)	(3.48)	(1.59)	(2.88)	(1.66)
N	10131	34646	13680	46588	10013	34304	3349	31052	5024	14962
Treated Firms	1059	5891	1036	5733	1055	5868	355	5352	611	2697
Untreated Firms	1592	3227	1556	3136	1583	3208	612	2875	804	1526
Treated Clusters	37	90	37	90	37	90	37	90	37	90
Untreated Clusters	44	44	44	44	44	44	44	44	44	44
Dep. Var. Mean	1.09	0.38	12.46	10.51	6.90	4.85	8.18	3.96	8.36	8.25

Note: This table reports robustness checks on the DiD estimation. The first two models use alternative outcomes: (i) the year-over-year change in the fixed asset stock measured net of accounting depreciation normalized by sales in the pre-policy period and (ii) $Ln(K_t) - Ln(K_{t-1})$ where K is measured at historical cost from the tax returns. The remaining columns use the baseline outcome: $Ln(K_t) - Ln(K_{t-1})$ where K is measured net of accounting depreciation. The third model includes linear time trends at the three digit industry level. The fourth model removes firms that ever claimed the RD super deduction. The fifth model excludes Small and Micro-profit Enterprises (SMPEs). Standard errors are robust clustered at the three digit industry level. Coefficients and Dep. Var. Mean are scaled by 100. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A2: **Dynamic Difference-in-Difference**

A: **Outcome:** $Ln(K_t) - Ln(K_{t-1})$

B: **Outcome:** $Ln(K_t)$
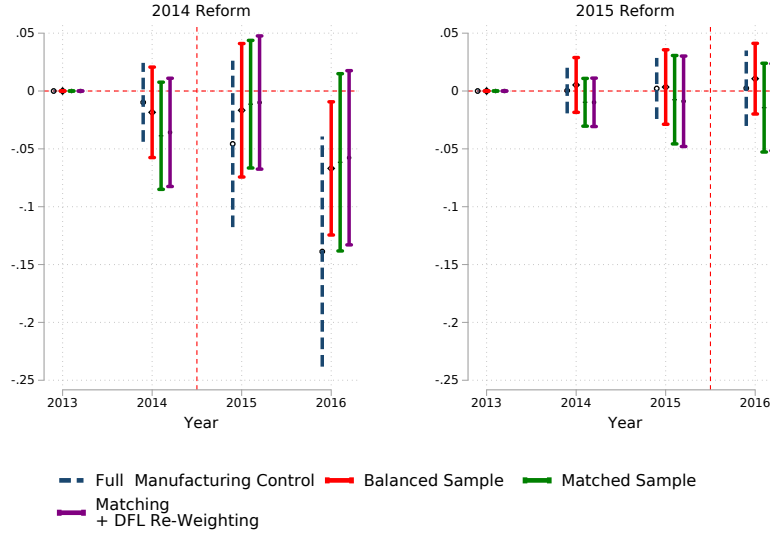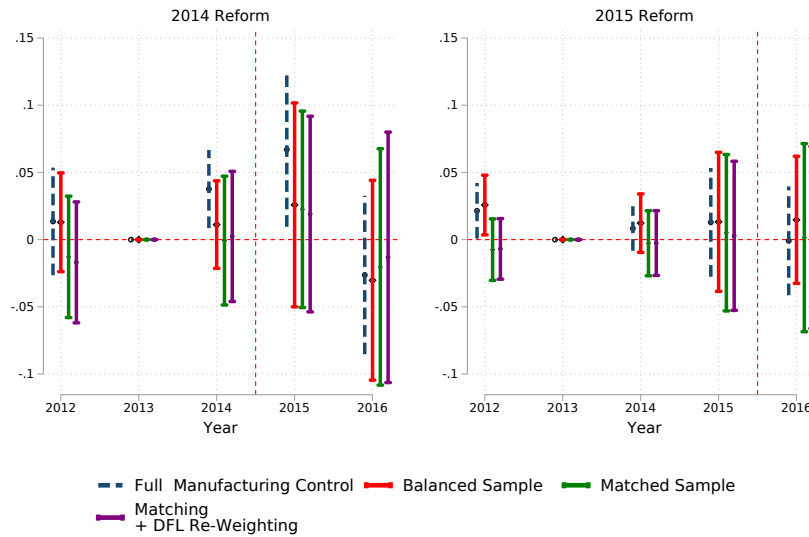
Note: This figure plots the estimated dynamic DiD coefficients β_s and 95 percent confidence intervals from the following specification: $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 0} \beta_s \times 1\{t = s\} \times D_i + \epsilon_{i,t,k}$ where time t is normalize to zer0 in 2013. The Manufacturing Control series restricts the control group to manufacturing industries. The final three series (a) restrict to firms present for all seven years, (b) restrict the matched sample as described in Section 3, and (c) restrict to matched sample and DFL re-weight. Our baseline results presented in the main text correspond to specification (c). Standard errors are clustered at the three digit industry level. Panels A and B presents results for the outcomes $Ln(K_t) - Ln(K_{t-1})$ and $Ln(K_t)$ respectively. Outcomes are winsorized at the 1 and 99% level in each year and treated/control group.

A.5   A Combined Dynamic DiD Specification

We center time around policy implementation and then pool the two treatment groups together to run a single dynamic difference in difference. The pooled specification is as follows:

$$Y_{it} = \alpha_t + \alpha_i + \gamma D_i + \sum_{s \neq 0} \beta_s \mathbb{1}_s \times D_i + \epsilon_{it} \tag{9}$$

Where as before, α_t and α_i are year and firm fixed effects. D_i denotes being in a targeted industry. We define event time s as time centered around policy implementation. So $s = 0$ corresponds to 2013 for the 2014 targeted industries and 2014 for the 2015 industries. Figure A3 shows the results.
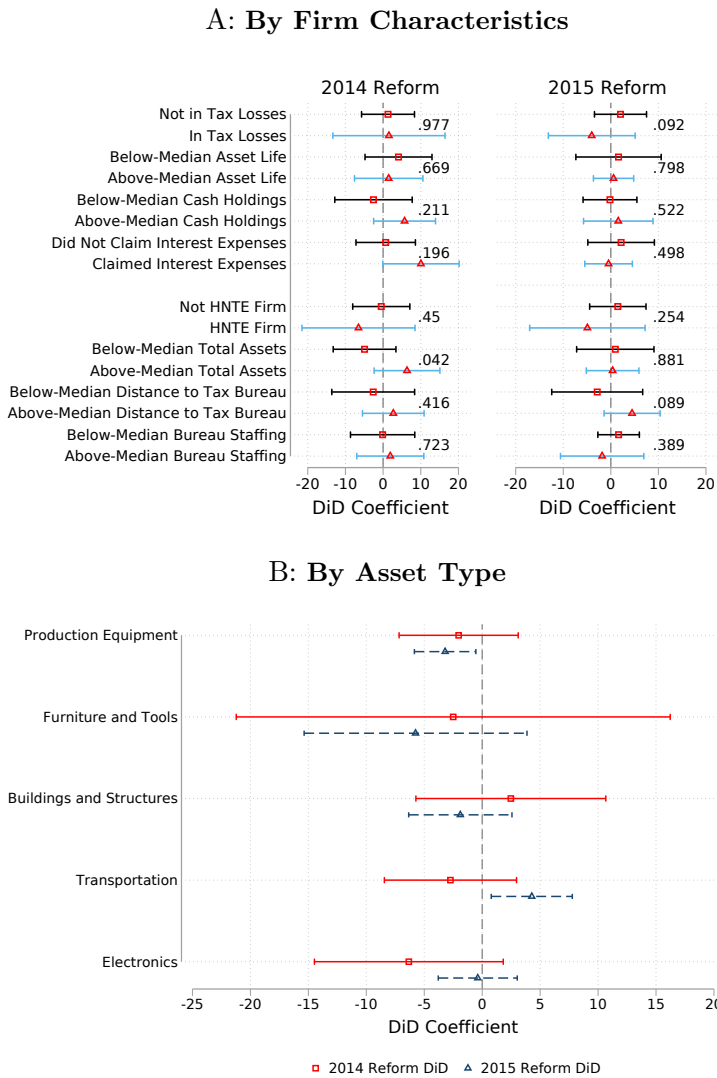
Figure A3: **Pooled Event Study**

A: **Outcome:** $Ln(K_t) - Ln(K_{t-1})$ B: **Outcome:** $Ln(K_t)$
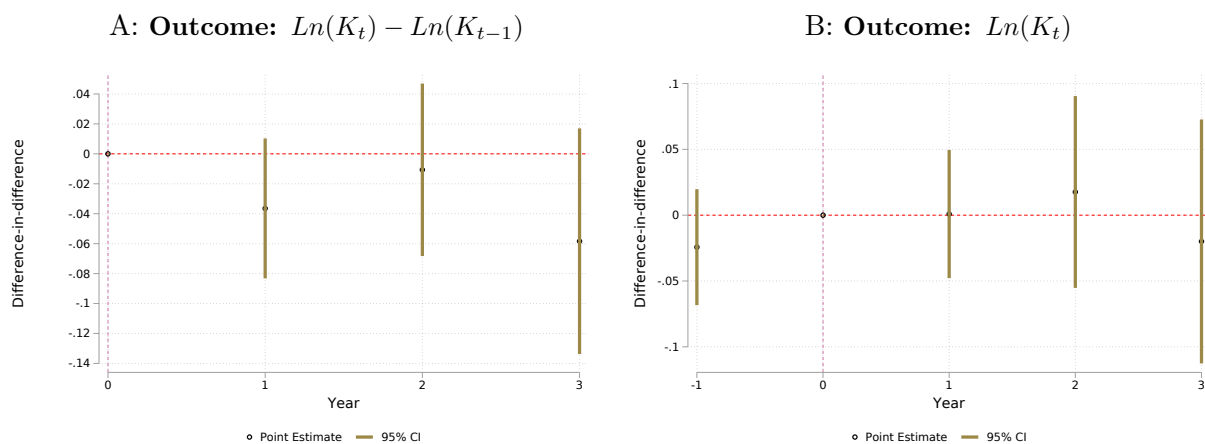
Note: This figure present estimates of β_s from equation (9) and 95% confidence intervals, using the baseline sample described in Section 3.

A.6 Additional Heterogeneity Results

Figure A4: **Heterogeneous Responses to the Reform? (Using Outcome $Ln(K_t)$)**

A: **By Firm Characteristics**


B: **By Asset Type**

Note: Panel A plots estimates of β and 95% confidence intervals from equation (3) for subsets of the sample. For each of eight different firm characteristics X, we split the sample into above and below median when X is continuous, or by $X = 0$ and $X = 1$ when X is binary. The first four characteristics proxy for a firm's readiness to invest, and the how beneficial AD is for the firm. The last four characteristics proxy for a firm's informational awareness and sophistication. P-values for the test of the treatment effects being different are displayed to the right of the confidence intervals. Panel B plots estimates of β for the full sample of firms but estimated separately for asset class. The outcome is $Ln(K_t)$. In panel A, K is the firm's fixed asset stock measured at net of accounting depreciation from financial returns. In Panel B, since net-of-depreciation measures are not observed at the asset level, K is measured at original cost from tax depreciation returns. Estimates are scaled by 100 to be consistent with the estimates reported in Table 4.

A.7 Size Heterogeneity and Very Large Firms

Figure A5 reproduces the size heterogeneity estimates from Figure 3, except using revenue rather than total assets to measure firm size. The results are very similar. Next, this section discusses the estimated treatment effect of the largest 5% firms, based on firms' pre-treatment revenue. The table at the bottom of Figure 3 shows that the largest 5% firms in the 2014 treated and matched control groups include 128 firms, and the largest 5% firms in the 2015 treated and matched control groups include 444 firms. The revenues for the smallest firms in these groups are CNY 270 million and CNY 233 million for the 2014 and 2015 treated firms, respectively. As a reference point, a manufacturing firm is officially classified as "large" in China if it has more than CNY 400 million in revenue and as "medium" if revenue lies between CNY 200 million and 400 million.

Figure A5: **Responsiveness by Pre-Reform Revenue Quartiles**

A: **Outcome:** $Ln(K_t) - Ln(K_{t-1})$ B: **Outcome:** $Ln(K_t)$
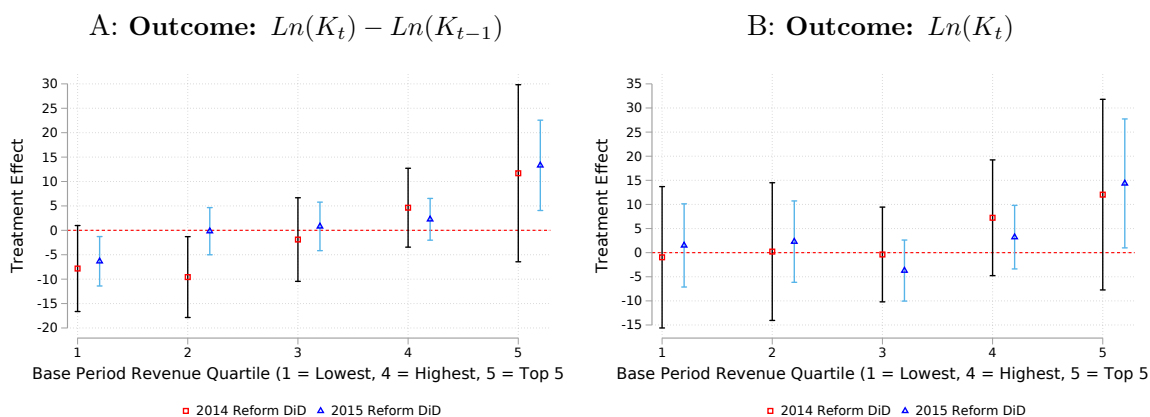
Quartile	Left Endpoints		Number of Firms	
	2014	2015	2014	2015
1	0.04	0.08	664	2295
2	4.88	4.73	672	2299
3	13.22	11.60	656	2252
4	39.81	33.90	651	2230
5	270.39	232.25	128	443

Note: Panels A and B split firms into four quartiles based on pre-treatment average revenue (1 = smallest, 4 = largest). Each then plots estimates of β from equation 3 and 95% confidence intervals for each quartile. The fifth bin (5) plots β for the largest 5% of firms. Estimates are scaled by 100 to be consistent with the estimates reported in Table 4. K_t is fixed assets net of accounting depreciation. The table lists the quartile's left-endpoints, in 1,000,000CNY, and the number of firms in each.

In Figure A6, we report estimation results for the dynamic response to AD, while controlling for firm and year fixed effects. For the 2015 AD policy, targeted firms in the top 5% show increased $Ln(K_t) - Ln(K_{t-1})$ in 2015, relative to control firms in the top 5% (Panel B). For both 2014 and 2015 cohorts, a larger divergence between the targeted and control

large firms appears in 2016. The pattern is qualitatively similar when the outcome variable is $Ln(K_t)$ (Panels C and D). Figure A6 also validates the parallel trend assumption. If we expand the regression sample to the largest 10% of firms, however, the positive investment response of the treated firms becomes smaller and loses statistical significance.

It thus appears that the largest 5% of firms in our sample of treated firms may have increased investment post AD. Figure A7 shows that these firms also have higher rates of claiming AD benefits than the full sample. However, several considerations caution against interpreting the results as purely causal. For instance, the greatest response of the 2015 targeted large firms is observed in 2015, even though AD for the 2015 treated industries was announced in September 2015. Such a quick response may not seem plausible (given adjustment costs, the speed of decision-making at large firms, etc.), and also stands in contrast with the delayed response of the 2014 treated firms. Moreover, the declining investment observed in the top 5% control firms during 2015-2016 may also be driving the positive treatment effect of top 5% treated firms.

Figure A6: **Growth of Fixed Assets Net of Depreciation**

A: **2014 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

B: **2015 Policy Change:** $Ln(K_t) - Ln(K_{t-1})$

C: **2014 Policy Change:** $Ln(K_t)$

D: **2015 Policy Change:** $Ln(K_t)$

Note: This figure restricts the sample to firms in the top 5% of the revenue distribution as of 2013. Panels A and B plot trends in $Ln(Kt) - Ln(K_{t-1})$ while Panels C and D plot trends in $Ln(K_t)$. In all four, K is fixed assets net of accounting depreciation as reported on balance sheets. Trends are plotted separately for the matched treated and control firms. The control firms are those in non-targeted manufacturing industries and matched as described in Section 3. For each panel, we estimate a dynamic version of equation (3) as $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 2013} \beta_s \times \mathbb{1}\{t = s\} \times D_i + \epsilon_{i,t,g}$. We then plot α_t for the control group and $\alpha_t + \beta_t$ for the treatment group and the associated 95% confidence intervals. The vertical red lines indicate the timing of the AD policy announcement and implementation.

Figure A7: **Comparing Full Sample Claiming Rates to Largest 5% Claiming Rates**
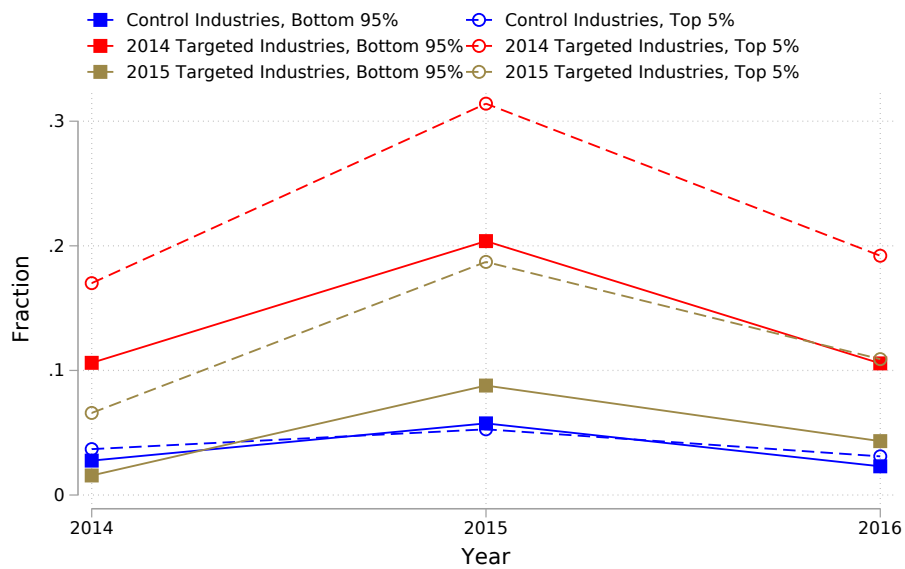
Note: This figure plots the fraction of firms, in each year, with positive AD deductions reported on their tax return, among the set of firms that had at least one year-over-year increase in fixed assets net of depreciation during the AD years. The sample is split into non-targeted industries, 2014 targeted industries, and 2015 targeted industries. Firms in these industries are then further separated into either the top 5% or bottom 95% of the firm revenue distribution.

A.8 Results from Orbis Data

Figure A8: **Growth of Fixed Assets Net of Depreciation from Orbis**

A: **Trends: 2014 Targeted Industries**   B: **Trends: 2015 Targeted Industries**

Note: This figure plots trends in $Ln(K_t) - Ln(K_{t-1})$, where K is fixed assets net of accounting depreciation, from Orbis. For each panel, we estimate a dynamic version of equation (3) as $y_{i,t} = \alpha_t + \alpha_i + \sum_{s \neq 2013} \beta_s \times \mathbb{1}\{t = s\} \times D_i + \epsilon_{i,t,g}$. We then plot α_t for the control group and $\alpha_t + \beta_t$ for the treatment group and the associated 95% confidence intervals. The vertical red lines indicate the timing of the AD policy announcement and implementation.

A.9 Comparison to Fan and Liu (2020)'s Analysis of 2014 Reform

The empirical setup in Fan and Liu (2020) differs from ours in the following ways:

1. **Data:** They use the National Taxpayer Survey Data (NTSD). Our data come from administrative tax records and financial statements for a single province. This results in differences in sample composition, policy group studied, time period covered, and outcome measures.

2. **Sample Composition:** Theirs is a national sample of mostly large and medium firms, corresponding roughly to the top size quartile of our provincial sample.

3. **Reform Studied:** Their NTSD sample ends in 2015. They therefore study only the 2014 treatment and only have one post-period. We study both the 2014 and 2015 treatments and have two post periods for the former. The treatment group in the 2014 treatment is approximately 1/5th the size of the broader 2015 treatment group in our data and in the nationwide Orbis data.

4. **Outcome Measure:** The NTSD contains reports of investment expenditure. Chinese corporate tax returns do not require taxpayers to report investment expenditures and therefore our data does not contain this measure. We instead measure changes in asset

57

stock recorded on balance sheets and tax returns. Their outcome measure is $Ln(I_t)$ and therefore they drop all observations with $I_t = 0$.[43]

5. **Control Group:** They use all non-targeted firms as the control group. This includes non-manufacturing industries. We instead use only non-targeted manufacturing firms, which are much more similar to the treated industries.

6. **Specification:** In their DiD analyses, in addition to including firm fixed effects, they interact year dummies with pre-treatment "average firm income, profit margin, and cash over asset share" in industry bins. We use firm fixed effects and match on pre-treatment covariates.

We address each of these differences below.

Sample composition: The NTSD sample is dominated by large and medium firms. Fan and Liu (2020) appear to report average sales of CNY 200 million, which would correspond to the top quartile of our sample. In addition, they report that larger firms in their sample of already-large firms responded more strongly to AD. In our heterogeneity analysis of investment responsiveness in Section 3.3 and Appendix A.7, we show that treated firms in the top quartile of our sample showed responsiveness by one measure of investment ($Ln(K_t)$), and the top 5% showed responsiveness by both $Ln(K_t)$ and $Ln(K_t) - Ln(K_{t-1})$. The different size composition of our samples, therefore, may be important to explaining the differences in our results.

Policy Studied and Timing: They study the 2014 treatment using a single post-treatment year (2015), whereas we use both 2015 and 2016 as the post-period. Panel A and C in Figure 1 show that using just 2015 as the post-period leads to a larger treatment estimate among the full sample; in the treatment group both $Ln(K_t) - Ln(K_{t-1})$ and $Ln(K_t)$ increase in 2015 relative to the control group, but then converge back towards the control group in 2016. Panel A of Figure A8 likewise shows a similar pattern using the Orbis nationwide data. This pattern would cause our treatment effects to look smaller than Fan and Liu (2020).

Difference in Outcome: Because they report estimates using investment expenditure, $Ln(I_t)$, the treatment effects are interpreted as percent changes in investment. To provide an approximate comparison to our estimates, assume that firms were in a steady state before the treatment, such that $I_{t-1} = I$ and $K_{t-1} = K$, and that S is zero. Starting from the steady state, what happens if investment increases by $X\%$? Setting $K_{t-1} = K$ and $I_t = I \times (1 + X/100)$ in equation (4) (from the asset accumulation model in Section 3) and re-arranging:

$$\frac{K_t}{K} - 1 = \frac{I}{K} \times (1 + X/100) - \gamma \tag{10}$$

---

[43]This generates potentially severe censoring.

For a given percent change in the investment level, the growth of the asset stock $\frac{K_t}{K} - 1$ depends on steady state (pre-treatment) level of $\frac{I}{K}$. Substituting K and I into equation (4) and re-arranging, one can show that $\frac{I}{K} = \gamma$ assuming $S = 0$. Therefore:

$$\frac{K_t}{K} - 1 = \gamma \times X/100 \tag{11}$$

Fan and Liu (2020) estimate that X was approximately 10 for the 2014 treatment industries in the first year after AD's implementation. If the rate of accounting depreciation is 10%, then a 10% increase in I_t relative to pre-treatment I increases K_t by approximately 1% relative to pre-treatment K. This is within the confidence intervals reported in Table 4.

Difference in Empirical Approach: For comparison purposes, we present results using our data but with their specification and control group. Figure A9 plots the estimated dynamic DiD coefficients β_s and 95 percent confidence intervals from the following specification:

$$y = \alpha_t + \alpha_i + \sum_{s \neq 2013} \beta_s \times 1\{t = s\} \times D_i + \gamma X_i \sum_{s \neq 2013} \times 1\{t = s\} + \epsilon_{i,t,k} \tag{12}$$

where α_t and α_i are fixed effects for year and firm respectively. X_i represents bins of the industry-level average pre-treatment taxable income, profit margin (business profits over business revenues), and cash over fixed assets. The Full Sample series uses all non-targeted firms in the control group including service industries. This is closest to Fan and Liu (2020)'s setup. The Manufacturing Control series restricts the control group to manufacturing industries. The third series shows our baseline result described in the paper for comparison, which excludes $\gamma X_i \sum_{s \neq 2013} \times 1\{t = s\}$ but matches firms on similar pre-treatment covariates. When using $Ln(K_t) - Ln(K_{t-1})$ as the outcome, all three approaches deliver similar results, as shown in Panel A. Panel B shows that the full-sample Fan and Liu approach produces a positive treatment effect on $Ln(K_t)$, but with a substantial increase starting in 2014, which is difficult to interpret as causal since the AD policy was not announced until fall 2014. Using the Manufacturing sample and their specification, there is positive treatment effect in 2015, consistent with their results, but the treatment effect dissipates by 2016. Fan and Liu (2020)'s estimates would only capture the 2015 spike.

Figure A9: **Comparison to Fan and Liu (2020) Approach**

A: **Outcome:** $Ln(K_t) - Ln(K_{t-1})$

B: **Outcome:** $Ln(K_t)$
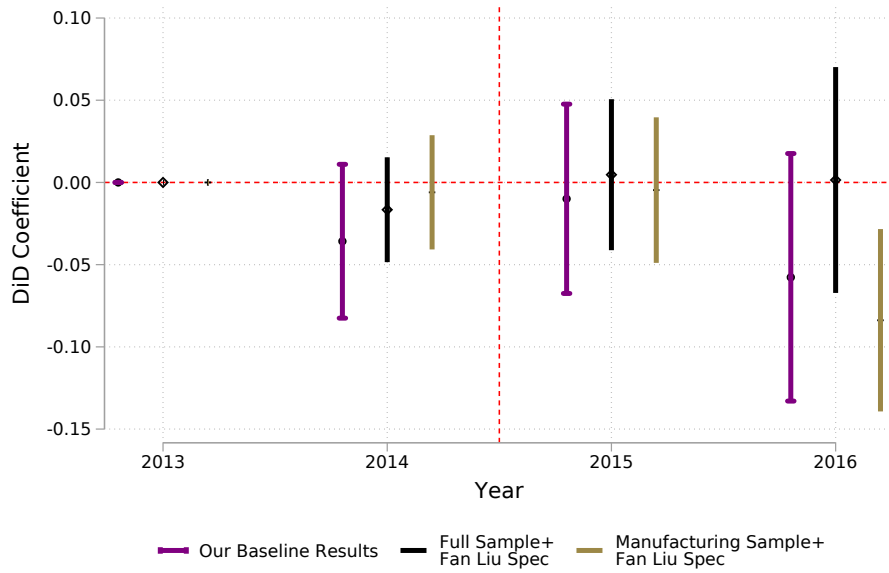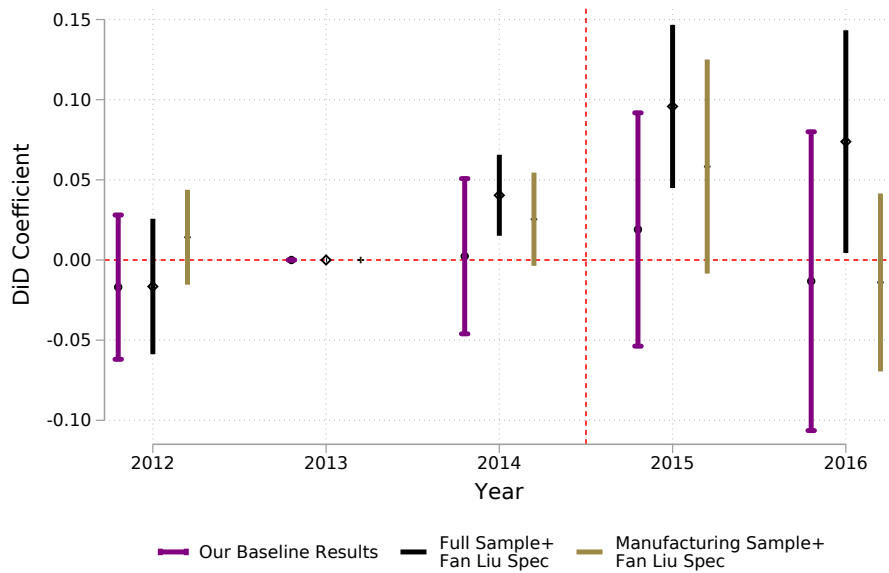
Note: This figure plots dynamic DiD estimates of β_s from equation 12 and 95% confidence intervals. See text above for details.

B   Additional Take-Up Results

Table B1: **Claiming Rate of AD by Industry**

	Claim Rate
Agricultural and sideline food processing industry	0.101
Food manufacturing	0.110
Textile industry	0.113
Textile and apparel, apparel industry	0.132
Leather, fur, feathers and their products and footwear	0.134
Wood processing and wood, bamboo, rattan, palm, grass products industry	0.063
Furniture manufacturing	0.159
Paper and paper products industry	0.118
Printing and recording media reproduction industry	0.094
Culture, education, beauty, sports and entertainment products manufacturing	0.116
Chemical raw materials and chemical manufacturing	0.160
Pharmaceutical manufacturing	0.212
Chemical fiber manufacturing	0.120
Rubber and plastic products industry	0.135
Metal products industry	0.120
General equipment manufacturing	0.120
Special equipment manufacturing	0.278
Automotive Manufacturing	0.136
Railway, marine, aerospace and other transportation equipment manufacturing	0.201
Electrical machinery and equipment manufacturing	0.152
Computer, communications and other electronic equipment manufacturing	0.275
Instrumentation manufacturing	0.300
Telecommunications, radio and television and satellite transmission services	0.164
Internet and related services	0.101
Software and information technology services	0.105
Total	0.140

Note: This table plots the percent of firms within each industry that claimed accelerated depreciation at least once between 2014 and 2016.

Table B2: **Present Value of Claimed and Unclaimed Tax Savings (1000s CNY)**

Treatment Group	Not Claimed			Claimed		
	Mean	Median	Total	Mean	Median	Total
2014 Treatment Group	75.53	3.77	176371	47.59	2.24	26937
2015 Treatment Group	57.97	4.84	543734	44.64	1.83	48117

Note: Table B2 shows estimated tax savings from accelerated depreciation (AD) for investment, separately according to whether AD was actually claimed for the investment. All values are expressed in 1,000 CNY. Investment is defined as the year-over-year change in the value of assets in asset class k ($K_{k,t} - K_{k,t-1}$). A firm is deemed to have claimed AD on the investment if the year-over-year change in AD is positive: $AD_{k,t} > AD_{k,t-1}$. Tax savings are calculated as $\tau_{i,t} \times (NPV_{k,AD} - NPV_{k,Regular})$, where $\tau_{i,t}$ is the tax rate the firm faces in year t, and the last two terms are the net present value of depreciation deductions under AD and regular depreciation, respectively, for asset class k. This calculation assumes (1) that firms keep the asset for the entire length of its useful tax life, (2) that firms are in taxable positions during the entire duration of the asset's life, (3) that the firm's tax rate does not change, and (4) if a firm did not claim AD on positive investment in year t, we assume it would not claim in later years. We restrict to firms in a taxable position in the year of investment. The base tax rate $\tau_{i,t}$ is calculated as firm i's income tax liability over its observed taxable income in the year of investment (t). This is 25% unless the firm is eligible for statutory reductions for SMPE firms ($\tau_{i,t} = 10\%$ or 20%) and HNTE firms ($\tau_{i,t} = 15\%$). Immediate expensing is assumed for imputed investment of value less than 5,000 CNY such that $NPV_{k,AD} = K_{k,t} - K_{k,t-1}$. The mean and median are calculated at the firm level after summing tax savings across asset types for each firm. The total sums the tax savings across all firm-asset class-year observations. Only observations in the treated years are included; 2014 to 2016 and 2015 to 2016 for the 2014 and 2015 treatment groups, respectively. Tax savings values are winsorized at the 1st and 99th percentiles for each treatment group before calculating the means, medians, and total amounts.

Table B3: **Correlations between Tax Administration Variables and Tax Compliance Proxies**

|	(1) Log(Firms / Bureau Staff)	(2) Log(Distance to Bureau)
Statutory Rate minus Cash ETR	-0.066	-0.049
	(0.047)	(0.044)
Fixed Assets / Revenue	-0.009	0.002
	(0.038)	(0.035)
Business Profits / Revenue	-0.229**	0.285***
	(0.095)	(0.094)
N	15146	14637
Three-Digit Industry FE	Yes	Yes
Prefecture FE	Yes	Yes
Mean STR - ETR	0.01	0.01
Mean (Fixed Assets / Revenue)	0.43	0.42
Mean (Business Profits / Revenue)	-0.02	-0.02

Note: This table provides additional results related to the concern that the distance between the firm and the tax office, and the ratio of firms to tax bureau staff, may be correlated with firms' tax non-compliance. We examine the correlation between the tax administration variables and three proxies for firms' non-compliance tendency. Our conjecture is that firms whose effective tax rates (ETRs) are much lower than the statutory rates may be less compliant; firms with more tangible fixed assets may be more compliant since it is easier for the tax bureau to verify their assets; and firms with a higher ratio of profits to revenue should be more compliant. We do not find the gap between ETR and statutory rate, or the ratio of fixed assets to total assets, is significantly associated with our two tax administration variables. While we find that firms located in areas where the ratio of firms to tax bureau staffs is higher report lower profit/revenue, those located further away from the tax bureau report higher profit/revenue. Overall, we do not find systematic evidence that firms located further away from tax offices, or firms facing a higher ratio of firms to tax bureau staffs, are more likely to avoid tax.

Table B4: **Predictors of Claiming: Accounting Resources**

|	Pr(Claim | Invest)			
	(1)	(2)	(3)	(4)
Log(# Accountants / # Workers)	-0.42			0.18
	(0.44)			(0.51)
Ln(# Accountants / # Firms)		-0.68***		-0.69**
		(0.25)		(0.29)
Ln(# Accounting Firms / # Firms)			-0.01	-0.32
			(0.42)	(0.44)
Ln(# Firms / # Tax Administrators)	-1.65***	-1.97***	-1.89***	-2.00***
	(0.35)	(0.34)	(0.35)	(0.38)
Log(Distance to Tax Bureau)	-0.63**	-0.83***	-0.76***	-0.83***
	(0.30)	(0.32)	(0.26)	(0.32)
Firm Characteristics	Yes	Yes	Yes	Yes
Two-Digit Industry FE	Yes	Yes	Yes	Yes
Prefecture FE	Yes	Yes	Yes	Yes
N	24555	24555	34374	24555
Claim Rate	6.30	6.30	6.85	6.30

Note: This table reports the estimated average marginal effects of regional accounting resources on the probability of claiming AD conditional on having purchased eligible investment using the probit model described in Section 5. Claiming and investment are defined as in Table 5. The measures of accounting resources are described in Section 5.3. Coefficients are scaled by 100. All continuous covariates are winsorized at the 1st and 99th percentiles within each year for each of the two treatment groups. Standard errors are computed using the delta-method and are clustered at the three-digit industry level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure B1: **New AD Claims as % of Total Claims**
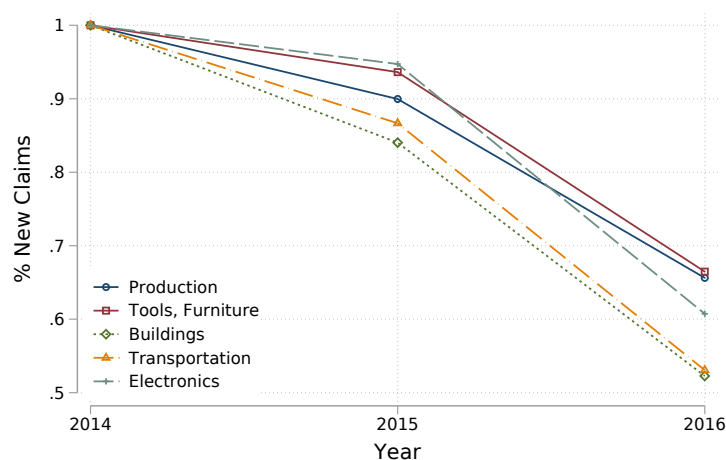
Note: This figure plots the percentage of first-time AD claims in total claims by asset type during the period 2014-2016.

Figure B2: **Distribution of Investment by Claimed Status**
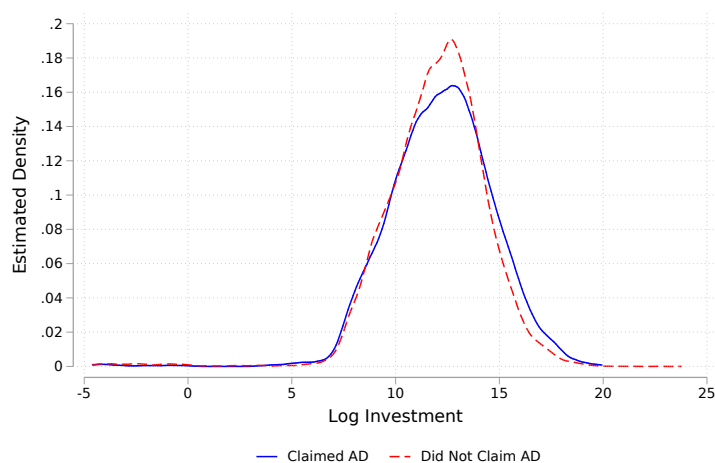
Note: This figure plots the distribution of investment (in logs) separately by whether AD was claimed on that investment. Investment is defined as the year-over-year change in the value of assets in asset class k ($K_{k,t} - K_{k,t-1}$). Observations with $K_{k,t} - K_{k,t-1} \leq 0$ are excluded. A firm is deemed to have claimed AD on the investment if $AD_{k,t} > AD_{k,t-1}$. The density is estimated using the Epanechnikov kernel and a bandwidth of .23 and .43 for the unclaimed and claimed distributions respectively.

Figure B3: **Tax Savings Across the Firm Size Distribution**
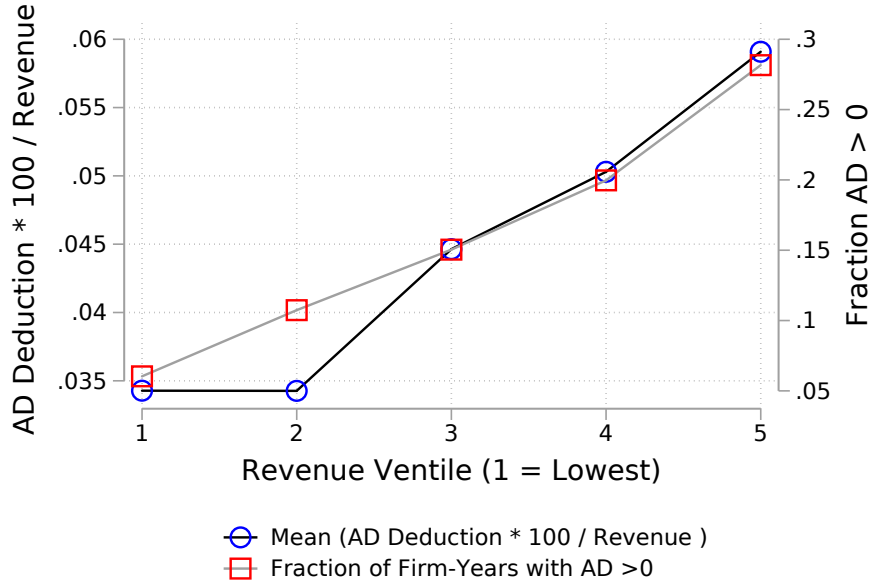
Note: This figure plots $\frac{1}{N}\sum_{i,t} AD_{i,t} \times 100/\text{Revenue}_{i,t}$ in each firm revenue ventile, to illustrate the size-gradient of AD tax deductions.

Figure B4: **Distributions of Tax Bureau Characteristics**

A: **Distance to Tax Bureaus**   B: **Firms Per Tax Administrator**
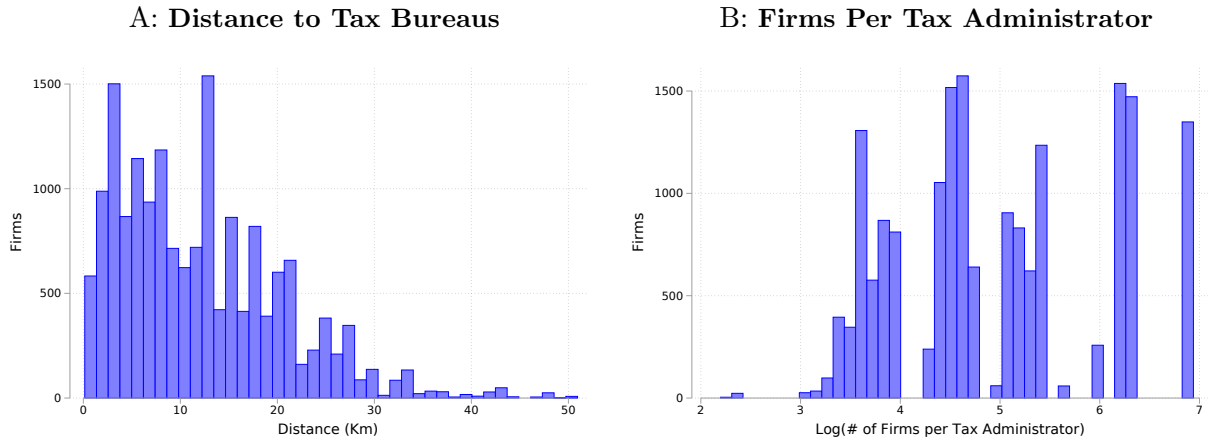
Note: These figures plot the distributions in 2013 of firms' distance to their local tax bureau and the number of firms the bureau oversees divided by the bureau's number of employees. Only firms in the targeted industries are kept.

66

C Selection Model

This section provides further details for the selection model specification used in Section 4. In this model, a firm in time t chooses to invest in assets of type k if their latent payoff function $(U_{i,k,t})$ is positive. The firm claims AD on that investment if the net benefit of doing so $(NB_{i,t,k})$ is positive. Both $U_{i,k,t}$ and $NB_{i,t,k}$ are functions of observables $X_{i,t}$:

$$I_{i,k,t} = 1 \text{ if } U_{i,k,t} = \beta X_{i,t} + \eta_{i,k,t} > 0 \tag{13}$$

$$C_{i,k,t} = 1 \text{ if } I_{i,k,t} = 1 \text{ and } NB_{i,t,k} = \gamma X_{i,t} + \epsilon_{i,k,t} > 0 \tag{14}$$

We are interested in how covariates $X_{i,t}$ predict the likelihood of claiming AD conditional on having eligible investment: $\frac{\partial P(C_{i,k,t}=1|I_{i,k,t}=1)}{\partial X_{i,t}}$. If we simply restrict to firms with investment $(I_{i,k,t} = 1)$ and correlate take-up decisions with $X_{i,t}$, selection bias arises if the idiosyncratic errors $\epsilon_{i,k,t}$ and $\eta_{i,k,t}$ are correlated (Heckman, 1979). As an example, assume that firm size increases the latent payoff of investing, and that the error terms are positively correlated. In this case, only small firms with idiosyncratically higher investment payoffs will invest. Due to the positive error correlation, these firms will also be more likely to claim, making small firms appear more likely to claim as well.

To account for this, we model the distribution of the error terms and their correlation. Error terms $\epsilon_{i,k,t}$ and $\eta_{i,k,t}$ are assumed to be jointly normally distributed allowing for non-zero co-variance. We then estimate β and γ by maximum likelihood. The resulting log likelihood contribution for firm i, in year t, in asset-class k can be written as

$$
\begin{aligned}
ln(L_{i,k,t}) &= (1 - I_{i,k,t})ln[P(I_{i,k,t} = 1)] + I_{i,k,t}(1 - C_{i,k,t})ln[P(I_{i,k,t} = 1, C_{i,k,t} = 0)] \\
&\quad + I_{i,k,t} \times C_{i,k,t} \times ln[P(I_{i,k,t} = 1, C_{i,k,t} = 1)] \\
&= (1 - I_{i,k,t})ln[\Phi(-X_{i,t}\beta)] + I_{i,k,t}(1 - C_{i,k,t})ln[\Phi(X_{i,t}\beta) \\
&\quad - \Phi(X_{i,t}\beta, X_{i,t}\gamma, \Omega)] + I_{i,k,t} \times C_{i,k,t}ln[\Phi(X_{i,t}\beta, X_{i,t}\gamma, \Omega)]
\end{aligned}
$$

where $I_{i,k,t}$ and $C_{i,k,t}$ are indicators for investing and claiming as defined in Section 5.1.2, Ω is the covariance matrix of $\epsilon_{i,k,t}$ and $\eta_{i,k,t}$ as in Equations (4) and (5), and Φ the cumulative distribution function for the normal distribution.

Omitting subscripts for simplicity, $Pr(C = 1 \mid I = 1)$ can be written as follows:

$$
\begin{aligned}
Pr(C = 1 \mid I = 1) &= P(X\delta + \eta_2 > 0 \mid X\beta + \eta_1 > 0) \\
&= \int_{-X\beta}^{\infty} \Phi((X\delta + \eta_1\Omega_{1,2}/\Omega_{2,2})/\sigma_{2|1})\phi(\eta_1; \sigma_1)d\eta_1 \Big[\frac{1}{\Phi(X\beta/\sigma_1)}\Big]
\end{aligned}
$$

And differentiating with respect to X, we have:

$$\frac{dPr(C \mid I)}{dX} = \left[\Phi((X\delta - X\beta\Omega_{1,2}/\Omega_{2,2})/\sigma_{2|1})\phi(-X\beta; \sigma_1)(\frac{dX\beta}{dX}) + \right.$$

$$\int_{-X\beta}^{\infty} \frac{d\Phi((X\delta - X\beta\Omega_{1,2}/\Omega_{2,2})/\sigma_{2|1})}{dX} \phi(\eta_1)d\eta_1 \left] \left[\frac{1}{\Phi(X\beta/\sigma_1)} \right] - \right.$$

$$\left[\int_{-X\beta}^{\infty} \Phi((X\delta + \eta_1\Omega_{1,2}/\Omega_{2,2})/\sigma_{2|1})\phi(\eta_1; \sigma_1)d\eta_1 \right] \left[\frac{\phi(X\beta/\sigma_1)(\beta/\sigma_1)}{\Phi(X\beta/\sigma_1)^2} \right]$$

To aid in the identification, we include three variables in the investment $(U_{i,k,t})$ equation which are excluded from the claiming equation: the firm's lagged cash holdings, an indicator for whether a firm claimed interest expenses on their prior year's tax form, and the lagged growth rate of their fixed asset stock. Lagged cash holdings and interest expenses proxy for the ability to finance investments in year t, but should not affect the payoff of claiming AD. We include the firm's prior year measure of $Ln(K_t) - Ln(K_{t-1})$ as investment has been shown to have some degree of serial correlation (Cooper and Knittel, 2006).[44] In column (1) of Table C1, we report the estimated marginal effects on the probability of investment in the current year. The dummy variable indicating interest expense deductions, firms' lagged cash holdings, and the firm's prior year's $Ln(K_t) - Ln(K_{t-1})$ all positively predict current investment and are jointly significant at the 1 percent level, partially validating them as effective excluded variables. The last two columns of Table C1 show the average marginal effects of each predictor on the probability of claiming conditional on investing. As described in the main text, results are nearly identical to a simple probit model.

---

[44]The lagged measure of $Ln(K_t) - Ln(K_{t-1})$ would violate the exclusion restriction if we defined claiming as $AD(t) > 0$, as prior year's investments can entail AD claiming into the present. Defining claiming as $AD(t) > AD(t+1)$ controls for the concern.

Table C1: **Predictors of Claiming: Firm-Level Characteristics**

	Selection Model		
	(1)	(2)	(3)
	Pr(Invest)	Pr(Claim | Invest)	
In Tax Loss Before AD	-7.11***	-2.69***	-2.73***
	(0.32)	(0.45)	(0.43)
Ln(Tax Loss Stock $_{t-1}$)	-0.20***	-0.09***	-0.08**
	(0.03)	(0.03)	(0.03)
Ln(Total Assets $_{t-1}$)	4.21***	0.39***	0.42***
	(0.21)	(0.15)	(0.14)
High and New Technology Enterprise	7.04***	2.27***	1.83***
	(0.62)	(0.61)	(0.59)
$Ln(K_{t,k} - K_{t-1,k})$		0.12***	0.09**
		(0.04)	(0.04)
Interest Deduction $_{t-1} > 0$	2.41***		
	(0.34)		
$Ln(K_{t-1}) - Ln(K_{t-2})$	3.78***		
	(0.38)		
Ln(Cash Holdings$_{t-1}$)	1.04***		
	(0.11)		
Two-Digit Industry FE	Yes	Yes	Yes
Prefecture FE	Yes	No	Yes
N	171846	171846	171846
Mean of Outcome Var.	28.32	6.36	6.36
Excluded Variables Joint P-value	0.000	0.000	0.000

Note: Column (1) reports the estimated average marginal effects of firm-level characteristics on $Pr(I_{i,t,k} = 1)$. Columns (2) and (3) report the estimated average marginal effects of firm-level characteristics on the probability of claiming AD conditional on having purchased eligible investment ($P(C_{i,t,k} = 1|I_{i,t,k} = 1)$). Claiming and investment are defined as outlined in Section 4. Coefficients are scaled by 100. All time-varying explanatory variables are measured in year $t-1$. Ln(Tax Loss Stock $_{t-1}$) is the natural logarithm of (Tax Loss Stock $_{t-1}$ +1) to account for zeros. All continuous covariates are winsorized at the 1st and 99th percentiles within each year. Standard errors are computed using the delta-method and are clustered at the three-digit industry level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$